

V.R.Sikkes

Estimating arrival times of inland ships



Estimating arrival times of inland ships

-

By

V.R.Sikkes

in partial fulfilment of the requirements for the degree of

Master of Science
in Transport Engineering

at the Delft University of Technology,
to be defended publicly on Tuesday February 19, 2019 at 2:00 PM.

Supervisor: ir. MB Duinkerken
Thesis committee: Prof.dr. RR Negenborn, TU Delft

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

Specialization: Transport Engineering and Logistics

Report number: 2019.TEL.8314

Title: **Estimating arrival times of inland ships**

Author: V.R.Sikkes

Title (in Dutch) Schatting van de aankomsttijden van binnenvaartschepen

Assignment: Masters thesis

Confidential: no

Supervisor: ir. MB Duinkerken

Date: February 19, 2019

Summary

Inland shipping is an important part of the Dutch economy, 34% of all transport is done by inland shipping. Just like trucks have uncertain travel times due to possible traffic jams and road works, inland ships have uncertain travel times due to locks, water levels and other ships. For both modalities, Rijkswaterstaat is the organization which controls and maintains the Dutch infrastructure. Rijkswaterstaat is part of the Dutch Ministry of Infrastructure and Environment and is responsible for the design, construction, management and maintenance of the main infrastructure facilities in the Netherlands. Which include among others the inland waterways and the management of inland shipping. The management of inland shipping is now done from many locations where objects and sectors are being operated individually¹.

Rijkswaterstaat wants to centralize the management of inland shipping and has started the project CBB (Corridor gericht Bedienen en Begeleiden) which, when translated, stands for Corridor oriented Operating and Guiding. The goal of this project is to centralize the management of inland shipping. The inland waterway system is divided into corridors. These corridors represent important shipping routes. The idea of CBB is that a ship can sail over a corridor faster and more efficient than it would do today due to better planning of objects and a better knowledge of the situation on the waterway. In order to achieve this, the CBB project will centralize the management of the inland waterways. The goal is to have 1 location for each corridor from where all the objects and sectors are being operated. Next to centralizing the operation, the CBB project is also creating and updating IT systems in order to implement corridor management. These systems help the operators with their jobs and facilitate the skipper in giving him better service and better information. For example, a new system is being developed to track ships. Other systems, like data systems about water levels and waterway characteristics are being updated. Another new system is the Trajectplanner. This system is a tool which helps both the skipper and Rijkswaterstaat with managing inland shipping. It does this by using data from other systems and using that data to simulate the entire journey the ship will be making. With this simulated data it makes schedules for locks and bridges which the ship will be passing and will make prognoses of the occupancy rate of the inland waterways. With these prognoses and schedules, the skipper can adapt his speed and/or route in order to sail more efficient.

For the scheduling and prognoses to be made, the Trajectplanner calculates and estimates ETA's (Estimated Time of Arrival) of the ships it has in the system. These ETA's are calculated for each point and ship in the system. The Trajectplanner uses these ETA's to schedule ships into the lock procedures and to let operators know which ships will arrive at certain points in time. These ETA's are therefore very important for the correct functioning of the Trajectplanner. For the Trajectplanner, the Dutch waterway network was modelled with dots and lines which are similar to a network of nodes and arcs. These arcs represent part of the waterway and the nodes represent points on the waterway. When a ship is starting his journey, the Trajectplanner predicts the route the ship will take. This route is predicted along the nodes and arcs. Each node the ship passes has an ETA (Estimated Time of Arrival) for that ship. These ETA's are calculated/predicted by the Trajectplanner. To calculate the ETA for the next node, the Trajectplanner uses the ETA it calculated for the previous node and adds the cost of the arc between the nodes. For arcs without objects on them, the cost consists of is the sailing time

¹ Objects are bridges and locks. Sectors in the context of this document mean the VTS (Vehicle Traffic Service) sectors that are managed by Rijkswaterstaat. These sectors are areas on the water where the traffic (Inland ships and big recreational vessels) is guided through the sector. These sectors are often either busy and/or are on a junction between waterways.

for that arc which depends on the type of ship, the type of waterway and if the ship is loaded or not. For arcs with an object on them, the cost consists of the sailing time for that arc plus the time it takes to pass the object. As the Trajectplanner uses the ETA from the previous node to calculate the ETA for the next node, it is important that these ETA calculations are accurate. This is not yet the case and therefore this research focused on the cause behind these inaccuracies and finding ways to improve these calculations and simulations.

The analysis of this problem was done in two ways. First a literature review was performed in order to see if any research was done at similar problems. Also, a search was conducted to other systems that had the same goal. The literary review revealed reports which used the same type of model (network of arcs and nodes) for similar purposes, but also literature was found about other modalities which also dealt with uncertain arrival times. These were however not applicable to this problem because the time scale at which the reports worked was either much smaller (seconds) or much larger (hours). The review did however give inspiration to tackle the problems of this report about how to model the Dutch inland waterway system and how to define the accuracy of a simulation system. The second analysis was performed on the current system, the Trajectplanner. This analysis focused on the workings behind the Trajectplanner and on the output of this system. The conclusion of this analysis was that the locks in the simulation system caused big errors to occur in the calculation of ETAs. The combination of the two analysis revealed that the error caused by the locks was caused by the way the Dutch waterway network and its objects are modelled in the Trajectplanner. Designing a new network model became therefore one of the improvements suggested in this report. Other improvements were a new lock model and a memory for the Trajectplanner as the simulation which the Trajectplanner performs is done every 6 minutes without knowledge of previous simulations. To test these improvements, a simulation environment was build where ships were generated and sailing from two sides towards and past a lock. The movements of these ships and the lock procedures were simulated here. Within this simulation environment, an estimator was build which estimated the ETA of all ships at different locations. As this estimator did the same job as the Trajectplanner, the effectiveness of the improvements were tested within the estimator component.

In an ideal situation, the ETAs the Trajectplanner gives are constant and don't change between simulation runs. As mentioned above, this is not yet the case. After the analysis that were performed on the system itself and literature, it was concluded that the main problem lies in the network model that is used by the Trajectplanner. The way the network is modelled is in such a way that it is not truly a node an arc network. This caused the biggest problem, namely the Trajectplanner not being able to know when ships are near, in or past a lock. The first improvement addressed this by modifying the network to a true node and arc network. This was done by adding extra nodes before and after the lock and making the lock an arc by itself. This caused the estimator to know when ships were past the lock and improved the KPI. The next step was to add an extra attribute to the ships which told the Estimator at which time the ship went into the lock. This 'lock awareness' gave the Estimator extra information about when the ships would be sailing out of the lock again. The validation and improvement experiments were all done with the same set of ships. This meant that the travel time of the ships was the same for every experiment. The resulting graphs that were the output of the simulation tool were therefore easy to compare. The goal of the improvements was to make the ETA calculations more constant than the Trajectplanner does now. This is something which was seen in the resulting graphs. This was also seen in the KPI's from these experiments. The mean value over all the ships didn't always improve from experiment to experiment, but the median did. It was also seen that both improvement ideas did improve the KPI. This shows that the network model which is used by the Trajectplanner could be changed to improve the simulation and calculations. It also shows that an extra attribute to a ship could further improve the simulation and calculations.

The simulation tool for this research was a very simplified version of the Trajectory planner. But even this simplified version shows improvement with the new ideas. It is recommended that, to further research and improve the Trajectory planner, a more detailed simulation tool is made, or a tool which could be implemented inside the Trajectory planner to monitor and test improvements real-time. Other improvements which could be tested would be a memory for the Trajectory planner, especially for the schedules it makes for objects. A suggestion would be to not change or make a new schedule if all ships in that schedule are still able to make the RTA. This way, the skipper can use the RTA. Another recommendation would be to have the Trajectory planner have a built-in KPI checker which can monitor the system on its KPI but also be used as a warning for something to go wrong. For example, when the ETA calculation suddenly and unexpectedly keeps rising, it might signal something to be wrong.

Summary (in Dutch)

De binnenvaart is een belangrijk deel van de Nederlandse economie, 34% van alle getransporteerde goederen gaat via de binnenvaart. Net zoals vrachtwagens last hebben van files en weg opbrekingen, hebben binnenvaartschepen last van lage waterstanden en sluisen. Dit zorgt in beide gevallen voor onzekere reistijden. Voor beide modaliteiten is Rijkswaterstaat de organisatie die zorgt voor de aanleg, onderhoud en beheer van de infrastructuur in Nederland. Voor de binnenvaart betekent dit de binnenwateren en de aansturing van de binnenvaartschepen. De aansturing van de binnenvaart wordt nu gedaan vanaf veel verschillende locaties waar objecten en sectoren individueel beheert/aangestuurd worden².

Rijkswaterstaat wil het beheer en aansturing van de binnenvaart centraliseren en is daarvoor een groot project gestart genaamd CBB(Corridor gericht Bedienen en Begeleiden). Het binnenwater netwerk is opgedeeld in corridors. Deze corridors zijn gelinked aan belangrijke scheeproutes. Het idee achter CBB is dat een schip makkelijker over een corridor kan varen door betere informatie en een beter planning. En om tot dit doel te komen streeft het CBB project om 1 corridor te beheren en aan te sturen vanuit 1 centrale locatie. Naast de centralisatie van de operatie is CBB ook bezig met het updaten en vernieuwen van IT systemen om zo corridormanagemnt te implementeren. Deze systemen zullen gebruikers (begeleiders, sluismeesters etc.) helpen bij hun taak en zullen zorgen voor een betere service naar de schipper. Er wordt bijvoorbeeld een nieuw systeem ontwikkelt om schepen te volgen. Andere system, zoals data systemen over waterhoogte en karakteristieken van de waterwegen, worden geüpdate. Een ander nieuw systeem is de Trajectplanner. Dit systeem helpt zowel de schipper als Rijkswaterstaat. De Trajectplanner doet dit door gebruik te maken van de data uit andere system en deze data te gebruiken om de bewegingen van een binnenvaartschip in het Nederlandse binnenwaterennetwerk te simuleren. Met deze gesimuleerde data kan het systeem planningen maken voor de bruggen en sluisen om zo te berekenen hoe lang een schip over z'n reis zal doen. Ook helpt het Rijkswaterstaat te voorspellen wanneer het druk zal worden op de waterwegen. Met deze planning kan een schipper zijn snelheid of route aanpassen en zo efficiënter varen.

De Trajectplanner berekent en schat de ETA's (Estimated Time of Arrival) van alle schepen op het netwerk. Deze ETA's worden berekend voor elk punt en elk schip in het system. De Trajectplanner gebruikt deze ETA's om schepen bij sluisen in te plannen en om begeleiders te laten weten welke schepen eraan komen. Deze ETA's zijn daarom belangrijk voor het functioneren van de Trajectplanner. Het netwerk van de Nederlandse binnenwateren dat de Trajectplanner gebruikt bestaat uit punten en lijnen en lijkt erg op een netwerk van knopen en zijden. De zijden in dit netwerk beschrijven de waterwegen en de knopen beschrijven punten op de waterweg. Wanneer een schip z'n reis start zal de Trajectplanner de route voorspellen die het schip zal nemen. Deze route zal lopen over het netwerk van knopen en zijden en elke knoop die het schip zal passeren krijgt een ETA voor dat schip. Deze ETA's worden berekend door de Trajectplanner. Om de ETA voor een volgende knoop de bereken gebruikt de Trajectplanner de ETA van de vorige knoop en telt daar de reistijd van

² Objecten staat voor de kunstwerken zoals bruggen en sluisen. Sectoren, in de context van dit document, staan voor de VTS(Vehicle Traffic Service) sectoren die beheert worden door Rijkswaterstaat. Deze sectoren zijn gebieden op het water waar schepen (Beroepsvaart en grote recreatie vaart) begeleid worden door de sector. Deze sectoren liggen vaak op drukke delen van de waterweg en bij kruisingen.

de zijde tussen de knopen bij op. De reistijd hangt af van het type schip en of het geladen is of niet. Ook hangt het ervan af of er een object op de zijde ligt. Een object zorgt voor extra reistijd omdat het tijd kost om een object te passeren. Bij een brug kan de passage tijd nul zijn als het schip eronderdoor kan, maar een sluis zorgt vrijwel altijd voor extra reistijd. Omdat de Trajectplanner de ETA van vorige knoop gebruikt om de ETA voor de volgende knoop te berekenen is het belangrijk dat de ETA berekeningen nauwkeurig zijn. Dat zijn ze op dit moment niet en daarom ligt de focus van dit onderzoek op de analyse van deze onnauwkeurigheid en de verbetering van de ETA berekeningen.

De analyse van dit probleem is op twee manieren gedaan. Eerst is er een literatuur onderzoek gedaan naar onderzoeken van vergelijkbare problemen of onderzoeken met vergelijkbare modellen. Ook is gezocht naar vergelijkbare systemen. Het literatuur onderzoek wees uit dat er onderzoeken gedaan zijn met vergelijkbare modellen en ook met vergelijkbare problemen. Ook is er literatuur gevonden vanuit andere modaliteiten waar onzekere aankomsttijden een rol speelde. Deze waren niet altijd toepasbaar op dit probleem omdat ze of met een te kleine tijdschaal werkte (seconden) of met een te grote (uren). Het literatuur onderzoek heeft wel inspiratie gegeven over hoe het netwerk te modelleren en hoe de nauwkeurigheid van het systeem te definiëren. Het tweede deel van het onderzoek is gedaan naar de huidige Trajectplanner en de data die eruit komt. De conclusie die uit deze analyse kwam was dat de sluisen in het systeem grote onnauwkeurigheden in de berekeningen veroorzaakte. De combinatie van de twee analyses liet zien dat de onnauwkeurigheid die kwam van de sluisen deels te maken had met de manier waarop het Nederlandse binnenwater netwerk was gemodelleerd. Het ontwerpen van een nieuw netwerk model was daarom een van de verbeteringen die uit de analyse kwam. Andere verbeteringen die voorgesteld werden, waren een nieuw sluis model en een geheugen voor de Trajectplanner, de Trajectplanner begint namelijk elke 6 min volledig opnieuw zonder kennis van vorige simulaties. Om deze verbeteringen te testen is een simulatie omgeving gecreëerd waarin schepen van de ene naar de andere kant varen. In deze reis komen ze een sluis tegen waar ze doorheen moeten. Binnen de simulatie omgeving was ook een schatter gebouwd die moest schatten wanneer de schepen bij bepaalde punten aankwamen. De schatter had dezelfde taak als de Trajectplanner. De nauwkeurigheid van deze schatter kon worden getest om zo verbeteringen uit te testen.

In een perfecte situatie zal de Trajectplanner ETA's geven die elke simulatie run weer hetzelfde zijn. Zoals hierboven beschreven is dat nog niet het geval en leek het grootste probleem veroorzaakte te worden door het netwerk model dat de Trajectplanner gebruikt. Dit model is namelijk geen echt knopen en zijden netwerk. Dit zorgde ervoor dat de Trajectplanner niet precies wist wanneer een schip vlakbij, in of voorbij de sluis is. De eerste verbetering lost dit op door het netwerk aan te passen naar een echt knopen en zijden netwerk, waarbij zijden een stuk van de waterweg zijn, of een object. Dit zorgt ervoor dat de Trajectplanner beter weet wanneer een schip in of voorbij een sluis is. De volgende stap was om de schepen een extra eigenschap mee te geven welke de schatter vertelde wanneer het schip de sluis in was gevaren. OP basis hiervan kon de schatter beter berekenen wanneer het schip er weer uit zou varen. Het valideren en testen met de simulatie omgeving werd gedaan met dezelfde set schepen die dezelfde reis maken en dus elke keer dezelfde reistijd hebben. De resultaten waren hierdoor goed te vergelijken. Het doel was om de ETA output constanter te krijgen dan wat de Trajectplanner nu levert. Dit is is gelukt en dat kon gezien worden in zowel de grafieken en de KPI die uit de experimenten kwamen. De gemiddelde KPI over alle schepen verbeterde niet bij elke stap, maar de middelste waarde wel. Ook zorgde elke verbetering voor een verbetering van de KPI. Dit laat zien dat een nieuw netwerk model en een extra eigenschap van een schip de Trajectplanner kunnen verbeteren.

De simulatieomgeving die was gebouwd voor dit onderzoek was een erg versimpelde versie van de Trajectplanner. Maar zelfs deze versimpeling laat zien dat de nieuwe ideeën voor verbetering kunnen zorgen. Om de Trajectplanner verder te verbeteren en te onderzoeken wordt het aanbevolen om een meer gedetailleerde simulatie te bouwen, of dat er in de huidige Trajectplanner een manier komt om nieuwe ideeën te testen. Andere verbeteringen zouden bijvoorbeeld een geheugen voor de Trajectplanner zijn die ervoor kan zorgen dat zolang alle schepen een planning kunnen halen, er geen nieuwe planning gemaakt wordt. Hiermee kunnen schipper hun reis dan gaan aanpassen. Een andere idee is om een ingebouwde evaluatie tool te maken zodat de ATA's en ETA's achteraf vergeleken kunnen worden. Ook is het hiermee mogelijk om problemen eerder te signaleren als bijvoorbeeld de ETA van een schip onverwachts blijft stijgen.

List of abbreviations

ETA – Estimated Time of Arrival

RTA – Requested Time of Arrival

ATA – Actual Time of Arrival

CBB – Corridor gericht Bedienen en Begeleiden

VTS – Vehicle Traffic Service

Contents

Summary	7
Summary (in Dutch)	11
List of abbreviations	15
1 Introduction	19
1.1 Inland shipping	19
1.2 ETA's	24
1.3 Problem description	25
1.4 Research goal + question	25
1.5 Research approach	26
2 Ship movement modeling and ETA prediction in literature	27
2.1 Similar systems	27
2.2 Similar models	28
2.3 Inaccuracy in prediction	30
2.4 Conclusion	32
3 KPI's	35
3.1 KPI's	35
3.2 Literature	36
3.3 Users	36
3.4 Proposed KPI	36
3.5 Conclusion	38
4 Current prediction and planning system	39
4.1 Trajectplanner	39
4.2 Estimating the ETA	40
4.3 Conclusion	41
5 KPI of the Trajectplanner system	43
5.1 System output	43
5.2 Analysis of the output	44
5.3 Results	50
5.4 Accuracy of the current ETA estimation (Trajectplanner)	51
5.5 Influence of a lock	52
5.6 Analysis conclusion	54
6 Alternative models for ETA prediction	55
6.1 Inaccurate ETAs	55
6.2 Network model	57

6.3	Lock model	62
6.4	Memory.....	62
6.5	Conclusion.....	63
7	Evaluation method	65
7.1	Programming language and modules.....	65
7.2	Simulation environment	65
7.3	Output.....	72
7.4	Conclusion.....	72
8	Evaluations.....	73
8.1	Experimental setup.....	73
8.2	Verification of the simulation.....	76
8.3	Validation of the simulation.....	77
8.4	Experimental plan.....	82
8.5	Design 1	82
8.6	Design 2	85
8.7	Conclusion.....	88
9	Conclusion	91
9.1	Subquestions.....	91
9.2	Main question	92
9.3	Overall conclusion	92
9.4	Recommendations.....	93
10	References	95

1 Introduction

This report describes the process of finding a way accurately predict the arrival times of inland ships. For this purpose, an existing model is analysed and improved on. This chapter will introduce inland shipping in the Netherlands and give background information about Rijkswaterstaat, the organization where this research was performed. It also explains the lay-out of the report and the research it describes.

1.1 Inland shipping

Inland shipping is an important part of the Dutch economy, 34% of all transport is done by inland shipping (1)

Just like trucks have uncertain travel times due to possible traffic jams and road works, inland ships have uncertain travel times due to locks, water levels and other ships.

For both modalities, Rijkswaterstaat is the organization which controls and maintains the Dutch infrastructure.

1.1.1 The inland waterway system

The system discussed in this chapter is the Dutch inland waterway system. This system of waterways consists of natural and man-made rivers. To connect these rivers and to regulate the flow of water, objects like locks have been constructed. Skippers and their ships are the users of this system. This report focusses on the professional skipper or the commercial inland shipping industry.

Figure 1-1 shows a map in with all the inland waterways displayed as different coloured lines, the different colours meaning the different type of waterways.

A few facts about this system are stated below:

5046 km of waterway, of which approx. 4800 km is suitable for the transportation of goods. The main 'high ways' between Amsterdam, Rotterdam, Nijmegen which go the Belgium and Germany consist of approx. 1400 km of waterway. 347 locks (3) of which 84 are operated by Rijkswaterstaat (4) and 2644 bridges (3) of which 1092 are operated by Rijkswaterstaat (5)




Scheepstypen		Bureau Voorlichting Binnenvaart	
Klasse			380 x
Vb	Grote tanker Lengte 135 meter - breedte 21,80 meter - diepgang 4,40 meter - laadvermogen 9.500 ton		
Va	Autoschip Lengte 110 meter - breedte 11,40 meter - diepgang 2,00 meter - laadvermogen 530 auto's		60 x
III	Containerschip Kempenaarsklasse Lengte 63 meter - breedte 7 meter - diepgang 2,50 meter - laadvermogen 32 TEU		16 x
Va	Standaard containerschip Lengte 110 meter - breedte 11,40 meter - diepgang 3,50 meter - laadvermogen 200 TEU		100 x
Vb	Groot containerschip Lengte 135 meter - breedte 17 meter - diepgang 3,50 meter - laadvermogen 500 TEU		250 x
Va	R0-rotschip Lengte 110 meter - breedte 11,40 meter - diepgang 2,50 meter		72 x
Vib	Koppelverband (schip met duwbak) Lengte gemiddeld 185 meter - breedte 11,40 meter - diepgang 3,50 meter - laadvermogen 6.000 ton		240 x
Vib	Koppelverband (schip met schip) Lengte gemiddeld 185 meter - breedte 11,40 meter - diepgang 3,50 meter - laadvermogen 6.000 ton		240 x

Figure 1-1 Left: Map of the Dutch waterway system, source: Rijkswaterstaat (6)/Right: Type of inland ships (7)

Locks are an important aspect within the waterway system. Different waterways can be connected through locks and locks are also used for watermanagement. For example, if the waterlevel on one side of the lock is to low, the lock can stay opened in order the let the water rise again, or the lock can stay closed so the waterway has time to 'refill'.

