

10 CREATING A TRAFFIC DATABASE

10.1 Elements required for creating a new traffic database

The maritime traffic is divided into two main groups: the **route bound ships** and the **non-route bound** (or random) **ships**. The route-bound traffic consists of merchant vessels and ferries sailing along the shortest route from one port to another. The non-route bound traffic contains vessels that mainly have a mission at sea, such as fishing vessels, supply vessels, working vessels and pleasure crafts.

Within the SAMSON-model these two groups are modelled in a different way.

The route bound traffic is modelled with shipping routes. Each route is composed of a number of subsequent links. A link is the straight line between two waypoints. The lateral distribution on each link describes the use of the link by the ships. The number of movements on each link is determined to route all ship movements of one year crossing the North Sea. For this purpose all ship movements of one year are bought from Lloyd's Marine Intelligence Unit (LMIU). This database of LMIU contains only the information that a ship has left port A with port B as destination. All movements are acquired for the movements of which port A and/or port B is located in the Baltic, North Sea, Irish Sea, European Atlantic Coast, and the Mediterranean. These are a slightly more than 1 million movements. For each movement the IMO-number is given for linking with the database with the shipping characteristics.

The way how the routes are determined is described in 10.1.1.

The non-route bound traffic cannot be modelled in the same way as the route-bound traffic is, because the movements are not included in the movements database of LMIU. The behaviour of these non-route bound ships at sea is very different. A non-route bound ship does not sail from port A to port B along a clear route, but from port A to one or more destinations at sea and then usually back to the port of departure A. The behaviour of these ships at sea is most unpredictable. Fishing vessels, for example, usually sail from one fishing ground to another during one trip.

Therefore the traffic image of the non-route bound traffic is modelled by densities of ships in a so-called grid cell. The data for this database for the North Sea is extracted from the VONNOVI-observations between 1998-2001, see [16]. This was the last period that this kind of observations has been carried out. In future, other sources are required to update this data. The AIS-data can deliver a part of this data, because most work vessels and suppliers have AIS. However, the largest part of the non-route bound is fishery and these ships have mostly no AIS on board, but this can change because a Directive for mandatory AIS on board of fishing vessels above a certain length is being prepared by the EC.

10.1.1 From movements to database

The modelling of the route bound traffic is based on the assumption that the vessels sail from one port to another via the shortest possible route, taking into account all rules and traffic measures. The base for the determination are the waypoints. The connection line between two waypoints is a link. It is assumed that ships assigned to a link follow this link from the start point to the end point. The link cannot be left between these points. This property is used to create additional course change points (often indicated with an X) within a link to describe the only route through a bending fairway with one link.



Normally when the waypoints are in open sea, the link between two waypoints is sailed in a straight line.

The determination of the route for a ship is different from the determination of a route on land.

The road network is very broad, but on the other hand there are only a few other network points that can be reached from one point. These links are predefined.

That is different at sea, because there are no physical limitations. Only the depth can be a limitation why another waypoint cannot be reached with a straight course. And ships will use this freedom and will follow as soon as possible straight lines from the origin to the destination.

The first maritime traffic database developed within the COST301 project was also built upon a predefined network, but at that time the ports were joined to a regional level. It described the main routes but did not describe the routes to and from individual ports with sufficient detail.

Later, the traffic database was determined on a port to port base. This meant that the network with links had to be extended considerably. As it is not possible to check all individual voyages of ships, it was difficult to observe whether strange routes were found. A missing link between two waypoints can result in strange routes.

At that time, that the route network had to describe the maritime traffic in more detail, it was decided to follow a different approach that would lead to a more accurate database with less human effort, because it is nearly impossible to avoid mistakes in a bunch of data that has to be prepared. The result of the new approach was a route generator that finds a route from port A to port B by a chain of waypoints. When all routes are determined, the required minimum network of links can be derived from all connected waypoints.

The required input for the route generator consists of:

- Waypoints with coordinates;
- Dedicated links;
- Forbidden lines;
- Traffic separation schemes;
- Movements, with the number of ships sailing from A to B per ship type and ship size.

A description follows of the purpose of each input file.

10.1.1.1 Waypoints (file: POINTSDE.NEW and WAYPOINT.DAT)

The file WAYPOINT.DAT contains all waypoints where ships will change course. In open sea ships will keep their course as long as possible. Therefore, the waypoints are located near the entrance of ports, pilot stations, in traffic separations schemes, buoys indicating a fairway and at the corners of wind farms. Further a number of waypoints are located at some distance from the coast to indicate the coastal route.

There are two types of waypoints; the "free" waypoints and the "dedicated" waypoints. These waypoints are dealt with differently by the route generator. A free waypoint is accessible from each other free waypoint or via a predefined link from a predefined link. However, a dedicated waypoint is only accessible by a predefined link to that waypoint. The benefit of this approach is illustrated with the routes to and from the Western Scheldt shown in Figure 10-1 and Figure 10-2.

