CHAPTER TWO

LITERATURE REVIEW

2.1. INTRODUCTION
Simulation is a powerful tool available to decision-makers responsible for the design and operation of complex processes and systems. It makes possible the study, analysis and evaluation of situations that would not be otherwise possible. In this chapter we review the basics of simulation, tools for performance improvement by simulations, basics of logistics and terminal systems. We also present a detailed review of relevant literature related to container terminals, rail yards and airport terminal systems.

2.2. BASICS OF SIMULATION
Law and Kelton(1991), Banks et al(1995) give good introduction to the art and science of simulation. According to Shannon(1998) simulation is defined as the process of designing a model of a real system and conducting experiments with this model for the purpose of understanding the behavior of the system and/or evaluating various strategies for the operation of the system. Thus it is critical that the model be designed in such a way that the model behavior mimics the response behavior of the real system to events that take place over time.

The term's model and system are key components of our definition of simulation. By a model we mean a representation of a group of objects or ideas in some form other than that of the entity itself. By a system we mean a group or collection of interrelated elements that cooperate to accomplish some stated objective. One of the real strengths of simulation is the fact that we can simulate systems that already exist as well as those that are capable of being brought into existence, i.e. those in the preliminary or planning stage of development.
2.2.1. Advantages and Disadvantages

Simulation has a number of advantages over analytical or mathematical models for analyzing systems. The basic concept of simulation is easy to comprehend and hence often easier to justify to management or customers than some of the analytical models. In addition, a simulation model may be more credible because its behavior has been compared to that of the real system or because it requires fewer simplifying assumptions and hence captures more of the true characteristics of the system under study.

Other advantages include:

- We can test new designs, layouts, etc. without committing resources to their implementation.
- It can be used to explore new staffing policies, operating procedures, decision rules, organizational structures, information flows, etc. without disrupting the ongoing operations.
- Simulation allows us to identify bottlenecks in information, material and product flows and test options for increasing the flow rates.
- It allows us to test hypothesis about how or why certain phenomena occur in the system.
- Simulation allows us to control time. Thus we can operate the system for several months or years of experience in a matter of seconds allowing us to quickly look at long time horizons or we can slow down phenomena for study.

2.2.2. Simulation Concepts

Although there are several different types of simulation methodologies, we will limit our concerns to a stochastic, discrete, process oriented approach. In such an approach, we model a particular system by studying the flow of entities that move through that system. Entities can be customers, job orders, particular parts, information packets, etc. An entity can be any object that enters the system, moves through a series of processes, and then leaves the system. These entities
can have individual characteristics which we will call attributes. An attribute is associated with the specific, individual entity. Attributes might be such things as name, priority, due date, required CPU time, ailment, account number etc.

As the entity flows through the system, it will be processed by a series of resources. Resources are anything that the entity needs in order to be processed. For example, resources might be workers, material handling equipment, special tools, a hospital bed, access to the CPU, a machine, waiting or storage space, etc. Resources may be fixed in one location (e.g. a heavy machine, bank teller, hospital).

The essence or purpose of simulation modelling is to help the decision-maker solve a problem. Therefore, to learn to be a good simulation modeler, one must merge good problem solving techniques with good software engineering practice. Following steps are mentioned in Carson(2002) for a successful simulation study.

1. **Problem Definition.** Clearly defining the goals of the study so that we know the purpose, i.e. why are we studying this problem and what questions do we hope to answer?

2. **Project Planning.** Being sure that we have sufficient and appropriate personnel, management support, computer hardware and software resources to do the job.

3. **System Definition.** Determining the boundaries and restrictions to be used in defining the system (or process) and investigating how the system works.

4. **Conceptual Model Formulation.** Developing a preliminary model either graphically (e.g. block diagram or process flow chart) or in pseudo-code to define the components, descriptive variables, and interactions (logic) that constitute the system.

5. **Preliminary Experimental Design.** Selecting the measures of effectiveness to be used, the factors to be varied, and the levels of those factors to be investigated, i.e. what data need to be gathered from the model, in what form, and to what extent.
6. Input Data preparation. Identifying and collecting the input data needed by the model.

7. Model Translation. Formulating the model in an appropriate simulation language.

8. Verification and Validation. Confirming that the model operates the way the analyst intended (debugging) and that the output of the model is believable and representative of the output of the real system.

9. Final Experimental Design. Designing an experiment that will yield the desired information and determining how each of the test runs specified in the experimental design is to be executed.

10. Experimentation. Executing the simulation to generate the desired data and to perform sensitivity analysis.

11. Analysis and Interpretation. Drawing inferences from the data generated by the simulation runs.

12. Implementation and Documentation. Reporting the results, putting the results to use, recording the findings, and documenting the model and its use.

Further discussions of these topics in simulation can be found in Banks et al(1995), Carson(2002,2003), Kelton(1995,1996). These well established methodology was used by us also in our work.

2.3. SIMULATION FOR DESIGN AND OPERATION IMPROVEMENT

Several case studies, briefly described here, illustrate the power of and benefits from simulation in improvement of operations. Simulation can profitably be applied to manufacturing system design during any or all stages of the production system life cycle - the conceptual design phase, the detailed design phase, the launching phase, or the fully operational phase (Ülgen and Upendram,1997).

The flexibility of simulation permits its application to a wide variety of manufacturing problems, such as capacity planning, machine and personnel scheduling, inventory control, and job routing (Martinich 1997). However, the
Combination of other competing objectives (e.g., reduction of material handling costs, high resource utilization, and low variance of output production) and stochastic variation required the analytical power of simulation. Simulation has a long and strong track record in analysis of manufacturing systems whose complexity and interaction of components defy closed-form methods (Clark, 1996).

Simulation had already become an accepted tool for improvement of manufacturing productivity in this context through documentation of previous successes and availability of training, as advocated by (Williams and Sadakane, 1997). Furthermore, simulation is most profitably used not as a "one-shot" technology for addressing questions during process design, but as a continuous improvement tool throughout the lifetime of the manufacturing process (Nelson, 1994). Application of simulation study in material requirements planning [MRP] is described in (Dittrich and Mertens, 1995). Swamidass and Winch (2002) present simulation and modelling as one of soft technologies included in a list of commonly accepted Advanced Manufacturing Technologies (AMT).

