

Newcastle University e-prints

Date deposited: 7 December 2009

Version of file: Author final

Peer Review Status: Peer reviewed

Citation for published item: Marinov, M and Viegas, J. A Simulation modelling methodology for Evaluating Flat shunted yard operations. **Simulation Modelling: Practice & Theory, 2009 (17), 6.**

Further information on publisher website:

http://www.elsevier.com/wps/find/journaldescription.cws_home/622330/description#description

Publishers copyright statement:

The definitive version of this article, published by Elsevier, can be found at the above citation

Use Policy:

The full-text may be used and/or reproduced and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not for profit purposes provided that:

- A full bibliographic reference is made to the original source
- A link is made to the metadata record in DRO
- The full text is not change in any way.

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

**Robinson Library, University of Newcastle upon Tyne, Newcastle upon Tyne.
NE1 7RU. Tel. 0191 222 6000**

A Simulation Modelling Methodology for Evaluating Flat-shunted Yard Operations

Marin Marinov * and José Viegas

*Instituto Superior Técnico - CESUR at Technical University of Lisbon,
Av. Rovisco Pais, 1049-001 Lisboa, Portugal*

Abstract

Provided, in this paper, is a simulation modelling methodology for analysing and evaluating flat-shunted yard operations. Created and implemented is a yard simulation model using a computer package for event-based simulation, SIMUL'8. The main idea behind the simulation modelling approach is to simulate the flat-shunted yard operations dividing the yard into segments so that the behaviour of each segment can be described and analyzed separately. The simulation model takes the shape of queuing network. The components of the queuing network are interconnected queuing systems that interact and influence one another, so that the global impact of freight train operations is captured. The queuing systems replicate the preliminary specified segments that consist of a set of work centres (i.e., servers) and/or storage areas (i.e., capacity limitations). This modelling approach allows us to study the processing capabilities of the yards. In the shape of "Case Study", we demonstrate how the proposed approach is implemented in terms of "real world" case.

Key words: Yards, Rail freight transportation, Event-based simulation

1. Introduction

In the traditional rail freight systems, unlike block trains and shuttle trains, a single freight car or block of freight cars does not move usually on one freight train directly to its demand destination. Instead, the freight car moves on various freight trains. This process can be specified by schedules. To some extent, the schedules indicate the connections between freight trains that the freight car must be part of in order to arrive at its demand destination at the appointed time. These connections in the rail networks, take place at the yards. If a connection between two planned freight trains in a yard fails, the client does not receive his freight at the appointed time, the yard queue materializes because freight cars are left behind awaiting some next possible connection, the yard limited physical capacity is reduced, the yard limited processing capability is reduced and the yard personnel encounters difficulties to serve the next freight trains, costs (both operating costs and waiting costs) for the rail freight service provider are on the increase, and the service provided deteriorates. Consequently, the yards are facilities that play an important role for the execution of the freight transportation service by rail and their importance must not be neglected.

1.1 Motivation. The railway freight operator under this study is Comboios de Portugal, CP – Carga, still the main Portuguese rail freight service provider. CP Carga schedules in advance its freight trains at planning management level. These schedules consist of information for the movement of freight cars in the shape of freight trains over the rail network subject to demands of origins and

* Corresponding Author. Tel. + 351 21 841 63 26
Email Address: marinov@civil.ist.utl.pt (Marin Marinov)

destinations, arrivals and departures of freight trains as well as reassembling in the yards. Next, the prepared schedules are submitted to the operations for execution.

At planning level, the planner tries to satisfy required demands for transportation according to the preliminary planned freight trains, historical data, and production scheme in operation as well as available slots in the actual timetables among the passenger trains in a case of extra freight trains. The planning process is totally performed manually and to a significant extent depends upon the experience and the skills of the planner. Moreover, the planning process methodology does not explicitly involve the estimation of the maximum processing capabilities of the yards. There is no instrument for it. Therefore, some suggestions for freight train/freight car movement are infeasible with respect to the maximum operational capacities of the yards.

On the other hand, what is observed in the front-line operations with the CP freight trains is that the operations fulfil the daily services with little regard to what was planned. The yard personnel say things such as: *“the superiors consider the shunting to be executed for 20 - 30 minutes, we cannot perform it because only the break test takes about 20 minutes...moreover, they planned too many freight cars to stay in the yard, there is not enough space, we need lines to execute our work, therefore in every opportune case regardless of the plan we send freight cars away to ensure space for the incoming freight trains!”* In response to this situation the planning personnel say things such as: *“we planned well what was required by the commercial department but the operation did not execute it as we planned”* (Marinov 2007, p. 12)!

Therefore, Planning and Operations at CP Carga have shown a low confidence in each other. This awkward situation appears to be because at the planning stage the maximum processing capabilities of the CP yards are not estimated and considered explicitly and therefore in many cases the operations encounter difficulties to produce what is scheduled. This situation motivated us to provide an adequate methodology accompanied with a reliable tool for estimating the maximum processing capabilities of CP Cargo yards.

1.2 Objectives. The objectives of this work are to provide an adequate methodology accompanied with a reliable tool for estimating, analysing and evaluating the performance capabilities of rail yards using an appropriate approach for the purposes of a rail freight operator. The provided methodology is envisaged to be used at planning management level of freight services by rail.

1.3 Paper Outline. This work includes a paper addressed at A Simulation Modelling Methodology for Evaluating Flat-shunted Yard Operations and is presented in a cumulative fashion, each section dependent to a certain extent on the material in the preceding sections.

The rest of this paper is organized, as follows: in Section 2 short definitions for yards are provided followed by a brief literature review focusing on analytical queueing models and simulations in Section 3. Next, a presentation of the event-based simulation language, SIMUL 8 and its attributes used for the objectives of this paper are given. A simulation modelling methodology for evaluating flat-shunted yard operations is provided in Section 5, followed by a detailed case study demonstrating its application in Section 6. We wrap up with synthesis and conclusions in Section 7.

2. Yards

Yards take a very important part in railway freight operation. Their behaviour is a subject of difficult analysis and deserves careful examination. Depending upon the yards' performance the quality of the railway freight network operation may be ameliorated or may be aggravated. To some extent, the yards dictate the freight train movement on the network. They concentrate a significant amount of resources (both static and dynamic) in order to receive, accommodate and send freight train compositions for the sake of provided service. In other words, the yards execute reassembly processes with freight trains based on the delivery requirements of the customers. The yards, however, are characterized with a limited number of tracks, meaning their performances are

confined by bounds. These performance bounds identify the yards processing capability and when they are stretched the yards malfunction. This phenomenon contributes to rail freight services of poor quality.

In general, there are three types of yards (Konstantinov 1969; Tasev 1969; Tasev and Karagyozev 1983; Raikov 1986; Kraft 2002; Rhodes 1998, 2003; Wegner 2003; Kumar 2004):

- 1) Flat-Shunted Yards by definition these are the yards in which the tracks lead into a flat shunting neck at one or both ends of the yard where the freight cars are pushed and pulled by a shunting engine which sorts them into the assigned track;
- 2) Hump Yards the largest and most effective classification yards with the largest shunting capacity (often several thousand cars a day). These yards are built with a hump: a lead track on a hill (hump) over which the cars are pushed by the engine. Single freight cars, or some coupled freight cars in a block, are uncoupled just before or at the crest of the hump and roll by gravity into their destination tracks in the classification bowl (the tracks where the freight cars are sorted). The speed of the cars rolling down from the hump into the classification bowl must be regulated because of the different natural speed of the freight cars (loaded or empty, heavy or light freight, number of axles) and the different filling of the tracks (whether there are presently few or many freight cars on it);
- 3) Gravity Yards these yards operate similarly to hump yards with the difference that the whole yard is set up on a continuous falling gradient and there is less use of shunting engines.

The railway operator under study, which is CP Carga – “Comboios de Portugal, Carga”, performs only flat-shunting operations. Currently, all the yards in Portugal are flat-shunted yards, and therefore our methodology is explicitly dedicated to flat-shunted yards.

In processing freight trains at yards, there are sets of operating processes to be fulfilled at different areas. More precisely, there are sets of operating processes on receiving freight trains at arrival area; sets of operating processes on shunting at reassembling area(s); sets of operating processes on freight trains bound to leave at departure area(s); sets of operating processes on serving locomotives at workshop area(s); as well as sets of operating processes on storing freight cars at “Rip Tracks” area(s). Therefore, we assume that flat-shunted yard consists of the following areas:

- Arrival Yard (AY) a number of tracks within the yard limits where freight trains arrive;
- Shunting Zone (SZ) a number of tracks within the yard limits where freight cars are shunted;
- Departure Yard (DY) a number of tracks within the yard limits from where freight trains depart;
- Workshop (WS) a number of tracks within the yard limits where the locomotives undergo maintenance and stand while waiting for their new assignment;
- Car Yard (CY) a number of tracks within the yard limits where freight cars not in current use are stored. This is also called “Rip Tracks”.

Figure 1 shows these areas graphically.

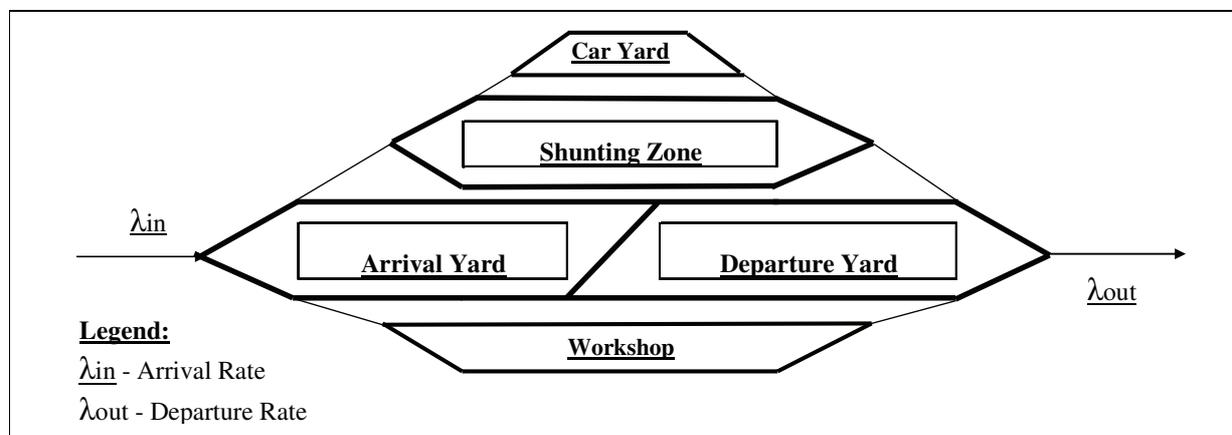


Fig. 1 A Simplified Drawing showing Flat-shunted Yard Areas

3. Literature

In the literature, analytical queueing models and simulations have been generally used in analysing and evaluating the complex yard behaviour. When using analytical queueing models the concept is that one decomposes the yard being examined into subsystems involving the different operations performed and corresponding physical and human resources (e.g., subsystem dedicated to receiving freight trains and inspections; subsystem dedicated to shunting; subsystem dedicated to departure of freight trains; etc.). After having this done one replicates and analyses the behaviour of each subsystem by a limited class of known queueing systems assuming they operate in *steady state* (Peterson 1977a, b). For more information on queueing systems, the interested reader is referred to Lee (1968), Cooper (1981), Gross and Harris (1985), Hall (1991). These queueing systems are classified by: arrival process; service process; number of servers; and maximum queue size. Generally speaking, this method has a pedagogical foundation and quickly provides insights on the behaviour of a queueing system. There are a few queueing systems, however, that operate with exact formulas (e.g., $M/M/1(m)/\infty(b)$ - Poisson Arrivals, exponential service times, a single service (or m - servers), infinite capacity (or b - buffer size);). If an exact formula does not exist, approximations for computation of the measures of subsystem performance are used (Shore 1988a, b; Karagyoov 1990a, b).

Another class of analytical queueing models is a network of queueing systems. This issue is not yet resolved in a satisfactory way. In the literature only few “*products*” on queueing networks are found, but with descriptions that are quite limited in application. These are:

- Open Queueing Network;
- Closed (circuit or cyclic) Queueing Network;
- Queueing Network Analyser.

The closed queueing network products describe systems characterized with a finite number of potential customers, i.e. the number of customers is fixed. Scientifically, those models are only important when the number of potential customers is relatively small. Otherwise, in the case of a very large number of potential customers, they are accepted to be infinite and the system is treated as an open queueing network product. If the arrival process is accepted to be Poisson, the service times are accepted to be exponential, and the buffer sizes are accepted to approach infinity, then one is dealing with a classical exponential queueing network product of Jackson (Jackson 1963). In fact, the customer arrival process is not always plausibly described with a Poisson distribution and the usage of Jackson network is inappropriate. In these cases one better uses the queueing network analyser which is said to provide a fast solution for large networks with fixed routing probabilities. The arrival process is formulated as a renewal process defined by the first two moments with independent and identically distributed inter-arrival times. The service process must also be defined by the first two moments (Whitt 1983). Generally speaking, the main idea behind this tool is to solve the traffic equations and then decompose the network into single $G/G/m$ queues and solve those individually. This tool employs mainly approximations.