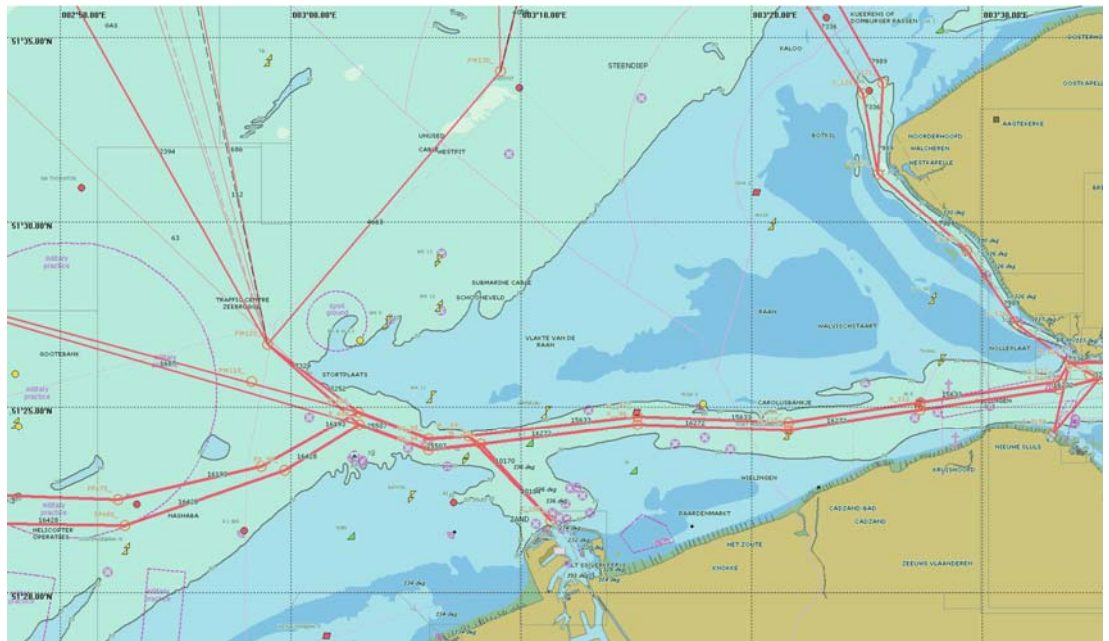


Figure 10-1 Entrance of the Western Scheldt

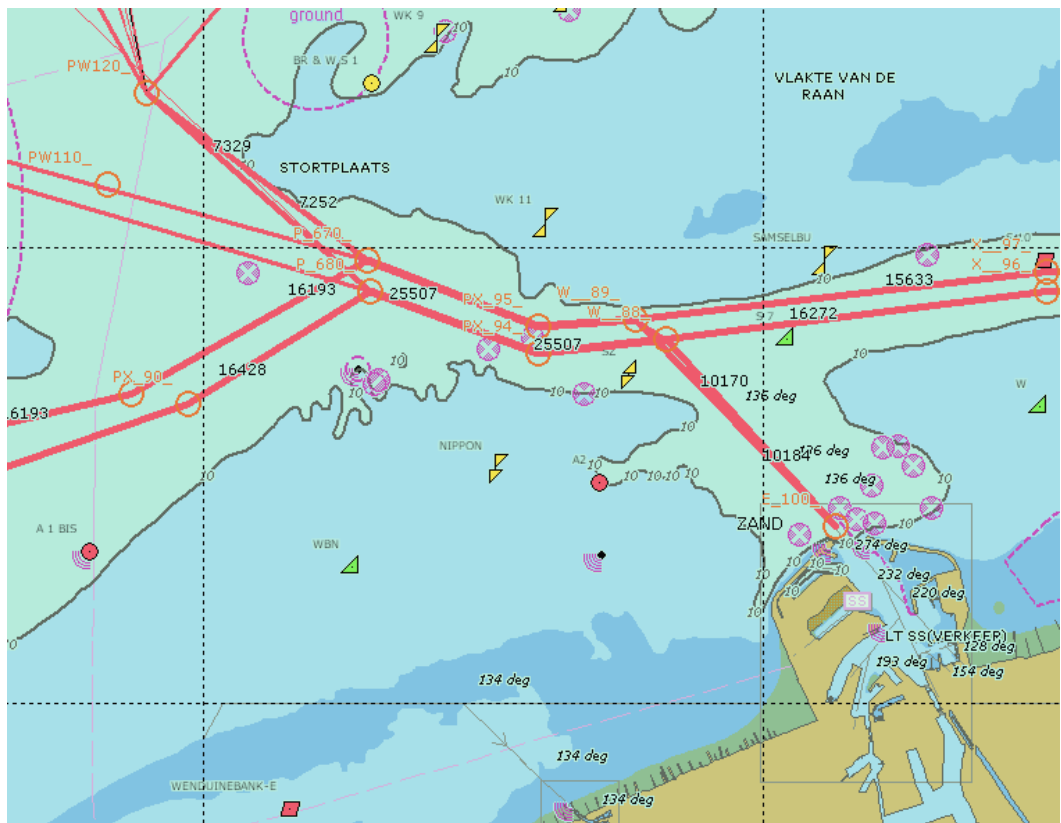


Figure 10-2 Zeebrugge



The ships entering and leaving Zeebrugge leave or enter the main route to and from the Western Scheldt, there is no other choice. The entrance of Zeebrugge is coded with E 100 (the E was originally used for end point in the first COST 310 database). The waypoint is coded as a dedicated waypoint. This means that all possible links from and to this waypoint have to be defined. The route from sea uses the link W 88 – E 100 and the link from Zeebrugge to sea uses the link E 100 – W 89. All other links are excluded. In case E100 was not a dedicated waypoint, the chosen route could have been E 100 – PW120 directly, which is not possible due to the depth. The points PX 94 and PX 95 are also defined as dedicated routes. P 670 and P 680 are the first free waypoints. This means that the links from PX 95 – P 670 and P 680 – PX 94 are still required, connecting a dedicated waypoint with a free waypoint, but from there the ships are free to go or come from. For example, outgoing ships can go to WP120 and from there to diverse northern directions, see Figure 10-1.