Material handling, the art and science of moving, storing, protecting, and controlling material between value adding operations, is one of the most complex and economically important function within a manufacturing system. Specifically, examples of simulation applications to material handling abound in the literature, such as optimization of operating policies for an automated material handling system (Dallari et al. 1996), evaluation of a distribution center tow-line material handling system (Bakst, Hoffner, and Jacoby, 1996), configuration of a material delivery system with dolly trains (Jeyabalan and Otto, 1992), development of dispatching rules for multiple-vehicle automatic guided vehicle [AGV] systems (Lee 1996), and improvement of a pull-strategy in the order-picking area of a distribution warehouse (Alicke and Arnold, 1997).
2.4. TECHNIQUES OF PERFORMANCE IMPROVEMENT USING SIMULATION

With the continuing advances in computer technology, simulation is increasingly used as a decision making tool. Most real-world systems are complex and computing values of performance measures and finding optimal decision variables analytically is very difficult and sometimes impossible. Computer simulation is frequently used in evaluating complex systems and optimizing responses.

A number of techniques and approaches have been proposed to solve the simulation optimization problem. There are several survey papers that discuss foundations, theoretical developments and applications of these techniques (Meketon, 1987; Jacobson and Schruben, 1989; Safizadeh, 1990; Azadivar, 1992; Fu, 1994a; Andrad´ottir, 1998; Swisher et al., 2000). The simulation optimization techniques discussed in these papers are listed in Figure 2.1 (Tekin and Sabuncuoglu, 2004).

![Figure 2.1: Simulation optimization techniques](image)
Local optimization problems are discussed in terms of discrete and continuous decision spaces. In a discrete space, decision variables take a discrete set of values such as the number of machines in the system, alternative locations of depots, different scheduling rules or policies, etc, and in a continuous space, the feasible region consists of real-valued decision variables such as order quantity and reorder quantity in inventory problems, release time of factory orders, etc.

Two most popular methodologies for the class of problems are: (i) ranking and selection; and (ii) multiple comparison procedures. For reviews of these two classes of techniques, one can refer to Bechhofer et al. (1995) and Goldsman and Nelson (1998). Other methods (e.g., random search, Nelder-Mead simplex/complex search, single factor method, Hooke-Jeeves pattern search) can operate in the infinite parameter space.

A great amount of work has been done with problems that have a continuous decision space. Below we discuss the most common methods from the literature.

Response surface methodology: Response Surface Methodology (RSM) is a class of procedures that: (i) fit a series of regression models to the responses of a simulation model evaluated at several points; and (ii) optimize the resulting regression function.

A survey of the RSM research from 1966 to 1988 is given in Myers et al. (1989). Box and Draper (1986) has an extensive discussion on response surfaces and experimental designs. Kleijnen (1998) discusses the use of statistical designs for what-if analysis in simulation and emphasizes how RSM combines regression analysis, statistical designs and the steepest descent (ascent) method to optimize a simulated system. Factorial and fractional factorial orthogonal designs are the best known first-order designs for RSM (Montgomery, 1991). Composite and rotatable designs are the most useful second-order designs (Montgomery and Evans, 1975). Ramberg et al. (1991) relate the orthogonal arrays advocated by Genichi Taguchi to classical experimental designs and use Taguchi’s techniques in the construction of mathematical metamodels for RSM.
Some researchers use RSM with other methods such as gradient-based techniques, quasi-Newton methods, and simplex experimental designs (Safizadeh and Signorile, 1994; Joshi et al., 1998). There are many examples of its real-time implementations. Kleijnen (1990, 1995) presents optimization of a decision support system of a Dutch company via RSM. Shang and Tadikamalla (1993) investigate a computer-integrated manufacturing system of an automated printed circuit board manufacturing plant and implement RSM to maximize output.

RSM provides a general methodology for optimization via simulation. Compared to many gradient methods, RSM is a relatively efficient method of simulation optimization in the number of simulations experiments needed. Another advantage is that it uses well-known statistical tools. The main drawback of RSM is its computational requirements if applied blindly.

**Gradient-based methods:** Four methods in the simulation optimization literature that are used for estimating gradients of the response: (i) finite difference estimates; (ii) perturbation analysis; (iii) frequency domain analysis; and (iv) likelihood ratio estimates.

Perturbation Analysis (PA) was introduced by Ho et al. (1979) in the context of a buffer allocation problem in serial production lines. PA, when applied properly to models that satisfy certain conditions, estimates all gradients of an objective function from a single simulation run. There are two classifications of PA: Finite Perturbation Analysis (FPA); and Infinitesimal Perturbation Analysis (IPA). FPA is designed for discrete parameters and is an heuristic that approximates the difference in a performance measure when a discrete parameter is perturbed by one unit. IPA is used to obtain derivatives of continuous parameters and estimates all partial derivatives from a single run by keeping track of related statistics of certain events during a run.

Since PA performs well in simple discrete-event dynamic systems which can be modeled as queueing networks, there are a number of papers on the applications of PA to queueing systems; i.e., Ho et al. (1984), Ho (1985). Wardi et al. (1991),
Chong and Ramadge (1993) and also Fu and Hu (1994). PA has been widely used for optimizing manufacturing systems of interest. Donohue and Spearman (1993) determine the most profitable capacity configuration for a production line by using PA. Yan et al. (1994) use PA to develop algorithms to approximate the optimal threshold values in a manufacturing system with two tandem machines. Heidergott (1999) uses smoothed PA to optimize threshold values of repair times in a maintenance model.

Various metaheuristics have been suggested for simulation optimization. Such methods include genetic algorithms, simulated annealing, tabu search, and neural networks. Although these methods are generally designed for combinatorial optimization in the deterministic context and may not have guaranteed convergence, they have been quite successful when applied to simulation optimization.

Evolutionary Algorithms (EAs) are heuristic search methods that implement ideas from the evolution process. As opposed to a single solution used in traditional methods, EAs work on a population of solutions in such a way that poor solutions become extinct, whereas the good solutions evolve to reach for the optimum. Recently, there has been an increasing interest in using EAs in simulation optimization because they require no restrictive assumptions or prior knowledge about the shape of the response surface (B"ack and Schwefel, 1993). Biethahn and Nissen (1994) identify alternative combinations of EAs in simulation optimization and discuss how they differ from traditional optimization methods. In general, an EA or simulation optimization can be described as follows: (i) generate a population of solutions; (ii) evaluate these solutions through a simulation model; (iii) perform selection, apply genetic operators to produce a new offspring (or solution), and insert it into the population; and (iv) repeat until some stopping criterion is reached.