On the other hand, simulation provides more realistic results when one deals with a certain level of variability, which is to a certain extent the present case. Generally speaking, in the freight train movement there are factors such as: stochastic behaviour, different structures, equipment and work technology at railway facilities, compound delays, and ripple effects subordinated to non-stationary arrivals track blockage, train breakdowns, scheduled and unscheduled interruptions, etc. All these factors make the development of analytical models to analyse and evaluate processing capabilities in terms of yards extremely difficult (Dessouky and Leachman 1995; Katchaunov et al. 1998).

To make the best use of simulation, when analysing the complex yard behaviour (or any other complex system) one better operates with a specific simulation tool created for this particular purpose, as in Germany (Pachl and White 2003) or the software package VIRTUOS (Klima 1997, 2001, Kavicka 2000).

When no specific yard simulation software is available or appropriate, one needs to analyse, choose and adapt an existing simulation tool for this purpose. A simulation language that has been used in examining terminal behaviour is General Purpose Simulation System (GPSS). GPSS is a process-oriented simulation language that combines sequence of events into single subroutines called blocks (Nadel and Rover 1967, Karagyozev 1983, Katchaunov et al., 1998, Razmov 2004, Ivanov 2005). We are also aware of a class project on hump yard simulation performed by Harrod (2003). The subject of this independent project is Queensgate yard - one of North America's largest hump switching yards. Harrod has studied Queensgate yard performances by using Arena Simulation Tool. Arena employs an object-oriented design for entirely graphical model development. Simulation analysts place graphical objects—called modules—on a layout in order to define system components such as machines, operators, and material handling devices. The core technology of Arena is the SIMAN Simulation Language (Takus and Profozich 1997). After creating a simulation model graphically, Arena automatically generates the underlying SIMAN model used to perform simulation replications.

It should be noted that CP Carga aims to run its freight trains on a strict fixed schedule basis, but since the operations cannot fulfil these schedules, as was argued above, the concept of strict fixed schedule becomes quite improvised one, bearing all the disadvantages of this awkward situation (i.e., service of poor quality, delays all over the network, low utilization rates of the assets, huge diseconomies of scale, and the like). Therefore, for our purposes required was an approach accompanied with a tool that would be able to mainly replicate this rail freight system in two situations, as follows:

- 1) *What happens when the schedules are fulfilled?; and*
- 2) *What happens when the schedules are not fulfilled?.*

Furthermore, the planning department of CP Carga fulfils its schedules in terms of inbound/outbound freight trains. In order to replicate this particularity and next evaluate CP schedules, it was required that one should work in terms of inbound/outbound freight trains and not precisely focusing, for instance, on organization of shunting processes and how it is influenced subject to changes in the freight train compositions. Thus, this study required appropriate apparatus for system replication, performance analyses and evaluations at mezzoscopic level. According to the best of our knowledge, a yard performance evaluation study of this type has never been conducted in Portugal before. This is the very first step at this front in terms of CP Carga. The only apparatus for system replication, performance analyses and evaluations at mezzoscopic level made available to us at the time was SIMUL 8 and therefore for the objectives for this discussion, the (visual) simulation modelling experiments conducted here were created and implemented using a computer package for event-based simulation, SIMUL 8. The main idea behind the simulation modelling approaches is to simulate the yard behaviour by dividing the yards into segments so that the behaviour of each segment can be described and examined separately. The simulation model takes the shape of queueing networks. The components of the queueing networks are interconnected queueing systems that interact and influence one another, so that the global impact of freight train operations is captured. The queueing systems replicate the preliminary specified segments that consist of a set of work centres (i.e., servers) and/or storage areas (i.e., capacity limitations). This modelling approach allows us to analyse and evaluate the processing capabilities of yards subject to yard resources involved.

4. Simulation Language SIMUL 8

A powerful tool to study general processes and operations is SIMUL 8 simulation language (Shalliker and Ricketts 1997). SIMUL 8 is a computer package for discrete-event simulations that allows us to create visual animation models of a wider range of queueing systems by drawing functional objects and attributes directly on the program screen. When the system being examined is properly modelled a simulation experiment can be undertaken. The flow moving over the

queueing network is shown by animation so that the appropriateness of the created simulation models can be more easily assessed. When the structure of the model is confirmed a number of trials are run. Measures of performance are described statistically.

Next, we concentrate on the usage of SIMUL 8's computer package below.

4.1 The SIMUL 8 Building Blocks. In simulating queueing systems by SIMUL 8 there is a set of building blocks to be described and used. Each block is characterized with specific, irreplaceable properties. The accuracy of the models and the representation of the real systems being simulated, depend on the accurate description and positioning of the blocks and on the right linkages between them. The linkages identify the service path of the "*work items*" which are the objects to be processed by the system being scrutinized. Building blocks of first importance in SIMUL 8 environment are, as follows:

Work Entry Point – this is a generator of work items. The work items are served by the system. In other words this is where the customer arrives in the queueing system to be served. The customer may be of many different natures. In the present case, in general, the customer is a freight train. The Work Entry Point is characterized by an arrival pattern. The arrival pattern may be subordinated to a familiar theoretical distribution or the researcher may use a time dependent distribution. We look at time dependent distributions in a later step of this section.

Work Centre – this is where the customer is being served by a server. The complex queueing systems consist of a set of work centres through which the service is executed in specific order in accordance with customer's processing path. Every work centre is characterized with inbound traffic, service pattern and outbound traffic. The outbound traffic is to be routed to other objects in a variety of ways, i.e. to next work centre, or to a "*buffer*" (Storage Area) waiting for next operation, or to leave the system when the full service is completed. The service duration pattern is to follow a particular probability distribution.

Storage Area – this is found as buffers or queues. Briefly said, this is where the customer is held while waiting to be served. The queues can be controlled by their capacity, by the service discipline as well as by the performance of the server to which the queues belong.

Work Exit Point – this is where the customer leaves the system and the service is declared as "finished".

4.2 Processes. To set up any process at the technological (sub)systems a resource or set of resources is required. A resource could be a human, a machine, an operator, a brigade, an employee, an order, etc. or just a single signal. So the resources are items in the simulation model which are necessary for fulfilment of the service at the work centres. The resources are usually specified and identified by particular properties and priority rules for their use. This strongly depends on the case being studied. Speaking in general terms, there are two types of resources, as follows:

Individual resources which are responsible only for one task; operating only at one work centre;

Shared resources which may perform tasks at several work centres. Shared resources are also defined as "floating resources" because at their work they float between the work centres that require them.

In SIMUL 8 environment one may operate with these two types of resource as the resources' assignments are easily set and controlled by the "Resource Dialog Box" menu.

4.3 Routing. Routing is the way by which the building blocks are connected with one another in order to manage the sequence of services to the customers. These are the linkages of the whole structure and in the simulation experiment they are to follow the throughput line of the service. Simply the linkages are made up by Work Flow Arrows which indicate the default path for the customer moving through the technological (sub)systems and storage areas of which the simulation model consists.

An important feature is that when there is no buffer between two successive technological (sub)systems, but there is a direct link between them, by default the customers do not move from the first technological (sub)systems to the second until the second technological (sub)systems is ready to serve them.

4.4 Technical Aspects of Running Simulations. An issue of first importance is that the simulation is performed for a certain period of time, according to plausible initial conditions, and for a sufficient number of experiments i.e., simulation trials. This is the art of making simulation, rather than the way of building simulation models.

By default the simulation starts off with no customers in the system being analysed. Such a case might be applicable in some occasions, but not to all. In most transportation systems e.g., there is a never-ending stream of vehicles. Therefore, if the simulation starts off with no customers in the system one creates an idle period which might not face the reality and yields unrealistic results for a while. This issue is known in queueing theory as *transient behaviour* and is advisable to be avoided in order to allow the system under analysis to approach a steady state. SIMUL 8 allows this by inserting a “Warm-up Period”. Such an option sets a number of time units in starting off the simulation and permits the customers to become spread throughout the system.

Further when one conducts a simulation experiment one works with instances. Every second instance provides hardly ever the same results as the previous one. So there is a level of variation in a simulation experiment one should be aware of. In reality there is variation from day to day, from month to month, from semester to semester, from shift to shift etc. In fulfilling the simulation experiment one makes “*replications*”. It should be noted, however, that the adopted term for this notion in SIMUL’8 language is “run”. Nevertheless, we shall stick with the notion “*replication*”, since it is commonly used in the literature. And so, each replication is characterized with a proper set of random numbers. It is called “Random Sampling”. Every random sampling yields a different set of results. Consequently, to investigate the level of variation, it requires a number of replications for each time period of one and the same simulation model to be conducted but with different random samples. The level of variation is easily investigated in SIMUL 8 environment by conducting multiple replications (called “*trial*”) automatically through “Conduct trial” option.

4.5 Interruptions (Scheduled and Unexpected). In simulation of queueing systems one assumes that the arrival pattern is a “given”. But in some railways, for the sake of the service the systems operate on schedules, meaning appointed customer arrival times and appointed customer departure times. To put it briefly, it means that the systems at providing their service are to meet a deadline. In doing so, in order to meet an appointed deadline the systems in their schedules insert waiting and idle times. Moreover, depending on the service process, maintenance requirements and safety reasons, there are stoppages in operations known in advance. The aforementioned intervals have by nature timing dimensions and are found as “*scheduled interruptions*”. Another class of interruptions is the *unexpected*. Unexpected interruptions are caused by unpredictable factors seen in service perturbations, schedule discrepancies, breakdowns, and technical failures. Undoubtedly, a good simulation model should be capable of replicating such interruptions. In SIMUL 8 environment the tools which describe the interruptions are the distribution of time between interruptions and the distribution of length of interruptions. To set and control interruptions in simulating by SIMUL 8 one operates with the Efficiency Dialog Box.

4.6 Time Dependent Distribution. Time Dependent Distributions help us analyse arrival patterns characterized with *predictable variability*, meaning queues occur in kind of a statistically predictable fashion. The product form of predictable variability is a *predictable queue* and the systems are to operate in quasi-steady state. In quasi-steady state the arrival rate $\lambda(t)$ is not a fixed value, but a function of time and yields the expected number of customers to arrive over a certain time interval. It is mathematically described as follows:

$$E(\text{customers arrivals between time a and b}) = \int_a^b \lambda(t) dt \quad (1)$$

The expected number of customers from time a to b is $\Lambda(t)$ and is computed by integrating the arrival rate $\lambda(t)$ from time a to b. Since the arrival rate $\lambda(t)$ is a function of time, which is not constant, consequently the absolute utilization level $\rho(t)$ must also be a varying function.

The investigation of systems that operate in quasi-steady state falls within the “*phase*”- analysis because they pass through different phases which are identified by different arrival rates and absolute utilization levels. This technique is quite powerful and necessary for detecting awkward situations, such as “peak” and “idle” hours. Those awkward situations can be clearly recognized when the queues start to materialize. In SIMUL 8 environment the predictable variability is overcome by making the customer arrival time dependent over a regular work cycle.

4.7 Measures of Performance (MOPs). In SIMUL 8 the measures of system performance are assessed per building block, per resource as well as for the system as a whole.

For a Work Entry Point the computed measure of performance is “Number of work items entered”.

For a Storage Area the computed MOPs are, as follows:

Number of Work Items in Storage - this is the queue length and its pattern can be seen in a graphic window;

Queueing Time demonstrated in two sets – the first set is to all work items; the second set is only to those work items which had to queue. The distribution of queueing times can be seen in a graph window as well.

For a Work Centre the computed MOPs are:

Number of Work Items (currently in the centre; completed; average; minimum and maximum);

The Percentage of Time for which the Work Centre is either awaiting work, working, blocked or stopped. It can be plainly visualized by a pie chart.

For a Work Exit Point the computed MOPs are work-complete results, as follows:

Number of Work Items completed;

Times in System (average, minimum and maximum as well as standard deviation). The distribution of time in the system can be depicted by a histogram.

For a Resource, the computed measure of performance is its utilization in percentage which can also be demonstrated by a graph.

4.8 Set of Results. Generally speaking, when modelling and simulating with SIMUL 8 there are two different sets of results that one may obtain. These are, as follows:

- set of results at the end of a replication
- set of results of a trial

The *set of results at the end of a replication* demonstrates what happened during the conducted replication on either average and/or as a limit values, or in other words these are the relevant values that were found over the length of conducted replication.