The WAYPOINT.DAT file contains also other possibilities to influence the routeing process. During the routeing 8 different structures are defined, originally linked to the ship size. Later this base was left because the draft of the ships of the same size class of a container ship is less than the draft of an oil tanker of the same size. Now it is possible to indicate for each type and size with 1 to 8 the network that has to be used (in file STRUKTUU.DAT). The networks in which the waypoint may be called are indicated with 8 digits "12345678" in the first 8 columns of WAYPOINT.DAT. For example a "78" can only be used for network 7 and 8. This indication is given for the waypoints in the Eurogeul.

Another reason to forbid the use of a waypoint is the cargo on board of a ships. Since June 1997 laden oil tankers with cargo above 10,000 ton or chemical tankers with cargo above 5,000 ton are not allowed to use the Texel TSS and are mandatory to use the deep water route TSS. In total 8 main ship types are distinguished. In another string of 8 characters in WAYPOINT.DAT, it can be indicated whether a waypoint can be called by a certain ship type. For example the CTG letters (for chemical, tankers (oil) and gas are missing in the waypoints of the Texel, Vlieland and Terschelling TSS.

Another possibility to influence the routeing is to encourage or discourage the use of a waypoint by adding a certain value (penalty) to the distance when the waypoint is used. Two types of penalties are used:

- A small negative penalty is used to concentrate some traffic flows by combining them to one traffic flow.
- A large positive penalty is for example used for offshore platforms. The waypoints of offshore platforms are only mentioned for destination traffic as suppliers for all platforms and oil tankers for oil platforms. Other traffic may not use these waypoints. When a ship has the platform as destination, the penalty has always to be taken, this will not influence the route to the platform.

Warning:

Be aware that some negative values for different waypoints will not cause a closed connection between the waypoints with a negative value as total route length. In this case the routeing will go into an infinite loop because all the time, better routes are found due to the negative route length.

The next item in a fixed position (column 29), contains an indication of the point. In case of a ">" the waypoint is considered as an end point and will never be explored.



The next item in a fixed position (column 30), contains a "\$" when it is a dedicated waypoint and contains a blank for a free waypoint.

The next 6 positions (column 31-36) are for the code of the point (left aligned). The use of the correct number of blanks is important (W130 is different from W 130).

The next two items contain the position of the waypoint, thus equal to the position of POINTSDE.NEW and thereafter follows the description of the waypoint that is not used and only for the use to write comment.

10.1.1.2 Predefined links (file: LINKS.RAN)

For creating a database there are two ways to describe the outside world, namely:

- By defining all areas that are not used by shipping. This is done by defining lines that may not be crossed (FORBID.FIG, see 10.1.1.3);
- By defining dedicated waypoints and defining all possible links to and from a dedicated waypoint that may be used by shipping.

A combination of these two possibilities is the best approach for large sea areas. In the neighbourhood of ports it is easier to work with dedicated waypoints because some waypoints will always be used when calling a port. Because the port entrance is also located within the land mass, it creates the possibility to define the port within the forbidden area as a dedicated waypoint and predefine the only link to cross the border line of the land mass. In this case all ships that call the port will get the same penalty, thus the fact that the port is within the forbidden area does not influence the routing. It is not allowed to leave the area through another line, because the port is a dedicated waypoint.

10.1.1.3 Forbidden areas (file: FORBID.FIG)

During the routing process, the only criterion for finding the optimal route is the distance. The routing process can only be influenced by translating pros and cons into a distance that is added to the object function.

The main part of the traffic database describes the open sea, thus where ships are free to go where they want. At the start of SAMSON all possible routes were predefined. This approach is abandoned at the moment that a more detailed database was required. The new approach is a combination of free waypoints and dedicated waypoints. By using only dedicated waypoints, the first approach with only predefined links is back. For the new approach with free movements between all waypoints, additional input is required to describe the areas where no shipping is possible as land masses and shallow waters. This additional information is described in the file FORBID.FIG (see example in Table 10-1). Each line contains the following items.

- A "0" for the start of a new set of points or a "1" for a subsequent point in the set.
- The colour for plotting the line in CALMOD.
- The geographical position (latitude and longitude) in degrees and minutes in WGS84 coordinates.
- The fifth item is the penalty that has to be added to the distance as soon as a link crosses the line between the indicated point and the point defined in the preceding line of FORBID.FIG. This is only applicable for the second and following point of a set, thus when the line starts with a "1".
- The last digits, mostly "12345678" indicate for which networks the penalty has to be applied. In case of a blank on position of network 1 to 8, the penalty is not applied.



The penalties used for land masses are high, 500 nm or 5000 nm. There is not a large difference between these values, because ships may not cross land masses anyway. The difference has a historical reason because in the first years 500 nm was enough because the database was determined for the North Sea. Later, when also the Mediterranean was added the 500 nm was not enough to avoid links from a free waypoint in the North Sea to a waypoint in the Mediterranean, because the 500 nm was less than the detour around Spain. Then also 5000 nm was used to exclude these routes long routes.

Table 10-1 contains the representation of the Dutch coastline, that is also plotted in Figure 10-3. There is only one link with a penalty of 400. This is done for the offshore platform near Ameland, that is located south of the line (53°30'N,5°36'E) and (53°36'N,6°20'E). Suppliers visiting the offshore platform will now cross this line. In the originally routing these ships sail in straight line from Den Helder to the platform, because this was the shortest route. The penalty of crossing the border had to be taken once the best option was to do that at once near Den Helder. Decreasing the penalty on the border North of Ameland had as result that the suppliers crossed through this line and approached the platform in the correct way.