The most popular EAs are Genetic Algorithms (GAs) (Goldberg, 1989), Evolutionary Programming (EP) (Fogel, 1992), and Evolution Strategies (ES) (Schwefel, 1981). These algorithms differ in the representation of individuals, the
design of variation operators, and the selection of their reproduction mechanisms. Back et al. (1997) describe the purpose, structure, and working principles of these three well-known EAs. In general, each point in the solution space is represented by a string of values for the decision variables (i.e., each position in the string represents the decision alternatives regarding a parameter in the system). The use of appropriate crossover and mutation operators reduces the probability of trapping to a local optimum. The crossover operator breaks the strings representing two members of the population and exchanges certain portions of the strings to produce two new strings, where the mutation operator selects a random position in a string and changes the value of that variable with a prespecified probability.

2.5. SIMULATION OF LOGISTICS SYSTEMS AND SUPPLY CHAINS

One of the great strengths of simulation modelling is the ability to model and analyze the dynamical behavior of a system. This makes simulation an ideal tool for analyzing supply chains because supply chains can exhibit very complex dynamical behavior. For example, simulation has been used to demonstrate and study the bullwhip effect (i.e., the amplification of demand variation as demand signals move up the supply chain from the end customer—see Forrester 1958 and Lee et al. 1997) in the MIT Beer Distribution Game.

2.5.1. Strategic analysis

This involves taking a look at the various components involved in the process and selecting the best logistics process among the alternatives. These components, which are to be reviewed, are revealed during the first step. This may include revamping the entire process to assessing how a single component can be used more effectively.

2.5.2. Planning

This involves the assembling of a plan that outlines the mission and goals for the logistics function and the programs and activities to achieve these goals.
Logistics planning is an iterative process. The plans have to be redefined every year to improve the quality of performance.

2.5.3. Managing change

This involves effective management to implement enhanced ways of conducting business. The management should keep changing the plans in accordance with the change in the market and also coach the organization to effectively embrace this change.

A list of processes/activities modelled and represented in a logistics simulation model are given below:

- Order processing at the warehouse (manual, EDI)
- Push order
- Pull order
- Terminal operations at the plants, warehouse, and the customers
- Grouping and palletizing
- Ungrouping
- Transportation mode selection
  - At plant
  - Warehouse
- Handling shortages (or surplus inventories)
- Send an order message to another warehouse
- Movement of parts
- Movements of finished products
- Customer orders
- Customer locations
- Direct shipments

2.6. TRANSPORT TERMINALS

Transport terminals are key elements in the supply chains of industrial systems. Ideally, a terminal must be planned so as to ensure an acceptable level of service in terms of waiting time for trains and customers. However, the increase of transport demand in a network through time, and other unpredictable problems (such as delay in train or airplane arrival, variability in train sizes, breakdowns, etc.), can reduce the level of service. There are basically two ways to face up to...
this situation: to improve operational methods, and/or invest in new facilities. The second solution is usually much more expensive, so the analysis should begin by exploring operational methods. If the desired performance is not achieved, then the investment in new facilities should be considered. Literature related to three important types of terminal systems are presented here.

a) Studies related to Container terminal operations
b) Studies related to Rail terminal operations
c) Studies related to Airport terminal operations

2.6.1. Studies Related To Container Terminal Operations

Port Terminals are evolving very fast due to factors such as demands of ships, material handling technology, information systems and automation to provide shorter handling times for more and more cargo. The ports are service facilities that have 24-hour work time per day, 365 days working in a year, all-weather operations. Port facilities involve big investments and require special operative and management skills. Container terminals have become hubs of international trade; container terminals assume an important role in the distribution of goods around the globe. Seagoing vessels are unloaded, and containers stored in the stack area are loaded onto inland water vessels, railway trains and road trucks at container terminals. Literature in this field is related to the areas of container terminal systems, terminal logistics and optimization methods, ship planning processes and storage and stacking logistics. General information about technical equipment for container terminals can be found in engineering oriented journals as well as specialized outlets (see, e.g., http://www.porttechnology.org/).

2.6.1.1. Container terminal systems

Container terminal operations are becoming more and more important and critical in the field of logistics. Therefore, an ever increasing number of publications on container terminals have appeared in the literature. While we refer to many of them in the subsequent chapters, some deserve special mention due to some of their general perspectives. Decision problems at container
terminals are comprehensively described by Vis and de Koster (2003) (with some 55 references up to 2001). An overview of relevant literature for problem classes like arrival of the ship, (un)loading of a ship, transport of containers from/to ship to/from stack, stacking of containers, interterminal transport and complete terminals is provided.

Meersmans and Dekker (2001) present an overview of the use of operations research models and methods in the field of design and operation of container terminals with its decision problems on strategic, tactical and operational level.

Murty et al. (2003) describe various interrelated complex decision problems occurring daily during operations at a container terminal. They work on decision support tools and discuss mathematical models and algorithms.

Murty et al (2005) describe a variety of inter-related decisions made during daily operations at a container terminal. The ultimate goal of these decisions is to minimize the berthing time of vessels, the resources needed for handling the workload, the waiting time of customer trucks, and the congestion on the roads and at the storage blocks and docks inside the terminal; and to make the best use of the storage space. Given the scale and complexity of these decisions, it is essential to use decision support tools to make them. This paper reports on work to develop such a decision support system (DSS). They also discuss the mathematical models and algorithms used in designing the DSS, the reasons for using these approaches, and some experimental results.

Nam and Ha (2001) investigate aspects of adoption of advanced technologies such as intelligent planning systems, operation systems and automated handling systems for container terminals. They set criteria for evaluation of different handling systems and apply them to examples in Korea. Results show that automation does not always guarantee performance (e.g. higher productivity) – it depends on terminal characteristics such as labour costs.

