

The MARIDES project: Intelligent Chartering in the Maritime Industry

A. Giannopoulos, D. Kosmopoulos, T. Varvarigou, Y. Bouloudis,
C. Hatzithanassiou, G.Mourkosis, R. Varvarigou, A. Malamos

National Technical University of Athens,
Department of Electrical & Computer Engineering,
Division of Computer Science, Telecommunications Laboratory
Zografou 157 73, Greece

{ ang, dkosmo, dora, yboul, hachr, georgemr, rea, amalam }@telecom.ntua.gr

Abstract. This paper presents the research conducted for the MARIDES project, which aims to improve the chartering process in shipping companies. MARIDES (Maritime Decision Support) focuses on the development of the necessary decision support software tools, which operate in an integrated environment. As far as decision support is concerned, given a certain chartering order, the system will estimate all the possible solutions (all vessels that can serve the given order) and it will present them to the user sorted, according to the estimated profit. Apart from deterministic data, stochastic data is utilized as well and the system will take into consideration data such as the future market potential. This is achieved through estimation of the ballast voyage in order to load the next cargo after the completion of the current order¹.

Keywords : Decision Support System, Automatic Data Acquisition, Maritime Chartering Process.

1 Introduction

Waterborne transport efficiency is dependent on various factors such as the operating costs of the vessels, the charter selection, port time, route followed, costs of cargo handling, communications reliability and efficiency. The trading-chain involves numerous inter-linked, intricate stages and also a large number of mediators (sellers, buyers, charterers, ship-managers, ship-owners along with their corresponding brokers, services providers, etc). All these factors heavily influence the efficiency and cost in Shipping. Current policies in shipping procedures are rather simplistic and static, although the business processes, especially the ones concerning chartering, have a very dynamic nature. Decisions are usually taken on the spot and they are based mainly on experience and intuition. However it is possible to offer decision support, by introducing active, “knowledge based” systems, which will be using continuous incoming information gathering and will utilize historical and statistical data stored in databases.

Decision taking for chartering is heavily based on information that it is available through different means of communication, (phone, fax, Internet, e-mail or verbally). MARIDES is designed to improve and automate the collection, organization and presentation of all this information. The company should no longer rely solely on the ability of the human resources to exhaustively inquire and process all the relevant information.

The need for automation in chartering has not been ignored by the market. The products that have appeared so far in the shipping market can be divided into the *Standalone Systems*, the *Online Information Sources*, the *Electronic market places* and the *Hybrid* systems. The *Standalone Systems* such as [14], [17] are usually used for voyage estimation and time charter analysis without any particular intelligence. The *Online Information Sources* such as [12], [22], [24], [19] provide to subscribers a lot of market related time-variant information such as indices, fixtures, vessel positions, bunker prices etc. The amount of the information that is provided is huge and it requires a big amount of experience to evaluate. The *Electronic Marketplaces* like [27] [26] [20] [13] [16] [18] [15] host online auctions for members and matching of cargoes with orders. However they are not widely accepted yet because many ship owners or charterers are reluctant to expose their actual needs. The *Hybrid* systems are combinations of the above systems and such systems can be found at [21], [25].

MARIDES is an ambitious project, whose main goal is the development of a unified system that will provide and present to the end user all the crucial data about a given order but will also aid the user to make a decision considering the order itself, the possible alternatives and information about the environment which will influence the decision. The user is able to access, through a homogeneous environment, all the data that are crucial for the given order, and also to have access to any additional general information (such as vessel particulars etc.). According to the above classification the software developed belongs to the hybrid systems.

From the decision support point of view, given a certain chartering order, the system will estimate all the possible solutions (all vessels that can serve the given order) and it will present them to the user sorted, according to the estimated profit - optionally modified to cater for other factors such as the future market potential (see Figure 5 in section 4). The evaluation is based not only on deterministic data but also on stochastic

¹ This research work is supported by the European Community, under grant IST-1999-21129

data. As will be explained in section 3 the stochastic data used concern the possibility of the vessel to be chartered from the destination port and the route that the vessel should follow to load the next cargo.