Table 10-1 Definition of the land mass of the Netherlands in FORBID.FIG

Nederland					
0	10	5125.0	355.0	500.0	12345678
1	10	5135.0	330.0	500.0	12345678
1	10	5214.0	426.0	500.0	12345678
1	10	5228.0	434.0	500.0	12345678
1	10	5303.0	443.0	500.0	12345678
1	10	5319.0	456.0	500.0	12345678
1	10	5330.0	536.0	500.0	12345678
1	10	5336.0	620.0	400.0	12345678
1	10	5347.0	744.0	500.0	12345678
1	10	5040.0	800.0	500.0	12345678

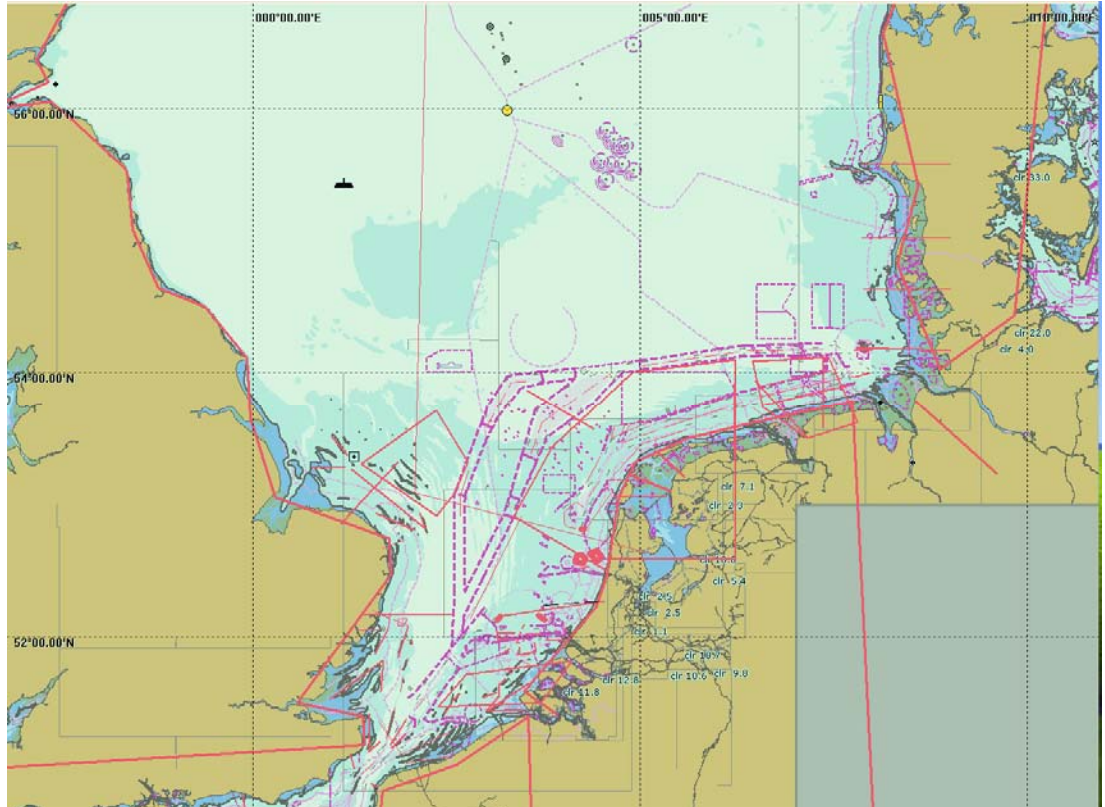


Figure 10-3 Plot of lines of FORBID.FIG

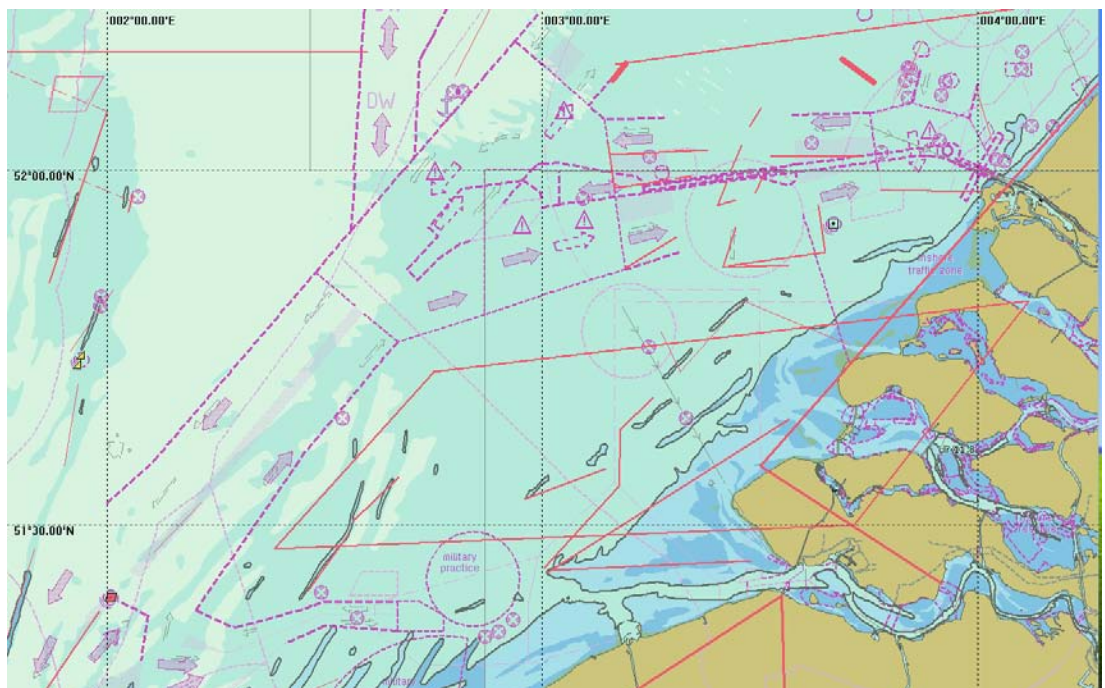


Figure 10-4 Lines of FORBID.FIG near Western Scheldt

10.1.1.4 Traffic separation schemes (file: STELSELS.DAT)

Traffic separation schemes (TSS) are defined at sea to create traffic lanes with one direction traffic. These traffic separation schemes are not mandatory, but it is not allowed to move against the direction, and the traffic lanes have to be crossed with perpendicular angles. This means the traffic will concentrate on these routes. But by the freedom of the sea, it is difficult to implement this behaviour in the routing process. The use of the TSS is implemented by giving a shortage on the travelled distance in a traffic separation scheme. Furthermore, the use of the waypoints of a TSS is limited. The perpendicular crossing of the TSS is difficult to implement. One has to be critical and define some lines in FORBID.FIG to avoid wrong crossings.

