# SIMULATION FOR POLICY EVALUATION, PLANNING AND DECISION SUPPORT IN AN INTERMODAL CONTAINER TERMINAL

Monaldo Mastrolilli, Nicoletta Fornara, Luca Maria Gambardella, Andrea E. Rizzoli and Marco Zaffalon
IDSIA - Istituto Dalle Molle di Studi sull'Intelligenza Artificiale
Corso Elvezia 36, 6900 Lugano Switzerland
E-mail: {monaldo,nicky,luca,andrea,zaffalon}@idsia.ch

## ABSTRACT

Different uses of a simulation tool in an intermodal container terminal are presented. Initially, the simulation model of the container terminal of La Spezia (LSCT) in Italy is described. Then it is shown its calibration and validation. The resulting model is used as a tool to validate alternative management policies, such as resource allocation and ship loading and unloading policies. The terminal management can use such policies for mid-term planning. Finally, the design of a decision support system to assist the management in real time decision making is discussed.

## INTRODUCTION

The advent of management information services and data processing greatly improved the ability of terminal managers to control the whole process, but still raw data has to be analyzed and treated to provide some insight on the performance of terminal operations. Simulation models and Operations Research techniques have proven to be a reliable and convenient tool to support the decision-makers in the daily operations in many cases (Hayuth et al. 1994; Blümel 1997; Bruzzone and Signorile 1997).
A simulation model of a terminal can provide a valuable tool for the management, especially to evaluate the performance of new policies (policy evaluation), to assess the effect of the implementation of these policies on the terminal state (planning), and to take operational decisions (real time decision support).
In this paper we explore these different uses for the simulation model, in particular with respect to resource allocation (RA) and loading and unloading (L/U) policies.
The first section is devoted to the description of the model structure; the following introduces its calibration and validation. In the third section it is discussed the integration of the RA module and the L/U scheduling module with the simulator, for the purpose of evaluating the computer-generated policies. The latter, computed on the basis of Operations Research techniques (Gambardella et al. 1996; Gambardella et al. 1998; Zaffalon and Gambardella 1998; Mastrolilli and Gambardella 1998) are embedded in the simulation model, which has been calibrated in the previous step. Their performance can be compared (starting from the same initial conditions and under the same input regime) with the performance obtained by the terminal management (historical data are used to feed the simulator).
The last section discusses the use of the simulation model as a mechanism to evaluate medium and long term planning decisions such as the space allocation policies. Furthermore it discusses the use of the simulation tool as a decision support tool if real-time data are available.

## THE SIMULATION MODULE

The architecture of the simulation tool is based on the partition of simulation objects between simulation agents and simulation components. In an intermodal terminal the two flows of information and of containers are present; the simulation agents use the flow of information to make decisions on how to direct the container flow.
The design of the present simulation tool is based on the object-oriented analysis and design paradigm (Booch 1994). Simulation agents and components are modeled as objects which store and exchange information on terminal inputs, states and outputs and which perform actions. The whole simulation is the result of the interaction of such single agents, each one endowed with *local* knowledge on its actions in response to other agents behavior (Zeigler 1984; Zeigler 1990) and to external events (like, trucks and trains arriving at terminal gate; ships arriving at terminal pier, etc.).

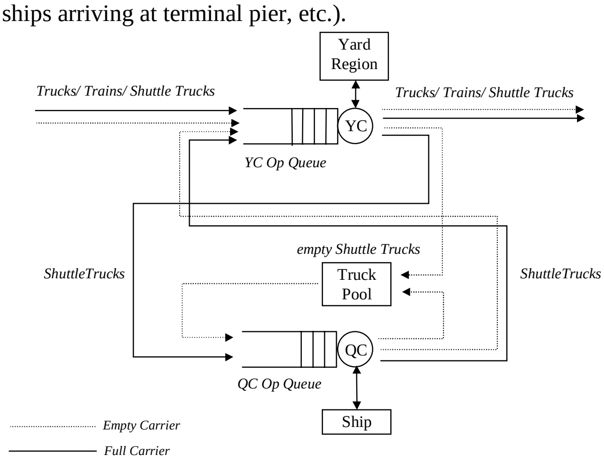


Figure 1. A representation of the flow of containers.