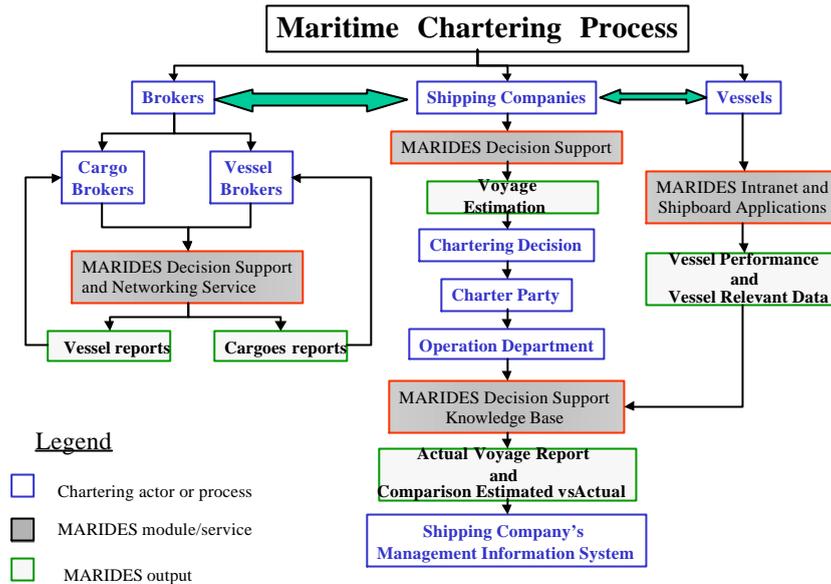


Figure 1 MARIDES incorporation in the maritime chartering chain.

The project addresses the chartering departments of shipping companies, by providing them with a system which creates an efficient way of work for the whole range of the company's staff, from its top management to selected fleet crew. Processing and availability of information in a unified representation form helps all the key players in the chartering process to find new opportunities of trade, improve the contractual conditions and reduce transaction costs. The combination of an enhanced communication intranet, a decision support tool and a knowledge base is a strategy that could lead a company to more effective decisions based on well weighed up criteria and not only on experts' intuitions. A diagram illustrating the way MARIDES is intended to be integrated in the chain of chartering actors and processes is presented in Figure 1.

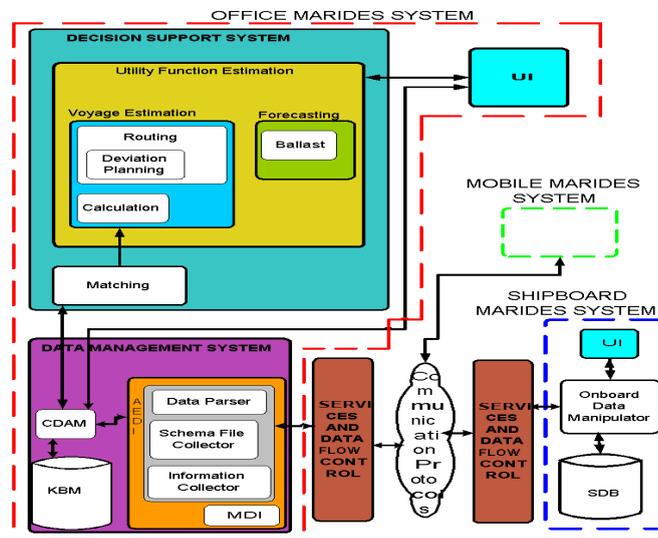


Figure 2 MARIDES Architecture Overview

2 MARIDES Architecture Overview

An overview of the MARIDES system is illustrated in Figure 2. The office system is the core of MARIDES and it comprises the Data Management System (DMS), which includes the central database (CDB), the Decision Support System (DSS) and the GUI. The mobile MARIDES system is a minimal version of the office system, for users that are not in the office. Its database is updated from the CDB.

The DMS collects, manipulates and stores data in order to make it usable by the algorithms and formulas of the DSS. It includes the CDB of the MARIDES system where all important data retrieved during the data manipulation process is stored. The specific data manipulation rules are stored in the knowledge base and the DMS automatically performs periodical accesses to remotely located data sites, in order to pro-actively seek and

update the content of the CDB. It also receives via e-mail incoming orders and reports from brokers in a predefined format.

The Shipboard MARIDES System (SMS) is the MARIDES module that is installed on-board and facilitates the transactional information exchange between the vessels and shipping company premises. Thus the chartering manager has a good overview of the chartering procedures regarding the whole fleet. The shipboard users are able to prepare and send their reports automatically or semi-automatically by using the SMS. All the reports from the office to the vessel and vice versa are stored in the Shipboard Database (SDB).