Four different types of automated container terminals were designed, analyzed and evaluated in a simulation model with very detailed cost considerations by Liu et al (2002). The performance criteria that are used in this study to evaluate and
compare different terminal systems are summarized as follows: Throughput: number of moves/hour/quay crane; throughput per acre; ship turnaround time: time it takes for a ship to get loaded/unloaded; truck turnaround time: average time it takes for a truck to enter the gate, get served, and exit the gate, minus the actual processing time at the gate; gate utilization: percent of time the gate is serving the incoming and outgoing container traffic; container dwell time: average time a container spends in the container terminal before taken away from the terminal; idle rate of equipment: percent of time the equipment is idle. The authors conclude that performance and costs of conventional terminals can be improved substantially by automation

2.6.1.2. Terminal logistics and optimization methods
The need for optimization using methods of operations research in container terminal operation has become increasingly important in recent years. This is because the logistics especially of large container terminals has already reached a degree of complexity that further improvements require scientific methods. The impact of concurrent methods of logistics and optimization can no longer be judged by operations experts alone. Objective methods are necessary to support decisions. Different logistic concepts, decision rules and optimization algorithms have to be compared by simulation before they are implemented into real systems.

Many of the problems in container terminal logistics can be closely related to some general classes of transportation and network routing problems (and therefore more or less standard combinatorial optimization problems) discussed comprehensively in the literature. Examples of these problems and some basic references may be given as follows: An early and very comprehensive survey on various types of routing problems is Bodin et al. (1983). For a recent survey on the vehicle routing problem (VRP) see Toth and Vigo (eds) (2002), arc routing problems are also considered in Dror (ed) (2000). The traveling salesman problem (TSP) asks for the shortest closed path or tour through a set of cities that visits every city exactly once. It is well explained in Lawler et al (eds) (1985);
more recent pointers can be found in Gutin and Punnen (eds) (2002). The rural postman problem (RPP), which is the problem of finding a least cost closed path in a graph that includes, at least once, each edge in a specified set of arcs, is considered in container terminal logistics by Steenken et al (1993). In the pickup and delivery problem a set of routes has to be constructed in order to satisfy a given number of transportation requests by a fleet of vehicles. Each vehicle has a certain capacity, an origin and a destination (depot).

The application of combinatorial optimization techniques has had little success in analyzing and increasing the performance of CTs (Hayuth and Fleming 1994). The complexity of the CT often requires complex models (combinatorial and non-linear), so the resulting models are extremely difficult or take too much time for solving problems (Hayuth and Fleming 1994). This motivated us to go for the use of simulation models.

2.6.1.3. The ship planning process

Ship planning consists of three partial processes: the berth planning, the stowage planning and the crane split.

Berth allocation: Before arrival of a ship, a berth has to be allocated to the ship. The schedules of large oversea vessels are known about one year in advance. They are transferred from the shipping lines to the terminal operator by means of EDI. Berth allocation ideally begins before the arrival of the first containers dedicated to this ship – on average two to three weeks before the ship’s arrival.

Berth planning problems may be formulated as different combinatorial optimization problems depending on the specific objectives and restrictions that have to be observed. Lim (1998) reformulates the problem as a restricted form of the two-dimensional packing problem and explores a graph theoretical representation. For this reformulation it is shown that this specific berth planning problem is NP-complete. An effective heuristic algorithm for solving the problem – applied to historical test data – is proposed.
Kim and Moon (2003) formulate a MIP-model for determining berthing times and positions of vessels in container ports with straight-line shaped berths. They develop a simulated annealing (SA) algorithm and show near-optimal results.

Stowage planning: In practice, stowage planning usually is a manual or offline optimization process using respective decision support systems (see, e.g., Shields (1984)). Most of the papers below describe research work applicable to enhance existing systems by appropriate optimization functionality.

Sculli and Hui (1988) investigate distribution effects and the number of different types of containers with respect to an efficient stowage in an experimental study. Performance of stacking policies is measured by volumetric utilization, wasteful handling ratios, shortage ratio, and rejection ratio. Results indicate that the number of different types of containers has the largest impact on these measures. Effects of stacking policy and maximum store dimensions are also significant.

Wilson and Roach (1999, 2000) divide the container stowage process into the two subprocesses and related subproblems of strategic and tactical planning level due to complexity of a stowage plan across a number of ports. They use branch and bound algorithms for solving the first problem of assigning generalized containers to a block in a vessel. In the second step a detailed plan which assigns specific positions or locations in a block to specific containers can be found by a tabu search algorithm. Good results (not always optimal) can be found in reasonable time. The same principles are described by Wilson and Roach (2001). They present a computer system for generating solutions for the decomposed stowage (pre-)planning problem illustrated in a case study. The authors present a GA approach in order to generate strategic stowage plans automatically. Initial computational experiments show effective sub-optimal solutions.

Simulation and online optimization in stowage planning is considered in Winter (2000), Winter and Zimmermann (1999). Especially in online settings as they are encountered in practice, waiting times of the cranes as well as congestions of
transport means below the cranes have to be minimized to avoid productivity reduction. Winter (2000) presents an integrated just-in-time scheduling model and algorithms for combined stowage and transport planning.

**Crane split:** Crane split allocates a respective number of cranes to a ship and its sections (bays) on hold and deck and decides on which schedule the bays have to be operated. Daganzo (1989) shows a MIP for a static crane allocation problem with no additional ships arriving during the planning horizon. It is exactly solved for small problem instances (i.e. small number of ships), and a heuristic procedure for larger problems is proposed. In addition, the dynamic problem is considered. In both models the berth length is assumed to be unlimited.

Gambardella et al. (2001) present a solution for the hierarchical problems of resource allocation — namely the allocation of quay cranes for (un)loading vessels and yard cranes for stack operations — and scheduling of equipment (i.e. (un)loading lists for each crane). Simulation results show reduction of equipment conflicts and of waiting times for truck queues.

Bish (2003) develops a heuristic method for minimizing the maximum turnaround time of a set of ships in the so called ‘multiple-crane-constrained vehicle scheduling and location problem (MVSL)’. The problem is threefold: determination of a storage location in the yard for unloaded containers, dispatching vehicles to containers and scheduling of (un)loading operations to cranes.

Park and Kim (2003) discuss an integer programming model for scheduling berth and quay cranes and propose a two-phase solution procedure. A first near-optimal solution for finding a berth place and time for each vessel and assigning the number of cranes is refined by a detailed schedule for each quay crane.