The TSS links are described in STELSELS.DAT of which an example is given in Table 10-2 for Dover Strait. This file can contain other non-TSS links that have to be dealt with in the same way as TSS-links during the routing process.

Each link of a TSS is described with one line, containing:

- The two codes of the waypoints.
- A factor that defines the advantage when taking this link. Thus 0.35 means that the routing process calculates with a length of $(1-0.35) \times$ (real length of the link).
- The use of the waypoints with a "*" or ".". In case of a blank the first waypoint can only be used coming from the direction backwards of the start point and leaving in the direction forwards the end point. In case of a "*" there is no direction limitation.
- "12345678" or blanks in it, to indicate which network may use this link.
- "OCOGFBUGFR" or blanks in it, to indicate which type of ships may use this link.

The meaning of the connection limitation is illustrated in Figure 10-5. This seems a little bit superfluous but it is not because otherwise the points can be used as waypoints on another route (red one of Figure 10-5) not using the TSS-link.

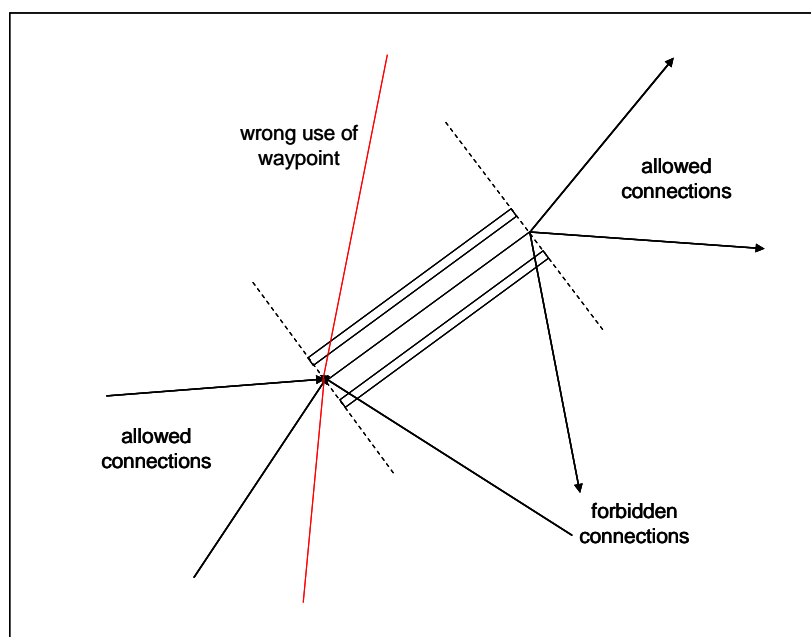
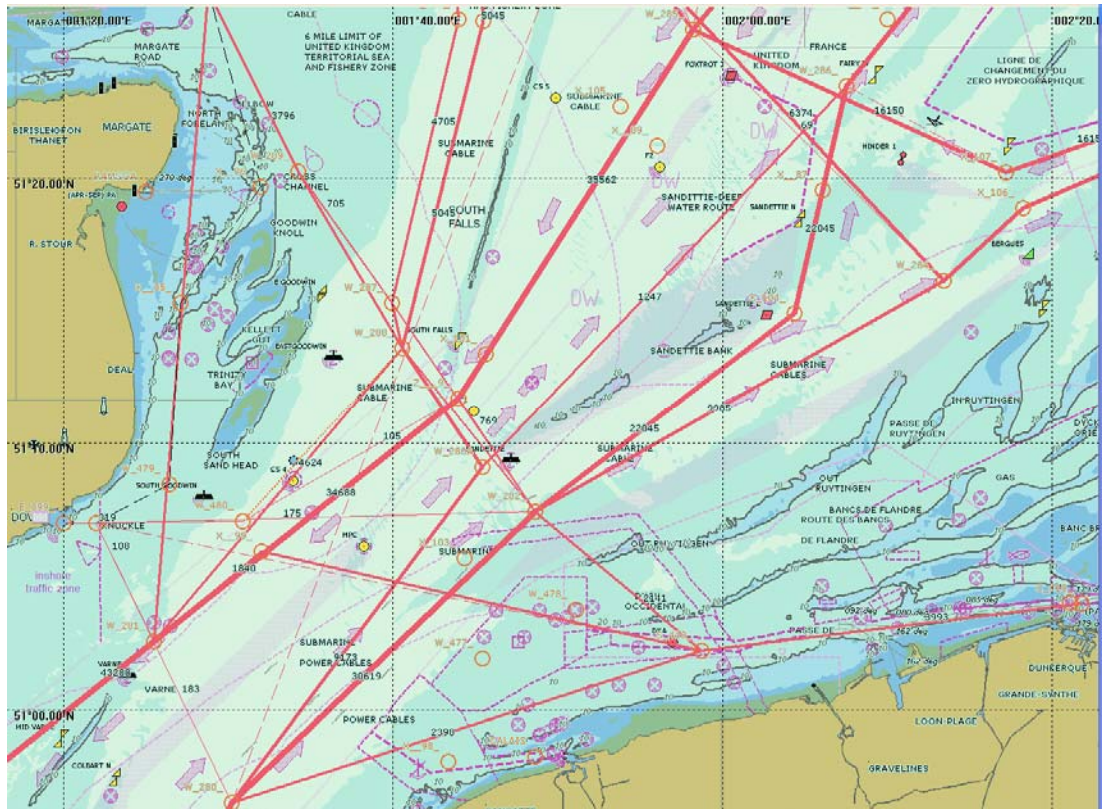