The system design was based on use-case analysis conducted making heavy use of end-user interviews. In order to make the novel functionality envisaged practical for the end-user, it was integrated into a platform including all standard office-automation services which the users expected from their computer. These fall into three broad categories:

- Voyage estimation (simple calculation of costs for a charter under consideration, its duration and so forth), port-distance calculation, ...
- Access to static data (e.g. vessel characteristics or port details), ...
- Communication with third parties (email, fax, telex), manual data input, ...

Integrating such functionality with the more specialised features (as described below) leads to the overall architecture depicted in Figure 2. Both the DMS and the DSS merge basic functionality with more advanced capabilities. For example, basic voyage estimation is a decision-support function, using modules such as routing. Utility estimation (described below) makes use of both voyage estimation and forecasting functions. Similarly, the intelligent automatic acquisition tools enrich the system database from which the user can view both old and fresh data by making similar, simple requests for information.

The end-user accesses MARIDES through a single, unified, fully integrated, task-driven interface (a couple example displays are shown below, in Figure 5). The upper-most interface consists of a web-like browsing tool which organises all activities according to the user's logic behind accessing and combining items of MARIDES functionality. Two panels to the right of the display, the task navigator and the context navigator, allow the user to exactly specify the desired task (whether involving calculation, display of data, both, or more) to be performed. The main section of the display is custom-designed for each task. It can display data in tabular or graphical format, require input for calculations, etc. – in general, any user-interface designed for a given task is placed here.

3 Automated Data Input

There are two kinds of data input tasks which have been improved, in MARIDES, from the purely manual tasks which they used to be to semi-automatic procedures: information extraction from any user communications received and active information acquisition from the internet. The solutions are semi-automatic in that any data which can be processed according to the rules which the system has learned is input automatically, while any unstructured or human-language-based information which is encountered results in the system paging the user to complete or, if necessary, fully perform the data entry.

User communications are via email, fax, telex and telephone, of which we process only emails – although use of a third-party OCR system for fax or telex would result in text files of identical format to the emails. Usually, these texts are almost entirely composed of application-related abbreviations. However, since the emails are written by humans, different abbreviations can appear for the same word, even in the same text. Also, the tabular format of the information exists for readability and is not of guaranteed structure. Sometimes the structure changes, especially when it comes from different sources. Natural language occasionally appears to comment on some fact, or the order of the abbreviations can occasionally be somewhat unpredictable. The system thus applies a template-based pattern-matching model of the communications structure to extract valuable information from incoming texts. The system's databases are automatically enriched with all information which is successfully extracted – this information can then be used by Decision Support components or viewed by the end-user as naturally as local data, without the need to browse lengthy communications. Note that the system is tolerant to unreadable texts or parts thereof. What cannot be processed is simply ignored as far as the input system is concerned and the user is optionally prompted to complete the task manually.

Automatic acquisition of information from the internet is also a hard task, but in terms of work necessary rather than in the design of the system. For each source (internet site) to be used, a wrapper must be written. This allows information to be extracted from pages which have been structured for presentation to human readers. Wrappers were designed for a number of important internet sites, allowing MARIDES to automatically learn fixture results, fuel prices, etc. Naturally, the user has the option to supervise and/or augment this information gathering process in the same way as with the processing of communications.

The value of this technology is that both data input systems can operate transparently to the user. Any incoming emails are automatically scanned, and the user can even choose to have them deleted after processing. Also, the internet acquisition tool need not be started by the user, but polls the sources used based on their

refresh times/rates. All information acquired is stored in the MARIDES database. Thus, the user receives information upon request, rather than upon its arrival. Of course, the user can view all communications and make any manual polling requests if desired.

4 Decision Support System (DSS)

This section presents the most important aspects of the design of the DSS subsystem. The DSS functionality consists of three separate modules, the third of which is presented in detail in this paper:

- Voyage estimation² calculations, performed according to standardised formulas used in the shipping industry
- Market monitoring module, which draws upon the data handling and input modules; this module involves informing the user of market data, which should affect his decision but is not normally considered by voyage estimation, such as the movements of competing vessels.
- Assessment of proposed charter business and main terms; this module makes considerable machine intelligence requirements, relying primarily on modelling and statistics, and is presented in detail in the following.