Figure 1 reports an example of the container flow in the terminal limited to a quay crane (QC for short; it is a crane used for loading/unloading a ship) and a yard crane (YC, that is a crane that moves containers for a given yard area). In the real terminal there are seven quay cranes, ten yard cranes and ten straddle carriers (the latter being alternative means, as compared to yard cranes, to move containers in the yard). Ships, trains and trucks entering the terminal have a loading list (the containers to import) and an unloading list (the containers to export). These lists are used by the yard and ship planners. The ship planner organizes the L/U operations of a ship. This requires a complex set of tasks to be considered, like allocating the work shifts for the QCs; computing the L/U policies in order to respect a ship's stability constraints; cooperating with the yard planner to assign

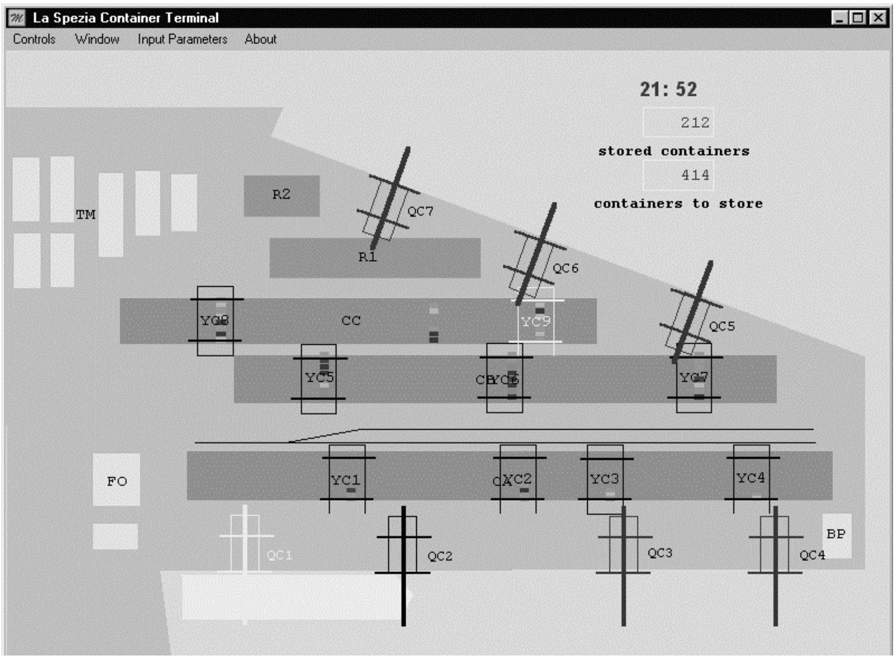


Figure 2. A snapshot of the simulation tool interface.

yard destinations to the containers to be unloaded and to require the containers to be loaded; put the quay cranes to work; supervise loading and unloading operations; collect statistics and evaluate performance.
On the other hand, the yard planner uses the lists of import and export containers (produced by the interaction with the ship planner) to build the schedule solving the job-shop problem associated with yard crane operations. This also implies a proper organization of container allocations on the yard in order to maximize yard crane performance, avoid crane deadlocks and minimize the time to access containers during storage and retrieval.
Besides this (high level) management performed by the ship and the yard planners, there are the local management decisions taken by "less intelligent" simulation agents such as cranes and shuttle trucks (every quay cranes has 3 shuttle trucks that bring containers from the QC to the yard and vice versa) that acts on the basis of a queue of operations to be performed (where an operation is a container movement). Figure 2 reports a typical screen-shot of the terminal during a simulation. A ship is moored on the west pier (north is to the left of the picture) and it is being unloaded by two quay cranes QC1 and QC2 (only QC1 is active, though). Containers are to be positioned on yard areas CA, CB, and CC. On these yard areas the yard cranes YC from 1 to 9 are working.