Figure 10-5 Connections to link through TSS

Table 10-2 TSS Dover Strait and some preferable links

TSS STRAIT DOVER NO					
W 270	W270G	.10	12345678	OCOGBUGFR	
W 270G	W 280	.10	12345678	OCOGBUGFR	
W 280	W 280A	.10	12345678	OCOGBUGFR	
W 280A	W 286	.10	78	OCOGBUGFR	
W 280A	W 286	-2.	123456	OCOGBUGFR	
W 280	W 282	.10	12345678	OCOGBUGFR	
W 282	W 284	.10	12345678	OCOGBUGFR	
W 282	W 286	.10	123456	OCOGBUGFR	
Links taken by small ships					
W2260G	W 391A	.10	*1234	OCOGBUGFR	
W 391	W2260G	.10	*1234	OCOGBUGFR	


Figure 10-6 Dover Strait

The example of Dover Strait is chosen for an elaboration because it contains some special aspects. All traffic through the TSS gets a distance reduction of 0.1 (10%). The link W 280A – W 286 is a very special link in the TSS because large ships have to follow the deep water passage while smaller ships have to sail a detour. This is solved by defining one line in STELSELS.DAT for large ships with “ 78”. These ships get a distance reduction of 10%. Because W 280A and W 286 are free waypoints that can be used by all ship type and size classes it has to be prevented that the smaller ships will take the link W 280A – W 286, because without distance reduction it will be shorter than the detour route. To prevent this use, not a reduction but an elongation is used. This is



defined in the line with the same link but now for “123456 “. The reduction factor is -2, meaning that the length of the link within the calculations will be $(1 - (-2)) = 3$ times the real length. Because this length is more than the detour, the ships will take the detour.

The last two lines are used to stimulate the use of these links for smaller ships “1234 “. The “*” now is used that the waypoints can be used for other routes, because these waypoints do not belong to a TSS.

10.2 Routeing process

The routeing process builds up a tree structure from the start point for a certain ship type and size. It contains all waypoints and for each waypoint it stores the minimum distance to reach that waypoint with the preceding waypoint. The process begins with the starting point, being the port of departure. From this point the distance to each point is calculated. This distance is stored with the preceding waypoint. Next, the waypoint with the smallest distance is chosen. From this new “exploring” waypoint, again the total distance to each other waypoint is calculated, thus the distance to reach the waypoint that is being explored plus the distance to the other waypoint. In case the calculated distance for a waypoint WP is less than the already stored distance for WP, a better route is found. The distance is changed, the preceding waypoint is changed and the waypoint is marked that the distance has been changed. The waypoints that have been changed have to be explored again.

Again and again the waypoint with the lowest distance that is not yet explored is explored. When the distance of any waypoint is not updated anymore during one cycle, the routeing process is finished and the optimal route to each destination from the starting point can be found by following the back pointers, thus the preceding waypoints. The process always leads to a result, except in cases where the start or end point, mostly dedicated waypoints, have no connections with free waypoints.

In the description here the distance is mentioned, but in fact it is more than a distance because the distance includes some extra items that are described in Chapter 10.1. The routeing process is explained with an example.

10.2.1 Example of the routeing process

The process is explained with the routeing tree that is shown in Figure 10-7. It contains far from all branches. Only the most important branches are included to describe the first steps of the routeing process.

The tree is built up for Zeebrugge (E 100). Figure 10-1 and Figure 10-2 contain a number of the waypoint codes of the example. The waypoints coded in red are “dedicated” waypoints thus the waypoints that only can be reached or left via predefined links. The black coded waypoints are free waypoints.

From the dedicated waypoint E 100, only two dedicated waypoints, W 88 and W 89, can be reached. The total distance to reach a waypoint is the value given in the waypoint box of Figure 10-7. W 88, located at a distance of 2.8 nm of Zeebrugge, is the waypoint with the lowest distance, not yet being explored. Thus W 88 is the next waypoint that is explored. There are two predefined links from W 88, one to W 128 to the Western Scheldt W 128, one to waypoint W 89 and one back to E 100 (of course skipped). The distance 3.2 already filled in for W 89 (from the direct link E 100 – W 89) is less than the distance via W 88 found in this step. Therefore the waypoint W 89 is



not updated with the new value. The light coloured waypoint boxes are found but not stored because there was already a better route in the tree.