Locks consist of lock chamber(s) with a certain length, width and depth which determine how many ships can go through the lock at a time. There are also extra rules for ships that carry dangerous materials when they enter a lock. Sometimes these ships have to go through the lock alone or with less ships than normal. An example, if a ship is carrying flammable goods, the ships needs to display a blue triangle ('Blauwe Kegel'). When the ship goes into the lock, it needs to be 10 meters away from other ships and cannot be in the lock when passenger ships are in the lock as well.

The users of this system consist of different type of ships. Figure 1-1 shows the ship types that sail the inland waterways. The figure also shows the class the ships belong to. This is the CEMT class which indicates the size of the ship and the type of waterway it needs to sail.

A few facts about the users are stated below:

8279 ships that sail under the Dutch flag (1). 368,733 million ton of goods transported (1). The Volkerak locks, the busiest lock in the Netherlands, had 106406 ship movements in 2017 (6)

1.1.2 Rijkswaterstaat and inland shipping

Rijkswaterstaat is part of the Dutch Ministry of Infrastructure and Environment and responsible for the design, construction, management and maintenance of the main infrastructure facilities in the Netherlands. Which include among others the inland waterways and the management of inland shipping.

The management of inland shipping is now done from many locations where objects and sectors³ are being operated individually.

1.1.3 CBB

Rijkswaterstaat wants to centralize the management of inland shipping and has started the project CBB (Corridor gericht Bedienen en Begeleiden) which, when translated, stands for Corridor oriented Operating and Guiding. The goal of this project is to centralize the management of inland shipping. The inland waterway system is divided into corridors, these corridors represent important shipping routes and are shown on the next page.

The idea of CBB is that a ship can sail over a corridor faster and more efficient than it would do today. This is due to better planning of objects and a better knowledge of the situation on the waterway. In order to achieve this, the CBB project will centralize the management of the inland waterways. The goal is to have one location for each corridor (figure 1-2) from where all the objects and sectors are being operated.

For Rijkswaterstaat, this will mean more efficiency when operating objects and better and easier communication between objects and sectors. When objects and sectors are going to be operated from the same location, it is easier for operators to communicate with each other and because multiple objects are operated from one location, one operator will be able to operate multiple objects. For the skipper this will mean a better service with better information about his journey and the situation on the water and near objects.

³ Objects are bridges and locks. Sectors in the context of this document mean the VTS (Vehicle Traffic Service) sectors that are managed by Rijkswaterstaat. These sectors are areas on the water where the traffic (Inland ships and big recreational vessels) is guided through the sector. These sectors are often either busy and/or are on a junction between waterways.

Next to centralizing the operation, the CBB project is also creating and updating computer systems in order to implement corridormanagement. These systems will also help make the work for operators easier and provide skippers better and improved information. For example, a new system is being developed that will track ships real-time. Other systems, like data systems about water levels and waterway characteristics are being updated.



Figure 1-2 Overview shipping corridors (6)

1.1.4 Trajectplanner

Another new system is the Trajectplanner. This system is a tool which helps both the skipper and Rijkswaterstaat with managing inland shipping. It does this by using data from many other systems and using that data to simulate the entire journey the ship will be making (Figure 1-2). With this simulated data it makes schedules for locks and bridges which the ship will be passing and will make predictions of the occupancy rate of the inland waterways. With these prediction and schedules, the skipper can adapt his speed and/or route in order to sail more efficient.

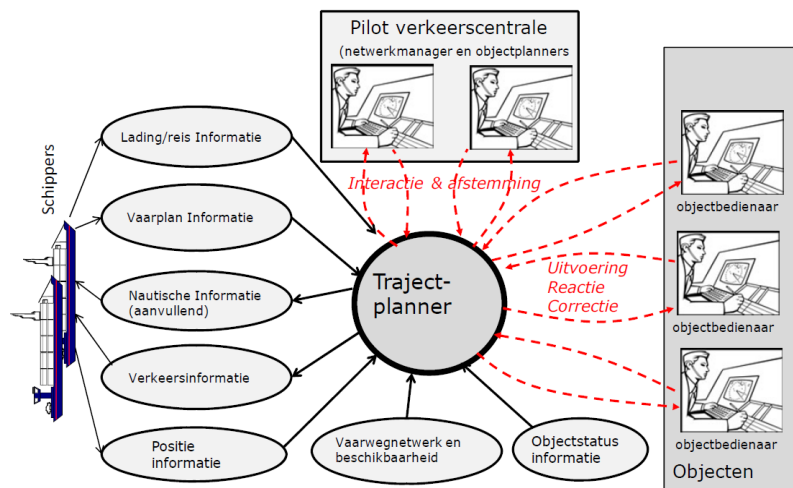


Figure 1-3 Trajectplanner and its connections (2)

For example, when a ship starts its journey, the skipper must give information about his cargo, destination and draft to Rijkswaterstaat. This information is then used by the Trajectplanner to predict the route the ship is going to take. When this is done, the Trajectplanner knows which objects the ship is going to come across. It can then schedule the ship in with the objects. In the case of bridges, if the ship is low enough and the bridge doesn't have to open, the ship can then sail through without stopping or slowing down. But when the ship encounters a lock, it has to go through the lock procedures. This in turn means that it must wait for the lock chamber to be ready and maybe ships that have to go in first. If it is busy at the lock, it can be that the ship has to wait for an hour to let other ships go through the lock. The Trajectplanner can schedule all ships into the lock procedures, so when it is busy at the lock, the Trajectplanner can tell the ship to slow down, so it doesn't have to wait when it arrives at the lock. Because the Trajectplanner has all this data it can tell the skipper when to arrive at a lock, but it also knows when it is through the lock and can therefore tell the skipper when he will arrive at his final destination.

The data which is calculated by the Trajectplanner is also used by Rijkswaterstaat to efficiently control their objects and corridors. When the lock operators know when ships arrive, they can reduce the amount of empty lock operations. And when corridor managers know when certain parts of the waterway will be busy with ships, they can take actions accordingly.

For the scheduling and predictions to be made, the Trajectplanner calculates and estimates ETA's (Estimated Time of Arrival) of the ships in the system. These ETA's are calculated for each point and ship in the system. The Trajectplanner uses these ETA's to schedule ships into the lock procedures and to let operators know which ships will arrive at certain points in time. These ETA's are therefore very important for the correct functioning of the Trajectplanner.

These ETA's are important for the skipper too. Currently the skipper adds an extra 30% time into his planning to compensate for uncertainty such as busy waterways, uncertain passing time of locks, bridges and unforeseen incidents. This means that at the destination, the next shipper has no real idea at what time the ship will be arriving. Therefore, either the shipper has to have his equipment, personnel and storage space ready for several hours, or the skipper arrives and there is no possibility anymore for him to unload.

With accurate ETA's the skipper can plan his journey better and more efficient, reducing waiting times and lower fuel consumption (when waiting, the skipper keeps the engine running).

For the skipper, the Trajectplanner helps to reduce the uncertainty in his planning. And with this reduction in uncertainty, more trips can be planned.

Therefore the goals of the Trajectplanner are:

- Accurate simulation of inland waterway systems and ships
- Correct scheduling for objects
 - No waiting times for ships at objects
 - More efficient sailing/ less time waiting with engine running
- Accurate arrival time at destination

To reach these goals, the ETA's the Trajectplanner calculates need to be accurate

1.2 ETA's

For the Trajectplanner, the Dutch waterway system was modelled with dots and lines which are similar to a network of nodes and arcs. This will be further discussed later in this report. These arcs represent part of the waterway and the nodes represent points on the waterway. Often these nodes lay on junctions but can also lay on a straight part of the waterway. In other words, a node can have 2 or more arcs connected to it. When a ship is starting his journey, the Trajectplanner predicts the route the ship will take. This route is predicted along the nodes and arcs. Each node the ship passes has an ETA (Estimated Time of Arrival) for that ship. These ETA's are calculated/predicted by the Trajectplanner. To calculate the ETA for the next node, the Trajectplanner uses the ETA it calculated for the previous node and adds the cost of the arc between the nodes. The cost of an arc consist of sailing time for that particular ship + the passage time of an object if there is one on the arc. If the arc has a lock as object on it, the passage time depends heavily on how many other ships need to pass the lock as well.

1.3 Problem description

As mentioned before, the ETAs which are calculated by the Trajectplanner are important. Because the Trajectplanner uses the ETA of one point to calculate the ETA for the next point, it is important that these calculations and simulations are accurate. At the moment, this is not always the case

The Trajectplanner simulates 6 hours into the future. The ETA it calculates for the node at the final destination at that time is not very accurate. The Trajectplanner keeps recalculating the ETA's during the journey of the ship and the closer the ship gets to the object or destination, the more accurate the ETA's become.

Inaccurate ETA's is a problem for single ships, but also for the other ships.

Single ships

If an ETA of a ship for one of the nodes is incorrect/inaccurate, all the points following will be incorrect/inaccurate. And when there is a lock on the route and the ETA for that lock is inaccurate, the schedule for that ship will be inaccurate. The inaccuracy which occurred at the first point will become bigger at the last point.

Multiple ships

If the schedule for a single ship is incorrect because of an incorrect ETA, the schedule for all the other ships that were in the same schedule will be incorrect. And so, if an ETA for one ship is inaccurate, ETA's for other ship will also become inaccurate.

1.4 Research goal + question

The goal for this research will be to analyse and find ways to improve the current system, the Trajectplanner

To reach the research goal, a research question was formulated:

- **How to accurately predict ETAs of inland ships**

In order to help find the answer to this question, several subquestions are formulated:

- 1) **Are there other methods/systems that predict the ETA of inland ships?**
- 2) **How to define the accuracy of the ETA prediction?**
- 3) **How does the current simulation system predict the ETA?**
- 4) **How accurate is the current prediction?**
- 5) **How to improve the current prediction?**
- 6) **How to evaluate a new prediction model?**
- 7) **How accurate is the new prediction model?**

1.5 Research approach

In order to reach the goal of this research the 'problem' was first analysed in two ways. A literature review was conducted in order to find similar systems like the Trajectplanner, or reports which dealt with similar problems (uncertain travel times/ETA prediction). This is discussed in chapter 2. In order to correctly do the second analysis, the accuracy of the ETA estimation for inland ships needed to be defined, so this was discussed in chapter 3. The second analysis was performed on the current system, the Trajectplanner. The ETA estimation and the output of the Trajectplanner were analysed in order to find the problem(s) and come to improvements. This is discussed in respectively chapter 4 and chapter 5.

From these analysis, new ideas and methods to improve the ETA estimation came forth, which is discussed in chapter 6. And in order to test and evaluate these, a simple system which was similar to the Trajectplanner was built in order to experiment with the new methods. This is discussed in chapter 7. After the above, conclusions could be made and recommendations for further research and development of the Trajectplanner were given in chapter 8.

1	Other systems and papers	Literature
2	Accuracy of the prediction	KPI definition
3	Workings of the current prediction	Analysis
4	Accuracy of current prediction	Analysis
5	Improved methods	Methods
6	Evaluation method	Simulation
7	Accuracy of improved methods	Evaluation

2 Ship movement modeling and ETA prediction in literature

In order to answer the first subquestion: '**Are there other methods/systems that predict the ETA of inland ships?**', a search was conducted to find similar (IT)systems and a literature review was performed to find papers with similar problems. This review was done with multiple aspects in mind. In the question, similar systems, mostly refers to the predicting side of the Trajectplanner, but it also refers to the model of the Dutch waterway network which is used in the Trajectplanner.

The literature research focussed on two aspects in the problem description in the previous chapter. The way the inland waterway system is modelled and the inaccuracy or uncertainty of the prediction of the ETAs. In order to find a good way to model the inland waterway system, papers with similar models were searched for. This was done in the field of inland shipping, but also in the field of other modalities. The inaccuracy part of this review focused on methods that are used to predict the ETA of vehicles (ships, cars, airplanes) and if they have inaccuracy problems and how they deal with them.

2.1 Similar systems

When this report talks about similar systems, it means IT systems that perform a similar task to estimating the ETA. In this case the search can't look for systems in other modalities because of the unique situation inland ships have when they encounter locks. For example, a car navigation system can't be easily modified to work for inland ships. Because cars don't encounter objects for which they have to wait to get through. Cars do have to wait for traffic jams, but they can often take another route which inland ships can't do.

There are several IT systems that help the skipper plan his journey. First there are several routing applications, but these mostly consist of accurate nautical maps of rivers and canals and don't estimate the ETA.

There is a ship tracking system called Marine Traffic (7) (8) which has a feature which estimates the ETA at the final destination. But this ETA prediction takes the ETA from the AIS system of the ship, or it estimates it based on the location, direction and speed of the ship. This could be fairly accurate for sea faring ships, but not for inland ships that encounter locks as it does not take these into account. There are also systems used by the government to simulate ships movements within the inland waterways(SIVAK and BIVAS). These systems are used to simulate the effects of policy changes, blockages or a new lock design. These systems do however simulate ship movements quite accurate and could therefore be used to estimate the ETA of ships. There are 2 aspects of these systems that wouldn't work towards the goal of the Trajectplanner, namely it is not possible to get data for 1 single ship, it is only the bigger picture the system looks into and both systems work with historical data and not live data. The last point is not bad in a general sense, but wouldn't work for the skipper that wants an accurate planning for his current journey. (9) (10)

2.2 Similar models

2.2.1 Inland shipping

The literature review first focussed on research that was done in the field of ships and inland shipping. A few interesting papers were found. In Caris et al. (2014) (11) research challenges were found within inland waterway transport. One of the challenges is the creation of a model which models the whole inland shipping system. Including waterways locks and harbours.

A recent study done by Li (2017) (12) Actually described a model like this, but this paper is discussed in the next section. Like Caris et al.(2014) (8) already mentioned, some components of the system have already been modelled. For example, Zheng et al. (2015) (13) described a model which plans the movements of a waterborne AGV (Automated Guided Vehicle) within a harbour. Their model took into account the time window in which a barge has to arrive to be loaded/unloaded. Multiple papers have been written about the simulation and/or improvement of the locks and their procedures. Verstichel et al. (2013) (14) described a way of placing ships in locks efficiently; Hengeveld (2012) (15) shows that by controlling the arrival times of ships, the waiting times at locks can be reduced; Visser (2016) (16) shows that by controlling the locks centrally, the efficiency of the whole system can be improved. The last 2 papers had to simulate the procedures of the locks in order to find the answer to their respective research questions.

Visser (2016) (16) also has to describe (a part of) the Dutch waterway system. This is done by modelling the system as nodes and arcs where the nodes represent the locks and the arcs the waterway between the locks. See figure 2-1

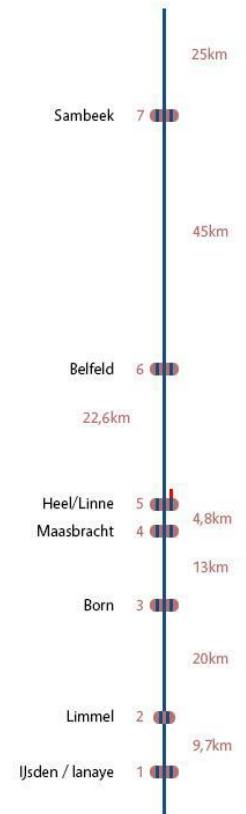


Figure 2-1 Network model from Visser(2016)

2.2.2 Other modalities

As mentioned, the review also looked at other modalities. To find relevant papers, the review looked at papers which described similar problems and/or systems, especially papers which uncertainty issues.

In plane routing and planning, there are multiple sources of uncertainty, like bad weather, arrival delay and mechanical problems. These are covered in Zhang et al. (2017) (17). They also had to model the system. In this case, the aircraft flying from airport to airport, but unlike ships, aircraft can fly straight to their destination, without encountering any crossings or junctions. So the model used in Zhang et al. (2017) (17) is not suitable to be used as a reference for this system.

Zhou et al.(2016) (18) describes a problem which exist in the routing and timetabling of trains. These routes have to be fast and safe at the same time. And like inland ships, trains don't always have choice in route. Another similarity is between the train stations and the locks. Trains have an arrival time, a platform and a departure time. Inland ships have an arrival time, a lock chamber and a departure time. All of these factors need to be taken into account when planning. Also the model used to describe the system of a train station can be converted to use as a model to describe a lock. See figure 2-3

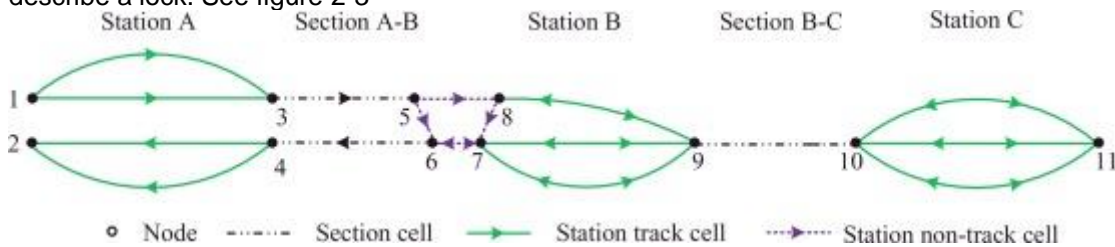


Figure 2-3 Schematic of a railway network from Zhou et al.(2016)

In the figure above, the nodes represent enter or exit points of a station or piece of track. In the case of a station, there can be multiple tracks between the enter and exit point. For a lock it would look similar

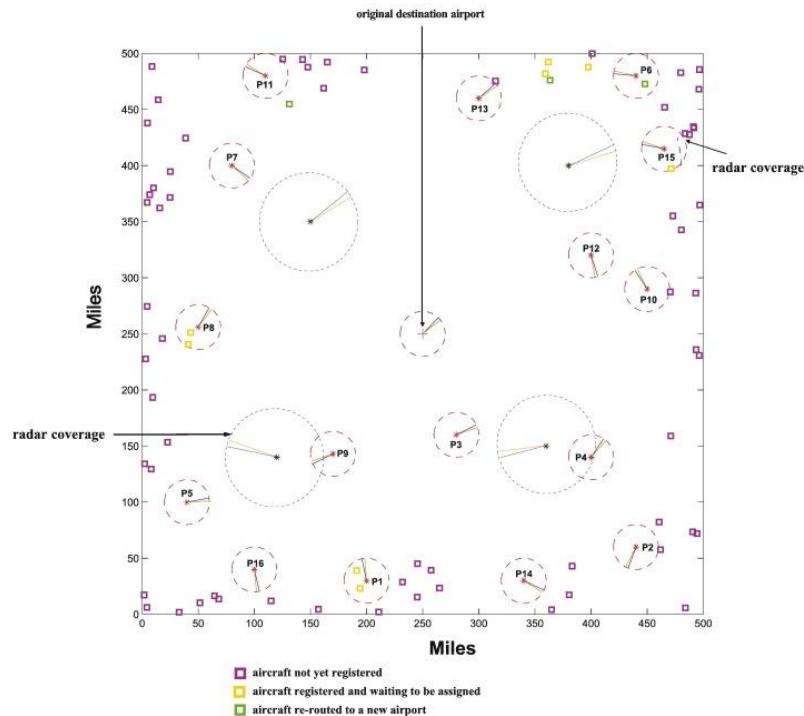


Figure 2-2 schematic of a network with airports from Zhang et al. (2017)

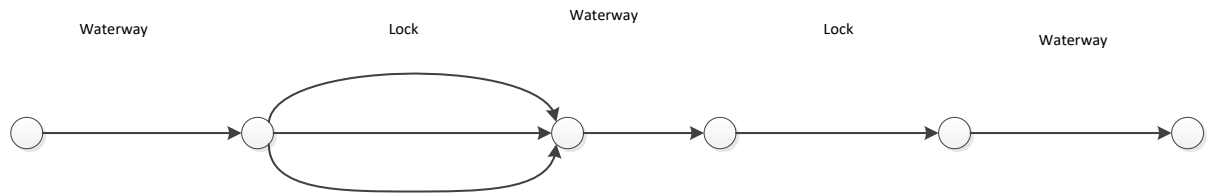


Figure 2-4 Schematic of waterway network

In the figure 2-4, the nodes represent again an enter/exit point and the arcs represent a piece of waterway. In the case of locks there can be multiple arcs which represent multiple lock chambers.

2.3 Inaccuracy in prediction

2.3.1 Inland shipping

The second part of the review focussed on the inaccuracy or uncertainty of the predicted ETA. A recent study was done by Li (2017) (12) On the dynamic traffic assignment model which, in the paper, is used to predict the extra waiting time at locks, when another lock is closed. It can also predict the extra waiting times at locks when there is a sudden rise in traffic. The paper models the entire inland waterway system with an arc and node network, where arcs are the waterways and nodes represent the junctions and objects. With this network, the model simulates and predicts the travel time of the ships.

The model used in Li (2017) (12) could be used as an inspiration for the model used in this report. It also shows that locks are an important factor in the total travel time of ships which could lead to the uncertainty factor in their journey and the inaccuracy in ETA prediction.

Parolas (2016) (19) Uses machine learning to predict the ETA of containerhips at the port of Rotterdam. The paper uses historical data about arrival times as well as position, speed, draft and weather data to predict the ETA of container vessels arriving at the port from sea. The biggest uncertainty in this case is the weather. The time frame in these journeys is often several days. The system and uncertainty aspect of this paper are not comparable with the system and uncertainty aspect in this paper. However, Parolas (2016) (19) does clearly show the importance of an accurate ETA. When a ship arrives, it needs space at the quey, cranes and personnel need to be available, barges or inland ships need to be ready to take the containers further on their journey etc.

2.3.2 Other modalities

As in 2.2.2, ETA's are important for planes as well. Roy et al. (2006) (20) describes the problem of accurate ETA prediction of airplanes near an airfield. The biggest problem in this paper is the noisy data received about the position and speed of the plane. The paper proposes a new filter and prediction tool to fix this. This is not comparable with the problem described in this paper.

ETA prediction is also important in public transport. Cathey et al. (2003) (21) uses the vehicle location data to predict if the vehicle will be on time at their next stop and if not, what kind of delay the vehicle has. Tan et al. (2008) (22) also use vehicle location data but uses this to predict the arrival at a signalized intersection in order to give the bus priority over the intersection. For both these papers it looks like the time scale is too small to compare it to inland shipping, but the error in ETA prediction in proportion to the travel time is roughly the same. See table 2-1. For both modalities, the error is one time component down (i.e. hour -> minute / minute -> second)

Table 2-1 Error versus total travel time.