The *set of results of a trial* demonstrates a result summary on average for conducted trial as well as the level of variability over the simulation experiment. In the result sheet the level of variability is assessed by computation of confidence intervals.

5. A Simulation Modelling Methodology for Evaluating Flat-shunted Yard Operations

The adapted concept for simulation of the operating processes with freight trains at flat-shunted yards targeting estimation of processing capabilities is to follow the throughput line of the yard being examined. In Figure 2, a detailed scheme of the throughput line of the flat-shunted yard is presented. The throughput line of the yard is identified by a set of interconnected consecutive and/or parallel operating processes at the different areas (i.e., technological systems) of the yard, discussion of which comes next:

i. There are n freight trains spread throughout the railway network that are to be served by the yard being examined. The freight trains are replicated as a flow and as we have seen in SIMUL 8 environment they are generated by tool called “Work Entry Point”. Depending on the objectives of the study the generation of freight trains may follow probability distributions if one deals with improvised operation, or may follow strict fixed pattern permitting some random fluctuation if one aims to analyse system performances under structured operation. A measure that one obtains is the number of freight trains generated for a certain period of time (i.e., the period included in the simulation experiment).

ii. After having generated the freight trains, they are located over the railway network. To represent this property in SIMUL 8 environment one creates a Storage Area characterized with infinite capacity, and hypothetically speaking this is where the freight trains stand while waiting to be served by the yard being examined. This is required because once the freight train is on the road, the train cannot be cancelled; the train can be delayed in cases of “*en route*” perturbations or an oversaturated yard that is not able to accommodate it immediately. Then the freight train is forced to wait outside of the yard limits. To observe this phenomenon, such a Storage Area is inserted. The measure of performance observed is the number of freight trains in the railway network that requires service by the yard being examined.

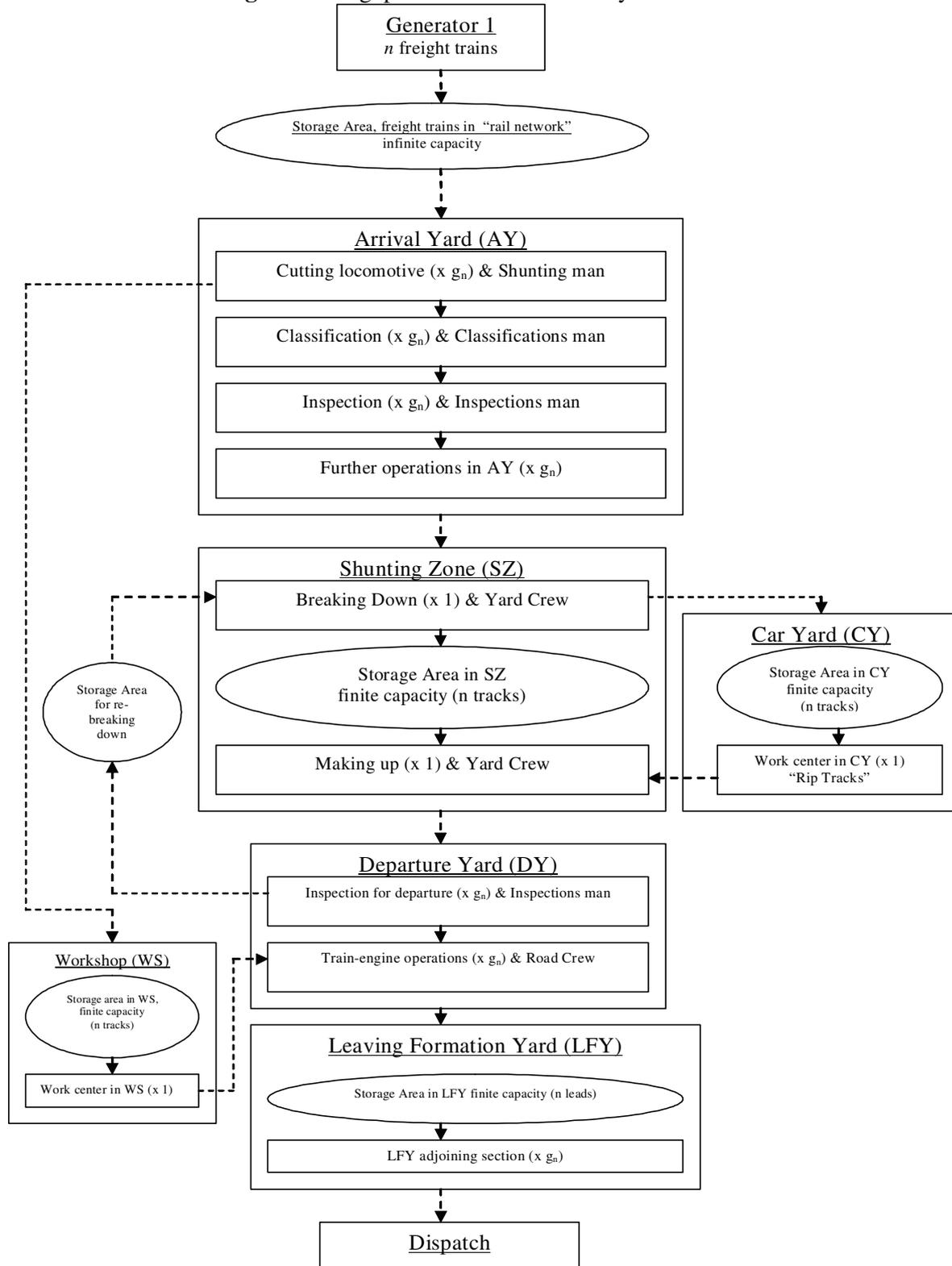
iii. Freight trains arrive individually, one by one on the Arrival yard tracks where a set of operations is fulfilled. One considers operations, such as: cutting the locomotive from the freight train composition; classification of freight cars; technical and commercial inspections; etc. In simulating with SIMUL 8, each operation is represented by a set of g equal Work Centres. The service patterns are to follow probability distributions. In fulfilment of each operation the responsible employee is required. The responsible employees for the simulation are set up by “*Floating Resources*” and one assigns e.g., for fulfilment of cutting the locomotive from the train composition a shunting man; for classification of freight cars a classifications man; for technical inspection an inspections man; etc. One obtains: Measures of *work centre* performance, such as: Number of freight trains completed; as well as the percentage of time for which a given work centre is either awaiting a freight train, working, blocked or stopped. This is visualized by pie chart. A measure of *floating resource*'s performance that one obtains is its utilization in percentages.

iv. After having completed the service at Arrival yard the freight train is prepared for breaking down and the freight train proceeds to the Shunting zone. The shunting zone is divided into two parts. The first part is dedicated to breaking down and this is where the rearrangement process begins. Technologically, breaking down means that the freight train is spread out in pieces on a limited number of tracks in order to be easily regrouped by the next technological subsystem, which is dedicated to making up of freight trains. Those freight cars found to have defects, or not in a current need are set off to “Car Yard” area for storage and/or repair. Later on when those freight cars are technically ready and/or there is need for them in the daily service they are to join the rearrangement process at the technological subsystem dedicated to making up of freight trains.

The whole set of operations in the Shunting Zone, involving “*Breaking down*” of freight trains, “*Making up*” of freight trains (as well as setting off to and picking up freight cars from the Yard Car), is fulfilled by Shunting Brigade, or simply said Yard Crew (or shunting crew). Thus, to replicate in SIMUL 8 environment the rearrangement process fulfilled within the Shunting Zone of the yard one inserts one Work Centre dedicated to breaking down process; followed by a Storage Area with finite capacity to replicate the number of tracks used for shunting of freight trains

(normally they are 2 or 3 in the practice of flat-shunted yards) and next one inserts another Work Centre dedicated to making up process. The service times for breaking down and making up freight trains are usually distributed according to Normal or Uniform Distribution with upper and lower bounds. However, if one leaves the model that way, the model will be somehow incomplete and the replication of the rearrangement process would not face reality.

Fig. 2 Throughput line of Flat-shunted yard



Let us be reminded that both breaking down and making up of freight trains are fulfilled by Shunting Brigade (or yard crew). Consequently, the execution of these two processes is dependent upon the performance of the Shunting Brigade seen in execution of a number of shunting operations employing moves of shunting engine (backward and forward from one track on another, empty and loaded pulling and pushing freight cars); coupling and uncoupling of freight cars and on/off the shunting engine; works with the switches and the like. To replicate this complex function of the Shunting Brigade within the Shunting Zone one inserts a *floating resource* involved to act in both the Work Centre dedicated to breaking down of freight trains and the Work Centre dedicated to making up of freight trains on a frequent basis. And so, this *floating resource* is here used to replicate the performance the Shunting Brigade (or Yard Crew, as used in Figure 2) and one is thus able to observe its utilization in percentage for the time period of the simulation experiment.

Next, the Car Yard is replicated as one Storage Area with finite capacity and one Work Centre, as its attributes are directly linked with the attributes of the Shunting Zone, i.e., the Work Centre dedicated to breaking down of freight trains and the Work Centre dedicated to making up of freight trains.

v. The next technological system is dedicated to the operating processes with freight trains at Departure Yard. At this technological system the freight train is prepared to depart. One considers two main sets of operations, such as: inspection for freight train departure; and freight train - engine operations. An *inspection for departure* consists of mechanical inspection of the freight cars and commercial inspection of the freight. *Freight train - engine operations* consist of arriving and putting locomotive on train composition as well as full brake test. Full brake test is a detailed inspection of the brake system. Brake pipe leakage must not exceed the norm. All of the brake rigging of each car must be properly secured and must not bind or foul. Retaining valves must be in operating condition. The brake cylinder piston travel of each car must be adjusted to within the limits prescribed for that specific equipment (White 2003a). For fulfilment of those two sets of operations one sets up *floating resources*, such as: inspections man and road crew. An important element is that during the technical inspection defects or other faults on the freight train composition might be detected. In such an awkward situation re-breaking down is required. Thus, the operating processes at the departure yard are represented by two technological subsystems. One subsystem dedicated to inspection for departure and another dedicated to train-engine operations.

vi. A crucial element that requires some additional care is the locomotive service at the flat-shunted yards. The locomotives are generally served in separate areas of the yard. Such areas are called Workshops or/and Roundhouses. The locomotives are specified by categories, manufacturing series, specific properties and technical characteristics. They are the engine of the freight train composition. In railway terms, each individual locomotive is called “unit”. An engine of freight train may be a single unit or several units coupled together in multiple where more power is needed for the sake of the freight transportation service. Coupled in multiple means coupling units to the locomotive consist, at either the rear or between operating unites. Other subject is that of the *shut down locomotive*. Locomotives that are shut down are known as *dead engines* and this usually occurs when a locomotive is required in some place on the railway network and cannot be utilized as a prime mover (prime mover means a main generator for pulling the freight train) because there are not enough freight cars to be pulled in this direction and the locomotive should make a *non-revenue run*. To economize such a non-revenue run, engines may be moved dead in freight trains between the freight cars. Such a locomotive cannot have the brake system operating and must be transported like a boxcar, cutting out the control of the brakes. However, in practice non-revenue runs cannot always be avoided.

For the objectives of this discussion, one considers that after having pulled the freight train composition into the arrival yard the locomotive is cut off from the freight train and sent to the technological system: “Workshop”. There the locomotive undergoes a mechanical inspection that consists of verifying the technical readiness of the locomotive, checking the fuel if diesel, water and other technical conditions. After having the inspection done the locomotive is prepared for its next assignment. These processes in SIMUL 8 are replicated, as follows:

-The workshop is characterized with limited capacity and that is replicated by inserting one storage area with finite capacity, which equals the maximum number of locomotives that the workshop can accommodate. Measures that one obtains are: number of locomotives entered in workshop; number of locomotives currently waiting for a mechanical inspection as well as queueing times which may be demonstrated by a histogram.