Figure 3 presents a schematic architecture of the third module:

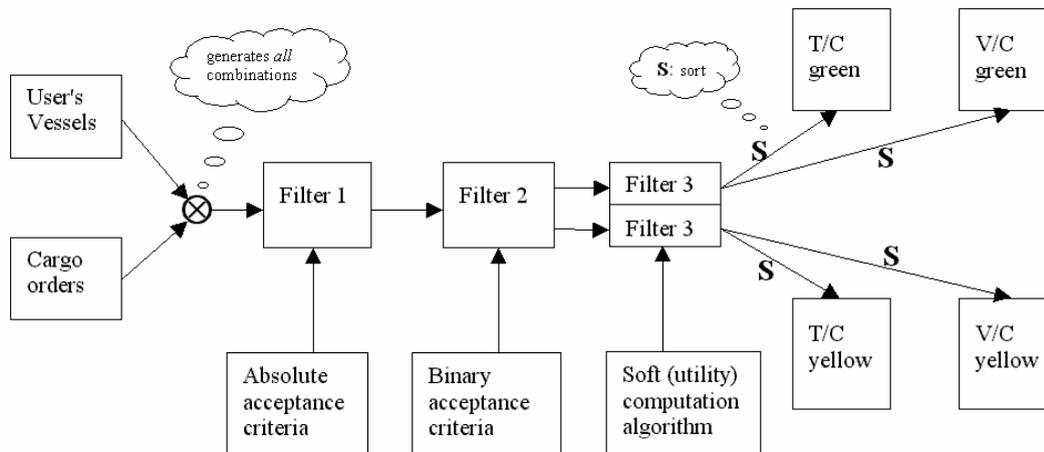


Figure 3 Flow chart for the DSS sub-component which assesses cargo orders.

The system's input consists of all possible combinations of 1) the set of incoming cargo orders and 2) the set of ships and their status. These combinations are then fed through the three Filters.

The system's outputs are ship-cargo matches organised according to the fashion in which they should be displayed to user: 1) in separate windows according to a general type and desirability classification and 2) sorted from best to worst according to the evaluated utility functions. Four windows are used, separating combinations which cannot be compared to each other except by a human expert: 1) Time-Charter and Voyage-Charter³ options are separated, since they offer a ship-owner essentially two different strategies of employment for his

² The term Voyage Estimation refers to standard calculations performed by Chartering managers, either by hand or with some simple form of automation, in order to make a basic prediction of the cost of a voyage based on known parameters (such as the current fuel prices) and well-understood mechanisms in shipping (generally simple ones; for example, "total fuel consumption" is "daily fuel consumption" times "days at sea"). Some more complicated parameters are generally simplified based on experience: "daily operating costs" is a constant derived from experience operating the vessel under consideration and includes all costs which might be considered "miscellaneous", such as acquiring provisions for the crew, fixing chance damages of a general nature, and so on.

³ Voyage Charter (V/C) refers to a contract where a vessel undertakes a single voyage under hire based on the quantity of cargo carried, while Time Charter (T/C) refers to a contract where a vessel is hired for a period of time independently of what voyages might be undertaken. Except for the basic difference that V/C is a contract of substantially smaller duration than T/C, it is also very important that under V/C the ship-owner faces any unexpected risks, since the fee is a constant for the voyage as planned in advance, while under T/C the ship-owner receives a constant fee for the duration of the charter and is not in danger from risks which might cause the voyages actually performed to deviate from the initial plan agreed upon. However, since T/C is a low-risk option for a ship-owner, the fees agreed upon are typically lower than what might be expected from a consideration of V/C fees.

fleet, and 2) “problematic” options⁴ (of each charter type) are also shown separately for the user to judge whether the appearance of a high rate might offer an unexpected opportunity.

As a black box, the module is a system which finds feasible matches of cargo orders with vessels and presents them to the user sorted according to their assessed desirability; matches which cannot be compared to each other but should be considered separately by the user are provided through separate output views.