Modality	Error	Total travel time
<i>Bus</i>	Seconds to minutes	Around an hour
<i>Inland ship</i>	Minutes to an hour	Hours

The error in both papers is mainly caused by traffic and weather, where for inland ships, locks have the most influence on their journey time.

The big difference is that weather conditions and traffic conditions are hard to predict and even harder to plan for whereas locks can have a schedule. And if the skipper sticks to that schedule, there should be no extra waiting time for the ship.

2.4 Conclusion

As an overview, the papers used in this chapter are put in an overview table.

Table 2-2 Overview of papers used

Paper	Node and arc network model	Uncertainty in arrival times	ETA prediction	Usable/Adaptable for inland shipping
Caris, Limbourg, Macharis, Lier, Cools. <i>Integration of inland waterway transport in the intermodal supply chain: a taxonomy of research challenges</i> . 2014.	Yellow	Yellow	Yellow	Red
Li, K. <i>Dynamic traffic assignment model for inland waterway freight transport</i> . 2017.	Yellow	Red	Red	Green
Zheng, Negenborn, Lodewijks. <i>Predictive path following with arrival time awareness for waterborne AGVs</i> . 2015.	Red	Green	Red	Red
J. Verstichel, P. De Causmaecker, F.C.R. Spieksma, G. Vanden Berghe. <i>Exact and heuristic methods for placing ships in locks</i> . 2013.	Red	Red	Red	Red
Hengeveld, J. J. S. <i>Optimization to reduce waiting times at locks</i> . 2012.	Red	Yellow	Red	Red
2016, T.H.A. Visser –. <i>Lock scheduling model for the series of locks in the Maas</i> . 2016	Yellow	Red	Red	Green
X. Zhang, S. Mahadevan. <i>Aircraft re-routing optimization and performance assessment under uncertainty</i> . 2017.	Red	Green	Red	Red
W. Zhou, H. Teng. <i>Simultaneous passenger train routing and timetabling using an efficient train-based Lagrangian relaxation decomposition</i> . 2016.	Green	Yellow	Green	Green
Parolas, Ioannis. <i>ETA prediction for containerships at the Port of Rotterdam using Machine Learning Techniques</i> . 2016.	Red	Green	Green	Red
Kaushik Roy, Benjamin Levy, Claire J. Tomlin. <i>Target tracking and Estimated Time of Arrival (ETA)</i> . 2006.	Red	Green	Green	Red
F.W. Cathey, D.J. Dailey. <i>A prescription for transit arrival/departure prediction</i> . 2003.	Yellow	Green	Green	Red
Chin-Woo Tan, Sungsu Park, Hongchao Liu, Member IEEE, Qing Xu, and Peter Lau. <i>Prediction of Transit Vehicle Arrival Time for Signal</i> . 2008.	Green	Yellow	Green	Red

Table 2-2 shows an overview of all the papers used in this chapter with coloured indicators next to which tells if the paper had certain subjects described in it and whether it is used as inspiration for this research. Green = yes, red = no, yellow = partly. For example, the report by K. Li(2017) has a yellow indicator at the 'Node and arc network model', which means that the report did use a node and arc network model, but a simple node an arc network model. What this means is that in a regular node and arc network model either the nodes or the arcs take time to pass. In the model used by K.Li, both the nodes and the arcs take time to pass.

Reports about inland shipping often didn't use a regular node and arc representation of the waterway, although Li (2017) (12) comes close.

Zhou et al.(2016) (18) described an interesting model which could easily be converted for use to describe an inland waterway system.

The model used by Zhou et al.(2016) (18) combined with the model used by Li (2017) (12) could be used to make a full model of the inland waterway system.

The paper by Li (2017) (12) also shows that locks are a big source of uncertainty. However the model used by Li assumes that ships will follow the route given by the model which would only be true in a perfect situation.

Parolas (2016) (19) showed the importance of an accurate ETA when multiple stakeholders come together in the transportation of goods.

Problems with the received data about position and speed or problems with traffic on the way will not be used in this paper because the error in location and speed compared to the size and travel times of the ships are assumed neglectable.

The answer to the subquestion: **Are there other methods/systems that predict the ETA of inland ships?** Is yes, there are similar systems and papers that have written about them. Although other models can be found in literature and in IT systems, none of them cover all the aspects of the Trajectplanner.

However, a combination of papers can be used to tackle the problems of the Trajectplanner. The inaccuracy or uncertainty of travel times has been seen in other papers, but not with the same aspects as inland shipping.

Aspects that were used as inspiration further in this report were: The data and parameters used in the SIVAK and BIVAK systems to calculate/simulate ship movements. The network models used by Visser and Li, but also the network model used by Zhou. The planning aspect with bus routing used Cathey and Tan was also considered as well as the definition of accuracy of the system used by Parolas et al. and Karbassi et al. This is further discussed in chapter 3.

3 KPI's

To analyse the current system, an improvement and ultimately answer the main question, it is necessary that a KPI (Key Performance Indicator) is formulated. This chapter will answer the subquestion: **How to define the accuracy of the ETA prediction?**

3.1 KPI's

Rijkswaterstaat uses KPI's for all their services, but because the Trajectplanner is not operational yet (it is running and working, but it is not yet used in the operation (guiding ships, lock operations, etc.)), there are no KPI's formulated yet.

This means that a new KPI needs to be thought of. And because the main question of this report is about accurate arrival times it is logical to include the estimated arrival time (ETA) in the KPI.

Just like a car navigation system, a system which simulates the movements of ships will simulate every couple of seconds or minutes. This means that the ETA to the destination will change if new events occur in the simulation. This also means that the calculated ETA can be tracked over time and with the knowledge of the actual arrival time (ATA), the accuracy of the ETA can be monitored (afterwards).

This would look something like this.

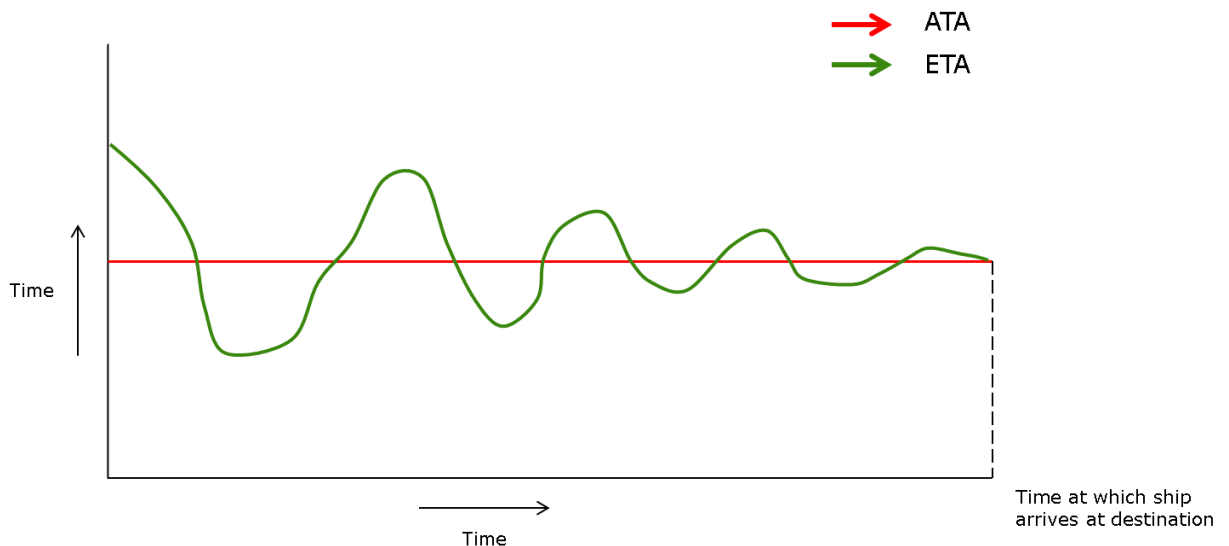


Figure 3-1 ETA's monitored through time with the ATA as comparison

In figure 3-1, the ETA and ATA are shown as lines. The ATA is only known after the ship has arrived but is drawn as a straight line in order to compare it to the calculated ETA. Both the horizontal and

vertical axes display a time value, but the vertical axis displays the calculated ETA and the horizontal axis displays the time in which the calculation took place.

3.2 Literature

As a big part of the literature review in chapter 2 was done on subjects which involved estimation of arrival times and the improvement thereof, it makes sense to check the same reports to see how they defined the amount of improvement their idea, models or methods had.

The two reports which discussed a public transport system (Cathey et al. (2003) (21) and Tan et al. (2008) (22)) defined their accuracy of the system by making sure that the traveller had a 90% certainty that their bus arrived within a certain time frame. This time frame gets smaller when the bus gets closer to the stop.

The research done by Li et al.(2017) (12) used the travel time, waiting time and lock enter time as KPIs.

Other reports like (Parolas et al.(2016) (19) and Karbassi et al.(2003) (23)) used the mean error, or mean squared error to define the accuracy of their method and to show improvements over other methods. The error in this case was always the absolute difference between the ATA and ETA at certain moments.

3.3 Users

As the KPI is an indicator to how accurate the system is, the users of that system, in this case the skippers and Rijkswaterstaat, need to be taken into account when defining the KPI

Skippers/Ships

The skippers will use the Trajectplanner to tell them at what time they will be at their destination. But they will also use it to determine if they need to speed up or can slow down in order to sail through a lock as efficient as possible. Because of this, they need a system (Trajectplanner) which gives accurate information from the start. And the information may not fluctuate too much, else it won't be useful for the skipper.

Rijkswaterstaat/Operator

Rijkswaterstaat and the operators of objects will use the information from the Trajectplanner to make schedules for the objects and to determine how busy it will become where and at what time in the future. To make schedules ahead of time, they also need accurate information from the Trajectplanner as soon as possible.

3.4 Proposed KPI

With the literature and the users in mind it is proposed that the KPI can be defined as an absolute difference between ETA and ATA

$$KPI = |ETA - ATA|$$

It is no use to the skipper that the ETA is accurate when he is nearly at his destination, so a dynamic KPI based on distance to the location would not work. The skipper needs an accurate ETA when he plans his journey, or when he is at the beginning of his journey. The ETA needs to stay accurate, also during the journey. If the skipper wants to use Trajectplanner the way it is intended (More efficient planning and sailing, slowing down or speeding up), the accuracy of the ETA needs to stay constant during the trip. It is therefore important that the ETA doesn't fluctuate to much. This fluctuation can be measured and presented by taking the median or average of all the KPI data points. However, with the skipper in mind, a KPI which consist of the following:

$$KPI = AVERAGE|ETA - ATA|$$

Has no use for the skipper. If the skipper needs to use this system, he needs to be able to check his ETAs at every moment. So a spike in difference between ETA and ATA would not give the skipper confidence in the system even if it is only one spike and the average difference is still very low. So with the skipper in mind, the KPI needs to be as low as possible, all the time. So a KPI for a system needs to be defined on the highest difference between ETA and ATA. The proposed KPI is therefore:

$$KPI = MAX|ETA - ATA|$$

This means that the KPI is based on the max difference between the ETA and ATA. This is done because a spike in the ETA calculations would cause errors throughout the journey but also for other ships. It is therefore important that this value is as low as possible.

3.5 Use of KPI

Further in this report it will be explained that the ships travel along a set of points which they pass. For each point on the journey, this KPI can be measured and calculated. As the ATA is only known after the ships has reached the point which is measured can the KPI be determined.

In this report, the KPI is used to analyse the system and any design alternatives. This means that the ETA is measured for several points during an analysis. This also means that the KPI for several points was determined. This was done in order to find the cause of the inaccuracy and to test whether locks have a big impact on the ETA prediction.

For a skipper other points on the journey will be interesting to know the accuracy of. For the skipper, the point before a lock is the most interesting, because he needs to adjust his speed to that prediction. For the next shipper, the ETA on the transfer point (destination of the skipper) is the most interesting. For this research, the largest value of the KPI was the focus point because that would indicate the biggest source of inaccuracy. The goal was then to get the KPI for this point as small as possible.

3.6 Conclusion

The KPI proposed in section 3.4 will be used for this research. It will be used during the analysis of the current system and to analyse any improvements. This KPI was formed by considering literature with similar subjects and the users of the system.

As mentioned in the previous section, the KPI can be determined for every point in the system after the ship has sailed past it, which makes the ATA known. For this research, the point where the KPI was determined for the system was the point where the KPI would (assumed to) be the largest.

The answer to the subquestion: **How to define the accuracy of the ETA prediction?** Is the maximum absolute difference between the ETA and the ATA on the point where this value is the largest, or

$$KPI_x = MAX|ETA_x - ATA_x|$$

Where x is the point where the values ETA and ATA are measured during a certain time period and where a smaller KPI is better.

4 Current prediction and planning system

The current prediction and planning system is the Trajectplanner which has been mentioned in the introduction. How this system works and **How does the current simulation system predict the ETA?** will be answered in this chapter. The code with which the Trajectplanner is build is confidential and can't be shown.

4.1 Trajectplanner

As explained before, the Trajectplanner is a system which calculates, simulates and predicts the movements of all the inland ships which sail on the Dutch waterways system and have an AIS transponder (A radio device which sends out the position, speed and ID of the ship). To simulate this, the Trajectplanner receives information from other systems and uses this do run the simulation. For example, the system receives GPS data for all the ships as well as their departure and destination points. To be clear, the Trajectplanner has no influence on the ship, the way a ship sails or the workings of objects. The RTA (Requested Time of Arrival) it gives out for ships and the schedules it gives out for objects, are suggestions for the skipper and the operator. The Trajectplanner does not make decisions.

The ships in the simulation of the Trajectplanner sail around in a model of the Dutch waterway system. Two figures of parts of this system are shown below.

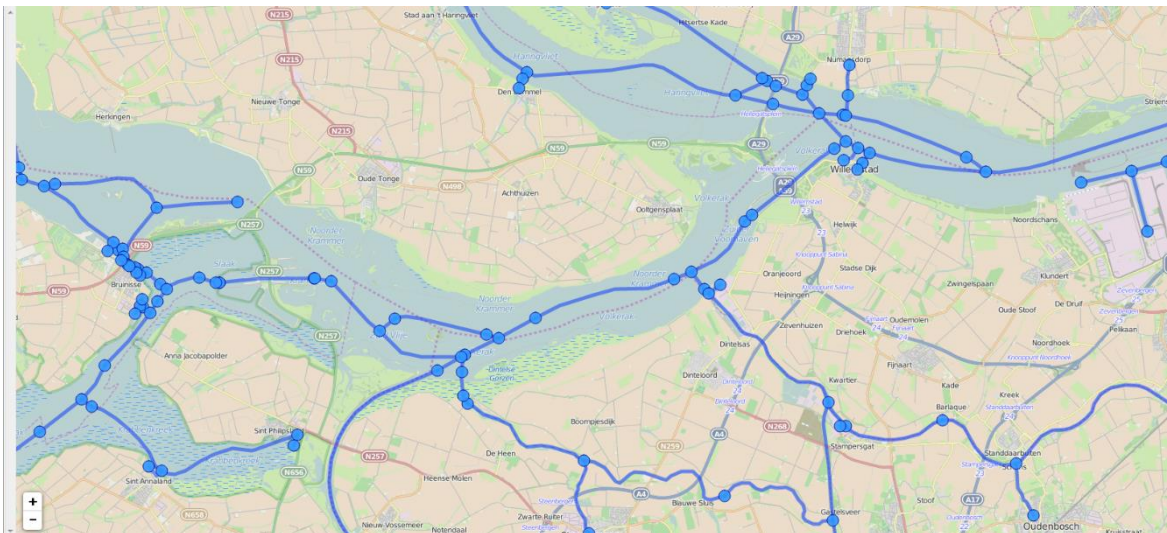


Figure 4-1 Image from Trajectplanner UI

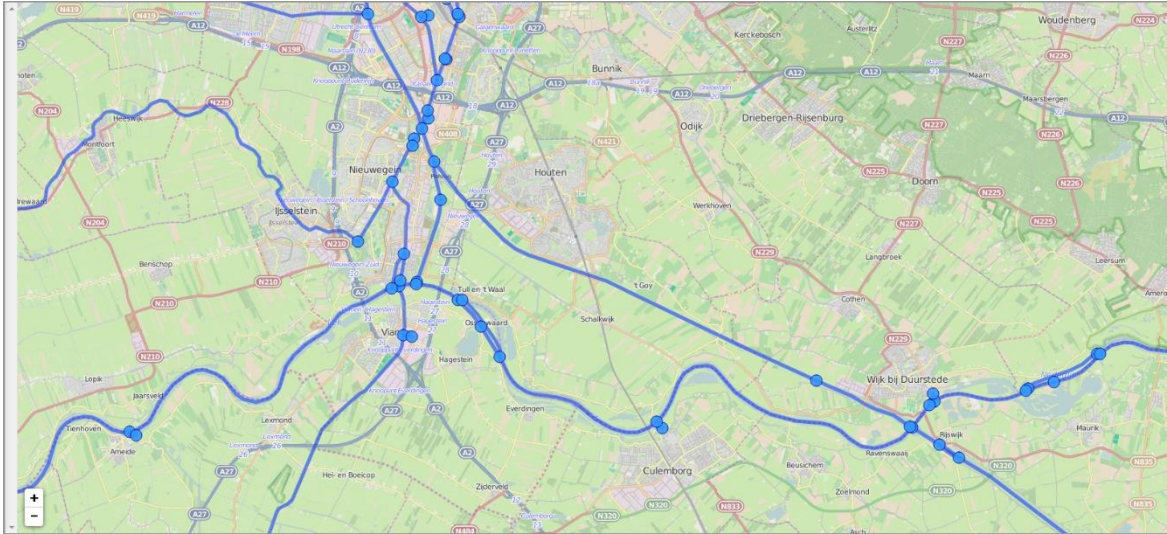


Figure 4-2 Image from Trajectplanner UI

In both figures above, part of the Dutch waterways system is displayed as used by the Trajectplanner. Figure 4-1 shows the area around the Volkerak locks and figure 4-2 shows the area around Utrecht/Nieuwegein in which the Beatrix, Irene and Marijke locks are positioned. The locks are not shown or indicated in these figures.

In both figures blue lines and dots can be seen. These lines represent the waterways on which ships sail and the points represent points on the waterway to where ships can sail to and from. It could be said that the lines are arcs which connect the nodes, in this case the blue dots. The system displayed in these figures is not entirely the same as a node and arc graph as some dots lay on the waterway without being a junction or beginning/end position.

The dots are used to represent junctions, begin/end positions like ports and docks. These dots don't represent objects like locks and bridges.

These dots are used to calculate ETA's. For every ship in the system, the Trajectplanner calculates the ETA's for each dot that ship will encounter. How the Trajectplanner does this will be discussed in more detail in the next section

4.2 Estimating the ETA

Every 6 minutes, the Trajectplanner simulates all the ships and their movements in the inland waters of the Netherlands which have AIS transponders on board. The skipper itself has to put in the destination. The Trajectplanner combines all the information it has and uses it to calculate the best route for the ship and uses this route to determine which objects the ship will encounter. And if needed, the Trajectplanner schedules the ship in a lock. It does this every 6 minutes and does not keep track of previous simulations. In other words, it starts over every 6 minutes.

The Trajectplanner uses an event-based simulation which means, certain calculations will only be done if an event is triggered. This has implications on the calculations for the rest of the journey. For some actions, this makes sense, for example, only when a ship's route will come across a lock within 6 hours will the Trajectplanner schedule the ship for the lock.

To calculate the best route for the ship, the Trajectplanner uses the Dijkstra Algorithm and for the best fit in a lock it tries multiple fits and checks which one has the most ships in it. It also checks whether it is still possible for all the ships in a certain schedule to make the requested arrival time at the lock. If not, a new schedule has to be made.

As is mentioned before, the Trajectplanner calculates the ETA for the next node by adding the travel time of the arc between the nodes to the ETA of the previous node. Simply said, there are two types of arcs, arcs without objects on them and arc with objects on them. The travel time for arcs without objects on them is easy to estimate. In this case, the Trajectplanner uses the type of ship, the type of waterway and if the ship is loaded or not, which tells the Trajectplanner the draft of the ship. With this information the Trajectplanner can determine the speed the ship will travel the arc and with the length of the arc it calculates the travel time.

If an arc has an object on it, the travel time is more difficult to estimate. If the ship can just 'sail through' the object, like an open lock or a bridge that is high enough for the ships the sail under, the travel time is estimated just like there is no object.

When the object can't be passed, like a lock or a low movable bridge, the travel time becomes more difficult to estimate. The Trajectplanner will have to schedule the ship in for the object. For a bridge, this is still relatively easy because the opening of the bridge is often depended on the traffic that is traveling over it. So the extra travel time becomes the extra waiting time the ship has to do when waiting for the bridge to open. Most of the ships don't encounter these kinds of bridges because the main sailing routes have bridges that are high enough to sail under. The main obstacle a ship will encounter is a lock. If a ship will encounter a lock on its route, the Trajectplanner will schedule the ships into the lock. Based on this schedule, the Trajectplanner gives out an RTA (Requested Time of Arrival) for the node before the lock. The Trajectplanner then assumes that the ship will be on time and can sail right into the lock. The passage time of the lock then becomes the time it takes the current ship and all the other ships that are in the same schedule to sail in, go through the lock procedures and then all sail out again. The travel time can then be estimated by adding the passage time of the lock to the travel time based on the length on the arc and the estimated speed of the ship. Other ships, current and water levels are not taken into account by the Trajectplanner.

The Trajectplanner does know the current position of a ship, so every 6 minutes, the ETA is updated. When a ship is on an empty arc (an arc without objects), the Trajectplanner knows where on the arc the ship is and uses this to calculate what fraction of the arc the ship still has to sail and uses this fraction of the travel time of that arc to adjust ETA estimations.

When a ship sails on an arc with an object, for example a lock. The Trajectplanner doesn't know when the ship has passed the lock. This is because the locks are not modelled in the network model. This is further discussed in chapter 5 where the output is analyzed

4.3 Conclusion

The answer to the subquestion: **How does the current simulation system predict the ETA?**

A table (table 4-1) is made to show what the Trajectplanner uses to estimate the ETA.