## CALIBRATION AND VALIDATION

Once the simulation model of the terminal is available, it must be checked in order to verify its capability of reproducing the real terminal behavior. This is accomplished in two steps, namely calibration and validation. In general, calibration means tuning the simulation model parameters in order to match as close as possible the simulation outputs with the data measured in the real terminal over a given time interval. A proper validation phase ensures that the result of the calibration is such that the simulator reproduces the reality under different conditions (hence, not only when the data from the calibration set are used). The validation step is carried out on a validation set of data. On such a set, the simulator is expected to act in a way that the real terminal is properly emulated. Of course no tuning is allowed at this time.
Notice that in the calibration and validation phase, the simulation module evolves using the same inputs and policies used by the terminal management over the calibration and validation periods. For this reason, the resource allocation and the L/U policies are the ones adopted by terminal managers during the periods when calibration and validation data were collected.
In the sequel, we give a short example of validation and calibration of the simulator. All the data used in the present paper are taken from the database of the LSCT. The focus is on a period of two weeks, from 5/11/1998 to 5/24/98. The data describe the activity of the La Spezia container terminal in great detail, in such a way that every container movement can be found. In particular, the database reports the resources (yard cranes, quay cranes, and straddle carriers) that move the container, the time of the operation and the origin and destination of the movement, described by three coordinates each (bay, row and tier both on the ship and on the yard). Furthermore, the database tracks the activity of every transport means that enters and leaves the terminal (ships, trucks, trains).
From this wealth of data, some working shifts for calibration and some other for validation were extracted. In particular, it is considered the period that starts at 1 am on the 5/11/1998 and ends after 10 shifts of 6 hours each (the advantage of such a choice is that no ships are at the terminal at the selected starting time, and the system is therefore in a natural *initial empty state*; finding the initial empty state is often one of the major problems for the simulation practitioner).
The L/U process of each quay crane is examined. For this purpose, the number of moved containers by each quay crane in

each shift is monitored. The first four shifts constitute the calibration period and the remaining shifts (from 5 to 10) are used for the validation.
The calibration parameters are,

- for each quay crane, the mean time and its standard deviation ($\mu_{qc}$ and $\sigma_{qc}$) needed to move a container from the stowage to the shuttle truck below the crane and vice versa;
- for each yard crane, the mean time and std. dev. ($\mu_{yc}$ and $\sigma_{yc}$) needed to move a container from a yard position to the shuttle truck waiting beside the crane, and vice versa; a linear function *n* that represents the time required to retrieve a container which is stored below a pile of *n* containers;
- for the straddle carriers, the mean time and std. dev. ($\mu_{sc}$ and $\sigma_{sc}$) required to pick a container from a shuttle truck and store it in the destination yard.

The speed of cranes (both yard and quay ones) in moving from bay to bay are known and assumed deterministic. Also the mean time and std. dev. of travel of a shuttle truck from origin to destination are given. The arrivals of ships, trains and trucks are deterministic since they are read from the LSCT database. Remember that a truck carries at most two containers, while trains bring an average of 40 containers and ships vary from hundreds to thousands of containers. Their identification numbers are read from the database and then simulation evolves, loading and unloading these transport means.
After performing a few experiments it is clear how the simulation system is sensitive to parameter changes; we set $\mu_{sc}$ equal to $\mu_{yc}$ and $\sigma_{sc}$ to $\sigma_{yc}$, thus assuming equivalent performances for straddle and yard cranes. We also set the standard deviations to 20% of the mean values, thus three calibration parameters ($\mu_{yc}$, $\mu_{qc}$ and the time needed to access a container in a pile) are left to tweak.
A set of eleven experiments (simulations) was performed. Each experiment is identified by a different combination of the three calibration parameters.
Table 1 reports the results (only for Q6 and for three experiments, Sim1, Sim2 and Sim3, for lack of space).

Table 1. Calibration of Q6.

| Shift | LSCT | Sim1 | Sim2 | Sim3 |
|---|---|---|---|---|
| 1 | 127 | 137 | 140 | 135 |
| 2 | 168 | 149 | 140 | 107 |
| 3 | 102 | 120 | 128 | 140 |
| 4 | 108 | 99 | 97 | 123 |

The simulation combinations correspond to the values of the parameters reported in table 2.

Table 2. The calibration parameters.

| | $\mu_{qc}$ (min/cont) | $\mu_{yc}$ (min/cont) | $t_a$ (min) |
|---|---|---|---|
| Sim1 | 2 | 1.75 | 0.8 |
| Sim2 | 2.2 | 2 | 1.2 |
| Sim3 | 2 | 1.75 | 1.75 |