**2.6.1.4. Storage and stacking logistics**

Stacking logistics has become a field of increasing importance because more and more containers have to be stored in ports as container traffic grows continuously and space is becoming a scarce resource.
Cao and Uebe (1995) propose a tabu search based algorithm for solving the transportation problem with nonlinear side constraints – a general form of the problem of assignment of storage positions for containers with minimized searching and/or loading costs and satisfaction of limited space and other constraints.

Kim (1997) investigates various stack configurations and their influence on expected number of rehandles in a scenario of loading import containers onto outside trucks with a single transfer crane. For easy estimation regression equations are proposed.

Kim et al. (2000) formulate a dynamic programming model for determination of the storage location of export containers in order to minimize the number of reshuffles expected for loading movements. The configuration of the container stack, the weight distribution of containers in the yard, and the weight of an arriving container are considered. For real-time decisions a fast decision tree is derived from the set of optimal solutions provided by dynamic programming.

A GA-based approach for minimizing the turnaround time of container vessels is described by Preston and Kozan (2001). The problem is formulated as an NP-hard MIP-model for determining the optimal storage strategy for various schedules of container handling (random, first-come-first-served, last-come-first-served). Computational experiments show that the type of schedule has no effect on transfer time if a good storage layout is used. Changes of storage area utilization in the range of 10–50% result in linear changes of transfer time.

Zhang et al. (2003) study the storage space allocation problem in a complex terminal yard (with inbound, outbound and transit containers mixed). In each planning period of a rolling-horizon approach the problem is decomposed into two levels and mathematical models. The workload among blocks is balanced at the first level. The total number of containers associated with each vessel and allocated to each block is a result of the second step which minimizes the total distance to transport containers between blocks and vessels. Numerical
experiments show significant reduction of workload imbalances and, therefore, possible bottlenecks.

2.6.1.5. Transport optimization

Quayside transport: Li and Vairaktarakis (2001) address the problem of minimizing the (un)loading time for a vessel at a container terminal with fixed number of internal trucks (not shared among different vessels). An optimal algorithm and some heuristic algorithms are developed for the case of a single quay crane. Effectiveness of the heuristics is shown by analysis and computational experiments. The case with multiple identical quay cranes is not solved, but the complexity is analyzed.

Bish et al. (2001) focus on the NP-hard vehicle-scheduling-location problem of assigning a yard location to each import container and dispatching vehicles to the containers in order to minimize the total time for unloading a vessel. A heuristic algorithm based on an assignment problem formulation is presented. The algorithm's performance is tested in computational experiments.

Meersmans and Wagelmans(2001) consider the problem of integrated scheduling of AGVs, quay cranes and RMGs at automated terminals. They present a branch and bound algorithm and a heuristic beam search algorithm in order to minimize the makespan of the schedule. Near optimal solutions are obtained in a reasonable time. A beam search algorithm and several dispatching rules are compared in a computational study under different scenarios with similar results. The study also indicates 'that it is more important to base a planning on a long horizon with inaccurate data, than to update the planning often in order to take newly available information into account'.

Carrascosa et al. (2001) present multi-agent system architecture to solve the automatic allocation problem in container terminals in order to minimize the ships' docking time. The paper focuses on the management of gantry cranes by a 'transtainer agent'. This work is framed into a project to the integral management of the containers terminal of an actual port. The independence of
subsystems obtained from a multi-agent approach is emphasized. (The approach is also described by the same group of authors in Rebollo et al. (2000).

The landside transport: The landside transport is split into the rail operation, the truck operation and the internal transports. A common means of operation is to allocate a given number of vehicles to each sphere of operation appropriate to the workload expected. A more advanced strategy is to pool the vehicles for all these working areas.

The problem of assigning jobs to straddle carriers is solved with linear assignment procedures combining movements for export and import containers. Steenken et al. (1993) deal with the optimization for the rail operation and internal moves. Different algorithmic approaches are used to solve the routing problems, as they can be found in machine scheduling, for solving the travelling salesman problem, the rural postman problem, etc. Both solutions were implemented in a real time environment and resulted in considerable gains of productivity. Results and architecture of implementation are presented in Steenken D (2003). Kim et al. (2003) discuss approaches and decision rules for sequencing pickup and delivery operations for yard cranes and outside trucks, respectively.

Crane transport optimization: Another field of application of optimization methods are the transports of gantry cranes operating in stacks. Kim and Kim (1997) present a routing algorithm for a single gantry crane loading export containers out of the stack onto waiting vehicles. The objective is to minimize the crane's total transfer time including set-up and travel times. The model's solution determines the sequence of bay visits for pick-up operations and the number of containers to be picked up at each bay simultaneously. The developed algorithm is named 'efficient' and shows solutions to problems of practical size 'within seconds'. In a more detailed paper (Kim and Kim, 1999) the same algorithm is used for solving the MIP of a 'practical problem of a moderate size'. The load sequence of individual containers within a specific bay remains undetermined.
Zhang et al. (2002) describe the dynamic RTG deployment problem with forecasted workload per block per planning period (4 hours).

### 2.6.1.6. Simulation systems

In recent years, simulation has become an important tool to improve terminal operation and performance. Three types of simulation can be distinguished: strategical, operational and tactical simulation. Strategical simulation is applied to study and compare different types of terminal layout and handling equipment in respect to efficiency and costs expected. It is mainly used if new terminals are planned or the layout or the equipment of existing terminals has to be altered. Strategical simulation systems allow for easy design of different terminal layouts and employment of different types of handling equipment. The chief goal of strategical simulation is to decide on terminal layout and handling equipment which promises high performance and low costs. To match reality, simulation systems allow to design realistic scenarios or to import data of existing terminals.

Operational simulation is applied to test different kinds of terminal logistics and optimization methods. It has achieved growing acceptance at least at large terminals. Terminal operation and logistics at large terminals are already very complex and the effect of alternative logistics or optimization methods has to be tested with objective methods. Therefore, optimization methods are tested in a simulation environment before they are implemented in real terminal control and steering systems. Tactical simulation means integration of simulation systems into the terminal’s operation system. Variants of operation shall be simulated parallel to the operation and advices for handling alternatives shall be given especially if disturbances occur in real operation. Real data of operation then have to be imported and analyzed synchronously to the operation. Because of this ambitious requirement, tactical simulation is seldom or only partially installed at container terminals.