Table 4-1 Parameters in ETA estimation

Parameter that are used to estimate ETA	Parameter that are not used to estimate ETA
Ship class	Current
Current waterway class	Other ships on waterway
Speed (depended on ship and waterway class)	Waterlevel
Empty/Loaded	
Draft (depended on ship class and empty/loaded)	
Object on arc	
Travel time of arc	
Passage time of object (depended on schedule for object)	

The Trajectplanner uses the parameters above together with a network model consisting of nodes and arcs (not a true node and arc network) to simulate the movement of all inland ships and with that simulation it will estimate the ETAs for all ships. It needs to ben pointed out that within this simulation, the Trajectplanner only knows that a ship has passed a lock when that ships reaches the next node on its journey.

5 KPI of the Trajectplanner system

In order to determine **How accurate is the current prediction?**, the output of the system was analysed. As mentioned in chapter 3, multiple points were 'followed' so for each point a KPI was determined.

5.1 System output

The analysis was performed with the UI which is also available for the operators. This means that the data which is shown by the Trajectplanner is the same data that is used by the operators and systems like IVS Next.

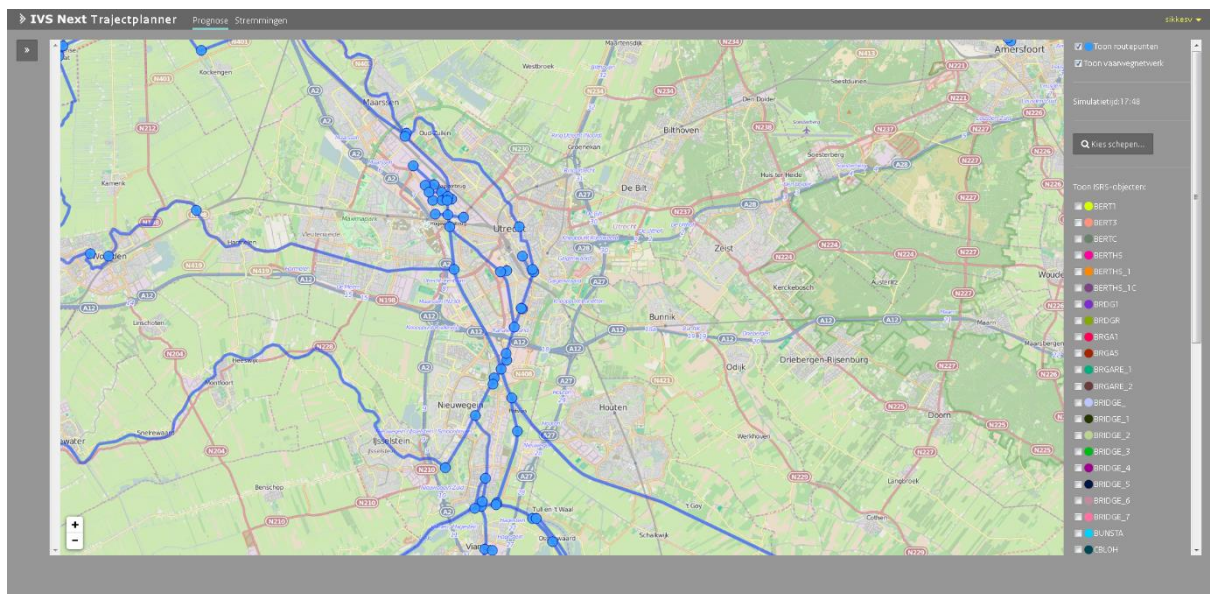


Figure 5-1 UI of the Trajectplanner

As can be seen in figure 5-1 and as mentioned in the previous chapter. The waterway is modelled as blue dots and blue lines. The lines represent parts of the waterway and the dots represent junctions or points on the waterway. For each dot the Trajectplanner calculates the ETA's for all the ships that pass that dot. These dots don't have any characteristics other than being a point for which an ETA is calculated.

The data which the Trajectplanner shows is shown in figure 5-2.

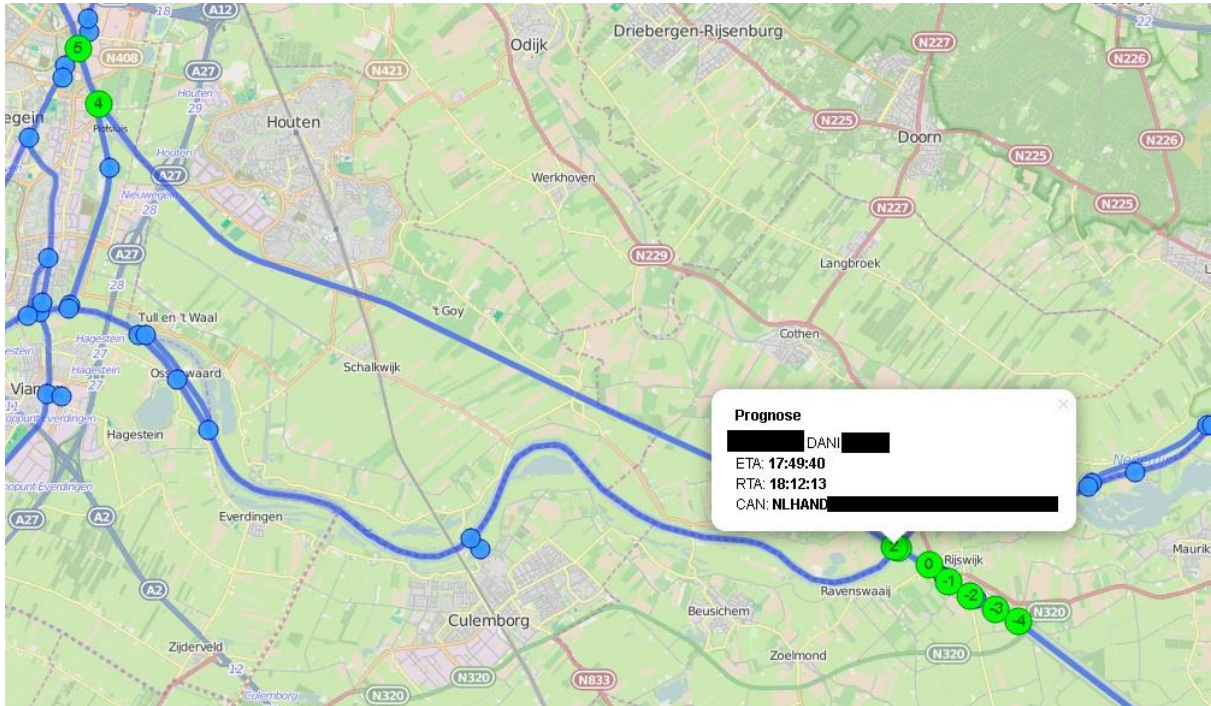


Figure 5-2 Data shown by the Trajectplanner

For normal dots the only data the Trajectplanner gives is the ETA for that point, for dots with a lock or bridge behind it, the Trajectplanner also gives an RTA (Requested Time of Arrival) for when the ship is requested. This RTA indicates at which time the ship has to be there in order to go through the lock or past the bridge. This RTA value can be used by skippers to decide if they can slow down or have to speed up.

The data which is shown in figure 5-2 is used to analyse the Trajectplanner. The Trajectplanner simulates every 6 minutes which means that every 6 minutes, new data is available.

In order to perform an analysis, a ship is chosen to follow for a couple of hours. The routes these ships take can be different, but often there is a lock within the route to see how this effects the ETA's. The ships are also followed in real time via AIS which is also used to see how busy the waterway is and other factors that can influence the simulation. During the analysis, comments are written down about speed, other ships on the waterway, location in front of a lock etc.

5.2 Analysis of the output

In order to start an analysis a ship is chosen and 2 or more points are chosen along the route to track. Every 6 min (or when new data is available) the ETA's and RTA's for every chosen point are recorded. These data points are put into a graph in order to see how the ETA fluctuates over time. An example with explanation is given on the next page.

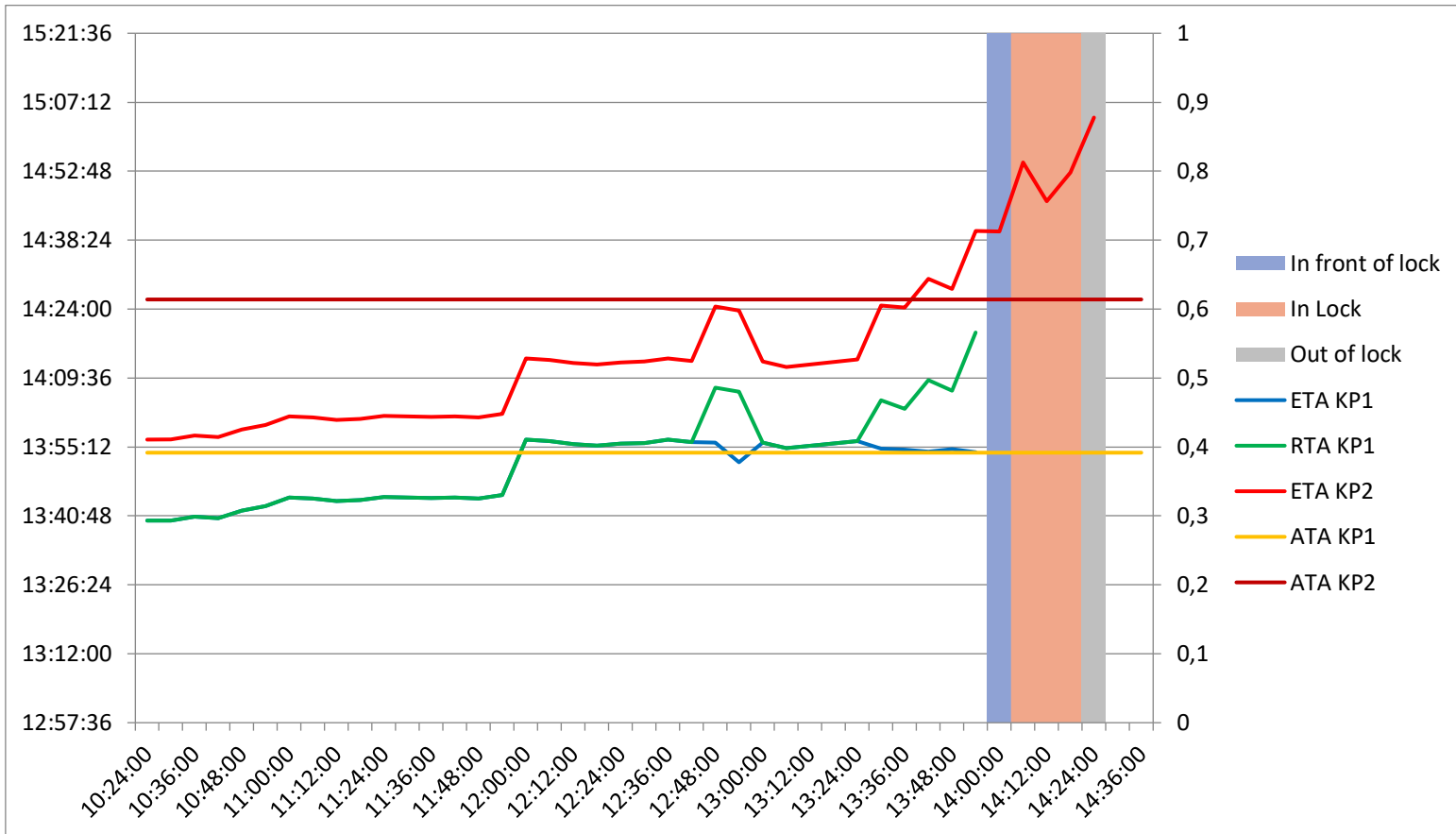


Figure 5-3 Graph showing ETA and ATA data from an analysis

On both the X and Y axis time is displayed. On the Y-axis the different ETA's are plotted and on the X axis, the simulation time is displayed. So in figure 5-3, the ETA for KP2 at 12:00:00 was 14:13:41

The ATA's (Actual time of arrival) are displayed as a horizontal straight line in order to clearly see the difference between the ETA and the ATA. The ATA itself is of course a single value only known after the ships has arrived at the node. But to make the difference visibly clear, the straight line is shown. The figure 6-3 above shows the ETA, RTA and ATA of 1 ship passing 1 set of locks. This means that before the ship arrives at the first point it doesn't encounter other locks or bridges which can influence its sailing.

In other analysis, a route is chosen where a ship encounters multiple locks and/or busy intersections in order to see how other locks can influence the ETA of different points. The bars in the graph show when the ships is before, in and out of the lock. This is to help analyse the graphs and to show what the ETA data does when a ship is near a lock.



Figure 5-4 Location of KP1 and KP2

The analysis shown in figure 6-3 is of a ship passing the Beatrix locks on 19-9-2017. In this analysis KP1 is the node south of the Beatrix locks and KP2 is the node north of the Beatrix locks. See figure 5-4 The lock itself is not shown in the figure but sits in between KP1 and KP2. When the analysis was started the ship was sailing on the Lek from Rotterdam. The ship would not encounter any objects until the Beatrix locks.

What can be seen in figure 5-3 is that the ETA and RTA for KP1 are equal for most of the time. What also can be seen is that the RTA for KP1 and the ETA for KP2 have an almost fixed distance between them.

ETA KP1, RTA KP1 and ETA KP2 all make a jump around 12:00:00. There was no explanation found for the 'jump' itself, but after the jump the ETA for KP1 becomes accurate within +/- 2 minutes.

The ETA KP2 keeps at a set distance from RTA KP1 and once the ship is near the lock, both the RTA KP1 and the ETA KP2

keep rising. Even when the ship is 'in' and 'passed' the lock, the ETA KP2 keeps rising.

As the Trajectplanner is not yet operational and thus not actively used by skippers, the RTA is not used in the other analysis.

Other analysis of ships sailing through the Beatrix locks are shown and discussed in Appendix B

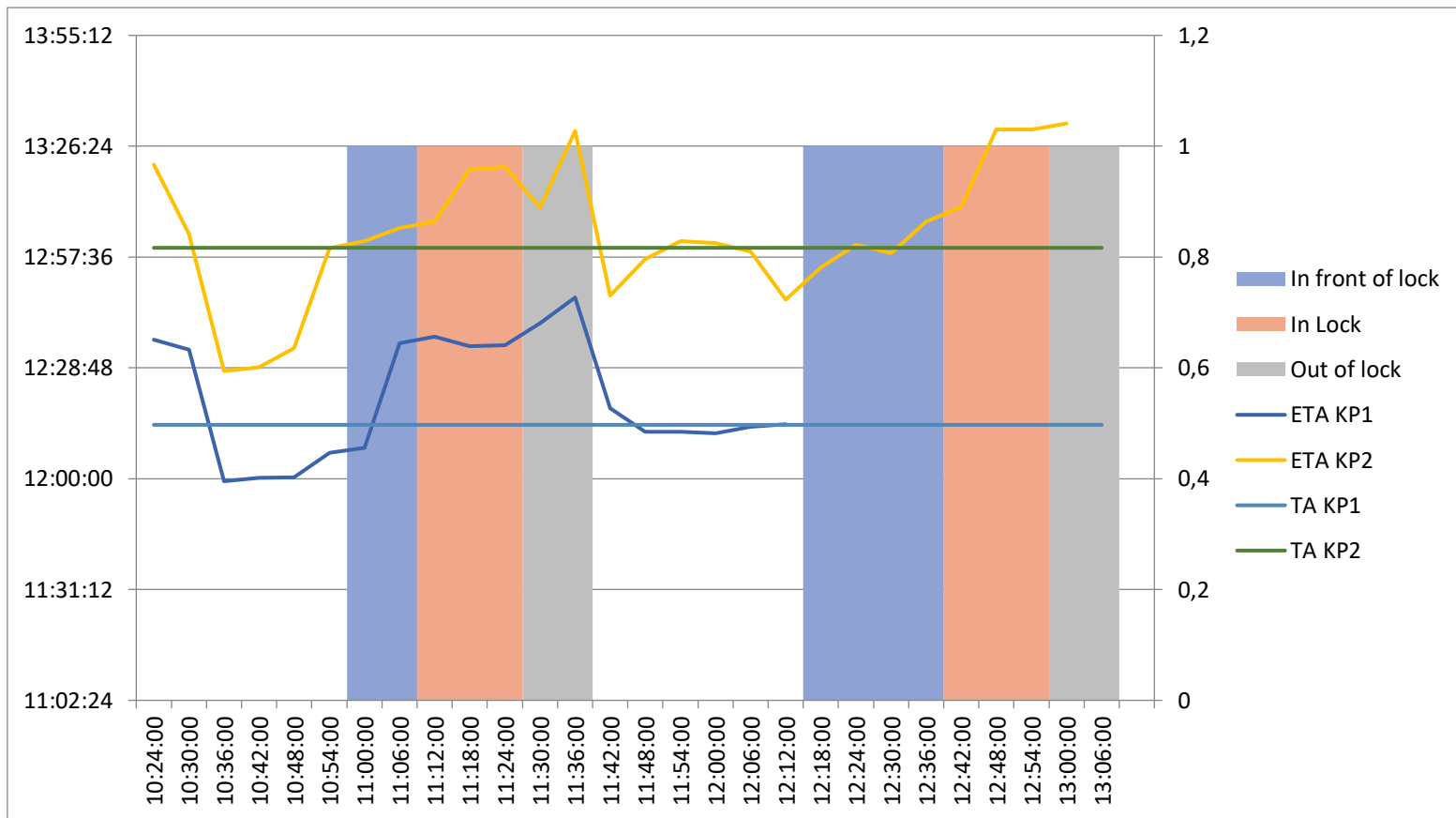


Figure 5-5 Graph showing ETA and ATA data from an analysis

Figure 5-5 represents the analysis of a ship passing through 2 locks, namely the Prinses Irene locks and the Prins Bernhard locks. The KP's that are mentioned in the figure above are the KP's north (KP1) and south (KP2) of the Bernhard locks. See figure 5-6



Figure 5-6 Location of KP1 and KP2

The figure can be confusing because there are 2 sets of bars indicating when the ship is before, in and after a lock. This is to indicate what the influence is of a set of locks before the first KP (KP1). To be clear, the ship passes through the Irene locks before it reaches KP1. And it is clearly visible that when the ship is in the first lock, around 11:08:00, the ETA's for both KP's fluctuate a lot. When the ship is out of the first lock, around 11:42:00, the ETA for KP1 becomes accurate within a few minutes. The ETA for KP2 keeps fluctuating and rising, even when the ship is in the Bernhard locks.

The analyses that were discussed above were somewhat short. Longer analysis were performed but they showed the same thing as in that the ETA started fluctuating when the ship is near a lock. If a ship sails on a waterway without objects to pass, the ETA is accurate within a few minutes. And because the Trajectplanner is not fully operational yet, it could happen that there was no data at all for a while, sometimes a few minutes, sometime a few hours. After talking to the developers, the cause of this was that when the Trajectplanner does not receive all the data for all the ships in the system, it can't run the simulation. This meant that some analyses only have a few data points which weren't enough to draw conclusions from.

Two examples of the situation described above are shown in Appendix B figure B-3 and B-5

Because the Trajectplanner receives data from several systems, it can also happen that the data is outdated. This happened during a long analysis of a ships journey. The route the Trajectplanner predicted was completely wrong because the destination the Trajectplanner received was wrong. Luckily the Trajectplanner got the right destination and made a new route which was analysed. This route had no locks or other objects on it for which the ship had to be scheduled. The graph of this analysis is shown below.

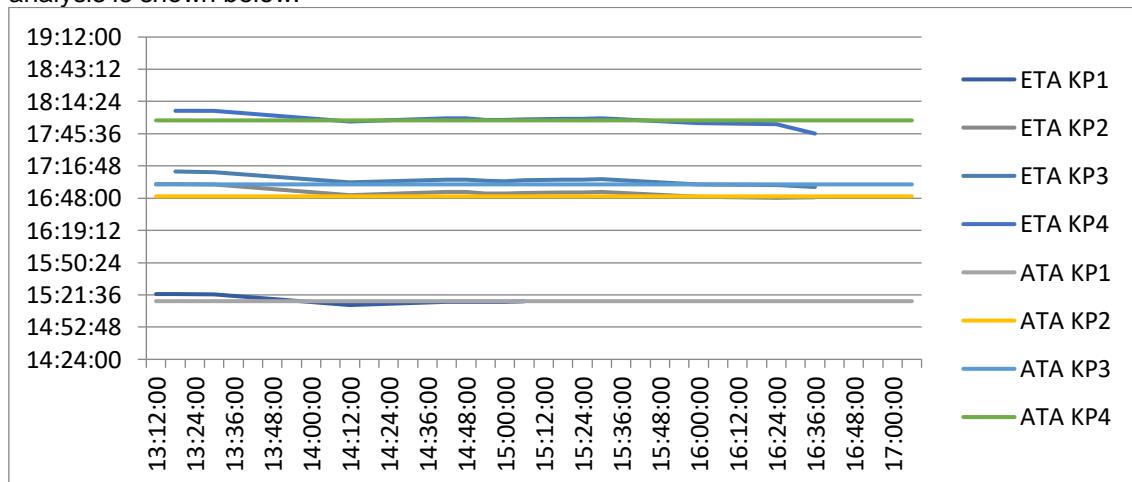


Figure 5-7 Graph showing ETA and ATA data from an analysis

The graph shown in figure 5-7 perfectly shows that when there are no objects on the route of the ship, the Trajectplanner is accurate within a few minutes. The ETA's for all 4 KP's stay almost constant and didn't fluctuate much. Also, the sudden rise in ETA when the ship comes near a KP did not occur

5.3 Results

After the analysis of these graphs, several conclusions can be drawn.

The different figures show that when a ship is near a lock, the ETA for the point after the lock keeps rising, as if the Trajectplanner keeps scheduling the ships into the lock without knowing the ship is already near, or even in the lock.

When there are no objects between the ship and the KP that is analysed, the ETA for that KP is accurate within a few minutes and doesn't change very much.

The RTA for a point keeps changing which indicates the schedule or planning changes every time.

This makes sense as the Trajectplanner keeps no record of previous simulations.

Because of the first point made above, if a ship encounters multiple locks, the ETA for the end point will fluctuate a lot. This was seen when a ship was analysed with multiple locks on the route.