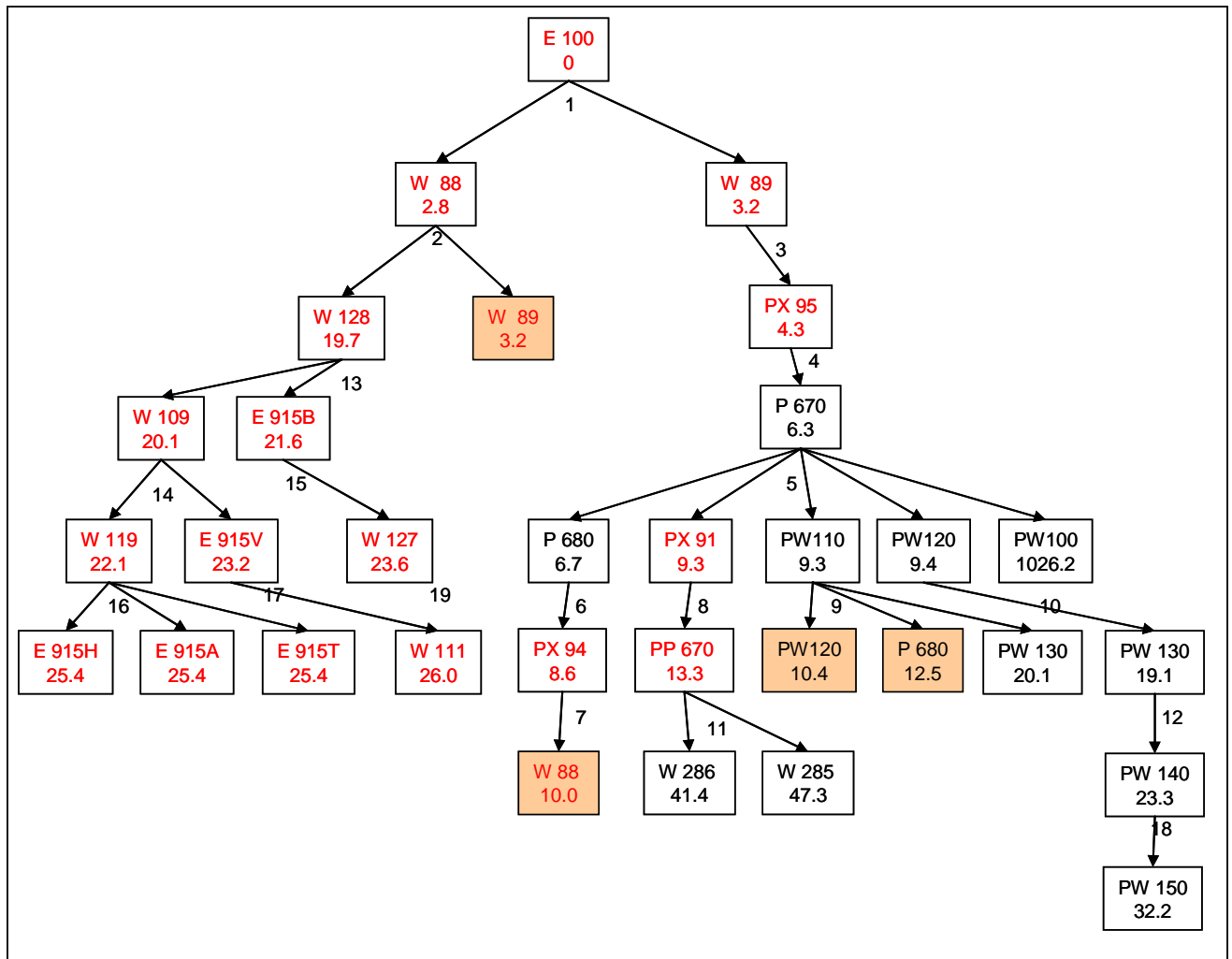


Figure 10-7 Routing tree

For each new step the waypoint with the lowest distance is taken. In this way the tree of Figure 10-7 is built up. In step 5, where the free waypoint P 670 is explored, the tree contains many hundreds of branches. Only the five nearby points are presented. The very large value for PW100 is caused by crossing a forbidden area, which is dealt with in 10.1.1.3.

The example needs still one remark. In step 9 the waypoint PW130 can be reached with a distance of 20.1 nm. In step 10 the waypoint PW130 can be reached after 19.1 nm. This means that a better route is found. The stored value of PW130 is updated with the new distance and a new preceding waypoint. After this change the way back of waypoint WP130 will be: WP130 > PW120 > P 670 > PX 95 > W 89 > E 100 and not anymore WP130 > PW110 > P 670 > PX 95 > W 89 > E 100.

At the end of the routing process all waypoints of POINTSDE.NEW contain the minimum distance to reach that waypoint and the preceding waypoint. By following the way back via the preceding waypoints, the route to each destination can be extracted from this tree. The present tree from Figure 10-7 contains already the routes to the



destinations in the Western Scheldt, namely for Breskens E 915B, Vlissingen E 915V, Hansweert E 915H, Terneuzen E 915T and Antwerpen E 915A. The destinations Hansweert, Terneuzen and Antwerpen are implemented as a waypoint in the Western Scheldt. The routes to the ports through the Western Scheldt are not further described, because it was outside the scope of the North Sea database.