-The service at the workshop (roundhouse) is represented by a Work Centre. The service times are to follow probability distributions. In CP Carga terms, the responsible employee for checking the locomotive technical readiness between subsequent assignments is the engineer who will drive the locomotive. If there are faults ascertained and the locomotive is not in technical readiness to fulfil its next assignment, the locomotive might need a heavy repair. In such an awkward situation, the locomotive is withdrawn from service and is pulled dead to a terminal that has a repair-shop. Measures obtained are: Number of locomotives currently in the workshop as well as Locomotives served by the workshop.

vii. Every freight train requires an engine. To equip a freight train with locomotive(s), “train-engine operations” are needed. These operations consist of arriving and putting a locomotive(s) on a train composition as well as a full brake test. Theoretically speaking, in the technological subsystem dedicated to train-engine operations, the locomotive and already assembled freight cars are set together. The freight train is finally composed. In order to replicate such a detail in SIMUL 8 environment one activates the discipline “Collect” in “Routing in” menu. Measures of performance observed are as follows: Number of freight trains completed per technological subsystem; Number of freight trains currently in the technological subsystem; Queueing times as well as the percent of time for which the technological subsystem is either awaiting a freight train and a locomotive to be merged; working; blocked; or stopped. This is also displayed by pie chart.

viii. After having completed the operations at the Departure yard the freight train is ready to leave. Then, the freight train needs permission, meaning a slot on the assigned adjoining section. Giving a slot depends upon the organizational pattern of the movements of all categories of trains as well as the traffic rules over the assigned adjoining section. Consequently, the freight train might wait for a while before leaving. Since every freight train has a particular assignment, the train is to leave the yard on a defined lead. One replicates this by inserting: one Storage area characterized with finite capacity that equals the number of leads; and a Work centre to replicate the execution of operation on leaving of freight trains.

ix. As soon as the freight train leaves, its service at the flat-shunted yard ends. In order to declare the simulation design of yard as “completed” one inserts a Work Exit Point. This is an ending technical aspect in simulating with SIMUL 8 and might be defined as a “Dispatch”. Results that one might observe in Work Exit Point have the same meaning as measures of yard’s performance and these are: number of freight trains completed for the time period of the simulation experiment as well as time in yard. One is able to display such results by a histogram showing the distributions of time of freight trains in the yard being examined.

In implementing this methodology one creates a simulation model replicating flat-shunted yard behaviour on the basis of which runs simulation trials with the purpose of analysing and evaluating the throughput level of the yards through a set of measures of performance (MOPs). Also, depending on the number of freight trains to be served and their arrival patterns at the yard being examined one identifies the “saturation” point at which one or more of the technological (sub)systems of the yard under study might be blocked. This corresponds to the maximum processing capacity of the flat-shunted yard subject to technical characteristics, resources involved, and work technology incorporated at the production scheme of the yard.

6. Case Study “Gaia Flat-shunted Yard”

6.1 Description. The flat-shunted yard Gaia is located in North Portugal and serves the following rail lines: Linha do Douro, Linha do Leixoes and partly Linha do Norte, Linha do Milho with feeders Linha de Guimarães and Linha de Braga. Layout of Gaia yard is sketched in Figure 3.

Tracks I and II are the main lines of the station “Vila Nova de Gaia” where the passenger service is executed. Tracks I and II are not used for shunting purposes. Usually, the tracks III, VII (IV), VIII (V) and IX (VI) are used for either receiving and/or departing freight trains. These 4 tracks specify both the Arrival yard and the Departure yard of Gaia. Track X (VII) is usually used for storage of electrical road locomotives. The Shunting zone incorporates 4 electrified tracks (i.e., tracks XI (VIII), XII (IX), XIII (X) and XXVII) and tracks indicated from 1 to 10 which are not electrified. When it is necessary that freight cars stand and wait further processing within Gaia, those freight cars are usually stored on some of the tracks from 1 to 10. So, under some conditions some of those tracks may form a small Car yard. This small car yard is required from time to time and is not physically separated from the shunting zone where a shunting crew is in operation. Therefore, we consider such an operation as a part of the shunting. Furthermore, there are 2 pieces of track that are not electrified known as “CIMPOR” and “Tegão Areia”. CIMPOR is used for parking diesel road locomotives while waiting for their next assignments. “Tegão Areia” is usually used for shunting.

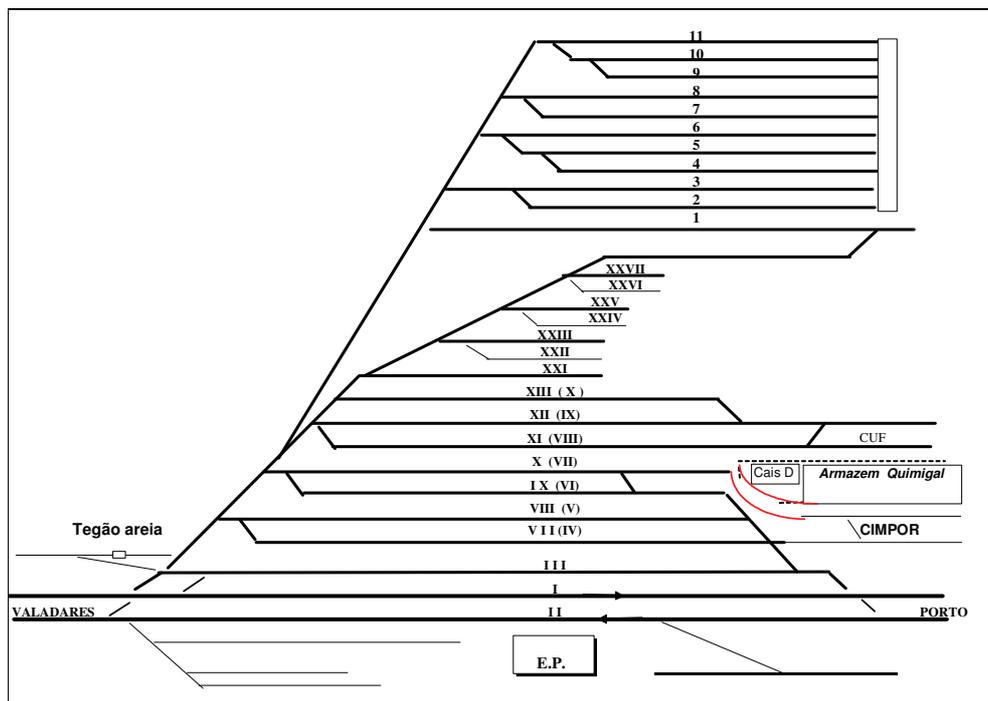


Fig. 3 Layout of Flat-shunted yard “Gaia” / *Source:* Dprpt. of Planning and Control, CP Carga

Track number 11 is deteriorated and cannot be used for service. The pieces of tracks from XXI to XXVII are leads to an old workshop, which is deteriorating now, unfortunately. The pieces of tracks from XXI to XXVII are used for either shunting and/or storage.

There is one shunting crew operating with one shunting locomotive (i.e., one engine in operation).

6.2 Throughput Line. The inbound freight trains arrive at the arrival yard of Gaia, where a shunting man is required to cut the road locomotive from the composition. After having cut the road locomotive, if it is a diesel, the locomotive usually goes to CIMPOR to wait for its next assignment, and if it is electrical, the locomotive usually goes to track number X (VII) to also wait for its next assignment. Next, classification of the freight cars is fulfilled by a classification man, i.e., the classification man walks around the composition to be shunted indicating the next destination of each freight car either empty or loaded. According to the indications provided by the classification man the shunting is executed by a shunting crew.

Before beginning the breaking down process, which is the first part of the reassignment process, however, the composition undergoes a technical inspection (i.e., the running conditions of each freight car are verified) by an inspection man. If defects or failures on freight cars are detected the

inspection man indicates these freight cars to be withdrawn. Usually, freight cars found having defects are stored temporarily on some of the tracks from 1 to 10 as sketched on Figure 3. Next, the breaking down process is fulfilled, meaning the composition is broken down into freight car groups on tracks of the shunting zone. The breaking down process is followed by the making up of the outbound freight trains, which is also called “*making up process*”. The making up of the outbound freight trains is fulfilled by the same shunting crew that breaks down the inbound freight train. The making up process consists of assembling a composition for an outbound freight train. Next, the assembled composition waits for a road locomotive to arrive and put on. In daily operation the locomotive drivers are responsible to verify the conditions of road locomotives, and therefore a technical inspection of each road locomotive is executed by the locomotive driver before every locomotive haul. If it is an electrical locomotive the driver verifies the pantograph positions and the indications on the cabin panel. If it is a diesel locomotive, the driver verifies the fuel, the water, the indications on the cabin panel etc.

In the meanwhile the freight cars are inspected. After having arrived and put the road locomotive on the composition, the full break test is fulfilled. If defects or failures on outbound freight train are detected, then re-breaking down is required. If problems are not detected the conductor of the road crew puts the beacon on the rear of the composition and the freight train is ready to leave the yard. The operation on leaving the yard is assisted with a REFER (i.e., the Portuguese Infrastructure Manager) worker and is fulfilled after having received a “green” signal given by the line dispatcher.

6.3 Modelling with SIMUL 8. In the real world, yards are facilities that rarely reach steady state. One of the reasons is because most freight railways tend towards strict fixed schedules. The railway freight operator under study is no different. The schedules are fulfilled in advance and consist of detailed information for the complex freight train operation. For the analytical modelling purposes, e.g., one considers fixed arrivals of freight trains at the arrival yard per hour. In practice, however, the freight train arrival rates vary over a regular daily cycle and are somehow predictable. This characteristic is very important to correctly specify the arrival pattern of freight trains at the yards and SIMUL 8 this is replicated as one makes the freight train arrivals time-dependent.

In the following section we conduct simulation modelling in order to analyse Gaia yard performances. Simulation is not a model. In general, it might be thought as a tool that provides information about the behaviour of the system under study by describing the way it works and calculating measures of its performance under deterministic or randomized scenarios. In order to obtain plausible outputs, one must very carefully specify and replicate the physical characteristics of the system, the true service times as well as the true input seen in the arrival process. We implement the simulation modelling methodology provided in Subsec. 4 above. Definitions presented there are used. We consider time-dependent freight train arrivals and predicted averages of the service times per technological subsystem obtained through observations. It should be noted that the railway freight operator under study (“CP Carga”) does not have available practical data about frontline freight train operations in yards. Therefore, in order to obtain the required data for this discussion, observations of real time frontline freight train operations during the night and during the day have been made. Next, all results obtained have been confirmed by the frontline yard personnas well as the planning staff. Thus, the simulation flat-shunted yard model developed is a process-oriented model implemented using the SIMUL 8 computer package. In order to scrutinize the behaviour of the closely studied yard, different alternative courses are examined. However, there is always an initial course that we called “Status Quo”.

Status Quo, Gaia Experience: The “Status Quo” characteristics of Gaia yard in SIMUL 8 environment are replicated as follows:

1) Generation and External Elements

Layout

Create a Work Entry Point called *Generator_Freight Trains* to generate freight trains

Create a Storage Area called *Rail NetworkQueue* to hypothetically replicate where the freight trains stand while waiting to be served by the yard

Create a Work Centre called *Arriving of Freight Train* to replicate the physical arrivals of the freight trains at Gaia arrival yard (i.e., this Work Centre is required to replicate the operations on preparing the route, e.g., all the switches to be put in order so that the freight train goes into the arrival area of the yard)

These three attributes in SIMUL 8 environment should look similar to Figure 4.

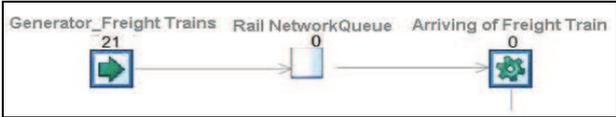


Fig. 4 External Elements' Layout in SIMUL 8

Freight Trains Generation

At the bottom of the Inter-Arrival Times options box of the Created Work Entry Point called *Generator_Freight Trains*, we select a (previously created) time dependant distribution from the list. The time dependent distribution consists of several arrival periods where each period is characterised with a different arrival rate of freight trains. This is to replicate that the freight train arrivals are subordinated to predictable transportation scheme i.e., a schedule

Storage Area

In the Storage Area dialogue box (Storage Bin Details) called *Rail NetworkQueue*, leave the capacity as Infinite to replicate that the freight trains stand somewhere in the network before being served by the yard

Work Centre

In the Work Centre Options dialogue box called *Arriving of Freight Train*, choose the Uniform distribution from the list and further set up lower bound equal to 1 minute and upper bound 2 minutes in order to replicate the physical arrival of the freight train at the arrival yard (it is assumed that the operations of preparing the route as well as arriving of freight train into the Arrival yard take from 1 to 2 minutes)

2) Arrival Yard

Layout

Create a Storage Area called *Arrival YardQueue* to replicate the limited number of tracks in Gaia arrival yard

Create a Work Centre called *Cutting* to replicate the operation of cutting the road locomotive from the train composition

Create a Work Centre called *Classification_Technical Inspection* to replicate the operation of classification and technical revision of the freight cars

Create a Work Centre called *Further Operations* to replicate further operations in arrival yard. In this Work Centre we consider the arrival of the shunting crew

SIMUL 8' attributes of Arrival yard should look similar to Figure 5.