4.1 *Filters 1 and 2*

Filter 1 chooses all feasible matches of cargo orders with vessels. Filter 2 separates fully acceptable cargo orders from others, which have some qualitative problem related to them. Normally a chartering manager would ignore the options in the worse set, unless a very expensive rate were offered. For example, it may be that certain vessel is not normally used to carry dangerous materials (although, physically and legally, it could do so) or that a given cargo order originates from a Charterer who is not trusted: these options would normally be dismissed without further consideration. Since it is not possible for a computer system to analyse how unusually high a fee would make such options worth consideration, they are separated from the rest, so that a single glance of the user at the top of their list will show whether any “dangerous” but profitable opportunities are available.

Filters 1 and 2 are an issue of Data Management (modelling and processing the appropriate data types). The criteria used are exactly specified (as illustrated below), although our increasing experience with the system’s employment in practice will lead to gradual improvements or additional options being added. Once the databases have been built, the appropriate results can be obtained with fairly straightforward SQL, some pseudo-code examples of which are given at the end of this section.

Basic Filter 1 criteria:

1. Ship can carry this type of cargo
2. Ship is allowed to travel (from origin) to destination (or within specified Trading Areas for T/C) (e.g. Australia does not accept old ships)
3. Ship is capable to travel (from origin) to destination (or within specified Trading Areas for T/C) (not the same as above, e.g. target port may be too shallow, use Loadport and Disport characteristics)
4. Ship is capable of travel for the desired duration (e.g. no need for dry-dock before end of Charter)
5. Ship is open for charter during the period required by the cargo order (from ship records and first Laytime and last Cancelling date)

Basic Filter 2 criteria:

1. Cargo type is to be avoided (re: cargo exclusions)
2. Voyage passes through regions which are to be avoided by user choice
3. Voyage passes through conditionally undesirable regions
4. Voyage passes through regions suffering from extremely bad weather
5. Charterer is not trusted
6. Broker is not trusted
7. Market is very low and duration of charter is long
8. Market is very high and duration of charter is short
9. Very high level of competition detected at and/or near destination port
10. User has specified preferences for voyage parameters which are not satisfied

Filter 1 criterion 2, example:

```
SELECT FROM my_ships × my_orders  
WHERE order_destination  
NOT IN ship_travel_restrictions
```

Filter 2 criterion 2, example:

```
SELECT FROM result_of_filter_1  
WHERE (index_for_vessel_type_considered < user_defined_market_threshold1)  
AND (||order_delivery_date - order_load_date|| > user_defined_time_threshold1)
```

4.2 *Filter 3*

This is the most heavily computational sub-module, which comes closest to decision *making* than any other one. Of course, the purpose here is not to supplant the user’s judgement, but rather to make as powerful an estimation as possible, and then to offer this estimation as an aid to the user. The result is an intelligent heuristic to assess the effect of a voyage’s destination on its desirability.

A forecasting module integrated with a standard Voyage Estimation would theoretically provide a more direct solution to the problem of decision support: with the exact parameters of future options accurately predicted (as might be hoped from research described in [1] [3] [11]) and varying for each possible choice, optimal planning could be achieved. However it is impossible for the accuracy of such forecasts to be adequate

⁴ see description of Filter 2 in the next section

for long-term planning (even just weeks ahead) in this application, since the parameters to be predicted are fuel prices, market conditions, political decisions and instabilities, details about global weather, etc.

MARIDES addresses the problem based on the idea that an *indirect* indication, a useful *feature* can be discovered, being an easily measurable characteristic which can be drawn from historical data, and also a reasonably “easy” to predict characteristic of future voyages. This feature need not be something normally accounted for in usual calculations, if only it depends on the important factors which should influence decisions and from which the effect of these factors can be estimated.

Filter 3 *augments* the normal Voyage Estimation procedure by including a prediction of the duration or distance of the Ballast Voyage⁵ which will *follow* the voyage being considered. This is an indirect measure of the desirability of travelling to the destination port of the current voyage. On the one hand, decision support follows directly: the longer the estimated future ballast, the more the estimated profits for the current voyage decrease; the measure is naturally in monetary units, although of course not comparable to the results of normal Voyage Estimation. On the other hand, prediction is a reasonable task: we need not predict fuel prices, for example, which would depend on immeasurable factors such as the influence of politics; ballast size is predicted based on location, season, type and size of ship, and the detection of patterns present in historical data is a reasonable statistical task. This discussion is summarised in list form as follows:

feature is feasible to predict:

- reasonable quantities of commonly available data required (for example [12] [22] [24] [19] provide good databases for this purpose)
- reasonable prediction accuracy possible (as reasoned above)

feature is a useful indication for decision making:

- depends on ship type and size and on location and seasonal considerations, easily expandable to more features
- strongly measures a basic desirability connected with voyages, that is, basic patterns in the length/duration of (future) ballast voyages, due to geography, trading patterns, and so on (a clear example is that a tanker delivering to Japan is condemned to a longer ballast voyage than another delivering to the USA, since there is no hope of reloading in Japan).
- in addition, the measure is not a simple cost-related feature taken from historical data; it is correlated with many factors (market fluctuations, bunkering, anything which will influence a charterer's decision-making) and so offers at least a partial measurement of all relevant information. Whereas such parameters could not be estimated directly, the correlated feature used is in fact summarising some of the *most important* results of their influence: the cost of a vessel travelling without cargo, as predicted based on real data, i.e., based on the way real chartering decisions were taken in the past.

The process of Ballast Voyage estimation involves the following steps:

1. Draw upon data for estimation (*pre-processing step*) – build a database with records of past ballast voyage distances from historical data, extracting information on ballast voyages from consecutive entries in fixture lists⁶, excluding problematic data (missing values or outliers which are likely to be due to incorrect reports or other errors)
2. Similarity matching – find the degree of similarity between available historical records and input; we use a utility-theory-inspired similarity metric [23] to find records which are useful for using as samples for a prediction for the unknown quantity. A basic additive form of this function will be used:

$$\begin{aligned}
 dist(\text{record1}, \text{record2}) &= \\
 &= dist(\text{port1}, \text{port2}, \text{size1}, \text{size2}, \text{time_of_year1}, \text{time_of_year2}) \\
 &= g_{ports} \{ dist(\text{port1}, \text{port2}) \} + f_{size} \{ dist(\text{size1}, \text{size2}) \} + \\
 &+ f_{time} \{ dist(\text{time_of_year1}, \text{time_of_year2}) \}
 \end{aligned} \tag{1}$$

with data-format-appropriate “distance”, or difference, measures used for each type-specific *dist* function. Additional factors could be included in a straight-forward manner, since the function is additive, if richer input data could be acquired for a later version of the system.

3. Data selection – choose which records should be used as a basis for prediction; the most similar results from the previous step will be used, but the number of records which are significant for the prediction must be

⁵ Vessels do not always travel loaded with cargo. It often passes that the loading port for one voyage is not the same as the discharge port of the previous one. The vessel must then travel without cargo, in other words “in *Ballast*”, between these ports. The costs of a Ballast Voyage are normally considered together with the costs of the normal voyage which commences from a port which is located away from the vessel's current location. The fact that traveling to a given port implies predictable probabilities of Ballast to possible next loading ports is not normally used for calculations (although humans may of course intuitively include such facts in their considerations).

⁶ fixture lists are lists which are made public and contain information on chartering contracts agreed upon for all vessels in the market

determined. We use statistical significance and goodness-of-fit criteria [2] [6] [8] [9] to make this decision. The Kolmogorov-Smirnov statistic [8] is especially useful to find the most appropriate probability distribution to use (this is also useful for the next step) by comparing actual data to possibly matching common⁷ [6] distributions:

$$Dn^+ = \text{MAX}(S(x[j]) - F(x[j])) = \text{MAX}(((j+1)/n) - F(x[j])) \quad (2)$$

where $S(x[j])$ is the cumulative distribution of the sample and $F(x)$ is any reference cumulative distribution. Dn^- is similarly defined, using MIN instead of MAX

4. Smoothing – estimate ballast distribution; the final prediction is based either by fitting a probability distribution to the data, or by taking the data itself to represent the probability density of the result, but in this latter case explicit smoothing must be used to reduce the effect of noise, outliers etc. in the historical data. The first case involves fitting the data to the distribution chosen from the step above, using standard maximum-likelihood techniques [10].
5. Utility calculation (*post-processing step*) – find utility of voyage having estimated ballast probability distribution. The utility of the voyage under consideration is calculated by a numerical integration based on:

$$U(\text{voyage}) = \int_0^{\infty} u(\text{TCE} | n\text{Bal}) p(n\text{Bal}) dn\text{Bal} \quad (3)$$