The simulation of harbour processes is normally based on stochastic discrete-event models, however combined simulation for some specific application is a growing sector (Nevins et al. 1998). Veenstra and Lang (2004) describes a
conceptual approach and presents a first study for analysing the economic performance of a container terminal design, using operational indicators. The study consists of the extension of an operational simulation model into a model allowing economic evaluation of the terminal in terms of cash flow generated. The concepts and preliminary results of the study are presented. The paper argues that the integration of the economic evaluation into the simulation model might give rise to problems with aggregation, but will also lead to the development of a potentially very interesting tool that can be used to assess advanced operational and financial strategies, such as dynamic pricing.

Gambardella (1996) present the first results in the development of a methodology to integrate simulation, forecasting and planning to support day by day and long term decisions for operators working in intermodal container terminals.

Gambardella et al (1998) discusses decision support system for the management of an intermodal container terminal. Among the problems to be solved, there are the spatial allocation of containers on the terminal yard, the allocation of resources and the scheduling of operations in order to maximise a performance function based on some economic indicators. These problems are solved using techniques from optimisation, like job-shop scheduling, genetic algorithms or mixed-integer linear programming. At the terminal, the same problems are usually solved by the terminal manager, only using his/her experience. The manager can trust computer generated solutions only by validating them by means of a simulation model of the terminal. Thus, the simulation tool also becomes a means to introduce new approaches into traditional settings.

Kulick and Sawyer (2000) mentions that simulation modelling that has been successfully used to analyze intermodal capacity issues for a wide variety of facilities. Simulation technology provides an analysis mechanism for large intermodal facilities that are difficult to duplicate with other methods due to the interaction of many variables.

Simulation of logistics processes at the Baltic Container Terminal (BCT) was performed using the Arena simulation tool (Merkuryev et al., 2000). The model
considers the terminal layout (with its two berths, container yards, roads, railway centre and In/Out gate), elements of the outside transport flows (ships, trains and trucks), internal transport (trailers, forklifts, quay and yard cranes), and information about simulation results as well: berth productivity and number of containers on a ship. The model allows productivity evaluation for the terminal equipment as well.

In designing container terminals one have to consider the choice for a certain type of storage and retrieval equipment by performing a feasibility and economic analysis. Vis(2006) compare, by means of a simulation study, the performance of manned straddle carriers and automated stacking cranes. As main performance measure, the total travel time required to handle storage and retrieval requests from both the sea- and landside of the terminal is used. It is concluded that automated stacking cranes outperform straddle carriers in a stack with a span width smaller than nine containers. From that point on straddle carriers reach a comparable performance.

2.6.2. Studies Related To Rail Terminal Operations
Many aspects of Railway operations have been studied by a number of authors. These include terminal operations, rail network optimization, freight movement, scheduling passenger trains and freight trains etc. There are many studies which consider intermodal terminals. Models in general include simulation as well as analytical models.

2.6.2.1. Simulating Rail Terminals
Klima and Kavicka (1996), used simulation to model marshalling yards in railway networks. The costly technology and high complexity of the operations performed require a great degree of coordination and control. Because of the intricacy of the system, the only suitable tool for evaluating conditions in this system is believed to be a simulation model. One of the features of the Klima and Kavicka model is the ability of the user to plan some standard activities such as interruption, termination, snapshots of the system state, etc., prior to initiation of the simulation run. Dessouky and Leachman (1995) present a detailed computer
simulation modelling methodology that can be used to analyze the increased traffic burden on rail track networks and delays to trains caused by congestion.

2.6.2.2. Intermodal Railroad Terminal Simulation

Intermodal terminals are critical components in the total intermodal freight transportation process, and their efficiency must be optimized if they are to remain competitive. In Ferreira and Sigut (1995), two different types of terminals are simulated: the conventional road/rail container transfer facility, and a proposed system named the RoadRailer terminal facility. Boese (1983) notes that the future demands that are to be placed on intermodal transportation systems will require substantial investments in existing and new terminal facilities. In order to optimize the operations of these terminals, computer modelling of these sites is imperative. The model developed by Boese has several program modules simulating different functions of the terminal in question. The simulation of the daily train operations reflects given cargo volume fluxes, types of load units, train schedules, selected rail operational strategies, and equipment capacities. The road counterpart utilizes a Monte Carlo simulation of the stochastic properties of truck arrivals at the terminal, according to different truck operating patterns. The core module simulates the single movements and actions of the transshipment equipment. A dispatch control module decides on the transshipment sequences prescribed by train operation and truck arrivals, while simultaneously trying to maximize equipment productivity and minimize truck waiting times. The presented simulation provides some information concerning terminal economies, operational strategies, and control systems. A trailer-on-flatcar (TOFC) terminal simulation model (TSM) is discussed in a paper by Golden and Wood (1983). This model provides information about productivity and throughput of trains and trailers at an intermodal facility using a detailed simulation.

Sarosky and Wilcox (1994) utilize a SLAMSYSTEM model to examine the feasibility of eliminating a terminal from Conrail’s intermodal network and shifting the remaining traffic volume to an alternate facility. Described in the paper is the
problem of optimal terminal size in the construction and operation of an intermodal terminal.

2.6.2.3. Simulating Truckload Trucking Networks
Research has been undertaken to examine the effects of hub and spoke (H&S) networks, similar to those utilized in less-than-truckload (LTL) and airline settings. See Taha et al. (1996), or Taha and Taylor (1994) for information about this problem, and for information about the HUBNET simulation tool developed for and employed in this analysis.

2.6.2.4. Terminal Operations and Capacity
Ferreira (1997) discuss the research and development of optimization and simulation tools in the operations planning of an Australian freight rail system. The author claims that the market share for rail freight is greatly determined by the level of service, especially in terms of transit times and the reliability of arrivals. These, in turn, are largely associated with track infrastructure design and maintenance schedules. Summarized in the paper are requirements for planning track maintenance and a description of a model to optimize the placement of sidings along a single-track corridor.

2.6.2.5. Non-Simulation Methods
Substantial literature discussing work-using techniques other than simulation to examine rail yards also exists. For example, Feo and Gonzalez-Velarde (1995) use a mathematical model to optimally assign highway trailers to rail car hitches in intermodal transportation terminals. An integer linear programming formulation that allows problems to be effectively solved by use of general-purpose branch-and-bound code is constructed.