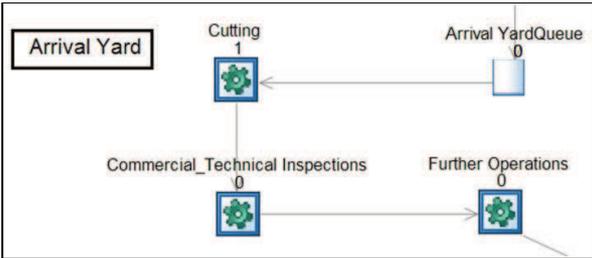


Fig. 5 Arrival Yard Layout in SIMUL 8

Storage Area

In the Storage Area dialogue box (Storage Bin Details) called *Arrival YardQueue*, change the capacity to 2 numbers (i.e., 2 tracks in the arrival yard of Gaia)

Work Centres

In the Work Centre Options dialogue box called *Cutting*, select the Uniform distribution with lower bound 1 minute and upper bound 2 minutes to replicate the execution of operation of cutting the road locomotive from the train composition. Floating resource required is a shunting man.

In the Work Centre Options dialogue box called *Classification_Technical Inspection*, choose the Uniform distribution inserting lower bound 15 and upper bound 20 minutes to replicate the operations on classification and technical inspections of freight cars. Note that these operations are fulfilled in parallel and jointly modelled. However, two floating resources are required, i.e., classification man and inspection man.

In the Work Centre Options dialogue box called *Further Operations*, choose the Uniform distribution with lower bound 1 minute and upper bound 2 minutes to replicate the arrival of the shunting crew before starting the breaking down of the composition

3) Shunting Zone

Layout

Create a Work Centre called *Breaking Down* to replicate the operation of breaking down the train composition

Create a Storage Area called *Shunting ZoneQueue* to replicate the limited number of tracks of shunting zone

Create a Work Centre called *Making Up* to replicate the operation of making up the train composition

Create a Storage Area called *Car YardQueue* to replicate the limited number of tracks in the car yard

Create a Work Centre called *Car Yard* to replicate the time of train compositions stand while waiting in the car yard

Model attributes replicating the Shunting Zone and the Car Yard should look like Figure 6.

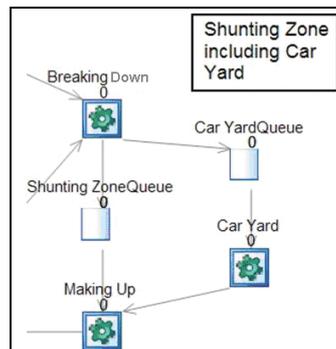


Fig. 6 Shunting Zone Layout in SIMUL 8

Storage Areas

In the Storage Area dialogue box (Storage Bin Details) called *Shunting ZoneQueue*, change the capacity to 7 to replicate the number of tracks of the shunting zone of Gaia

In the Storage Area dialogue box (Storage Bin Details) called *Car YardQ1*, change the capacity to 4 to replicate the number of tracks in the car yard of Gaia

Work Centres

In the Work Centre Options dialogue box called *Breaking Down*, choose the Uniform distribution with lower bound 15 and upper bound 25 minutes to replicate the operation of breaking down the freight train composition. Recourse in operation is the shunting crew.

In the Work Centre Options dialogue box called *Making Up*, set up the Uniform distribution with lower bound 5 and upper bound 10 minutes to replicate the operation of making up the train composition. Recourse in operation is the shunting crew.

In the Work Centre Options dialogue box called *Car Yard*, select the Exponential distribution with the mean equal to 500 minutes to replicate the time of freight train compositions stand while waiting in the car yard

4) Locomotive Workshop

Layout

Create a Work Entry Point called *Generator_Road Locomotives* to generate road locomotives

Create a Storage Area called *WorkshopQueue* to replicate the limited number of tracks in the locomotive workshop

Create a Work Centre called *Workshop* to replicate the operation of technical inspection of the road locomotives

These attributes should look similar to Figure 7.

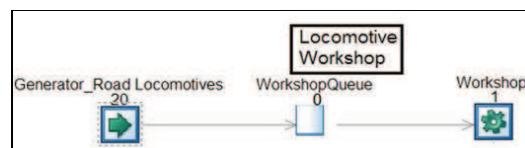


Fig. 7 Locomotive Workshop Layout in SIMUL 8

Road Locomotives Generation

It is assumed that the available road locomotive fleet is sufficient to satisfy the number of planned outbound train compositions but with a random component that is caused by schedule deviations and necessary inspections before ensuring the technical readiness of the arrived road locomotives. If a road locomotive is not in good condition for service, that locomotive is withdrawn and there is no guarantee that there is another road locomotive available to fulfil the haulage. Therefore, ensuring road locomotives is subordinated to stochastic factors. Our effort is not focused on road locomotive scheduling problems. However, one may wish to analyse the yard behaviour under the conditions of: 1) no shortages of road locomotive fleet or 2) shortages of road locomotive fleet

In the case of no shortages of road locomotive fleet, one may simulate with time-dependant fixed inter-arrivals where the number of required road locomotives equals to the number of scheduled outbound freight trains. In the opposite case, one may simulate with time-dependant exponential inter-arrivals. For our objectives the assumption “no shortages of road locomotive fleet” is considered as Status Quo course

Storage Area

In the Storage Area dialogue box (Storage Bin Details) called *WorkshopQueue*, leave the capacity as Infinite. It is assumed that there is sufficient capacity for accommodating road locomotives in the locomotive workshop of Gaia

Work Centre

In the Work Centre Options dialogue box called *Workshop*, set up the Uniform distribution with lower bound 15 and upper bound 35 minutes to replicate the operation of technical inspection of the road locomotive by the engineer

5) Departure Yard

Layout

Create a Work Centre called *Inspections* to replicate the operation of inspection of freight train composition before departure

Create a Work Centre called *Train_Engine Operations* to replicate the operation of arriving and putting road locomotive on freight train as well as the full brake test

Create a Work Centre called *Leaving the Yard* to replicate the physical departures of freight trains from the yard

The attributes replicating the Departure yard should look similar to the example provided in Figure 8.

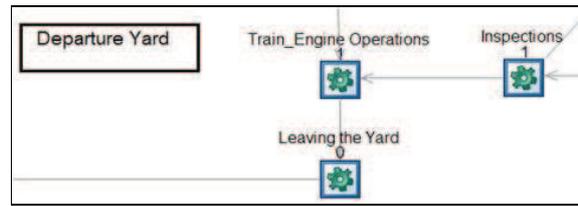


Fig. 8 Departure Yard Layout in SIMUL 8

Work Centres

In the Work Centre Options dialogue box called *Inspections*, set up the Uniform distribution with lower bound 10 and upper bound 15 minutes to replicate the operation of inspection of freight train composition before departure

In the Work Centre Options dialogue box called *Train_Engine Operations*, select the Uniform distribution with lower bound 15 and upper bound 20 minutes to replicate the execution of operations of arriving and putting locomotive on freight train as well as the full brake test. Make sure that from menu “*Routing In*” the discipline “*Collect*” is activated. This is necessary to replicate that the train composition can leave the yard only with a road locomotive

In the Work Centre Options dialogue box called *Leaving the Yard*, set up the Exponential distribution with the mean of 3 minutes to replicate the freight train departure from Gaia yard

6) Dispatch

Layout

Create a Work Exit Point called *Dispatch* to complete the simulation model

This attribute should be as shown in Figure 9.



Fig. 9 Dispatch Layout in SIMUL 8

7) Additional aspects of the yard simulation model

Floating Resources/Creation

Create a Floating Resource to set up shunting men

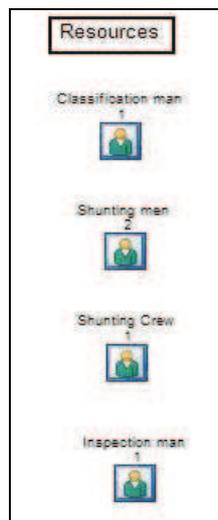


Fig. 10 Yard Floating Resources in SIMUL 8

Create a Floating Resource to set up a classification man
 Create a Floating Resource to set up an inspection man
 Create a Floating Resource to set up a yard crew

The SIMUL 8 attributes that replicate the floating resources in our yard simulation model should look like Figure 10.

Floating Resources/Functions

Set the Floating Resource “*Shunting men*” to do work in the Work Centre called *Cutting*. Number of this type of resources available: 2 (two) men

Set the Floating Resource “*Classification man*” to do work in the Work Centre called *Classification_Technical Inspections*. Number of this type of resources available: 1 (one) man

Set the Floating Resource “*Inspection man*” to do work in the Work Centres called *Classification_Technical Inspections* and *Inspections*. Number of this type of resources available: 1 (one) man

Set the Floating Resource “*Yard crew*” to do work in the Work Centres called *Further Operations*, *Breaking Down* and *Making Up*. Number of this type of resources available: 1 (one) crew

Interruptions

According to the operational regulation of CP yards, it is written that the shunting work must stop when any freight train is arriving in or leaving from the yard and can be restarted only when the arriving freight train has terminated its move completely or the leaving freight train has already left the yard limits. Therefore, we set up interruptions per arriving/leaving freight train at the Work Centres called *Breaking Down* and *Making Up*. This safety pause (in SIMUL 8 the term: “time of repair” is used) is replicated by Exponential distribution with the mean of 2 minutes

Warm-Up Period

In order to avoid the transient behaviour during the simulation of yard performance, we set up a Warm-Up Period of 1440 minutes (24 hours), meaning the results collection starts after this number of time units

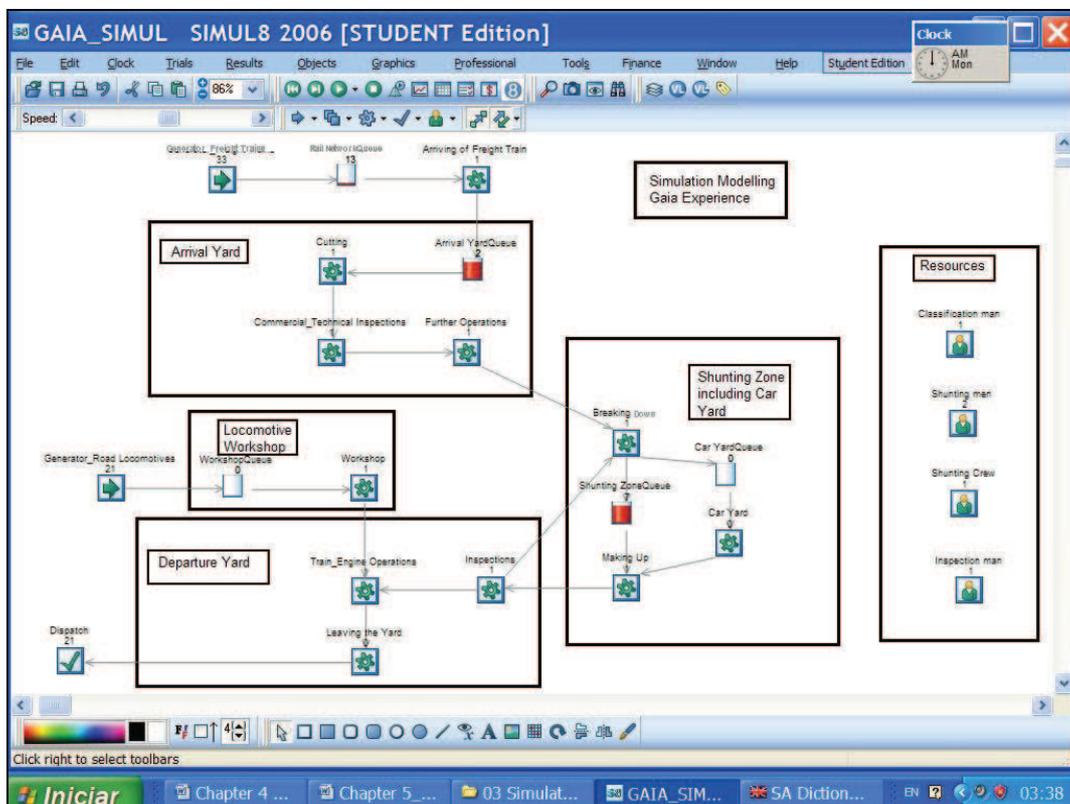


Fig 11 A SIMUL 8’ Animation Window in Modelling the Operating Processes with Freight Trains at Gaia Flat-shunted yard

A complete screenshot of the animation window of SIMUL 8 in simulating the operating processes with freight trains at Gaia flat-shunted yard can look similar to Figure 11. Note that the technological subsystems are properly named. The routing through technological subsystems is demonstrated by arrows and following the arrows' courses one can easily recognize the throughput line of the yard. The technological subsystems are grouped in a way that clearly shows which subsystem belongs to which yard area, how many they are and whether or not there is a storage area. Note that there is no Storage Area in Departure Yard of Gaia. This is because the making up of freight train and all the subsequent operations on the departure of freight train are fulfilled on the same track, meaning the tracks used for making up and departure operations overlap. This phenomenon requires no Storage Area in Gaia departure yard. The created floating resources that do jobs in some of the technological subsystems are grouped on the left (figure's projection - under the title "Resources"). As was mentioned, we are able to obtain and observe the levels of their utilization as well as other meaningful measures of yard performance under different courses and simulation trials, discussion of which begins below.