Factors like other ships, wind and current are not taken into account by the Trajectplanner but will be reflected in the current speed of the ship. This was confirmed by reviewing the programming code of the Trajectplanner.

5.4 Accuracy of the current ETA estimation (Trajectplanner)

To determine the accuracy of the current system, the KPI which was defined in chapter 3, was used on the analysis of the output of the system. As said in chapter 3, the KPI is defined by

$$KPI_x = MAX|ETA_x - ATA_x|$$

And for a data set or analysis of a ship's journey, this would give a table like this:

	KPI KP1	KPI KP2	KPI KP3	KPI KP4
Max KPI	00:12:34	00:32:05	00:28:19	00:31:28
Max KPI(seconds)	754	1925	1699	1888

Table 5-1 Example of KPIs from one journey taken by one ship

What table 5-1 shows is the KPI for the analysis that is shown in figure 5-11. Table 5-1 shows the KPI for each KP at maximum time difference between simulation time and arrival time.

This table shows that the MAX KPI across these 4 points was 32.05 min/1925 sec. This was for KP2. This table does not show the progression of the error, else the biggest KPI would occur at KP4. This could be because at the time when the error occurred at KP2, the schedule for the lock before KP4 could be empty, this means that the error at KP2 has less influence on the scheduling than when the schedule would be full of ships

In Appendix B, a table (table B-1) is shown with all the KPI's calculated from the analysis.

From this table, a few conclusions can be drawn:

- The max KPI shows that the Trajectplanner can estimate the ETA with an error up to an hour.
- It also shows that the difference does not stack. For example, if the max KPI at KP2 (KP after the first lock) is 40 min, it does not mean that the max KPI at KP 4 (KP after the second lock) is 80 min. The table shows that the max KPI is around 40 min for both locks.

The answer to the question: **How accurate is the current prediction?** For now is:

If the route of a ship has no locks, the Trajectplanner is accurate within a few minutes (+/- 5min, with exceptions)

If the route of a ship does have lock(s) on it, the accuracy of the Trajectplanner fluctuates with a maximum of about an hour.

5.5 Influence of a lock

So far, the analysis of the system suggests that the locks within the system are the biggest source of inaccuracy for the prediction of the ETA.

To confirm this, an analysis was performed on one lock for the duration of 6 hours and the ETA for all the ships that came through the lock were monitored. The results of this analysis are discussed in this section.

The analysis was performed by monitoring the north and south node at the Bernhard locks. See figure 5-8



Ships in both directions were followed and split into 4 graphs. The resulting graphs are displayed in Appendix B figures B-6 through B-9. The graph showing the ships towards the lock are coming from outside and don't encounter objects before the Bernhard lock. The graph showing the ETA for ships away from lock are ships that just went through the lock or are going through the lock.

From the previous part of the analysis it is expected that the ships that sail toward the lock have almost straight lines in the graph, meaning that their ETA is calculated correctly from the start of the analysis.

It is also expected that ships sailing away from the lock have a rising ETA when they come near the lock and the node.

The names in these figures are anonymised.

Figure 5-8 Location of South and North node

The ETA's and ATAs from this analysis were also compared to determine the KPI for each journey and each ship. These were then used to get the KPI averaged over several ships making the same journey. The table with these results is shown below. (Table 5.2)

Table 5-2 KPI table from lock analysis in seconds

	Mean	Median	Standard deviation
<i>KPI for ships sailing from the South, towards the lock</i>	3676	3717	1387
<i>KPI for ships sailing from the North towards the lock</i>	1863	1044	3461
<i>KPI for ships sailing from the South, away from the lock</i>	3249	2083	3232
<i>KPI for ships sailing from the North away from the lock</i>	3118	3086	467
<i>KPI for ships sailing towards the lock</i>	2769	2533	2729
<i>KPI for ships sailing away from the lock</i>	3193	2495	2416
<i>KPI for all ships sailing through the lock</i>	2986	2495	2550

Before this analysis was done, the expectation was that the ships sailing away from the lock would have a large KPI, because the lock would cause the greatest error.

What can be seen in the tables above is that this is not fully the case, in the case of ships sailing from the North the difference is clear, a median KPI of 1044 sec = 17 min towards the lock and 3086 sec = 51 min away from the lock. In the case for ships coming from the south, it almost the other way around. This could be caused by the fact that the ETA calculations done by the Trajectplanner can sometimes jump around a lot which can be seen in the graphs shown in Appendix B and in this chapter. The standard deviation shows that there is a lot of variance in the KPIs between ships. Only in the case of the ships sailing from the North away from the lock is it clear that they have a KPI of almost 1 hour

A few remarks after this analysis:

- Some lines stay close together, these are ships that 'sail in a group' and have therefore the same ETA predictions.
- The ships going through the lock and away from it, often have rising ETAs. This was expected after the previous analysis.
- The lines of ships sailing towards the lock aren't as straight as expected, but the ETAs still move towards the ATA or hover around it.
- The graphs show that there is quite some fluctuations in the ETA calculations for each ship and the KPI data shows there is a lot of variance in KPIs between ships as well

5.6 Analysis conclusion

After this analysis it is clear that locks are a big factor in the accuracy of the calculated ETAs. There are several factors that play a role in this problem

- The Trajectplanner doesn't know when a ship is in front, in or out of a lock and keeps planning the ship into the lock. What happens is that the ETA for the next node keeps going up with the passing time of the lock. This means that the Trajectplanner doesn't seem to know where on the arc a ship is when there is a lock on the arc.
- Because the Trajectplanner doesn't know what it has previously done it makes a new schedule every time it simulates, even if the previous schedule was good, the system doesn't know
- Missing or incorrect/old data is a problem for the Trajectplanner because the simulation cannot be done then. This is not directly the fault of the Trajectplanner, but it is a problem.

Section 5.4 gave a preliminary answer to the subquestion: **How accurate is the current prediction?**

This was after several analysis of different ships taking different journeys.

The analysis performed in section 5.5 give a better view of the accuracy of the system because it analysed one route (going both ways) with multiple ships.

The answer to the subquestion: **How accurate is the current prediction?** Would now become 00:49:46 or 2986 seconds (see table 5-2 value for all ships sailing through the lock)

6 Alternative models for ETA prediction

With all the information that was gathered in the previous chapters, there are some aspects where improvement of the Trajectplanner is possible. This chapter will explore and discuss those aspects. In order to fully understand the alternative models, the issues and problems that come with those will be explained.

This chapter will answer the subquestion: **How to improve the current prediction?**

6.1 Inaccurate ETAs

The reason why inaccurate ETAs are a problem for the correct functioning of the Trajectplanner are explained in chapter 1. If one looks at the analysis in chapter 5 it can be seen that the ETA for the point behind the lock can fluctuate and vary a lot. What this means is that all the ETAs for the points after will have the same fluctuations. If on the entire route, there is only 1 lock, these fluctuations are only a problem for the nodes after the lock and for the next shipper at the destination of the ship. But if there are more locks after the first lock. This becomes a problem for all the other ships that go through the other lock(s). This is because the schedule for a lock is based in the ETA's the Trajectplanner calculates. So when they start to fluctuate, the schedule which is made with those ETA's is of no use. This also means that, when ship 1 approaches lock 1 and the ETA for after the lock is fluctuating. And ship 2 is sailing towards lock 2 which is on the same route as lock 1 but further on the route, the schedule for lock 2 in which both ship 1 and ship 2 are present will be inaccurate (illustrated in figure 6-1). So this shows that when the ETA of one ship for one point start to fluctuate, multiple ships and multiple routes are effected by this inaccuracy.

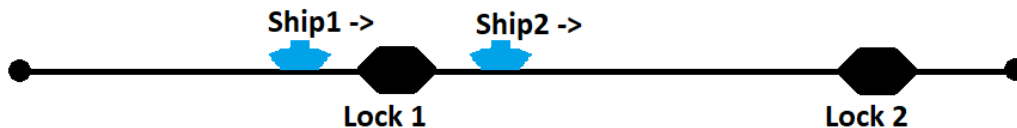


Figure 6-1 Illustration of two ships passing the same two locks

6.1.1 Ideal situation

In an ideal situation, all the ETAs that are calculated at the start of the journey are accurate and don't change much during the journey. This would mean that if the same analysis was performed as in chapter 5 all the lines would go straight (or hover around the ATA) as in figure 6-2



Figure 6-2 ideal output from ETA predictions/calculations

This would mean that the schedule that are made with these ETA's will stay the same and skippers can start to use the schedules and improve their efficiency.

A analysis that came close, was the one that didn't have any objects on it in chapter 5, figure 5-7

6.2 Network model

As was mentioned in chapter 4, the model which is used to model the Dutch waterway network consist of points and lines, which could be seen as a node and arc network. But, the way it is modelled it is not a true node an arc network, for instance, the objects are not modelled on the network, but rather exist somewhere on a line. This is the first area where improvements can be made.

6.2.1 Modelling the system

To describe this system, one can take an example from the previous chapters where it was shown that a system like this can best be described by nodes and arcs.

For our system we need to model the waterways and the objects. This can be done in two ways:

6.2.2 Model 1:

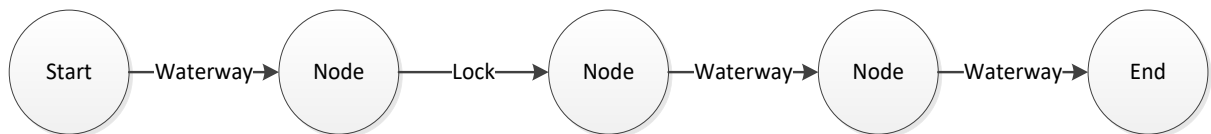


Figure 6-3 Model 1

In Model 1 the nodes represent points on the waterway, these points can be intersections or points before and after an object. The waterways and objects are modelled as arcs.

In this model the nodes represent points on the waterway with no other property than a position. Ships leave a node at the same time they arrive at it.

The arcs in this model can represent a part of the waterway or an object, like a bridge or lock. This means that the arcs can have the properties of a waterway (length, depth, width, etc.) or the characteristics of an object, for example in the case of a lock, the arc has properties like length and width of the chamber, x and y position in the chamber. For both type of arcs (waterway and object), it takes the ship a certain amount of time to pass the arc. This depends on the properties of the arc, but also on the properties of the ship. If a lock has multiple chambers, then this model could model them with multiple arcs. This can also be seen in Zhou et al.(2016) (15) which used arcs to describe the multiple railway tracks within a station

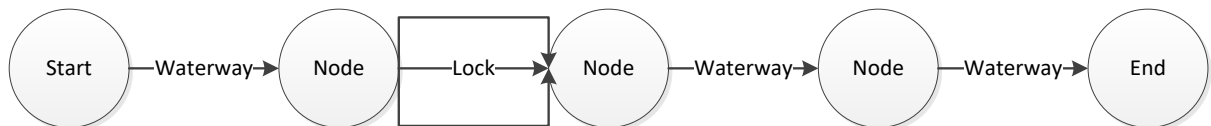


Figure 6-4 Model 1, lock with multiple chambers

This would give the advantage of giving each lock chamber their own properties.

By moving and showing the ship along the arc, one could see how long it would take the ships to reach the next node. In the case of waterways, it would show how far the ship still needs to go and in the case of a lock it would show where in the lock procedures the ships is.

The nodes in front and after the lock would give a clear trigger point of when the ship is entering the lock and when it leaves it again.

6.2.3 Model 2:

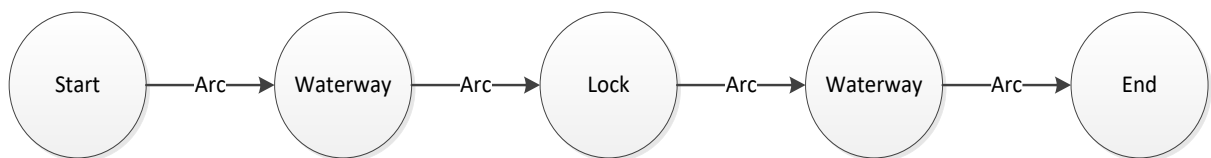


Figure 6-5 Model 2

In Model 2 the arcs represent the connections between different nodes and the nodes can represent a point on the waterway, or an object on the waterway. Where an object could mean a lock or bridge, but also a start or end point, like a harbour or dock. In this case it is the nodes which take time for the ship to pass. If a ship is on a node, it would mean it is in the lock, under the bridge or sailing on the waterway. In model 2 there would be no way of showing where on the waterway the ship is.

6.2.4 Chosen network model

After the research and analysis that has been done in the previous chapters, it is suggested that the Dutch waterway system is modelled as in Model 1. Where nodes are points on the waterway and arc represent the waterways and objects. With this model the different paths a ships can take, the different lock chambers a lock can have and the progress a ships has made on an arc can clearly be modelled as well as visualised.

A more detailed description of this model would be:

The Dutch inland waterway system is modelled with a network of nodes and arc. In this model, the nodes are a single point and don't take any time to pass. The arcs in this model can have different properties which determine the time it take to travel the arc.

As mentioned before, the inland waterway system consists of natural and man-made rivers and canals. These waterways connect ports to other ports or to other countries. A ship will enter this system of waterways from the port or from another country. If one generalizes this system, a ship will sail from port to port. Where port can mean anything from a big port like Rotterdam to a single dock which belongs to a factory.

A port in this case will be modelled as a node. The waterways and will be modelled by the arcs connecting these nodes. The Dutch waterway system consist of several rivers flowing to the sea. These rivers are connected by locks in order to regulate the water flow and levels. Other waterways have locks on them just to regulate the water levels on each side of the lock. This means that ships will encounter locks on a regular basis and that these are an important part of their planning. These locks will be modelled in this model by arcs. Other objects like bridges or open locks will also be modelled as arcs. Their travel time will be quite short. A waterway arc can have different lengths,

representing the length of the waterway. Object arcs will be quite short, with a node right in front and right after the object, representing the 'start' and 'end point' of the object.
 A fictional network of an inland waterway system is illustrated in figure 6-6

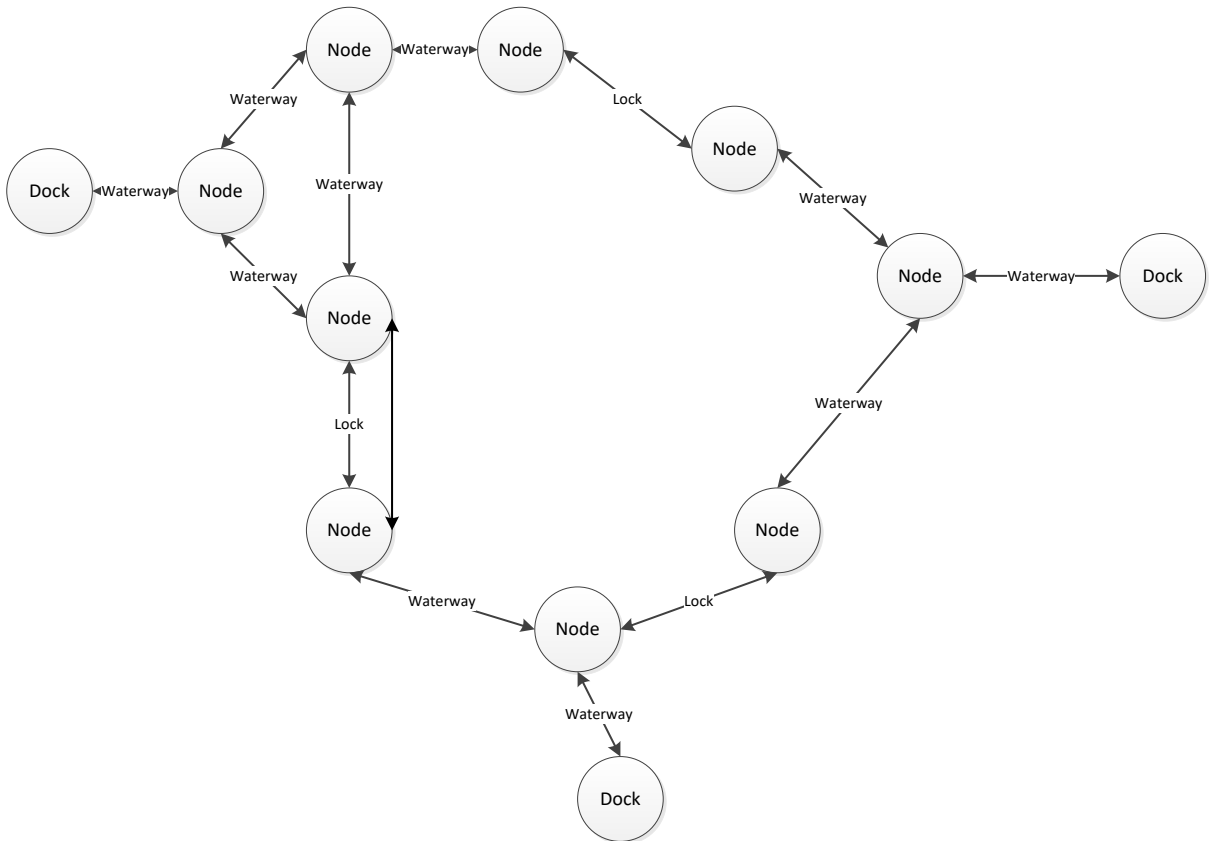


Figure 6-6 Proposed model

Figure 6-6 above is fictional network of docks, nodes, locks and waterways.

Nodes

The nodes in this model are 'timeless' and don't take time to pass. They indicate a start or end point and are used for intersections and before and after objects.

Arcs

The arcs in this model represent the waterways and objects and are bidirectional. Both of which take time to travel.

For waterways, this travel time depends on the type of ship and the type of waterway. The type of ship indicates its length and width; the maximum speed of the ship; the draft when it's loaded; the draft when it's empty.

The type of waterway indicates its length and width; the water level; the maximum sailing speed for ships.

For objects, this travel time depends on the type of object and the type of ship.

The type of ships indicates the same as for waterways plus when the ship carries dangerous materials, there are rules to follow when it comes to placing ships in locks with dangerous materials.

The type of object indicates what object it is (lock, bridge, open lock); if the ship can sail through (open lock, high enough bridge) or it has to wait; and whether there are other procedures (lock procedures).

It can be seen that one of the locks has two arrows. This indicates that that lock has two lock chambers. Each lock chamber can be operated independent of the other one. Although operators in real life will try to operate them parallel, so when one goes up, the other goes down. This way the water flow on each side of the lock stays the same.

6.2.5 Network model improvement

The network model used by the Trajectplanner as a representation of the Dutch inland waterways is somewhat incomplete. The main issue is that objects like locks and bridges are not modelled as an arc or node, but rather lay somewhere on an arc. This means that when an arc has an object on, the arc has the properties of that object, but also the surrounding water. This makes it difficult to estimate the arrival time at the next node because the object has no location on the arc and it is therefore impossible to know when a ship has passed the object. This is most likely one of the greatest causes of the ETA inaccuracy seen in chapter 5. In order to solve this and to improve the accuracy, the network model needs to be altered in such a way that objects are modelled as well. It is suggested in this section that this is done by modelling an object with 2 nodes and an arc between them. See figure 6-6. This makes sense, because the Trajectplanner already uses a network model where the arcs take time to travel and not the nodes. For locks, these nodes can lay on the position of the doors. This way, the RTA for such a point would make much more sense as it is the requested time at the door. Not the requested time for a node which can be kilometre away. For bridges these nodes can lay on the spot where ships would normally have to wait for the bridge to open. For bridges that can't open, a single node would suffice as it would only indicate the position of the bridge and would occur no extra travel time as with other nodes. Around the Beatrix lock, this would look something like in figure 6-7.



Figure 6-7 Left: Current network model/Right: New network model

The left picture shows the current network model used by the Trajectplanner and the right shows the proposed network model. The extra blue dots(nodes) indicate where the lock is situated. This can be used by the Trajectplanner to know if the ship is in front, in and passed the lock.

This should improve the ETA calculation in the following way. As can be seen in chapter 5, the ETA keeps rising when the ship is on the arc with the object on in. So if everything else stays the same, with as new network model, the ETA only starts rising when the ship is at the starting node of the lock. For example, if a ship travels the route passed the Beatrix lock going from south to north (bottom of the figure to the top of the figure). as shown in figure 6-7. In the current model, the ETA would start rising when the ship passed node 1 and wouldn't stop rising until node 2. With the new network model,

the ETA would start rising when the ship passed node 2, until node 3. This would still mean an inaccurate ETA, but far less than the current inaccuracy.

6.3 Lock model

The next improvement will be discussed with the assumption that the new network model as discussed in the previous section is implemented.

To estimate the ETA for nodes behind the lock, the Trajectplanner schedules ships into the lock. This way, it can estimate when the ship will be through the lock and sailing towards the next node. To make this schedule, the Trajectplanner takes into account the arrival time (estimated), the in- and outsail times of all the ships and the dimensions of the ship. The dimensions are important as they are the key component that determines how many ships will fit into the lock at the time. With all this information, the Trajectplanner makes a schedule and gives out an RTA for all the ships sailing towards the lock. But with the current system, once the ships has passed the node before the lock, this RTA is no longer needed and the Trajectplanner doesn't know where the ship is in relation to the lock. With the new network model implemented, the system would know that for example in figure 7-7 right. If the ships sails from 1 to 2 to 3 to 4, when the ships is passed node 2 and not yet passed node 3, the ships is inside the lock. This information does not help yet with estimating the exit time of the ship because if the system doesn't know when the ship entered the lock, it can not estimate the exit time. So an improvement would be an extra data point which tells the system when the ship entered the lock, so the system can then estimated when the ships will exit the lock. This should stop the rising of the ETA, once the ship has entered the lock.

6.4 Memory

As mentioned in chapter 4 the Trajectplanner doesn't remember data from previous simulations and starts over every 6 minutes. This means that the schedules made for all the ships are made every time the simulation runs. This is shown in chapter 5 were the RTA for the node before the lock changes a lot. An improvement should be that the Trajectplanner has a memory where it can store the schedule that it made. With this memory it could check if all ships can still make their scheduled times. And if so, no other schedule should be made. If this was the case, the schedule could be passed on to the skippers and they could make sure that they would make that scheduled time.

The memory could also be used for more improvements, like if the ETA for a specific node changes a lot from one simulation to the next, it could signal that something is wrong or has changed. This could be used by the Trajectplanner to check calculations, data or send a warning to someone who can check it out.