where: $n\text{Bal}$ is the next-voyage ballast duration, estimated as described, $u(x) = \sqrt{x+A}$, as studied by Cullinane [4] [5], $p(x)$ is the probability of x and TCE is the Time-Charter Equivalent⁸. The TCE is evaluated by a standard Voyage Estimation where the price charged is either that quoted by the cargo order, or, if there exists no quotation, the highest possible (*realistically* possible) derived from the last-done value⁹ and, when available, additional fixture information from fixture lists (last-last-done value etc). The default setting for a Voyage Estimation is a break-even rate computation, that is, what TCE must be charged in order for profit to be 0. Here, we require a TCE profit given that we charge the highest rate feasible. Last done values show the rates, which are currently agreed upon in the market, this of course being just an indication of what Charterers will accept to pay.

The data flow described here is depicted in Figure 4. In the figure, it is shown how an input voyage for consideration is compared to other voyages. Data available for comparison is in the first three columns, port, size and time. The value z (representing a distance to be travelled) is to be estimated using the records of ballast voyage distances $a, b, c, d, e, f, g, h, i, j$. Once the similarity of the input to each of the stored records is assessed, sorting these similarities gives the ordering of the ballast voyage distances from the most to the least useful. How many of these values to use, is determined by selecting enough to get a reasonable representation of the distribution in question; the CDF (cumulative distribution function) is depicted in this case. In Step 5, above, the numerical integration (not depicted) would result in a prediction of z .

5 Implementation

The MARIDES system is implemented for Windows 2000 or Windows XP. The databases of the system are based on Microsoft SQL Server technology. The internal communication with the database is performed through ADO recordsets. The GUI is implemented as Active X documents implemented using Visual Basic (see Figure 5). The overall organisation of the GUI is based on a custom platform called BOS (Business Operating System). The DSS utilizes SQL to access the database and the algorithms are programmed in C++ dlls.

The communication within the system components concerns mainly files and it is implemented through FTP or SMTP (attachment of XML files). The communication with the vessels is performed over a dial-up connection via satellite. Standard cryptographic techniques are used over all remote links to ensure confidentiality and integrity of the transmitted information.

⁷ normal, gamma, etc.

⁸ The Time Charter Equivalent is a standard way to describe a fee for a voyage charter by simply dividing the net fee by the estimated duration of the voyage. It is a simple normalisation and the chartering manager needs to intuitively understand the arising incongruences due to the differences between Time and Voyage charters (for example, the differences in risks explained above).

⁹ the last-done value is in turn derived from fixture lists (this is an issue of basic data input)

- [6] Fisz, M. (1963). *Probability theory and mathematical statistics*, 3rd edn. New York: Wiley.
- [7] Keeney, R., Raiffa, H. *Decisions with multiple objectives: Preferences and value tradeoffs*. New York: Cambridge University Press, 1999.
- [8] Kendall, M., Stuart, A. *The advanced theory of statistics*, 4th edn. London: C. Griffin, 1977.
- [9] Papoulis, A. *Probability, random variables, and stochastic processes*, 2nd edn. New York: McGraw-Hill, 1984.
- [10] Stark, H., Woods, J. W., *Probability, random processes, and estimation theory for engineers*, 2nd edn. Englewood Cliffs, N.J.: Prentice Hall, 1994.
- [11] Yao, J., Tan, C. L., A case study on using neural networks to perform technical forecasting of forex. *Neurocomputing*, vol 34, pp.79 – 98, 2000.
- [12] Baltic exchange: www.balticexchange.com
- [13] Bulknet: www.bulknet.com
- [14] Burmester & Vogel: www.voyage-estimation.com
- [15] Cargo Biz: www.cargobiz.com
- [16] Cargonow: www.cargonow.com
- [17] Danaos: www.danaos.com
- [18] Interbox: www.interbox.com
- [19] Intertanko: www.intertanko.com
- [20] Laycan: www.laycan.com
- [21] Levelseas: www.levelseas.com
- [22] Lloyds: www.seasearcher.com
- [23] MARIDES, www.marides.com
- [24] Maritime data: www.maritimedata.com
- [25] Net Ship Brokers: www.netshipbrokers.com
- [26] Ocean Connect: www.oceanconnect.com
- [27] ShipIQ: www.shipiq.com