2.6.2.6. Scheduling issues
On busy congested rail networks, random delays of trains are prevalent, and these delays have knock-on effects which result in a significant or substantial proportion of scheduled services being delayed or rescheduled. Carey and Carville (2000) develop and experiment with a simulation model to predict the
probability distributions of these knock-on delays at stations, when faced with typical patterns of on-the-day exogenous delays. These methods can be used to test and compare the reliability of proposed schedules, or schedule changes, before adopting them. They can also be used to explore how schedule reliability may be affected by proposed changes in operating policies, for example, changes in minimum headways or dwell times, or changes in the infrastructure such as, layout of lines, platforms or signals. This model generates a reliability analysis for each train type, line and platform. They also use the model to explore some policy issues, and to show how punctuality and reliability are affected by changes in the distributions of exogenous delays.

In scheduled (timetabled) transport systems (for busses, trains, etc.) it is desirable at the planning stage to know what effect proposed or planned changes in the schedule may have on expected costs, expected lateness, and other measures of cost or reliability. Carey and Kwiecifiski (1995) consider such effects here, taking account of the random deviations of actual times (or arrivals, departures, etc.) from the corresponding scheduled times. They also take account of various forms of interdependence (knock-on effects) between the timings (arrivals, departures, connections, lateness, etc.) of different transport units and formulate a stochastic model of such a complex transport system. (For generality, the underlying deterministic version of the model is consistent with versions of various existing deterministic transport models).

2.6.3. Studies Related To Airport Terminal Operations

The modelling of airport terminal operations has advanced significantly over the last 15 years (Tosic, 1992). Available models have improved in detail and fidelity, as well as “user friendliness”. As a result, their use as decision support aids or design tools in terminal development projects has been steadily increasing. Some existing models are “strategic” in nature sacrificing level of detail in exchange for speed and flexibility, while others are primarily “tactical” incorporating high levels of detail in data and system definition. Mumayiz (1990, 1997) and Tosic (1992) have presented exhaustive overviews on the
development of terminal simulation technology and on their applications to airport terminals.

Jim and Chang (1998) mentions that recent airport capacity studies have indicated that there is an imbalance in passenger terminal, airfield and airspace planning at many major airports. Traditionally, the emphasis has been on airfield and airspace development and analysis. Not much emphasis has been made on passenger terminal design. Therefore, there are many cases around the world exhibiting congestion problems at the airport passenger terminal as the number of air passengers continue to increase.

Following the analysis presented in (Transportation Research Board, 1987), landside elements may be subdivided into three classes:

- **Processing facilities**: they process passengers and their luggage.
- **Holding facilities**: areas in which passengers wait for some events (as the check-in opening for a flight, the start of flight boarding, etc).
- **Flow facilities**: the passengers use them to move among the landside elements.

2.6.3.1. *Ticket counter and baggage check-in*

Capacity of check-in processing facilities is judged by considering the average service time and by comparing the number of passengers in a terminal holding area with the size of that area. Some of the analytical models proposed in the literature for check-in counters belong in the class of Queuing Theory Models. This is also the case for most of the other processing facilities in airport operations. Lee proposed a pioneering application of M/M/n queuing systems to check-in procedures (Lee, 1966).

Newell initially proposed a deterministic approach (Newell, 1971). This model had a strong influence on further developments in this area, and it has applications in modelling several types of facilities where service is provided to individuals by a “processor” of some kind. Basically, this is a graphical model that computes approximately the total waiting time of passengers, given the
cumulative arrival function at the check-in counter and the service rate for each time period. This simple and effective model has also been extended (to representing more than one flight) in (Tosic et al., 1983). A simulation model based on the Monte Carlo method has been presented in (Tosic et al., 1983). Being a simulation model it needs detailed data, and provides quite realistic information on the behaviour of check-in counters.

2.6.3.2. Passenger security screening

Originating passengers must undergo a security screening operation. Sometimes transfer passengers also have to pass through security screening while moving to a connecting flight. For this reason, security-screening areas are often elements of queuing and delay for passengers. Both stochastic and deterministic queuing models have been proposed in the literature. Examples of application of the stochastic models are in (Rallis, 1958, 1963, 1967). In particular, the Copenhagen terminal building was analyzed by applying M/D/n queuing systems. Newell proposed a deterministic model by means of graphical analysis using cumulative diagrams of number of passengers versus aircraft departure time (Newell, 1971).

2.6.3.3. Gates

A lot of models for gate assignment have been proposed. Some of them take into account both the type of aircraft and the passenger walking distances. Basically, they are based on a gate assignment with first in - first out (FIFO) rule (Le et al., 1978) and (Hamzawi, 1986). Babic et al. proposed a method to minimize passenger walking distances by properly assigning aircraft to gates every day, taking into account passenger flows on that particular day (Babic et al., 1984). Mangoubi and Mathaisel incorporated transfer passengers in their formulation of the flight-to-gate assignment problem (Mangoubi and Mathaisel, 1985). Both approaches assume that a specific configuration is given so that walking distances are known and fixed, and, therefore, these models are appropriate at the tactical level. Wirasinghe and Vandebona proposed a long-term planning model (Wirasinghe and Vandebona, 1987). As for gate position requirements,
Bandara and Wirasinghe proposed a way for determining the gate position requirements based on a deterministic model (Bandara and Wirasinghe, 1989). Edwards and Newell investigated stochastic models of gate utilization (Edwards and Newell, 1969). Steuart proposed a different stochastic model (Steuart, 1974). Yan et al. (2002) proposes a simulation framework, that is not only able to analyze the effects of stochastic flight delays on static gate assignments, but can also evaluate flexible buffer times and real-time gate assignment rules. A simulation based on Chiang Kai-Shek airport operations is performed to evaluate the simulation framework.

2.6.3.4. Baggage claim facilities

Baggage claim is the most critical step of the inbound baggage system. The number of passengers waiting in the baggage claim depends on the rates at which passengers arrive from the gate and luggage is processed. In general, the maximum demand levels occur when larger aircraft arrive. The baggage claim area capacity can be measured considering the average time passengers must wait to retrieve their checked baggage and comparing the number of people in the claim area with the size of that area. The number of passengers claiming baggage must be calculated from schedule forecasts. In general, the linear dimension of the device is determined on the basis of the number of passengers, rather than of baggage, except in some cases in which baggage ratio is very high. The expected average time passengers have to wait for bags and the number of waiting passengers in the claim area can be computed by simple queuing models.