6.4 Results obtained by Simulation (General). In the case of strict fixed schedule the scheduled freight trains over the regular daily cycle ought to arrive at Gaia according to the pattern shown in Figure 12, below. The arrival pattern in Status Quo course of Gaia yard is then a time dependent distribution, which is shown in Table 1, below.

However, there are delays. There are cancellations and extra freight trains every day. What is scheduled is not fulfilled. CP Carga, the rail freight operator under study, encounters difficulties to fulfil its schedules. CP Carga formally tends towards scheduled operations. The practice shows something different, however. To some extent, CP Carga is dealing with an improvised uncontrolled freight train operation. The induced uncertainty of the freight train movement over the network is intimately connected with the incurred huge amount of long term average costs. This situation causes the company to experience a significant level of waste.

Fig. 12 Defined Time Dependent Distribution of Scheduled Arrivals of Freight Trains at Gaia

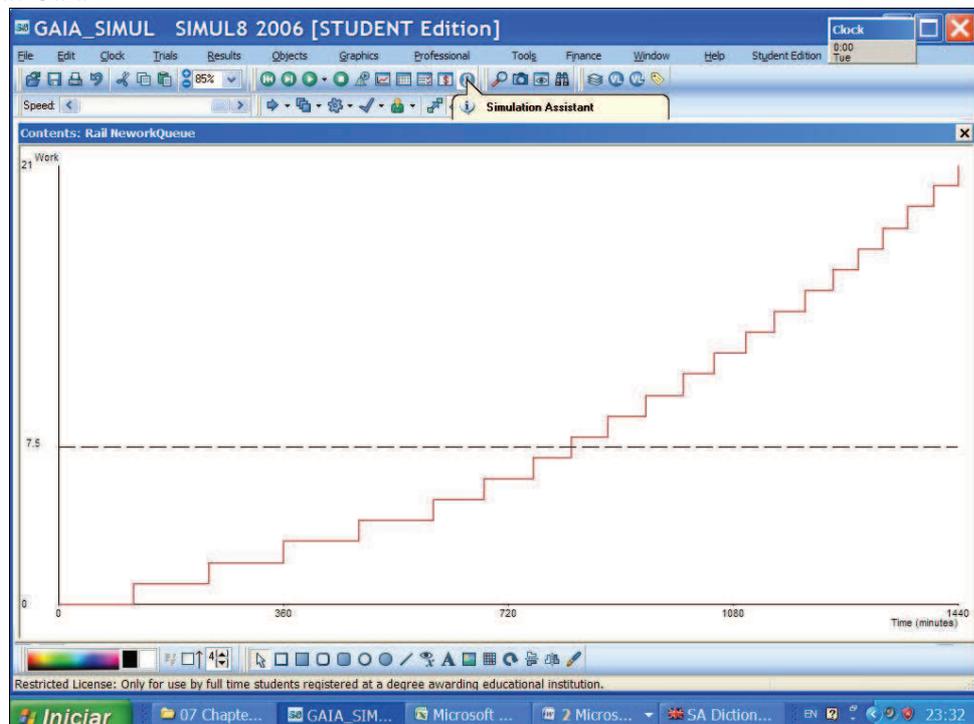


Table 1 Time-Dependant Freight Train Arrivals, Gaia Experience
Source: Dprt. of Planning and Control, CP Carga

Time Dependent Distribution	
Periods	Freight Train Arrivals
From 0:00 to 04:00	2
From 04:00 to 08:00	2
From 08:00 to 12:00	3
From 12:00 to 16:00	4
From 16:00 to 20:00	5
From 20:00 to 24:00	6

For example, in Figure 13, one observes randomly selected real world runs of Freight train No 51031 as well as its planned schedule. This freight train is scheduled to run at night from Pampilhosa to Gaia yards, a distance of 101 kilometres. The scheduled departure time of freight train No 51031 is 2:43 a.m. and the scheduled arrival time at Gaia is 3:56 a.m. The curve with bold and circlets shows the scheduled movement of the train. All other curves show movements out the schedule. It is beyond the scope of this study to analyse in detail the delays of CP Carga freight trains as well as the reasons of frequent occurrence of those delays. However, these delays are of importance for the accuracy of the yard simulation model. Needless to say, if the freight trains' arrivals are subordinated to a stochastic process, it is senseless to employ deterministic arrivals during the simulation as the model would not replicate real world performances. Simulating with deterministic arrivals is worth it if one aims at analysing yard behaviour under the conditions of strict fixed schedule. Therefore, for the objectives of this study we employ both time-dependant deterministic inter-arrivals and time-dependant stochastic inter-arrivals. The time-dependent deterministic inter-arrivals are replicated by Fixed distribution. The time dependant stochastic inter-arrivals are subordinated to Exponential distribution.

Fig. 13 Deviations from Schedule of Freight Train No 51031 Scheduled to run at Night from Pampilhosa to Gaia Yards

Source: "TrainOffice" – this sample was collected in February, 2005

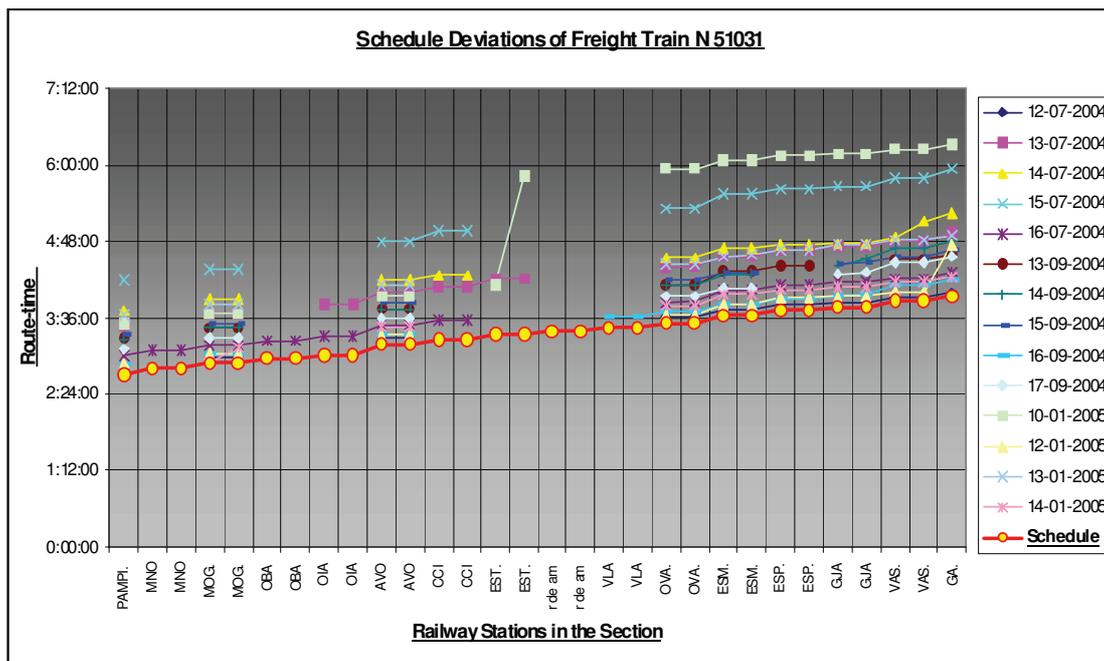


Table 2 Comparison of Two Trials through Measures of Yard Performance, Gaia Experience

Measures of Gaia Yard Performances, Deterministic Inter-arrivals vs. Stochastic Inter-arrivals				
Simulation Object (Attribute)	Performance Measures	Measure Units	Deterministic Inter-arrivals	Exponential Inter-arrivals
-	-	-	Average	Average
Generator_Freight Trains	Number Entered	Number of Trains	21	21.14
Rail NetworkQueue	Average Queue Size	Number of Trains	0	0.09
	Average Queueing Time	Minutes	0	4.95
Arrival YardQueue	Average Queue Size	Number of Trains	0	0.18
	Average Queueing Time	Minutes	0	11.29
Cutting	Average Use	-	0.03	0.23
	Waiting	Percent (%)	95.90	76.66
	Working	Percent (%)	2.19	2.17
	Blocked	Percent (%)	1.91	21.17
	Stopped	Percent (%)	0	0
Commercial_Technical Inspection	Average Use	-	0.27	0.35
	Waiting	Percent (%)	74.33	65.29
	Working	Percent (%)	25.54	25.21
	Blocked	Percent (%)	0.13	9.50
	Stopped	Percent (%)	0	0
Further Operations	Average Use	-	0.021	0.08
Breaking Down	Average Use	-	0.31	0.49
	Waiting	Percent (%)	67.44	50.59
	Working	Percent (%)	29.62	41.11
	Blocked	Percent (%)	0	5.35
	Stopped	Percent (%)	2.94	2.94
Shunting ZoneQueue	Average Queue Size	Number of Trains	0.012	1.60
	Average Queueing Time	Minutes	0.76	71.29
Making Up	Average Use	-	0.17	0.29
	Waiting	Percent (%)	80.47	68.12
	Working	Percent (%)	11.08	15.12
	Blocked	Percent (%)	5.51	13.82
	Stopped	Percent (%)	2.94	2.93
Inspections	Average Use	-	0.18	0.41
	Waiting	Percent (%)	81.5	59.26
	Working	Percent (%)	18.49	25.22
	Blocked	Percent (%)	0.002	15.52
	Stopped	Percent (%)	0	0
Train_Engine Operations	Average Use	-	0.75	0.94
	Waiting	Percent (%)	74.47	75.62
	Working	Percent (%)	25.52	24.38
	Blocked	Percent (%)	0	0
	Stopped	Percent (%)	0	0
Leaving the Yard	Average Use	-	0.036	0.049
	Waiting	Percent (%)	96.35	95.12
	Working	Percent (%)	3.65	4.88
	Blocked	Percent (%)	0	0
	Stopped	Percent (%)	0	0
Generator_Road Locomotives	Number Entered	Number of Locomotives	21	21.33
Dispatch	Average Time in System	Minutes	94.97	454.95
	Minimum Time in System	Minutes	81.35	165.21
	Maximum Time in System	Minutes	119.15	845.05
	St. Deviation of	-	10.30	204.96
Classification Man	Utilisation	Percent (%)	25.54	25.21
Shunting Crew	Utilisation	Percent (%)	44.34	60.10
Inspection Man	Utilisation	Percent (%)	44.03	50.43

Furthermore, if the freight train arrivals are subordinated to a stochastic process, then the generation of road locomotives is also subordinated to stochastic process, and therefore we set up

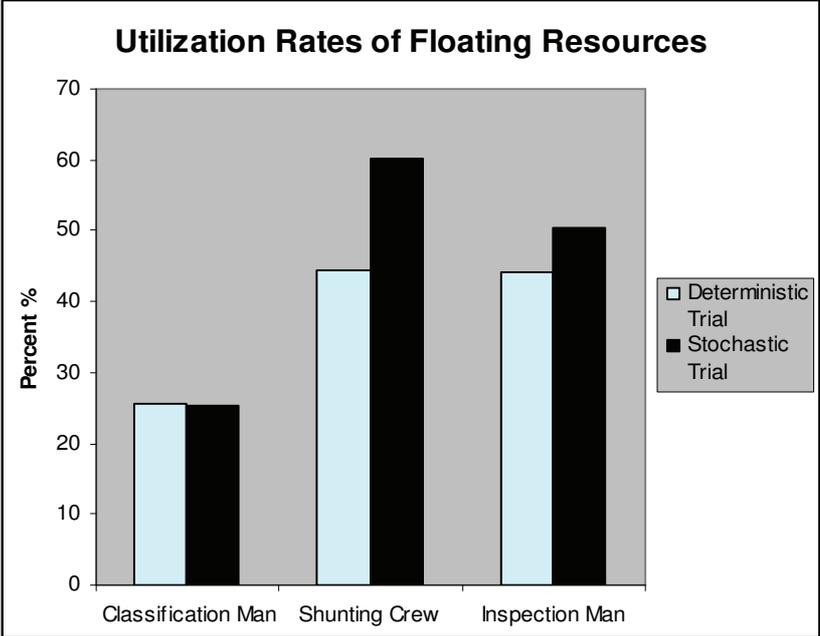
time-dependent Exponential inter-arrivals for “ensuring road locomotives” as well. Thus, two trials have been conducted as every trial consists of 1000 runs of 5 days per week in each run. For simplicity we call them Deterministic Trial and Stochastic Trial, respectively.

Note that this study is not on optimisation and therefore we employ simulation trials of 1000 (thousand) runs because our particular objective is to analyse the global impact on the yard performances under different circumstances as we are searching for a good set of rules leading to adequate conclusions for the tactical decision-making purposes.