Analogous results are produced for the yard cranes.
In order to establish a ranking among the different parameter combinations, we have devised an error measure that is based on the distance between the real number of containers moved by each quay crane and the simulated number, in each shift. The above error is then normalized with respect to the maximum number of containers which can be moved by a crane (in our case, this number is equal to 170 containers per shift; notice that the latter is also the maximum possible absolute error per crane and per shift, hence it is just a normalization constant). In this way we force the model to track the behavior of the quay cranes during each shift. These errors are then averaged over all the quay cranes and over the number of worked shifts, in order to obtain a scalar quantity as performance indicator. The formula adopted is,

$$J = \frac{\sum_{i=1}^{7} \sum_{t=t_i^s}^{\min(k, t_i^e - 1)} \frac{\left| n_{i,t} - \bar{n}_{i,t} \right|}{n_{\max}}}{\sum_{i=1}^{7} \left[ \min\left(k, t_i^e - 1\right) - t_i^s + 1 \right]}$$

where *i* is the QC index (there are seven quay cranes); *t* is the work shift index; $t_i^e$ and $t_i^s$ are respectively the initial and the final work shift for QC *i*; $n_{i,t}$ represents the number of simulated containers moved by quay crane *i* in shift *t*, while $\bar{n}_{i,t}$ is the amount of containers moved by the same crane at the real terminal; $n_{\max}$ is the maximum number of containers which can be moved by one quay crane during one shift (equal to 170); *k* is the final shift of the considered period (either calibration or validation). Notice that the comparison is not based on the last shift worked by a quay crane (this is the reason for the minimum operator in the formula), since it biases the error estimate. In fact, during the last shift, the simulated crane can only move the containers left from the previous working shifts, which are the results of the errors made there.

Table 3. Validation of Q3.

| Shift | LSCT | Sim1 | Sim2 | Sim3 |
|---|---|---|---|---|
| 5 | 0 | 0 | 0 | 0 |
| 6 | 66 | 66 | 66 | 66 |
| 7 | 84 | 110 | 133 | 144 |
| 8 | 96 | 117 | 131 | 120 |
| 9 | 84 | 37 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 |

The best set of parameters is found at the value of $J^*$ = 0.12, corresponding to experiment “Sim1”. The interpretation of such a value is that, on average, the simulator moves a number of containers (per shift and per quay crane) that is 12% far from the real one, as compared to the maximum possible distance (170 containers, i.e. 100% error). The validation of this model returns *J* = 0.18. Table 3 reports, for example, the result of the validation of QC 3.
The same argument can be applied with regard to the whole ships, instead of the single cranes, by extending the above formula in a straightforward way. In that case, it is observed a lower variability, where the relative error is $J^*$ = 0.08 for the calibration set and $J^*$ = 0.11 in the validation case.
The results deserve a few words of comment, since they seem susceptible to improvements. An important remark is that our

main objective is to obtain a model of the terminal to test alternative scheduling policies. For this purpose, the model, as a first approximation, does not handle unexpected and rare events such as crane breakdowns, and inexperienced operators. Unfortunately, such events happen in the real world terminal. On the other hand, it would not be fair to assess the feasibility of a scheduling policy against a worst case scenario, where terminal equipment works below the average performance. At this stage of the research we want to assess the policies where the only bottlenecks can be caused by resource competitions on the yard, that is by the creation of lengthy queues under the yard cranes, leading to quay crane starvation.

Once the scheduling policies are believed to be "safe" with respect to resource utilization (i.e., they do not tend to create these dangerous queues on the yard cranes), the robustness will be investigated on, by introducing these rare and unexpected events in the terminal model. Up to that point, the focus is on a terminal model which is a reasonable compromise between the ideal situation, where all the cranes work at full throttle, and the worst case scenario, where cranes and operators show marked drops in their performances.

## EVALUATION OF POLICIES

When the calibration and validation steps are successfully carried out, the simulator becomes a faithful model of the real terminal. As such, it can be used to evaluate alternative policies, as compared to the real ones adopted at the terminal. In the present case, the focus is on computer-generated policies, obtained by the application of algorithms from the operations research field. In particular, two main problems are addressed, the allocation of resources and the load/unload list generation.

The allocation of resources is the problem of deciding, on the basis of the predicted amounts of containers to be moved for ships, the number and type of resources that must accomplish such operations (Zaffalon and Gambardella 1998). In the LSCT case, yard cranes, straddle carriers and quay cranes (and the related manpower) are the available resources. Generally speaking, at the LSCT the resources are allocated once at the beginning of a week; the allocation plan is a list that reports the resources to be allocated for every work shift (i.e., 6 hours). The resources constitute the main expense of the terminal. For this reason, it is important that the plan allows all the containers to be moved while minimizing expenses (resources).