This memory would also help with monitoring the system as it could later check the ATA with all the ETAs it calculated.

The improvement in planning could stabilize the ETA estimation for the nodes behind an object. It would not directly improve the ETA calculation. Although at this point, stabilization of the ETA estimation would be an improvement.

The other things that can be done with a memory would not directly improve or stabilize the ETA estimation, but would help with analyzing the system and spotting problems.

6.5 Conclusion

This chapter looked into different aspects of the current simulation and planning system, the Trajectplanner and ways to improve the system with the knowledge of the previous chapters.

The answer to the subquestions: **How to improve the current prediction?** Is to implement the new network model and to give the system data about when ships enter the lock. With these improvements it is expected that the ETA estimations will improve.

The memory would probably help to stabilize the ETA estimation, but not directly improve it.

More improvements could be made, like taking into account more parameters when estimating the ETA (see chapter 4, table 4-1), but it is expected that the improvement mentioned in this chapter will have the biggest impact on the ETA estimation

7 Evaluation method

To evaluate whether or not the alternative models thought of in the previous chapter will actually improve the estimation of the ETA, a way of calculating the movements of ships on a network and estimating their arrival times needs to be found. This also answers sub-question: **How to evaluate a new prediction model?**

Unfortunately, it was impossible to test the improvements on the current system, the Trajectplanner. It is possible to calculate the passage of 1 ship through a lock, but when multiple ships sail passed the same lock in two directions, the calculations become very complicated because of all the dependencies between ships and the lock. It was therefore needed that a small simulation environment was build using discrete event simulation to simulate all the movements of the ships. How this simulation was made and works will be explained in this chapter. The experiments that were done to test whether the improvements actually improved the estimation are discussed in the next chapter.

7.1 Programming language and modules

This simulation is made in the programming language Python. Because Python is open source and free there are a lot of extra modules which are made for Python and can be used for free.

One of these modules is Salabim. Salabim is a discrete event simulation module which is made by Ruud van der Ham (24). This module is used to make the entire simulation.

Other modules which were used are:

XlsxWriter : used to write the simulation data to excel

Random : used to create 'random' numbers

7.2 Simulation environment

The simulation tool has to simulate the movement of inland ships sailing over a piece of waterway with a lock on it. To do this, discrete event simulation is used to create a simulation environment. The waterway is modelled as a straight line with the only property being its length. The lock is situated on this line. The lock and ships are modelled as components of the simulation. See figure 7-1



Figure 7-1 Schematic of the network model

Components can have different properties like size, location and many other properties which will be discussed later. Another important component within the simulation environment is the estimator. This component tries to estimate where the ships will be and at what time they will be at either A or B. It calculates and estimates the ETA for these points. These points can be defined in the simulation and will be modelled after a real-life scenario.

Other components that are used are: ship generator (generates ship on each side of the waterway), multiple queues, ATA monitor (monitors whether a ship has reached certain points and saves the ATAs) and a monitor that monitors different statistics of the lock.

The simulation simulates certain things accurately, other things are simplified and some things are not modelled. These are listed below.

Modelled accurately:

- speed of ships
- length of ships
- in- and outsail times of ships at the lock
- lock size
- service time of the lock

Simplified:

- area/capacity of the lock chamber
 - o The capacity of the lock chamber is modelled only with length
- The exact procedures at a lock
 - o The procedures are simulated as a constant length of time
- Placement and room between ships in locks
 - o The rules for ships carrying dangerous materials are not modelled
 - o The room between ships in a lock is neglected. For example, in a 300 m lock, 3 ships of 100 m will fit.

Not modelled:

- Influences on the sailing of ships
 - o Ships will have their own speed which will not be affected by other ships, wind, current or other influences.
- Small pleasure vessels
 - o Small ships and pleasure vessels that would have AIS in the real world are not modelled
- Curves and corners in the waterway
- Rules for watermanagement
- Draft and cargo of ships

7.2.1 Inspiration

The inspiration for this simulation came from the creator of the Salabim module. He has made simulations of a lock. He made the simulations in different ways, so his simulations were used as inspiration and reference for this simulation. Especially the ship and lock components are based on his programming. The components are somewhat altered to use in this simulation environment, but the working principle is the same as his simulation.

7.2.2 General description

When the simulation starts, ships are created at the two ends of the network. The ships start to sail until they encounter the lock. They will then want to go through the lock. After the ship went through the lock it will then sail on to the end. When it reaches the end, the ships will stop.

As long as the simulation runs, the lock is active. The lock checks whether there are ships waiting to go through the lock on both sides of the lock. When there are no ships, the lock will passivate itself. It can be made active again by a ship requesting to go through the lock.

When a ship wants to go through the lock, the lock will check whether the water is on the 'right' side. If not, it will go through the procedure empty. It will then open the doors and let the ship(s) in. When the lock is full or no more ships want to go through on that side. The lock will close the doors and start the procedures. When the procedures are done. The lock will open the doors on the other side and let the ships out. Then the process starts over again by the lock checking whether any ships want to go through.

During the simulation, an estimator is running as well. Simply said, this estimator will estimate the ETAs for all the ships in the system.

The above is a simple description of the simulation of the different components. What can be seen is that the ships, the lock and the estimator don't influence each other. The ships have full control over their own speed. The lock only gives permission to sail in or out, but does not actually tell the ships to sail in. The estimator only looks at the data from the ships and makes calculations based on that.

7.2.3 Components

In this section, the components: ship generator, ships, lock and estimator will be explained in more detail as they are the key features of the simulation.

7.2.3.1 Ship generators

The ship generators are the component that generate the different type of ships at different time intervals. There are two, one at each end of the waterway. They create ships with a random type which determines the ship's length, speed, in- and outsail times at the lock and their starting and end position. The generators wait a 'random' amount of time before creating the next ship. Each ship also gets some lists and a dictionary which are used to store different types of data in, in order to later analyse the journey of the ship.

Process of Ship generator

While the simulation is running, do:

Wait a random amount of time (Triangular distribution between 12 and 60 min, with an average of 15 min)

Create ship with:

Name *#Name the ship*

Side *#The side of the ships is the side of the generator, East or West*

Type *#The type is given at random where all types have the same probability*

Length

Speed

Insail

Outsail

Startpoint

Startlocation

Endpoint

Endlocation

Record

Multiple lists

Set eta1 through eta4 to 0

If side of ships is West

Node1=NodeA

Node2=NodeB

Node3=NodeC

Node4=NodeD

Else

Node1=NodeD

Node2=NodeC

Node3=NodeB

Node4=NodeA

7.2.3.2 Ships

As said above, the generators create the ships with all their attributes. These ships then start sailing. Each time unit in the simulation environment is seen as a second, so the speed the ships have is in m/s. The ships keep on sailing until they reach the lock, or their end point. Nothing in the simulation alters the ship's speed or location other than the ship itself. This was done on purpose as it is accurate to the real-world situation. When the ship reaches the lock, it will request the lock door to open and request space in the lock. When the water is at the 'wrong' level, the ship has to wait until the lock makes a 'switch' and releases the doors and space in the lock. When all this is done, the ship can sail in, wait for the lock to make the switch and sail out again. When the ship has sailed in, it will immediately request the other door to open. Of course, this can only be done when the lock releases it after the 'switch' has been made. The ship will then sail until the end point is reached. At this point in the simulation, the ship component will take all the data which is collected and write it to an excel file. The speed of the ships has a slight variation in it, which is implemented by normal distribution with an average of 0 and a standard deviation of 70% of the 'standard' speed.

Process of Ship

```
Start the record
Mode = Sailing to lock
Add self to allshipsq
While location is not the same as the lock door
    Wait 1 s
    New location becomes location + speed + random variation of percentage of speed
Location becomes the location of the door
Enter self in Key_wait(Side)
Mode = waiting to enter lock
If lock is passive, activate lock
Request room in the lockchamber and the door to be opened
Leave Key_wait(Side)
Mode = In Lock procedures
Enter Lockchamberqueue
Wait Insail
Location = 0
Release door
Request door open other side
Wait Outsail
Release door other side
Location becomes location of door on other side
Mode = Out of lock procedures
While location is not in range of endlocation (+/- 5)
    Mode = Sailing to end
    Wait 1 s
    New location becomes location + speed + random variation of percentage of speed
ATAend = now
Write all data from lists to excel file
```

7.2.3.3 Lock

As mentioned with the ship component, the ships control themselves. This is also true for the lock component. The lock constantly checks whether any ships want to go through on each side of the lock. When there is a ship waiting, the lock will release the doors and the lockspace and let the ships sail in. Both the doors of the lock and the space in the lockchamber are modelled as resources. The locks of the door are modelled as a resource with a quantity of 1 and the lockspace is modelled as a resource with a quantity with a value based on the length of an actual lockchamber. When ships arrive at the lock, they will request the door and the lockchamber. Because the door has a quantity of 1, only 1 component at the time can claim (use) it. This means that only 1 ship at the time can sail through the doors. The ship will request his length as a quantity to claim from the lockchamber. So when a ship is 100 m long, and the lock chamber is 300 m long, the ship will request a quantity of 100 from the lockchamber resource with a quantity of 300. This means that after the ships has gone into the lock, 200 m is left for other ships to fill. When the ships that were waiting have sailed in, the lock takes back the door and starts the lock procedures or 'switch'. When this is completed, the other door are released by the lock in order for the ships to sail out again. The whole process then start again with the lock checking the queue on the side it is now on.

Process of the lock

Request all the doors

While simulation is actives

If length of Key_wait(LockSide) == 0

If length of Key_wait(-LockSide) == 0

Become passive

Release the doors on the lockside

Wait 300 s

Request door on lockside with high priority

Release lockchamberspace

Wait Switchtime

Lockside become – Lockside

Release the doors on Lockside

Request doors on Lockside

7.2.3.4 Estimator

The estimator is the component which can be seen as the Trajectory planner of this simulation tool. It tries to estimate when ships will arrive at certain points. It does this by checking the position, the distance to the next node and the speed of the ship. The estimator is the key component in this simulation tool and different versions are used to test improvements. The key difference between these versions is what they use to estimate the ETA. The first version is used to replicate current Trajectory planner. It does this in a simplified way, which means that it will estimate the ETA, but it will not schedule ships into the lock procedures. This is done because after the analysis done in the previous chapters, the inaccuracy in the ETA estimation is thought not to be caused by the scheduling. The estimator does try to estimate the amount of time the ships have to wait before entering the lock. The basic process of the estimator is shown below. This process is highly depended on the version of the estimator that is used in the simulation environment as the improvements are made mostly in the process and code of the estimator.

Process of the estimator

While the simulation runs

For all ships in allshipsq

Depending on ship's location

Determine travel distance to node 1

Determine travel distance to node 2

Determine travel distance to node 3 (only with new network model)

Determine travel distance to node 4 (only with new network model)

Depending on ship's location

Basic

$ETA1 = now + \text{traveldistance} / \text{ship's speed}$
 $ETA2 = (ETA1 \text{ or } now) + \text{distance between node 1 and 2} / \text{ship's speed} + \text{switchtime}$

New network model

$ETA1 = now + \text{traveldistance} / \text{ship's speed}$
 $ETA2 = (ETA1 \text{ or } now) + \text{traveldistance 2} / \text{ship's speed}$
 $ETA3 = (ETA2 \text{ or } now) + \text{distance between node 3 and 2} / \text{ship's speed} + \text{switchtime}$
 $ETA4 = (ETA3 \text{ or } now) + \text{traveldistance 4} / \text{ship's speed}$

Lock awareness

$ETA1 = now + \text{traveldistance} / \text{ship's speed}$
 $ETA2 = (ETA1 \text{ or } now) + \text{traveldistance 2} / \text{ship's speed}$
 $ETA3 = \text{Waiting time} + \text{switchtime or Time left in lock}$
 $ETA4 = (ETA3 \text{ or } now) + \text{traveldistance 4} / \text{ship's speed}$

Add the ETAs to the ship's ETA lists

7.3 Output

The output of this simulation is mainly excel files. There are some lines of code which print some text and data to the console.

Every ship is a component itself and every one time unit in the simulation (which represents a second), the ship records its own position and mode. The mode tells something about what the ship is doing, for example the ship is sailing towards the lock, or is in the lock. The ETAs estimated for each ship are also recorded in the record of the ship itself.

All this data is stored in an excel file of the ship and is then used to make graphs of the different ETAs and then used to determine the KPI for each node.

7.4 Conclusion

The answer to the question **How to evaluate a new prediction model?** Is with a simulation tool. Ideally with the Trajectplanner itself, but because this is not possible, the simulation tool discussed in this chapter was made to evaluate new models.

This simulation tool in combination with the KPI defined in chapter 3 will be used in the next chapter to evaluate the new models

8 Evaluations

This chapter will describe the process of testing and experimenting with some of the alternative models from chapter 6. Each experiment will be discussed by first describing the experiment, the model that is used and then explaining the estimator that was used. After which the results are presented and finally an analysis of the result will be described. This chapter will answer the subquestion: **How accurate is the new prediction model?**

As in chapter 5, for each point on the network, a KPI can be determined. During this evaluations, the KPI was determined for each point in the network model.

8.1 Experimental setup

This section will explain different aspects of the experiments, such as an overview of the model which was used and which real world case was used, which improvements were tested and what experiments were done

8.1.1 Real world case

The scenario/case which was simulated was the Prins Bernhard lock and a stretch of waterway where this lock is situated.

The lock sits at the south end of the Amsterdam-Rijn canal and joins this canal to the Waal river.

The lock has about 25000 ships movements a year (25)

The length of the lock chamber = 350 m (26) and the service time is 720 s (=12 min)*

*The service time which was used is an average from several sources. Namely a reference in Li et al.(2017) (12), number of switches per hour (2) and following ships with AIS data.

The scenario which is modelled in the simulation is a stretch of waterway which is 4000 m long. A coordinate system is used with one axis. The lock will be situated on point zero, and ships will be generated from point -2000 (West) and point 2000 (East). This distance was chosen in so that the original nodes closest to the Prins Bernhard lock in the Trajectplanner model (figure 8-1) could be incorporated.

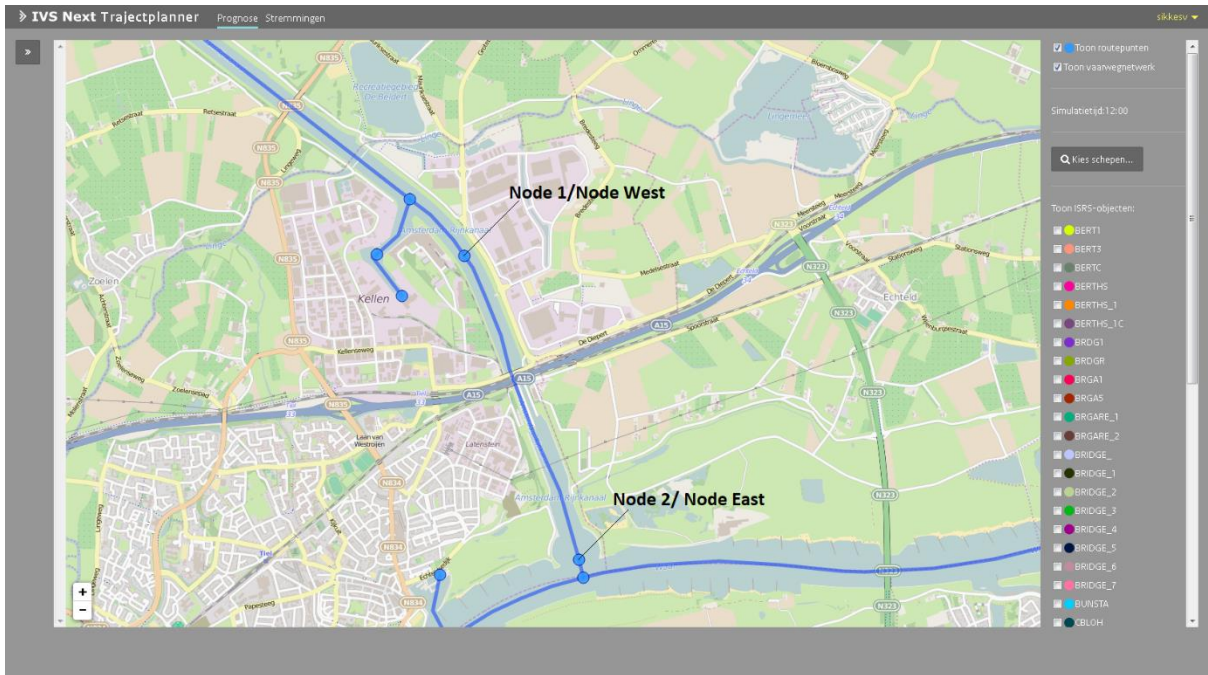


Figure 8-1 UI of Trajectplanner with location of Node 1/West and Node 2/East

8.1.2 Ships

The ships in the simulation are simulated after real type ships. This is also done in the Trajectplanner. The ships are modelled after the types mentioned in chapter 1.

The exact types were taken from the source code of the Trajectplanner.

For every type, an average length was used.

The speed, in- and outsailtimes were also taken from the source code of the Trajectplanner.

The end result was the following dictionary with ships, their length, speed in- and outsailtimes.

```
1. """
2. The types of ships that sail the Dutch inland waterways. These types are used
3. to define the speed, length in- and outsail times for every ship that is created.
4.
5. speed is in m/s
6. length is in m
7. in- and outsail is in sec.
8. (The time it takes for the ship to sail into the lock and to sail out)
9.
10. """
11. MO = {'type': 'MO', 'speed': '3', 'length': '20', 'insail': '123', 'outsail': '63'}
12. I = {'type': 'I', 'speed': '3', 'length': '39', 'insail': '167', 'outsail': '86'}
13. II = {'type': 'II', 'speed': '4', 'length': '53', 'insail': '219', 'outsail': '115'}
14. III = {'type': 'III', 'speed': '4', 'length': '70', 'insail': '170', 'outsail': '90'}
15. IV = {'type': 'IV', 'speed': '4', 'length': '105', 'insail': '222', 'outsail': '116'}
16. Va = {'type': 'Va', 'speed': '4', 'length': '123', 'insail': '244', 'outsail': '127'}
17. Vb = {'type': 'Vb', 'speed': '4', 'length': '180', 'insail': '246', 'outsail': '128'}
18. VIa = {'type': 'VIa', 'speed': '3', 'length': '123', 'insail': '230', 'outsail': '121'}
19. VIb = {'type': 'VIb', 'speed': '3', 'length': '190', 'insail': '252', 'outsail': '134'}
20. VIc = {'type': 'VIc', 'speed': '3', 'length': '270', 'insail': '306', 'outsail': '162'}
21. VIIb = {'type': 'VIIb', 'speed': '3', 'length': '195', 'insail': '306', 'outsail': '162'}
```

As mentioned in the previous chapter, the ship generators create ships with a 'random' type. This type was then used to define the speed, length etc of the ship. Each type had the same probability to occur in the simulation.

8.1.3 Estimators

As mentioned, the estimator in the simulation has the task that the Trajectplanner has in the real world which is estimating ETAs. Each experiment/test has their own version of this estimator. As the ships cannot be controlled by the Trajectplanner, they are not controlled by the estimator. The estimator purely has to estimate the time of arrival. One difference between the estimators and the Trajectplanner is that the estimators don't schedule ships into the procedures, they only estimate the waiting time based on the number and length of the ships arriving before a particular ship. This means that the estimator used in the simulation does not give out RTAs, but only gives ETAs. Before each experiment, the estimator used in that experiment is discussed in more detail.

8.2 Verification of the simulation

To verify the simulation of the ships the simulation environment was run with only 1 ship from either end and starting at the same time. To check all the calculations, the variation in speed was taken out and both ships were of the same type. This meant that both ships started at t=10, were of the type MO, so had a speed of 3 m/s, a length of 20 m, insail time of 123 sec and an outsail time of 63 sec. See section 8.1.2.

The lock in the simulation starts with the doors open on the West side, so the ships coming from the west can sail in without waiting. The lock waits a total of 300 sec for ship to arrive and sail in.

The ship is considered at the lock door or at the final destination when the location is within +/- 5 m of the door's location or final location. In the case, -180 +/- 5 m (west door), 180 +/- 5 m (east door), -2000 +/- 5 m (west end location), 2000 +/- 5 m (east end location).

So for the ships from the west side, the following calculation can be made.

$$\frac{2000 - 185}{3} = 605 \text{ is sail time from start to lock door}$$

$$123 + 300 - 123 + 720 + 63 = 1083 \text{ is insail + remaining waiting time + switchover time + outsail}$$

$$\frac{1995 - 180}{3} = 605 \text{ is sail time from lock door to end}$$

$$605 + 1083 + 605 = 2293 \text{ is total sail time from start to finish}$$

The ships from the east has the same calculations, but has to wait an extra 1083 s for the lock to open the doors on the east side. So a total travel time of 3376 sec.

With both ships starting at t = 10, their ATA at the end should be 2303 for the west ship and 3386 for the east ship.

Table 8-1 shows the results from the simulation run.

Table 8-1 Verification test results

	Name	Type		Start		End
wShip.0	wShip.0	MO	started at	10	and arrived at	2303
eShip.0	eShip.0	MO	started at	10	and arrived at	3387

It can be seen that the west ship (wShip.0) did exactly what was calculated. The east ship however shows an extra second of travel time.

After inspection of the ship's logs. It could be seen that for some reason, the simulation let east ship sail for 1 second longer at the location of the door and the end location. Because the ship had to wait anyway at the lock, this extra second did not matter, but it did at the end. This situation is shown in table 8-2. Table 8-2 shows parts of the log which were found in the actual excel files of the ships. These were meant to check the simulation and can be confusing.

Table 8-2 Part of log file from verification test

Time	Ship	Location	Status	Time	Ship	Location	Status
613	wShip.0	-191	Sailing to lock	614	eShip.0	188	Sailing to lock
614	wShip.0	-188	Sailing to lock	615	eShip.0	185	Sailing to lock
615	wShip.0	-180	Waiting to enter lock	616	eShip.0	180	Waiting to enter lock
2301.0	wShip.0	1989	Sailing to end	3385.0	eShip.0	-1992	Sailing to end
2302.0	wShip.0	1992	Sailing to end	3386.0	eShip.0	-1995	Sailing to end
2303.0	wShip.0	s ATA End	2303	3387.0	eShip.0	s ATA End	3387

The table shows the difference at the door's location (-180 and 180). For wShip.0 the step after -188 would be -185 which is in the -180 +/- 5 range, so the simulation puts it at the door. For eShip.0, the simulation lets the ship go to 185 before putting it at the door. It does the same thing when the ships arrive at their end location. This is where the 1 second difference between the calculation and the simulation comes from.