Thus, results obtained for measures of Gaia yard performances are presented in Table 2. At first glance, the significant difference is seen in the estimated Average Time in System, which has the same meaning as Average Expected Throughput Time. The Deterministic trial shows that the estimated average throughput time comes up to approximately 1 hour and 20 minutes. The level of variability in the arrival and service processes in the Deterministic trial is relatively low, which has a positive effect seen in a reduction of average time in queue, and hence one expects to obtain lower values of throughput times. The Stochastic trial, however, shows Average Time in System up to 6 hours. This may seem a bit shocking. In order to interpret this result we should look at the measures of Gaia subsystems’ performances observed in the two simulation trials.

The utilization rates of the floating resources, i.e., Classification man, Shunting Crew and Inspection man are shown in Chart 1. One observes that the utilization levels of Shunting Crew and Inspection man in the Stochastic trial are clearly higher than those in the Deterministic trial. This indicates that these dynamic resources were busier in the Stochastic trial. However, note that the utilization rate of the classification man remains unchanged, i.e. approximately 25 %. If the estimated utilization of the classification man is the same in the two trials, it means that this employee performed the same amount of work in either Deterministic or Stochastic trial.

Chart 1 Effect of Arrival Process on Utilization Levels of Floating Resources, Gaia Experience



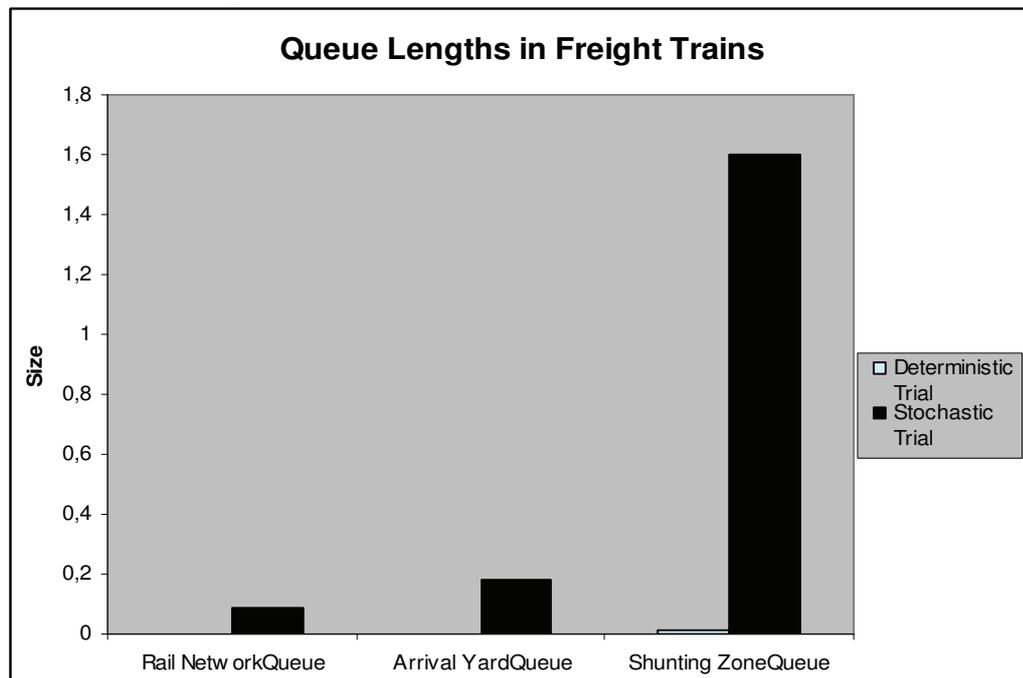
What does the increases in the utilization levels of Shunting Crew and Inspection men? It might be an increase of the number of freight train arrivals. That seems not to be the present case, however, because we do not observe differences on average at neither the number of arrivals nor the number of completed freight trains. Thus, the amount of work remains unchanged. In this situation, the utilization increases can be result from the following two reasons:

- 1) The freight trains required an extraordinary further processing to be fulfilled only by the shunting crew and the inspection man;
- 2) The variability in the arrival process.

One should not forget that this is a simulation experiment and its characteristics are controlled, to a certain extent. We did consider 2 % re-breaking down, which is not the reason for the observed increases in the utilization rates. Consequently, this at first look plausible increase in the utilization rates of Shunting Crew and Inspection man is caused by the significant variation of the freight trains' arrivals. The variability in the arrival process might increase the utilization of the yard resources but this phenomenon has a negative effect on the entire yard operation seen in the significant increases of queue lengths. Increases in queue length cause oversaturation in the yards and the yard personnel experience difficulties in execution of the yard operations. Such a situation is frequently also referred to as "*the system is blocked*". Uncontrolled improvised freight train movement requires a big size of the buffers in order to deal with the variability of freight train arrivals as well as departures. In terms of railway freight network operation the buffers are the yards.

In Chart 2, queue lengths observed in the two trials are given. The difference in the size of queue lengths between the two trials is apparent, especially the size of queue length in the shunting zones. Note that in the stochastic trial the average time in the shunting zone queue equals to approximately 71 minutes per freight train. This further indicates the frequent occurrence of blocking caused by the phenomenon that the whole set of subsequent yard operations is fulfilled on few tracks. In this case, searching for improvements by adding dynamic resources is senseless.

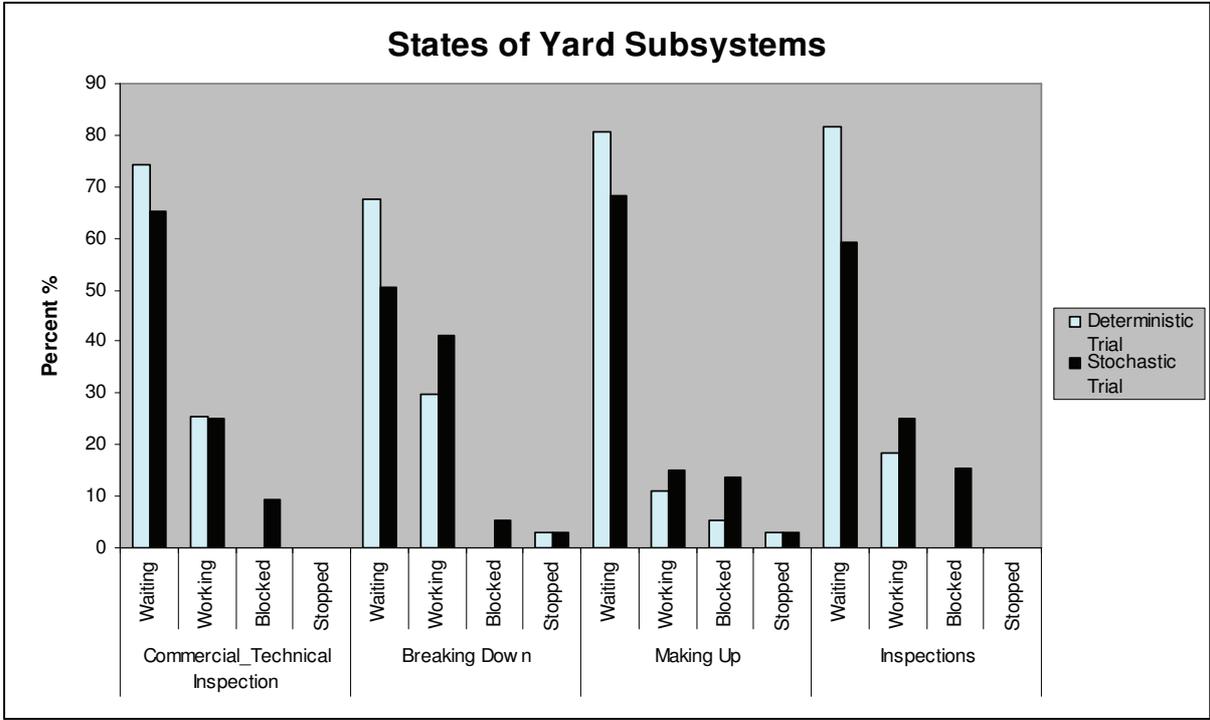
Chart 2 Comparisons of Queue Length Sizes, Gaia Experience



In SIMUL 8 environment, one is able to analyse the performance of the system under study through its four states during the simulation experiment. These states are: *Awaiting work*; *Working*; *Blocked*; and *Stopped*. Chart 3 makes comparisons between Deterministic trial and Stochastic trial as the figures are the states of Gaia subsystems in percentages. Note that the Gaia subsystems in the Stochastic trial experience higher percentage of time for which they were blocked. Moreover, note that these subsystems also experience higher percentage of time for which they were working. It is relative to the higher utilization rates of the shunting crew and inspection man, but in either case the

work processed remains the same and one bears witness only to increases in operating and waiting times (respect. operating and waiting costs incurred). In other words, these are undesirable increases in the variable costs for the company. The variable costs can be controlled by expedient planning and precise operation. They must be kept low, to the extent possible.

Chart 3 Percent of Time for States of Flat-shunted Yard Subsystems, Gaia Experience



On the other hand, in the Deterministic trial one observes that Gaia subsystems were not blocked. This phenomenon has occurred only in the technological subsystem dedicated to Making up of freight trains. This subsystem was blocked approximately 5% of the experimental time, which is caused by the fact that there are periods in which Gaia yard encounters difficulties to seamlessly accommodate and process the scheduled inbound freight trains. Most likely, this is the period from 20:00 to 24:00 in which 6 freight trains have arrived. It appears that this issue requires further examination also because of the fact that in the Deterministic trial Gaia subsystems demonstrate significantly high values of being in a state of “awaiting work”, which indicates ample idle times. The idle times entail high capital costs for the company and must be avoided.

6.5 Results obtained by Simulation (Changes in Arrivals/Traffic Rules). In the following section we examine Gaia yard performance through changes in the number of freight train arrivals (i.e., through changes in the traffic rules). It was argued above that in the case of few yard tracks no great improvements can be achieved by adding dynamic resources, meaning adding another yard crew or another shunting locomotive. The effect of such a decision will not be positive, just on the contrary. Therefore, we do not focus our effort on examining the effect on changes of dynamic yard resources. We wish to explore what is the maximum standing capacity of this yard subject to the utilization rates of the dynamic resources.

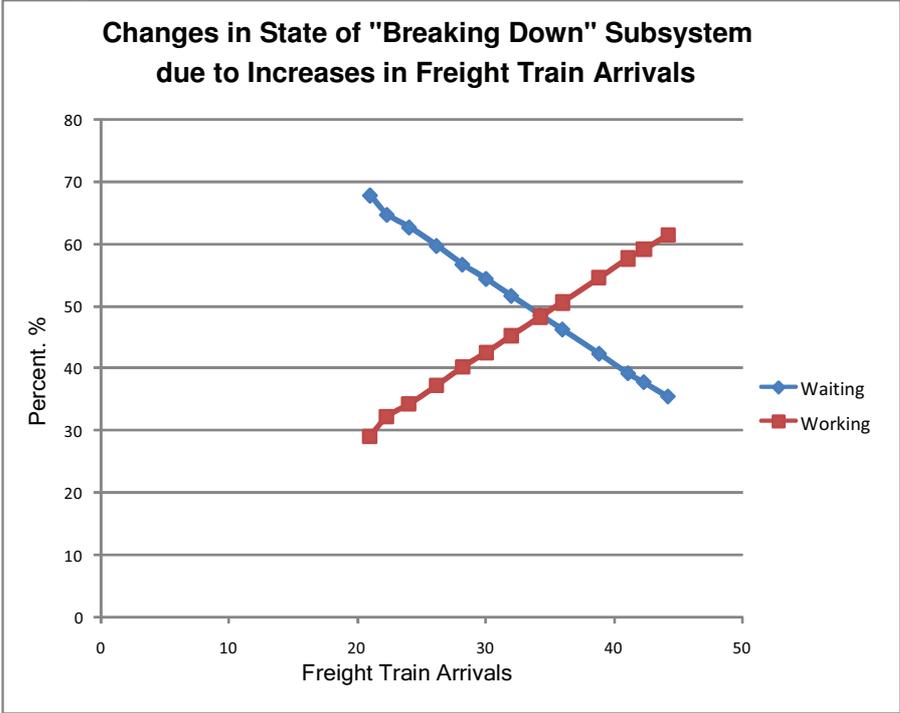
In this exploration, the set-up characteristics of the simulation yard model are as suggested in Table 3. Note that either Uniform or Exponential distributions for subsystems Cutting, Further Operations and Leaving the Yard are suggested. After having conducted a set of trials, the two distributions provide identical results for the performances of the three yard subsystems in question. For the purposes of this experiment Exponential distribution is employed. Thus, Gaia yard

performances versus changes in freight train arrivals are explored. Chart 4 examines the effect of increases in the number of freight trains arrivals on the state of the yard subsystem dedicated to Breaking down of freight trains. It verifies that increases in the freight trains to be processed results in increases in subsystem being in state of “working” and proportionally decreases in subsystem being in state of “awaiting work”.