The problem of creating the L/U list is hierarchically lower than RA (Gambardella et al. 1998). For this reason, it usually has a shorter time horizon and a more detailed view of the terminal. The L/U problem is mainly a scheduling problem. Its input data are the allocated resources, the container types, their amount and disposition both on ships and at the terminal yard. For every quay crane working on a ship at the terminal, the module must output an order list of containers to be loaded and unloaded. If the container is to load, the list indicates the location in the ship where it must be put; in the same way, it reports the final destination on the yard for the unloading containers.

The problems above are formalized and treated with operations research techniques. The RA problem is modeled as a network design (Zaffalon and Gambardella 1998; Gambardella et al. 1998). In other terms, the activity of the terminal in a time period is represented as a network of containers flow. The arc capacities are dimensioned according to the allocated resources. The latter are decision variables and the objective is the minimization of expenses. The complete problem is a mixed integer linear program (MILP) that is approximately solved by means of a commercial MILP solver implementing branch and bound. The L/U list problem is modeled by extending the formulation of a flexible job shop with setup times and limited resources. It is solved with local search, according to an original approach (Mastrolilli and Gambardella 1998). Empirical evaluations show that real instances of both the problems are well solved in a restricted amount of time (about 1 minute on a 133 Mhz pentium). This allows an effective interaction with the simulator to be realized. In this context, the simulator acts by requiring the resources to the RA module and the L/U policies to the L/U list generator module when needed. Then, it makes the terminal state evolve according to the information above and to external events (ship arrivals, train/trucks arrivals; the latter are generated and distributed according to one of the patterns studied by a forecasting module). The RA policy and the L/U policy are then successful if the simulation is able to carry out all the jobs for the period under study. Also the robustness of the policies can be checked; this is obtained by means of repeated applications of the simulation above (in fact, the replications are normally different since the simulation is non-deterministic).

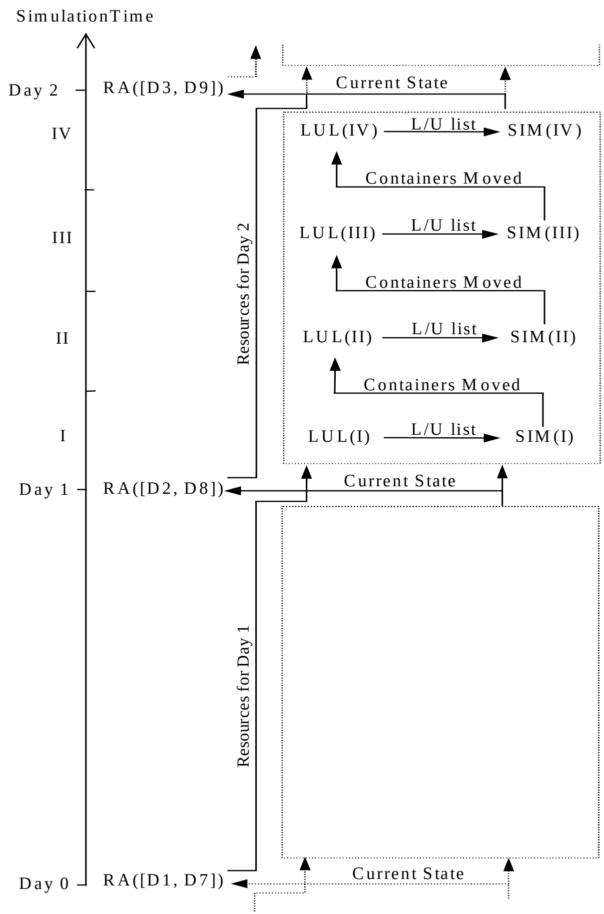


Figure 3. Architecture of the simulation-optimization phase.

Figure 3 reports the interaction of the three modules in our experiments: the RA module, the L/U list generator module

(LUL) and the simulator (SIM). The vertical axis represents simulation time. At the beginning of day 0, the RA module is ran, having as input the predicted state of ships in the period from day 1 to day 7. On its basis, the module is able to compute the resources to use in the same period (notice that RA must compute the resources to allocate with one day in advance, because this is the minimum span to book resources). Notwithstanding, only the resources computed for the first day (day 1) are used. The reason is that the reliability of the prediction used by RA quickly decreases with time. Therefore we chose to let the RA module have a time window of one week, while its output is used for the first day only.