As this does not impact the process of the lock or the process of the ships. The simulation is considered validated.

8.3 Validation of the simulation

In order to use the simulation for the testing of the improvements, the simulation had to be validated. This was done on 2 aspects of the simulation. The first aspect was that the ships took a realistic amount of time to pass the lock. And the second aspect was that the basic estimator had to show similar results to the current Trajectory planner.

8.3.1 Purpose

The purpose of the experiment is to validate the simulation of the ship movements. When the simulation is validated it can be used as a testing tool for improvements.

8.3.2 Validation data

The data that was used to validate the simulation was obtained by tracking the ships that went passed the 2 nodes mentioned above. This was done via the Trajectory planner and via AIS/GPS radar images (figure 8-2).

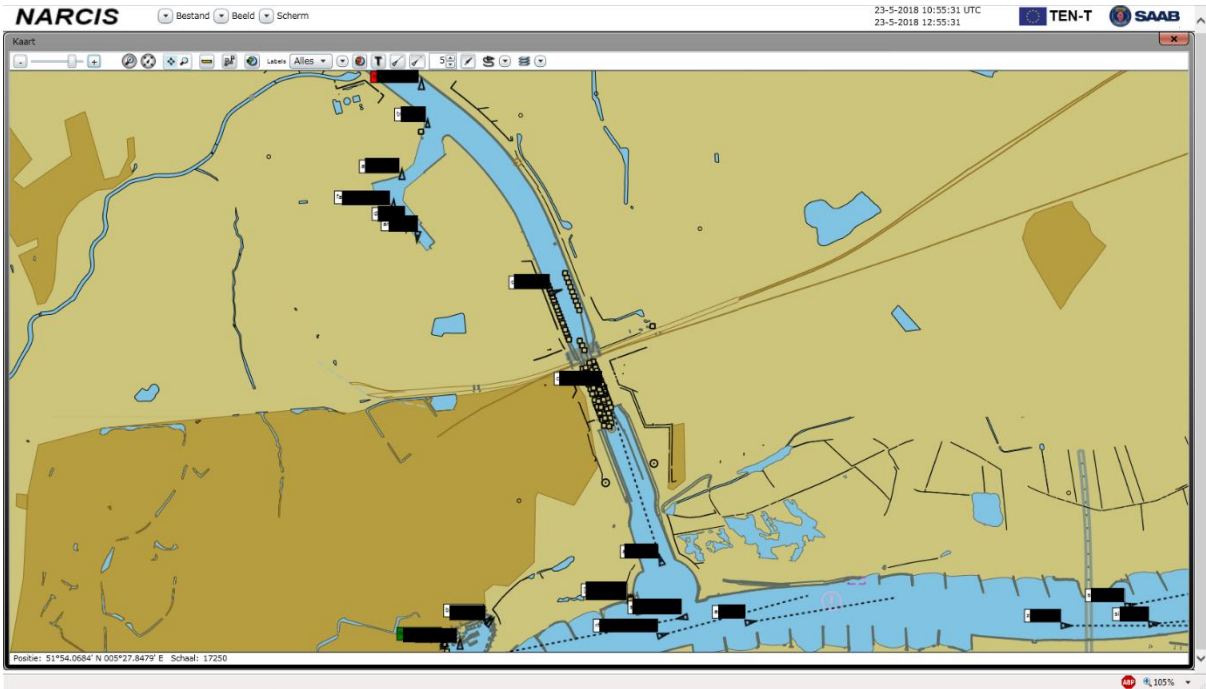


Figure 8-2 UI of NARCIS which was used to gather AIS/GPS data

By analysing this data, the ATA for several ships for the 2 nodes was acquired and could be used to validate the simulation. These radar images were also used to determine the service time of the lock. The ETA's for all these ships and the progressions through time was also analysed and used to validate the estimator in the simulation. This data can be seen in figure B-6 through figure B-9 as it was also used to analyse the Trajectplanner.

The type of ship was not taken into account with this analysis.

The amount of ships was not registered for this day. But via a different system, the amount of ships passing the lock in another time period was used to determine the maximum amount of ships per hour.

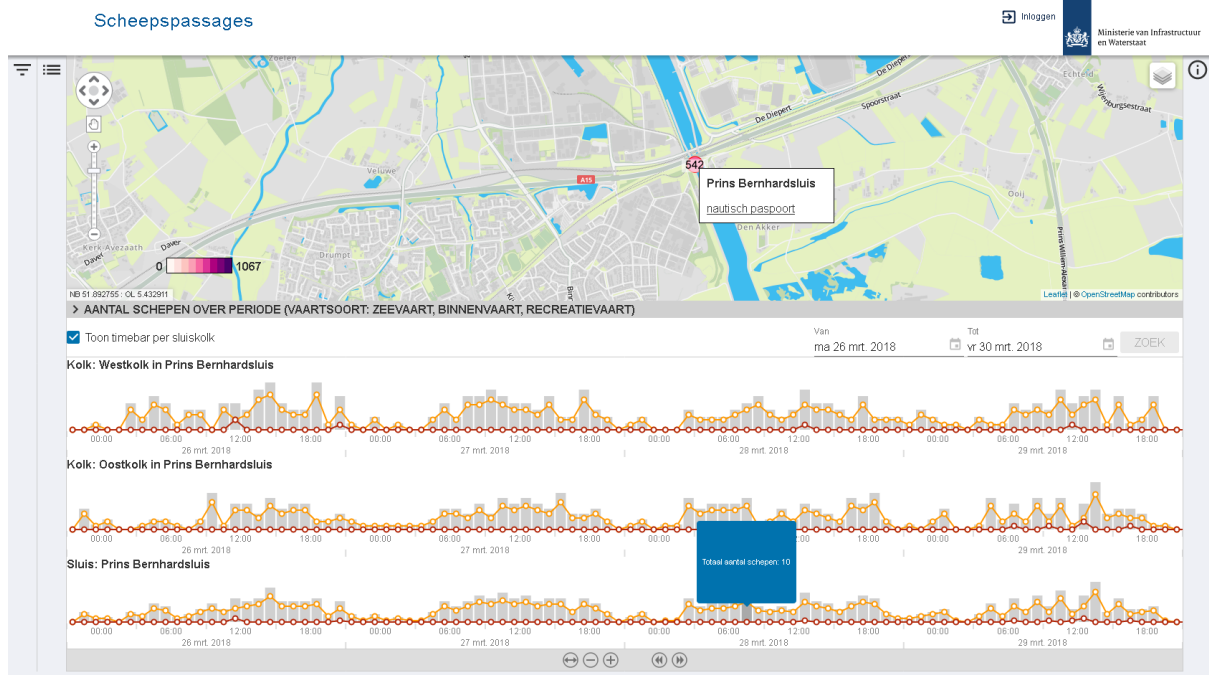


Figure 8-3 Data taken from SVM Lab(Rijkswaterstaat) about the Prins bernhard lock

The data shown in figure 8-3 was mainly used to determine the maximum amount of ships passing the lock per hour. This was used to determine the busiest time at the lock and use that for the simulation.

With this data the simulation was tweaked to represent the real-world situation.

8.3.3 Model

The model used in the validation test is a model with 2 nodes and a lock. The nodes are situated at -1250 and 1400. The lock is situated at 0. Ships are created at -2000 and 2000 so will have ETAs to 2 nodes. For each ship, independent of where they come from, node 1 will be the node before the lock and node 2 will be the node after the lock. See figure 8-4.

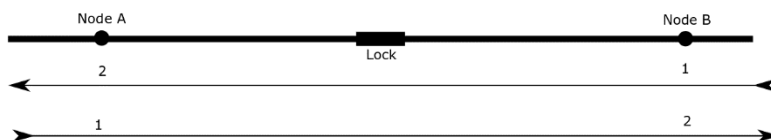


Figure 8-4 Simulation network model with 2 nodes

8.3.4 Estimator

The estimator used in the validation test is a version which has to replicate the current Trajectplanner. The network model used in this experiment has only 2 nodes (see figure 8-4) and replicates a part of the waterway around the Prins Bernhard locks mentioned in section 8.1.1.

For the estimator in this experiment, this means that when a ship is between 2 nodes with the lock on it, it doesn't know where the ship is anymore and it also doesn't know if the ship has gone through the lock. Only when the ship reaches the second node does the estimator know the position of the ship again.

To estimate the waiting time at the lock, the estimator uses the ETA of the current ship for node 1 and compares that with the ETA of other ships for that same node which travel in the same direction. The estimator assumes that ships which arrive first at the lock will go through first, so ships that arrive at node 1 before the current ship will go through the lock first. The estimator uses this information to estimate the waiting time and with this will estimate the ETA for node 2.

But, because the estimator doesn't know where the ship is when it passes node 1, it keeps estimating the total travel time until the ship has passed node 2.

The above is done in order to replicate the behavior seen in chapter 5.

8.3.5 Validation - Experiments

2 experiments were performed in order to validate the simulation. A sailing time experiment to check if the travel time between the 2 nodes was realistic. And an estimator test to see if the behaviour of the estimator resembles that of the Trajectplanner.

The 2 experiments are described and discussed in Appendix C, sections C.1 through C.5. Figure 8-5 shows 2 graphs of 2 ships that sailed in the simulation run. All the other graphs are shown in Appendix C section C.4 and C.5. For all the other experiments that were done, the ships and their journeys will be the same as in the validation experiments. This makes it easy to compare different design alternatives

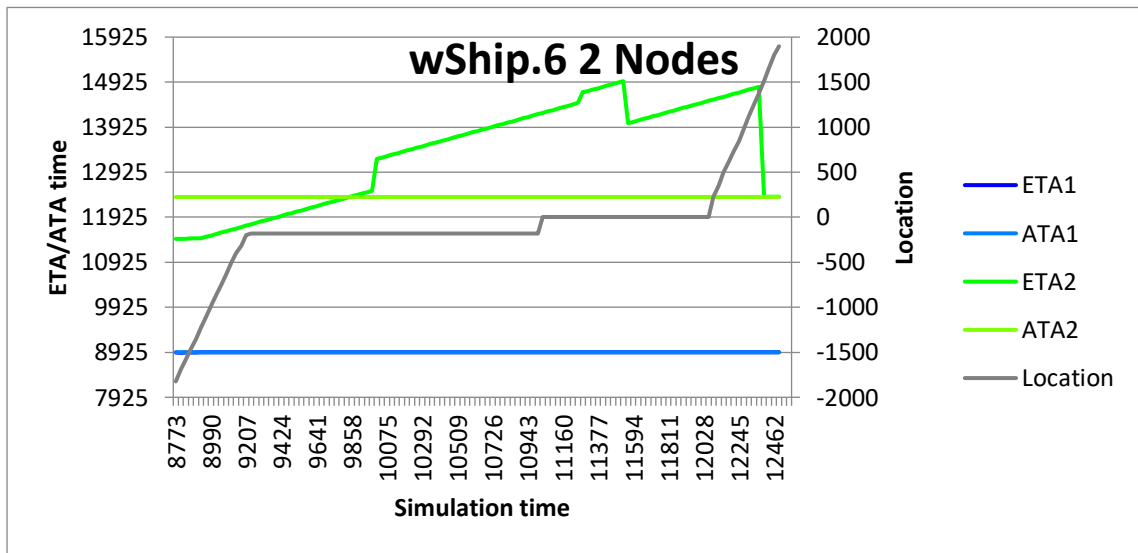
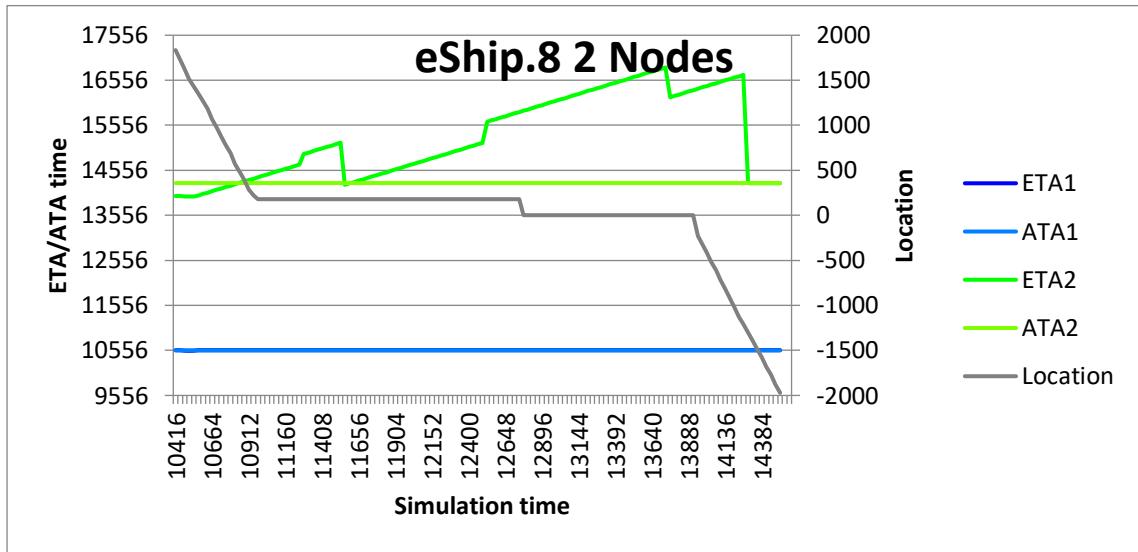


Figure 8-5 Graphs from experiment for 2 ships
 Table 8-3 Results from experiment for 2 ships

Ship	Node	Time in seconds
eShip.8 Max KPI	2	2123
wShip.6 Max KPI	2	2078

8.3.6 Validation - Results

With the results from the validation experiments, the simulation is considered validated. There are differences with both the real world and with the Trajectplanner, but those are explained by the simplifications and assumptions that were made in the simulation.

The result is that this simulation will be used to test improvements. These improvements will be made in different versions of the estimators. The graphs that are made by the simulation will therefore be different. The time it takes ships to sail from one end to the other will stay the same and with the same seed it is possible to make 1 on 1 comparison to the 'standard' system.

8.4 Experimental plan

With the verification and the validation, the simulation environment can be used to test whether the designs mentioned in chapter 7 would improve the arrival time estimation.

Two aspects that were discussed in chapter 7 are evaluated in this chapter, namely the new network model and the lock awareness model. The latter is evaluated in combination with the new network model as the lock model needs the extra nodes in order to function correctly. The experiments that were done are called Design 1 and Design 2, Design 1 uses the new network model and Design 2 uses the new network model + the new lock model.

Table 8-4 Experimental plan

<i>Experiment</i>	<i>Section</i>
<i>Design 1</i>	8.5
<i>Design 2</i>	8.6

The details for each experiment will be discussed in their respective section of this chapter

For every experiment in which an estimator is used, the output in the form of a graph will be shown of 2 ships. Also, the maximum difference between the ETA and ATA, or in other words, the MAX KPI will be displayed. This Max KPI is determined by taking the maximum value from all the points in the network. The graphs and data of the other ships can be found in Appendix C

8.5 Design 1

As the next test will be done in order to show the difference in results from the different estimators, the seeds will be the same as the last validation test. This way, the exact same ships with the same start and end time will be used.

The first improvement that need to be tested is the new network model. As mentioned before, the model now is not a true arc and node system. To make the current system a true arc and node system, extra nodes need to be added. Specifically, before and after an object. For a lock, this means the position of the doors, or just before the door. For a bridge, these points could also lay just before and after the bridge. The distance between the nodes would be very small.

8.5.1 Purpose

The purpose of this experiment is to see if the new network model already improves the ETA estimation without changing to much in how the estimator works.

8.5.2 Model

As this simulation only has a lock in it, the extra nodes will be placed just before the doors. As the lock is in position 0 and has a length of 360, the doors are located on -180 and 180. The extra nodes will therefore be placed on -200 and 200

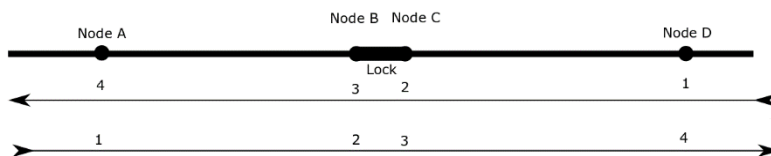


Figure 8-6 Simulation network model with 4 nodes

8.5.3 Estimator

The estimator used for this test was almost the same as in the first validation test. The big difference is the network model used by this estimator and illustrated in figure 8-6. The estimator still has no awareness of when the ship has passed the lock, but the arc where this happens is much smaller, namely between node 2 and 3 and not between 1 and 4 (if compared to the current network model). The estimator also makes use of the extra nodes by adjusting its waiting time estimation by using the fact that when a ship is passed node 2, the ship is already waiting to enter the lock. The rest of the estimation is done exactly the same as in the previous experiment.

8.5.4 Experiments

Test were performed with the extra nodes in the system but with the same estimator as in the validation test. The resulting graphs are presented below. The other graphs can be found in Appendix C section C.6

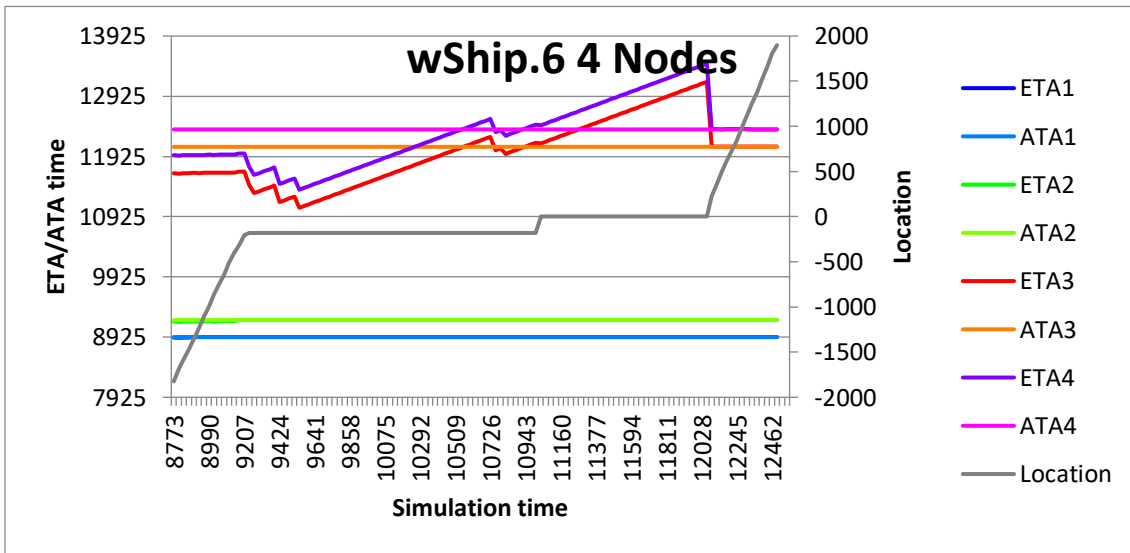
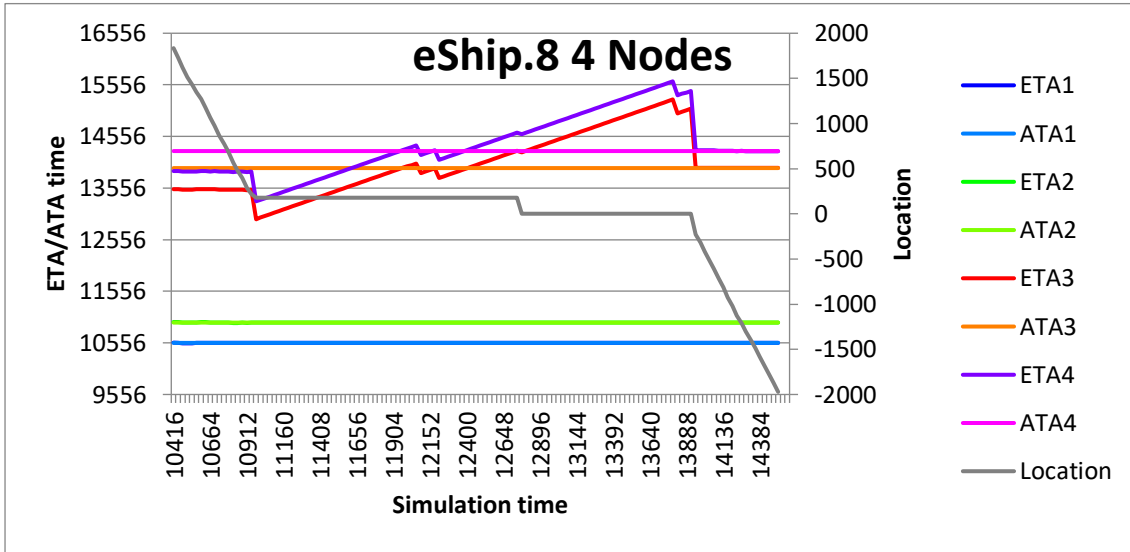


Figure 8-7 Graphs from experiment for 2 ships

Table 8-5 Results from experiment for 2 ships

Ship	Node	Time in seconds
eShip.8 Max KPI	4	1351
wShip.6 Max KPI	4	1389

8.5.5 Design 1 - Results

What can be seen in the figures (In this chapter and in the appendix) is that the ETA before the second node is quite stable and, in some cases, very accurate. But once the ship is passed the second node and waiting to go through the lock, the ETA for the third and fourth node starts rising again. Because the second and third node are quite close together, the rising stops quite fast and the ETA for the fourth node becomes accurate again.

8.6 Design 2

As previous tests showed. When the ship is in the lock, the ETA keeps rising. If the Trajectplanner is aware of the fact that the ship is in the lock, this might be avoided.

8.6.1 Purpose

The purpose of this experiment is to see what the influence is of the new network model plus the awareness of the estimator for when ships are in the lock.