10.3 Processing all movements

All movements that have to be represented by the traffic database are collected in the movements file. The base movements file is the file bought from LMIU once every four or five years. At this moment the last database acquired was from 2004. It contains a little more than one million records with ship movements in the format of IMO-number, sailed from port A with destination port B. This movement file linked to the ship characteristics database has been condensed to a movements file with the data items mentioned in the upper block of Figure 10-8. This file still contains nearly 80,000 records. During the creation of the database each record of this movement file is dealt with. First it is checked if the last created tree can be used for the new movement. This is the case when the port of departure A and the used network, that depends on the ship type and ship size, are still the same. In this case the route from B can be back traced and the movements can be added to the traffic database. If at least one of these two items is changed, a new routing tree is determined in the way that is described in Chapter 10.2. Thereafter the movements can be added to the traffic database. Next a new record from the movement file is read until all records are dealt with. In order to reduce the calculation time, it helps enormously when the movements are sorted in the sequence of network and port A. Now the number of times that the routing tree has to be determined decreases enormously. A programmed acceleration is that the routing tree is saved and restored when the same routing tree is requested, but this option works not as fast as a good sequence in the input file. Of course the traffic database will not depend on the sequence of the records in the movements file.

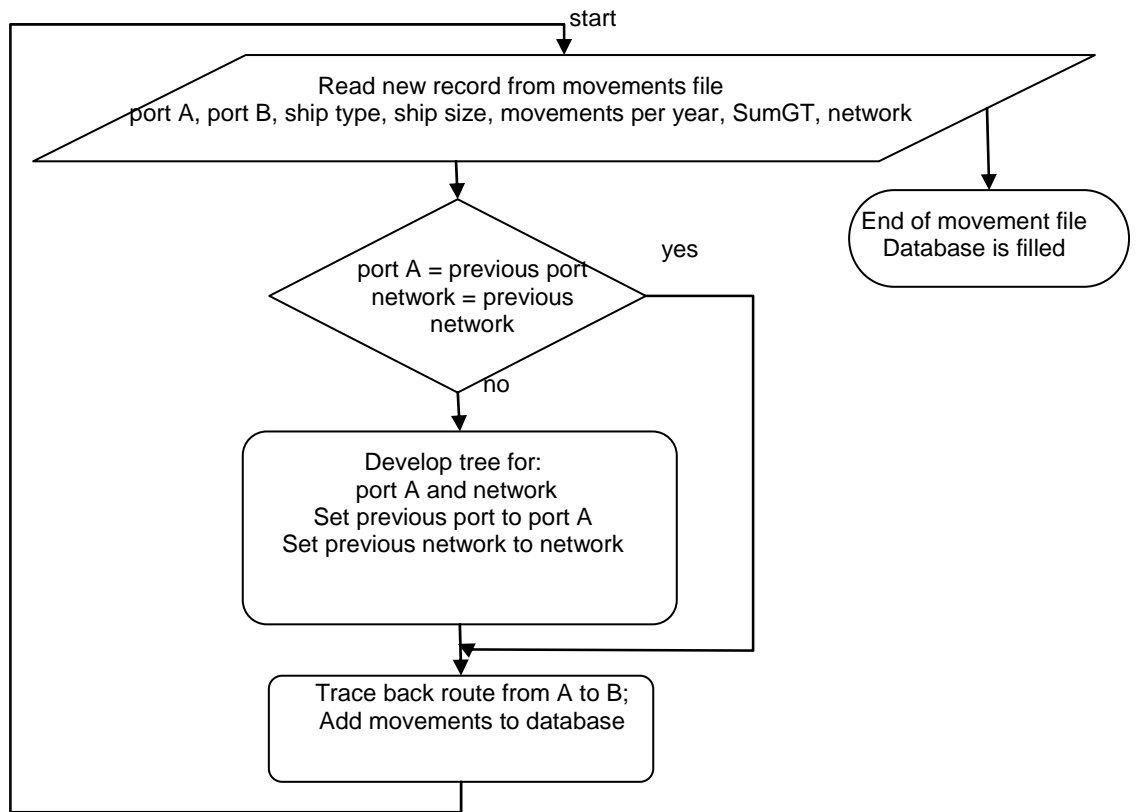


Figure 10-8 From movements to traffic database



10.4 Conclusions

The traffic database as result of the routeing process is the key stone for all calculations with SAMSON. Therefore it is a conscientious task. The elaboration in this technical report about the possibilities how to influence the routeing process, has made clear that creating a database is not a standard process. The best way to create a database depends on the complexity and size of the area and the maritime traffic. The user has a number of tools to model the outside world. The user has to be aware that when he changes something here, it can have unexpected changes elsewhere. This makes it necessary to test the result as much as possible before applying the traffic database within risk calculations.

The examples described in this report show that the outside world and behaviour of ships can be modelled in different ways. It is difficult to model some behaviour of ships in the routeing process. The experience is that the creation of a new traffic database is an interactive process. Based on the achieved results extra input is generated to steer the routeing process. After some iterations the traffic database is achieved that meets the requirements. Sometimes some imperfections remain, but these are acceptable because they do not really have impact on the results of a study.

10.5 Future work

At this moment the routeing process suffices. The problem is not the routeing process but more how to model the outside world. No special wishes exist at this moment.

Improvements seem to be possible for better and quicker generation and visualisation of the input files. Without any doubt this will result in a reduction of the time required for the creation of a new traffic database.