In the literature, mathematical queuing and simulation models have been developed to predict the arrival (of deplaning passengers and baggage) to baggage claim areas, and to forecast possible future conditions. In (Horonjeff, 1969) and (Barbo, 1967) a deterministic queuing model was developed to relate the arrival distributions of passengers (and the arrival distributions of baggage) to the number of passenger bags that are on the carousel at a given time. Browne et al. studied the baggage claim areas of the JFK airport in New York (Browne et
al., 1970). Their objective was to compute the expected maximum inventories of passengers and bags using inventory type models. Newell analyzed a baggage claim device and proposed a two queues system, one for passengers waiting for bags, the other for bags waiting for their owner (Newell, 1971). The problem was to estimate the number of passengers waiting in front of the devices for their bags. Tosic et al. proposed a Monte Carlo type simulation model to evaluate the elements of the baggage claim area (Tosic et al., 1983). In this model each passenger and all his/her bags are treated individually.

2.6.3.5. Passenger holding areas

Passenger holding areas are spaces where passengers move around and wait for flight departures and arrivals. These facilities include lobbies, gate lounges, transit passenger lounges, baggage claim area, the arrival area, the area set aside for ancillary facilities, etc.

The number of waiting passengers is a function of the number of aircraft served by the holding area, and their functional characteristics, including capacity and loading factors. The number of passengers simultaneously waiting in the terminal is also influenced by other important factors, such as passenger dwell time. The amount of time spent in a particular area, that is a fraction of passenger dwell time, is central to determine the number of simultaneous occupants of a given area (Odoni and de Neufville, 1992).

Dwell time is mainly caused by the amount of “slack” time that passengers spend in the various parts of the terminal building. This slack time is in turn allocated among the terminal holding areas. Clearly the loading, that is the number of simultaneous occupants, depends on the fraction of the slack time spent in that area. This discussion applies both to departing and transit; for arriving passengers the concept of slack time is less important because they try to leave the airport as soon as possible. Stochastic models for estimating dwell time are presented in (Odoni and de Neufville, 1992).

Ballis et al (2002) presents a simulation model that enables the investigation of charter passenger effects on air terminal facilities and enables the estimation of
the level of service offered. Some of the model's features can be easily implemented by use of spreadsheets. The paper concludes with a critical assessment of the results arisen in the master plan of two Greek airports where the simulation model was implemented.

2.6.3.6. Flow facilities
The total time spent by a passenger to cross the terminal building from its entrance point to the gate is the sum of the waiting and service times in the processing facilities plus the sum of the times required to move from a service station to another.

Large airport terminals with multiple gate positions necessarily involve large internal transfer distances. Mechanized circulation aids are commonly used to improve circulation in large terminal buildings. In airports with multiple terminal designs (e.g., Paris Charles de Gaulle), and remote satellites (i.e., London Gatwick), the distances can be so large that mechanized movement becomes essential. The terminal circulation component may be seen as a flow pedestrian problem and analyzed by using procedures and standards such as those suggested in (Transportation Research Board, 1987). The time required to travel from the curb to the gate is the most important measure of service level.

The arriving passenger flow is typically defined as a queuing network system with a series of processors, including gates, concourse, immigration checks, baggage claim systems, customs declaration, secondary examination, and lobby (FAA, 1988). Hence, the method needs to be capable of modelling tandem queues with multiple servers, probabilistic arrivals and services, pooled and separate queues, as well as various aircraft mixes. With the capability of handling various aircraft mixes, this model can also be applied in determining the impacts on domestic terminal operations for any larger aircraft.

Lozano et al. (2004) have developed a package that can simulate in detail passengers' traffic within the departures terminal of Málaga airport. The package performs a passenger-by-passenger "accelerated-time simulation" that considers at each step details like the class, flight and destination of each passenger. Once
the simulation has been performed, it can show, plot or give details about any queue in the terminal at any minute. Moreover, given a list of flights, it can also produce an enlarged list of flights with the same spectrum (models of planes, schedules, destinations). It has been implemented in the Computer Algebra System Maple 8.

Ray and Claramunt(2003) introduces a novel distributed computing environment designed as a simulation tool for the analysis of large and disaggregated data flows. The potential of the software is illustrated by a case study that simulates large people flows for different hall configurations of an airport terminal.

2.6.3.7. Capacity Estimation Models

A distinction between analytical and simulation models may be made based on the methodology used to compute capacity, delay or other such metrics(Bazargan et al., 2002). Analytical models are primarily mathematical representations of airport and airspace characteristics and operations and seek to provide estimates of capacity by manipulation of the representation formulated. These models tend to have a low level of detail and are mainly used for policy analysis, strategy development and cost-benefit evaluation (Odoni et al., 1997).

Monte-Carlo simulations have been used extensively to study the airport environment. This is a common simulation tool for sampling from cumulative distributions using random numbers until a steady state evolves. Given known or reasonable distributions, as the number of simulations increase, the results match the distributions and predict the likely outcome. This tool was used by Pitfield et al. (1998) to analyze potentially conflicting ground movements at a new airport proposed in Seoul, Korea. Pitfield and Jerrard (1999) uses Monte-Carlo simulations to estimate the unconstrained airport capacity – taking only safety requirements into consideration, and assuming all other factors such as air traffic management and control procedures and best pilot practices as “ideal” - at the Rome Fiumucino International Airport.
2.7. CONCLUSION

The above literature review helped us to understand the methodology for use of simulation for solving problems related to logistic terminals. It also helped us to understand the terminal systems, its structure, characteristics and problems. This was very useful when we undertook the study of real terminal systems for our work presented in next few chapters.

From the review of literature above it can be seen that mathematical models, heuristic models as well as simulation models are popularly used. Closer look shows that for small subsystems, which are, clearly defined mathematical models are more likely to be used. In case of larger systems or for ease of solving the problem, good heuristics are also popular. When complete systems have to be modeled simulation is found to be the best suited technique. We have therefore decided to use discrete-event simulation to build models of a few terminal systems and use them for problem solving. There is a need for development of common framework for simulation modeling of logistic terminals so as to be able to use minimum models to solve maximum variety of problems. The work on this is presented in the subsequent chapters.