8.6.2 Model

The network model used in this experiment is the same as with the previous experiment. This means that the new network model is also used in this experiment. See figure 8-6.

8.6.3 Estimator

To test whether it would improve the ETA prediction if the Trajectplanner would be aware of the fact that a ship is inside a lock, two aspects had to be added to the estimator in the simulation environment.

First the estimator needs to know when a ship is inside a lock. In the simulation environment, when a ship is inside a lock, its position becomes 0 (the position of the center of the lock). This information is used with this estimator to indicate that the ship has entered the lock.

The second aspect to be added is a form of memory or an extra attribute/characteristic of the ship, namely the time the ships entered the lock. When a ship has entered the lock (has position 0), the estimator will log that time for that particular ship and use it to estimate when the ship will be out of the lock again.

The estimation of how long a ship will be inside a lock is based on the average switchtime or service time of a lock. The estimator also used the outsail time of the particular ship but it does not take into account the outsail time of other ships that will exit the lock first. This is done to simplify the process of estimation as this experiment is done to show if it would improve the estimation at all.

The rest of the estimator works the same as in the previous experiment, the estimator which makes use of the extra nodes in the network.

8.6.4 Experiments

The resulting graphs are presented below. The other graphs can be found in Appendix C section C.7

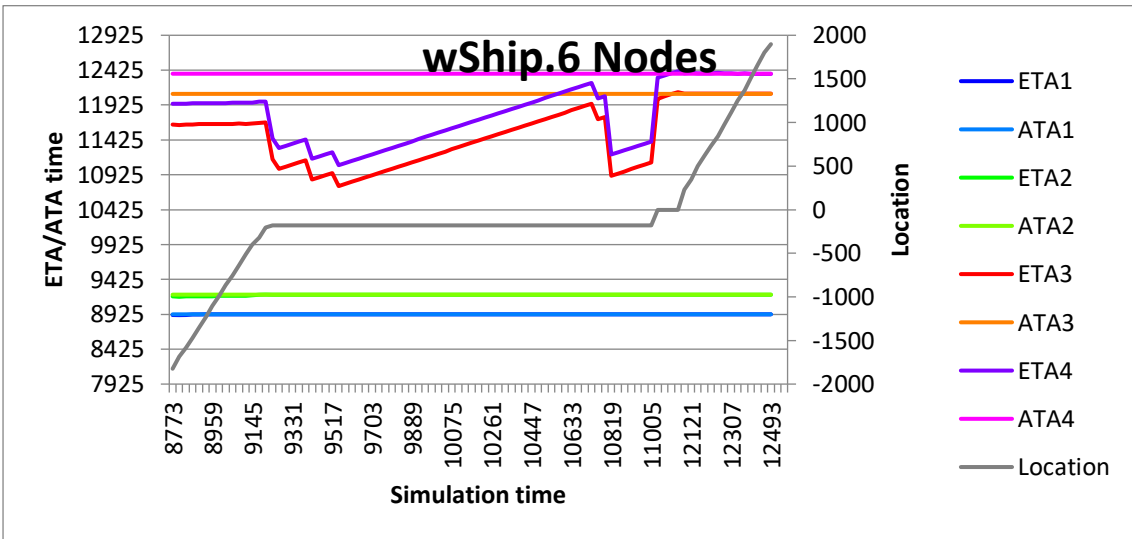
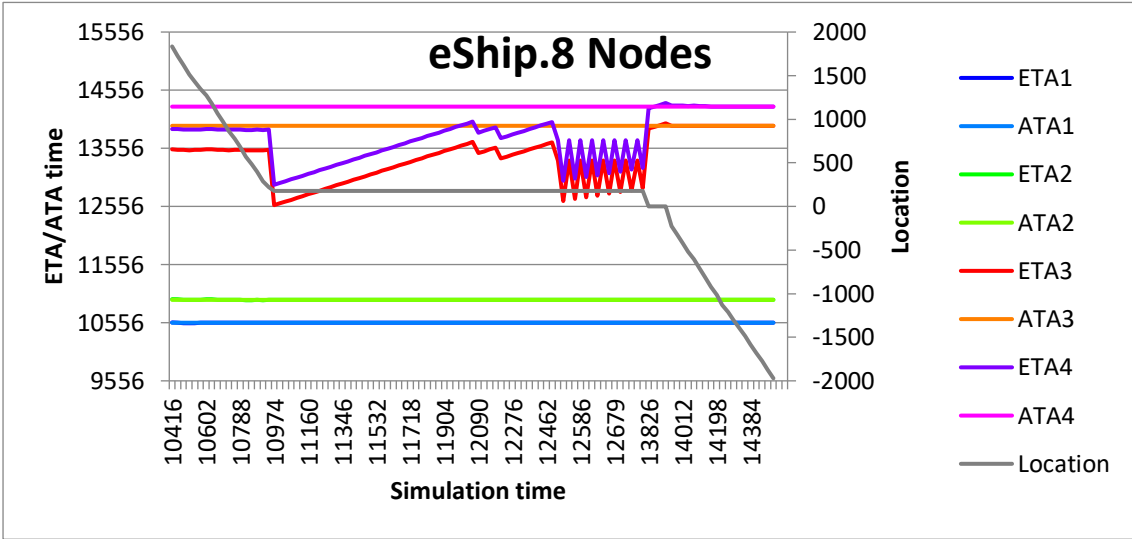


Figure 8-8 Graphs from experiment for 2 ships
 Table 8-6 Results from experiment for 2 ships

Ship	Node	Time in seconds
eShip.8 Max KPI	3	1364
wShip.6 Max KPI	4	1311

8.6.5 Design 2 - Results

What the graphs above show is that the ETA lines move more towards the ATA lines. In most cases they become more accurate with time instead of rising all the time. An important thing to notice is that when the ship enters the lock (location becomes 0), the ETA for node 3 and 4 becomes very accurate and stays that way. This means that the ETA estimation has less time to rise.

8.7 Conclusion

The table in Appendix C section C.8 displays all the max KPI's from the experiments done in this chapter for all the ships in the simulation. The average of all max KPIs over all ships is shown in table 8-7. Based on table 8-7, the mean value does improve, but not with every step. The median however does improve with every step.

Table 8-7 KPI from all ships in all experiments

	KPI for Trajectplanner	Validation of simulation	Design 1	Design 2
<i>Mean</i>	2986,244	2870,075	1493,517	1684,876
<i>Median</i>	2495	2719,333	1462,667	1447,75
<i>Standard deviation</i>	2550,245	688,954	268,562	688,065
Data type	Real world	Validation	New model	New model

The higher mean KPI over all ships shown in table 8-7 for Design 2 is not what was expected. After inspection it was seen that in the case of some ships, the KPI did improve, in other cases, it got worse, the cause was the fact that the estimator tried to estimate the time the ships still had to wait before entering the lock. Unfortunately, this caused some bad estimations in some cases. This can be seen when comparing the graphs in Appendix C6 and C7.

The data from table 8-7 and the graphs combined show that the improvements do improve the estimator and therefore would also improve the Trajectplanner. As an extra comparison, a graph was made (figure 8-9) combining the data of the same ship (wShip.6) making the same journey with two different estimators (Validation Design 2).

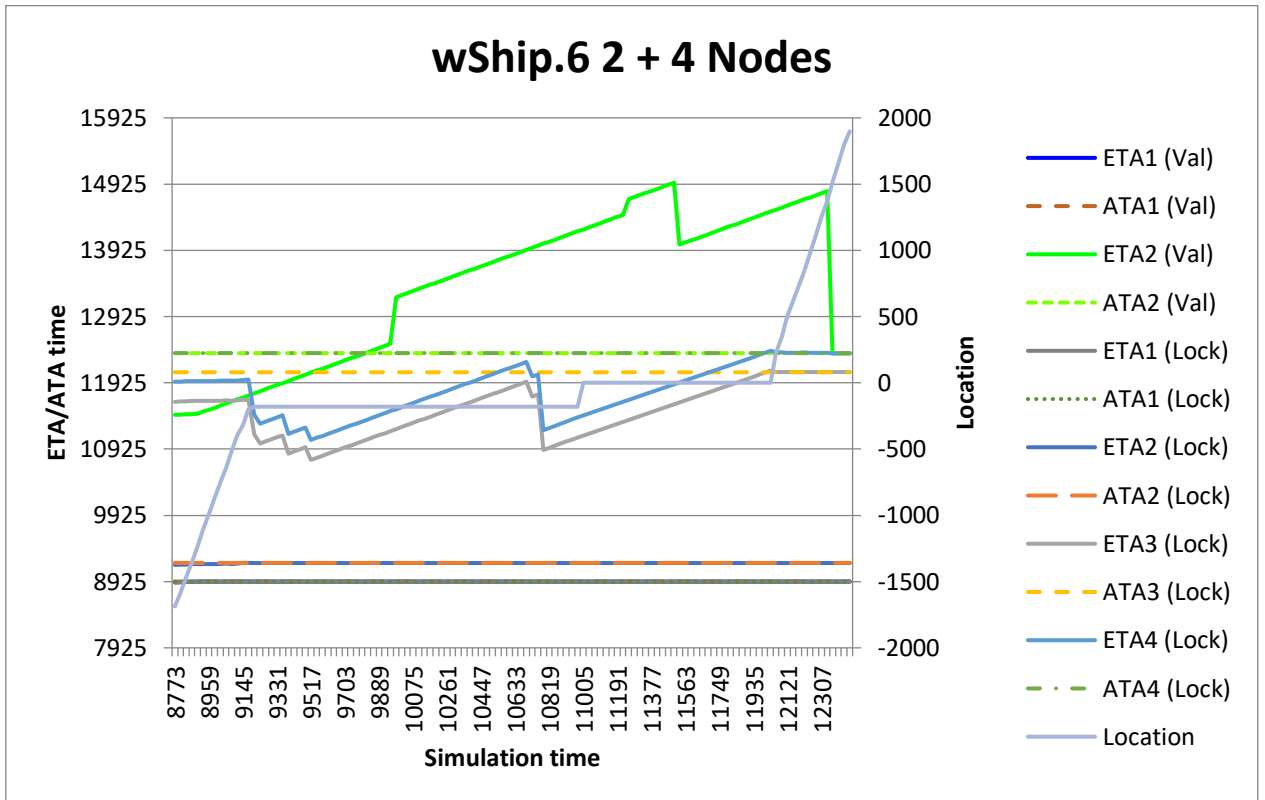


Figure 8-9 Combined graph of 2 experiments with the same ship

This graph was made by combining the graph for this ship from the validation test and the graph for the same ship from the Lock awareness test. It shows the ETAs from both tests where ETA1(Val) is the ETA for node 1 in the validation test and ETA1(Lock) is the ETA for node 1 in the Lock awareness test. The Validation test had a network with two nodes and the Lock awareness test had a network with four nodes. In this graph, Node 1 (Val) = Node 1(Lock) and Node 2 (Val) = Node 4(Lock).

The graphs show what is to be expected which is that the ETA1 for both tests is the same, and the ETA2(Lock) is very accurate. What the graph also shows is the difference between ETA2(Val) and ETA4(Lock). The location of the nodes for which these values are calculated is the same, but it can be seen that the ETA4(Lock) line stays closer to the ATA line than the ETA2(Val) line does. This graph is a good example of the more consistent ETA calculations.

The big difference between the estimators used in this simulation and the Trajectplanner is the planning module. The simulation did not schedule ships for locks, it looked how many ships were before another ship and estimated the extra waiting time. This estimation was based on the ETAs of those other ships. This is also why the ETA for some ships was 'jumpy'. In all tests, the estimator calculated differently based on the ship's position. And it can be seen in the last test that when the ship is passed node 2, the planning is no longer needed and the ETA becomes quite accurate, because the estimator knew that the ship was in the lock. With this the climbing of the ETA was also

reduced and, in some cases, completely gone. When the ship was actually in the chamber, the estimator knew this and could predict with high accuracy when the ships would sail out and when it would reach the next node.

The conclusion of these test is therefore that the extra nodes, the awareness of being in a lock and knowing when the lock was entered would improve the ETA calculation a lot. As can be seen in all the figures is that the waiting of ships before entering the lock takes up the most time and therefore the next step would be to improve the scheduling module in the way mentioned in chapter 9.

This answer to the subquestion **How accurate is the new prediction model?** Is twofold, with only the new network model as an improvement, the system becomes twice as accurate ($2986/1494=1.999$) With other improvements, the accuracy goes down a bit ($2986/1685 =1.772$).

9 Conclusion

This paper described the research that was done on the new system of Rijkswaterstaat, the Trajectplanner. The research was guided by a couple of sub-questions in order to answer the main question. Each of these sub questions will be discussed in this chapter.

9.1 Subquestions

Are there other methods/systems that predict the ETA of inland ships?

This question was answered in chapter 2. The answer was, yes there are similar systems, but not with the same functions as the Trajectplanner. Different aspects of the prediction came forth from IT systems and research papers, but no single system or paper had the same (or better) functionality of the Trajectplanner, which is predicting the ETA of inland ships were individual ships can be monitored or were ships can monitor their own ETA.

How to define the accuracy of the ETA prediction?

This question was discussed in chapter 3. The accuracy was defined by a new KPI. Because the system is not used in daily operation yet, Rijkswaterstaat didn't have a KPI for it yet. The new KPI was made with literature and the users of the system in mind. The KPI was defined by the maximum absolute value of the difference between the ATA and the ETA. As explained in chapter 3, the KPI could be used for every point in the system.

$$KPI_x = MAX|ETA_x - ATA_x|$$

This was partly done in the analysis of the current system and in the evaluation of the new designs. Also mentioned, the point for which the KPI is determined can be different. For example, for the skipper it is important to know that the ETA before a lock is accurate so he can adjust his speed accordingly. It is also important for the skipper, as well as for the next shipper, that the ETA is accurate at the final destination. For a researcher other points in the system are important depending on what is being researched. For this research, the points around the locks were mainly used for the determination of the KPI.

How does the current simulation system predict the ETA?

This question was discussed in chapter 4. The system works by simulating the movement of every inland ship and with that, the system can predict where ships will be and when they need to pass objects. To do this the system gets data about each individual ships and uses this to predict the path it will take and how long it will take to get from point to point.

How accurate is the current prediction?

This question was discussed in chapter 5. The current system has a KPI of 2986 sec which is roughly 50 min. This is an average taken over multiple ships and these ships sailed the same part of waterway. This value was measured near a lock where it was expected to be the worst. But this does show what kind of error the system could give.

How to improve the current prediction?

This question was discussed in chapter 6. The model currently used to represent the Dutch inland waterway system can be improved by making it a true node and arc network. This would help with improving the planning. Another improvement was a way for the system to know when ships entered the lock chamber, this would help with knowing when the ships would be out of the lock and with planning the rest of the journey.

How to evaluate a new prediction model?

This question was discussed in chapter 7. The way the new designs were evaluated is via a simulation. This was done because the calculations would be too complicated to do by hand as ships would interact with a lock. It was also necessary to simulate an 'Estimator'. Unfortunately it was not possible to test and evaluate with the current system.

How accurate is the new prediction model?

This question was discussed in chapter 8. The new model is more accurate and more consistent. The test showed what kind of impact the improvements had, but because the simulation did not focus on the scheduling of ships, it was discovered that, after the nodes and awareness of the lock, this would be the next challenge. The new design was almost twice as accurate. This was with the new network and with the lock awareness. The graph also shows a more consistent prediction.

9.2 Main question

How to accurately predict ETAs of inland ships

The answer to this question is by simulating the movements of all the ships over a network model of the system which is made up by a full node and arc network. Within this network locks are simulated by 2 nodes with an arc between them and have a variable passage time. This passage time can be calculated by scheduling the ships into the lock procedures and estimating the time it takes to get through the locks. Without having control of the ships, the simulation should assume that whoever is first at the lock, will go through first. With this and other rules, it should make an estimation of what time the ship is out of the lock again. The simulation needs to know where the other ships are and when they are in a lock chamber. It needs a memory which at least contains the entry time of ships into the lockchamber so it can predict the exit time. With all of the above it is possible to estimate the arrival times of ships. Within this research it was shown that with a new network model and a simple memory of the ships entering time could improve the performance by almost a factor of 2. This was with a simple simulation tool, so it is assumed that better results will be found when the new design is implemented and tested with the current Trajectplanner system.

9.3 Overall conclusion

As mentioned, this report analysed the Trajectplanner in order to find the cause of inaccurate ETAs. After an analysis of the current Trajectplanner backed by a literature review. It was concluded that the way the inland waterway network is not modelled correctly within the Trajectplanner. Therefore, an improved model was thought out and tested with a simulation. After the simulation it can be concluded that the improved model worked. But the simulation did not schedule ships the way the current Trajectplanner does. It can be seen that in the simulation, the waiting of ships before entering a lock

takes up the most time. Because of this, the scheduling becomes the next big problem after the network model and lock awareness issues have been solved.

The overall conclusion after the analysis and simulation tests is that there is quite some room for improvement of the Trajectplanner. Especially when the goal is to advise skippers to adjust their speed and to make inland shipping more attractive by giving accurate travel times.

9.4 Recommendations

One of the goals of the Trajectplanner is to inform skippers that they can sail faster or slower. And with that achieve more efficiency with accurate travel times and less laying still with engines running. In the bigger picture, this could improve the whole supply chain where inland ships are involved and make them more attractive to use.

To achieve this, the Trajectplanner needs to be improved, apart from the suggestions done in this report further study should take place.

The first step should be to implement the new model which models the inland waterway network. With this improvement it should be possible for the system to make better schedules but also to be aware whether ships are inside the lock or not. A second step would be to make a more accurate simulation tool, or a way to test new models, methods and components within the current Trajectplanner.

Other improvements could be found in:

- Improvement of the scheduling, as mentioned in chapter 4. The scheduling keeps changing every 6 min, this would not work for a skipper. For example, if a schedule is made it should stay that way as long as all the ships in that schedule are still able to make the RTA.
- A memory of passed simulations. With a memory, the Trajectplanner could check itself when the ATA for every node is known. It could also be used to flag big fluctuations in ETA calculations. This could indicate something is wrong with the system or with the ship.
- A test to see how much impact it would have when ships are scheduled into locks based on the Trajectplanner instead of the arrival time at the object. Now, ships enter the lock on the same order as they arrive at the lock. For the Trajectplanner to reach its full potential, it is needed that ships need to rely on a schedule they are given and not worry about being overtaken.

10 References

1. Binnenvaartcijfers. [Online] <http://binnenvaartcijfers.nl/>.
2. SVM Lab. [Online] Rijkswaterstaat. <http://131.237.92.226>.
3. [Online] <https://www.rijkswaterstaat.nl/zakelijk/zakendoen-met-rijkswaterstaat/rijkswaterstaat-als-opdrachtgever/programmas/sluzen.aspx>.
4. [Online] <https://www.rijkswaterstaat.nl/wegen/wegbeheer/bruggen/veelgestelde-vragen.aspx#vraag1>.
5. Rijkswaterstaat.nl. [Online] https://www.rijkswaterstaat.nl/apps/geoservices/geodata/dmc/vaarwegenkaart/productinfo/beschrijvende_documentatie/vaarwegenkaart_2013.pdf.
6. Rijkswaterstaat.
7. *Marine Traffic*. [Online] <https://www.marinetraffic.com/>.
8. *Marine Traffic ETA*. [Online] <https://help.marinetraffic.com/hc/en-us/articles/204215878-Port-Related-Events>.
9. Bivas Documentatie. *Bivas*. [Online] <https://bivas.chartasoftware.com/BIVASApplicatie/Documentatie>.
10. Bivas Reizen. *Bivas*. [Online] <https://bivas.chartasoftware.com/BIVASApplicatie/Documentatie/Reizen>.
11. Caris, Limbourg, Macharis, Lier, Cools. *Integration of inland waterway transport in the intermodal supply chain: a taxonomy of research challenges*. 2014.
12. Li, K. *Dynamic traffic assignment model for inland waterway freight transport*. 2017.
13. Zheng, Negenborn, Lodewijks. *Predictive path following with arrival time awareness for waterborne AGVs*. 2015.
14. J. Verstichel, P. De Causmaecker, F.C.R. Spieksma, G. Vanden Berghe. *Exact and heuristic methods for placing ships in locks*. 2013.
15. Hengeveld, J. J. S. *Optimization to reduce waiting times at locks*. 2012.
16. 2016, T.H.A. Visser –. *Lock scheduling model for the series of locks in the Maas*. 2016.
17. X. Zhang, S. Mahadevan. *Aircraft re-routing optimization and performance assessment under uncertainty*. 2017.
18. W. Zhou, H. Teng. *Simultaneous passenger train routing and timetabling using an efficient train-based Lagrangian relaxation decomposition*. 2016.
19. Parolas, Ioannis. *ETA prediction for containerships at the Port of Rotterdam using Machine Learning Techniques*. 2016.
20. Kaushik Roy, Benjamin Levy, Claire J. Tomlin. *Target tracking and Estimated Time of Arrival (ETA)*. 2006.
21. F.W. Cathey, D.J. Dailey. *A prescription for transit arrival/departure prediction*. 2003.
22. Chin-Woo Tan, Sungsu Park, Hongchao Liu, Member IEEE, Qing Xu, and Peter Lau. *Prediction of Transit Vehicle Arrival Time for Signal*. 2008.
23. Barth, Abdolreza Karbassi and Matthew. *VEHICLE ROUTE PREDICTION AND TIME OF ARRIVAL ESTIMATION*. 2003.
24. *Salabim*. [Online] <http://www.salabim.org/manual/About.html#>.
25. [Online] <http://binnenvaartcijfers.nl/aantal-passages-van-binnenvaartschepen-per-sluis/>.
26. Rijkswaterstaat. Vaarweginformatie. [Online] <http://www.vaarweginformatie.nl/fdd/main/infra/vin;jsessionid=8488CB9A18BA43FBF6CD6992B75E15FF.node1>.

27. [Online]

<http://www.bureauvoorlichtingbinnenvaart.nl/assets/files/downloads/scheepstypen2-5.pdf>.

28. Ioannis Parolas, Lóránt Tavasszy, Ioanna Kourounioti, Ron van Duin. *PREDICTION OF VESSELS' ESTIMATED TIME OF ARRIVAL (ETA) USING MACHINE LEARNING - A PORT OF ROTTERDAM CASE STUDY*. 2016.