The output of the RA module (at day 0) is passed as input to the box representing the interaction between the LUL and the SIM modules in a time period of one day (notice that the day is split into four shifts of six hours each). They use the allocated resources and the current state of the terminal (in terms of ship and yard states). In particular, at the first shift of day 1, the LUL module produces the L/U list for the ships currently at the terminal. Then the simulation tool emulates the terminal for the current shift, having as input the state of terminal, the allocated resources and the L/U list. At the end of the shift, the containers that have really been moved (by the simulator) are passed back to the LUL module in order to let it update its knowledge of the terminal state. On this basis, the LUL produces the list for the second shift, and so on up to the last shift of day 1. At the end of day 1 the current state of the terminal is passed to the RA module, which, also on the basis of the prediction for days 2 to 9, produces the allocations for day 3. The process is then repeated until the end of the simulation-optimization period.

To date the system is integrated according to the scheme above and the whole mechanism is under test. Empirical evidence from the current trials supports the confidence that the system is capable of producing significant policies for improving terminal management.

## DECISION SUPPORT

The ultimate goal of a simulation model is to be employed as a tool to take management decisions in a decision support framework. This is therefore the goal of the present research cooperation with the LSCT management. Model validation is specifically directed towards building a confidence in the model results. Once the terminal managers have acquired a feeling for the model capabilities and its limitations, they will be able to use it for different purposes like, evaluation of space allocation strategies (i.e., how to layout containers on the yard, with respect to different ship arrival scenarios); evaluation of resource allocation policies; evaluation of L/U policies.

The simulation model in a DSS framework is therefore acting as a test-bench for the policies that must be fed with real-time data. Data are defined available in real-time when it is a good representation of the real data with respect to the time frame needed to take a decision. In the LSCT case, a sampling interval of 24 hours may be sufficient to take reasonable decisions for resource allocation. It is possible that a considerably shorter sampling period might be needed for loading and unloading operations, while a time-frame of about one week is expected for space allocation policies.

Note that the simulation model is not strictly needed to produce the policies; with reference to figure 3, the real time data are fed into the policies at each decision step (such as Day 0, Day 1, ...) and a result is then available. The role of the simulation model is to produce a forecast of the effect of the suggested management action. Research and work on the subject of decision support is still on going and it will require an even greater interaction with the terminal management.

## CONCLUSIONS

This paper presents the part of our research work aimed at exploiting the representative power of the simulation model of the terminal. For this purpose, the problem of calibrating and validating the simulator is discussed and a simple example is presented. The applications of the validated simulation tool are then shown to be valuable in many different settings. The focus is especially on the evaluation of policies generated by optimization algorithms. Here the simulator acts as a test bench for assessing the effectiveness and the robustness of the policies in such a non-deterministic environment.

The combined system (i.e., optimization and simulation), once tested, can be the main component of a decision support system for terminal management.

## REFERENCES


Blümel, E. (ed.). 1997. *Managing and Controlling Growing Harbour Terminals*. Society for Computer Simulation International.

Bontempi, G., L.M.Gambardella and A.E.Rizzoli. 1997. "Simulation and Optimisation for Management of Intermodal Terminals." In *ESM'97* (Istanbul, Turkey, Jun.1-4). Society for Computer Simulation International, 646–652.

Bruzzone, A.G. and R.Signorile. 1996. "Genetic Algorithms and Simulation as Support for Planning a Port Terminal." In *Proceedings of the 8th European Simulation Symposium* (Genoa, Italy, Oct. 24-26). Society for Computer Simulation International, 616–620.

Gambardella, L.M., G.Bontempi, E.Taillard, D.Romanengo, G.Raso and P.Piermari. 1996. "Simulation and Forecasting in Intermodal Container Terminal". In *Proceedings of the 8th European Simulation Symposium* (Genoa, Italy, Oct. 24-26). Society for Computer Simulation International, 626–630.

Gambardella, L.M., A.E.Rizzoli and M.Zaffalon. 1998. "Simulation and Planning of an Intermodal Container Terminal." *Simulation, Special Issue on Harbour and Maritime Simulation*, 12 pages, to appear.

Hayuth, Y., M.A.Pollatschek and Y.Roll. 1994. "Building a Port Simulator." *Simulation*, 63, no.3: 179–189.

Mastrolilli M. and L.M.Gambardella. 1998. "A Neighborhood Function for the Flexible Job Shop Problem". Technical report 10-98. IDSIA, Lugano, Switzerland (Feb.).

Zaffalon M. and L.M.Gambardella. 1998. "A Network Design Approach to the Allocation of Resources in an Intermodal Terminal." Technical report 10-98. IDSIA, Lugano, Switzerland (Feb.).