Table 3 Possible Distributions for Yard Simulation Modelling Purposes

Distributions and Arrival Patterns for Yard Technological Subsystems		
Simulation Object (Attribute)	Dimension	Distribution
Generator_Freight Trains	Time-dependent Inter-arrivals	Uniform or Fixed
Arriving of Freight Trains	Service Times	Uniform or Fixed
Cutting	Service Times	Uniform or Exponential
Commercial_Technical Inspections	Service Times	Uniform
Further Operations	Service Times	Uniform or Exponential
Breaking Down	Service Times	Uniform
Making Up	Service Times	Uniform
Inspections	Service Times	Uniform
Train_Engine Operations	Service Times	Uniform
Leaving the Yard	Service Times	Uniform or Exponential
Generator_Road Locomotives	Time-dependent Inter-arrivals	Uniform or Fixed
Workshop	Service Times	Uniform

Chart 4 Effect of Increasing Freight Train Arrivals on State of Gaia Subsystem, Gaia Experience



Next, Chart 5 shows how the estimated average throughput time per freight trains increases with increases in the freight train arrivals. Furthermore, with increases in the arrivals it is expected that queues start to materialize. In terms of Shunting Zone Queue of Gaia this phenomenon is shown in Chart 6, where one observes increases in estimated values of average queueing time due to freight trains’ arrivals. Note that between 30 and 40 freight train arrivals, we start observing evident jumps in the estimated average queueing times.

Chart 5 Effect of Increasing Freight Train Arrivals on Estimated Throughput Times, Gaia Experience

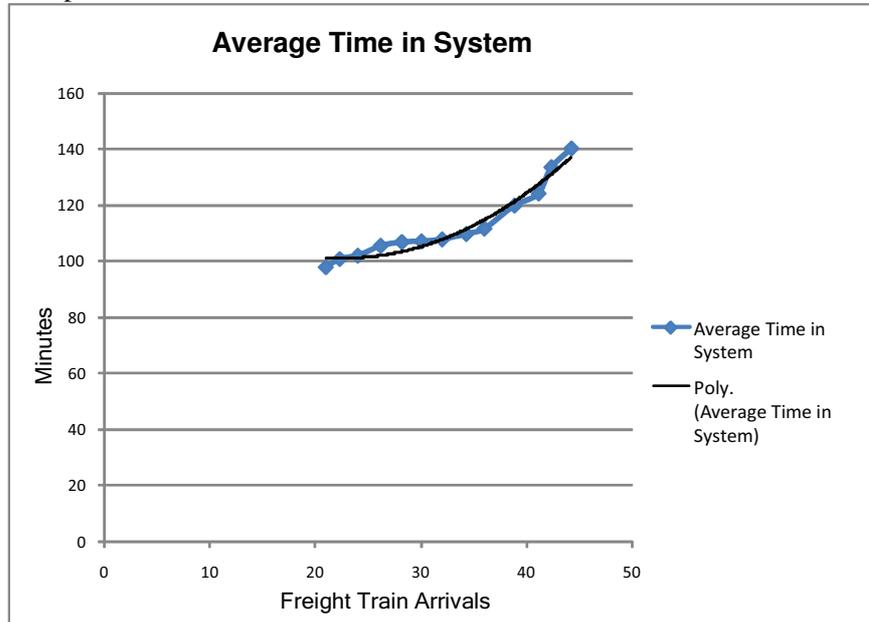
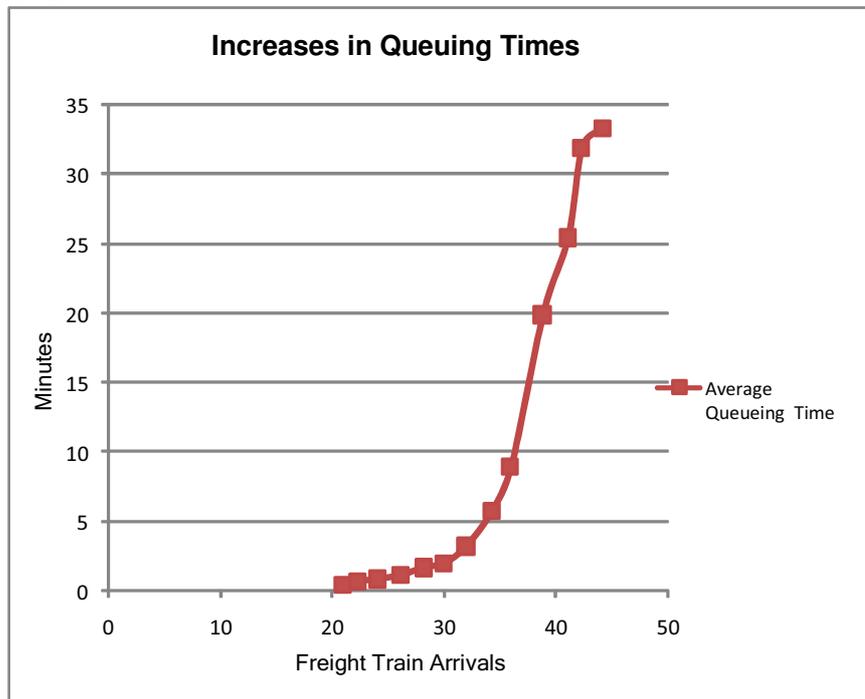


Chart 6 Increases in Queueing Times of Gaia Shunting Zone with Increases in Freight Train Arrivals



In order to experience a seamless-low-cost yard operation one better implements the rules of “*pull systems*” (see e.g., Hopp and Spearman 2001, 2004) and keeps the yard workload (just) below the threshold in which the queues apparently start to build up by controlling the number of inbound freight trains to be processed (i.e., the input that requires service by the yard). Thus, the freight trains will move through the yard subsystems “*unimpeded*”. The created simulation yard model in terms of Gaia Flat-shunted yard suggests us to specify an upper bound of the yard processing

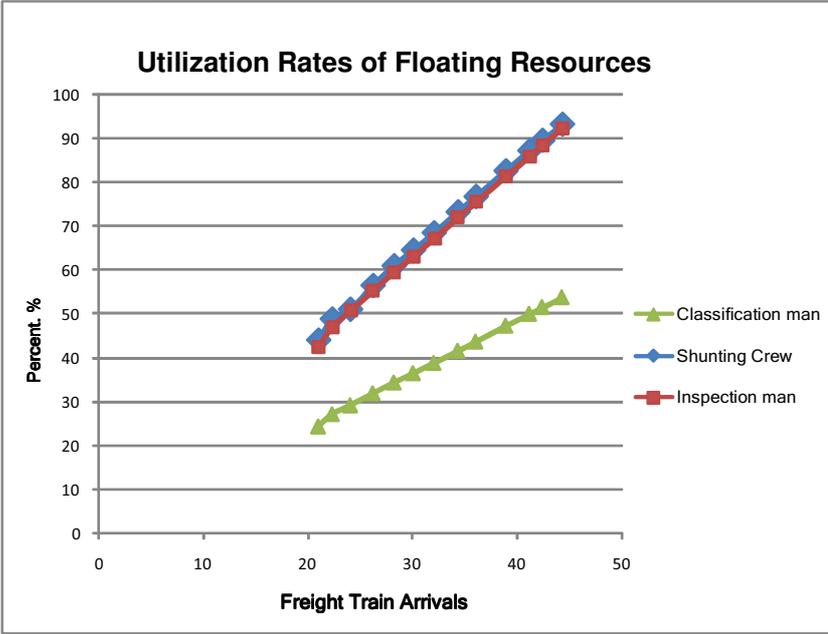
capability up to 33- 34 freight trains, before queues to start materializing progressively. Of course such an exercise should be done only when a steady low-variation scheduled railway freight operation is instituted.

Recalling the importance of the utilization levels of yard resources, we drew a chart in order to show the effect of increasing freight train arrivals on the utilization rates of Gaia floating resources. This is Chart 7 below, where one observes the utilization rate increases.

Note that the utilization rate of classification man is relatively low in comparison with the utilization rates of the shunting crew and the inspection man. At the optimal level of inbound trains (i.e., 33- 34) the utilization levels of these two resources fall within 69- 72%, which is very much in agreement with observations in the previous stochastic simulation trial, just the difference is that the number of freight trains performed during the stochastic trials are significantly lower (i.e., 21- 23).

Recalling the low utilization level of classification man, this would be of interest to rethink the work characteristics and obligations of this employee category. The classification man may execute additional works in aid of shunting crew as well as inspection man.

Chart 7 Effect of Increasing Freight Trains Arrivals on Utilization Rates of Floating Resources, Gaia Experience



7. Synthesis and Conclusions

A simulation modelling methodology for evaluation of flat-shunted yard performances is presented. A detailed discussion on how this methodology is implemented in terms of a “real world” case is provided. In implementing the here-proposed methodology one creates a visual simulation model using an event-based simulation language, SIMUL 8, in order to replicate the flat-shunted yard behaviour on the basis of which runs simulation trials with the purpose of analysing and evaluating the throughput level of the yard through a set of measures of performance (MOPs). Special attention to each attribute of the model is paid. The arrival process of freight trains is replicated by Time-dependent distributions. A SIMUL 8’ Animation Window in Modelling the Operating Processes with Freight Trains at Gaia is shown in Figure 11. A number of simulation trials have been conducted. Scheduled (disciplined) arrival patterns vs. Improvised arrival patterns for both freight trains and road locomotives have been examined. At each stage the advantages of the scheduled arrivals are shown. Also, the performances of individual yard subsystems are explored. Lastly, after having assumed disciplined arrival patterns (trains as well as locomotives) changes in the arrivals

are scrutinized. In general, the potential of the provided yard simulation model in SIMUL 8 environment is seen in Table 2 and Charts 1- 7. More specifically, Table 2 and Charts 1– 3 provide comparisons of the results obtained by different trials conducted through measures of yard performances. E.g., Chart 3 compares the percents of time for which yard subsystems were either working, awaiting work, blocked and stopped. On the other hand, Charts 4– 7 examine the effect of increasing freight train arrivals on the yard queues and utilization of yard floating resources. These charts provide important relationships involving traffic rules and yard performances. They can be used for adequate decision-making aiming at “*maximum production*” and “*optimal utilization level*” of yard resources “*subject to yard queues*”.

Acknowledgments: The authors wish to acknowledge the planning staff at Department of Operations and Control, CP Carga and the operational staff at Gaia yard for their support and contribution to this research. The authors also wish to thank the anonymous reviewers for generously giving their time to referee the manuscript of this paper and provide valuable comments and recommendations.

References

- Cooper R** (1981) Introduction to Queuing Theory – second edition, Florida Atlantic University, Elsevier North Holland, Inc
- Dessouky M** and **Leachman R** (1995) A Simulation Modeling Methodology for Analyzing Large Complex Rail Network, Simulation, Vol. 65, No. 2
- Gross D** and **Harris C** (1985) Fundamentals of Queuing Theory, New York: John Wiley and Sons, Inc,
- Hall R** (1991) Queuing Methods, for Services and Manufacturing, University of California at Berkeley, Prentice Hall International, Inc.
- Harrod S** (2003) Queensgate Yard - Arena Independent Project, QA 585
- Hopp W** and **Spearman M** (2001) Factory Physics: Foundations of Manufacturing Management, 2nd Edition, McGraw-Hill, Boston
- Hopp W** and **Spearman M** (2004) Commissioned Paper: “To Pull or Not to Pull: What is the Question?”, Manufacturing & Service Operations Management, Vol. 6, No. 2
- Ivanov K** (2005) Technological Process Projection in serving an Industrial Branch by Rail Operator, Master Thesis, VTU “T. Kableshkov”, Sofia, Bulgaria (written in Bulgarian)
- Jackson J** (1963) Jobshop-Like Queuing Systems, Management Science, Vol. 10
- Karagyozov K** (1983) Simulation Model of Marshalling Yard, Conference of an Applied Science, about Implementation of Electronic Computation Machines and Micro-Processors in Railway Transport, Varna, Bulgaria 21-22 (written in Bulgarian)

Karagyozev K (1990a) Approximate Method for Modelling of Open Queuing Network, International Conference Operations Research, University of Economics and Business Administration, Vienna, 28-31

Karagyozev K (1990b) On the Approximations of Moments and Operational Characteristics of GI/G/S Queues, International Conference Operations Research, University of Economics and Business Administration, Vienna, 28-31

Katchaunov T, Karagyozev K, Kuppenov D and Razmov T (1998) Simulation Modelling of Transport Processes, (Simulations with GPSS), VTU "Todor Kableshkov", Sofia, Bulgaria (written in Bulgarian)

Kavicka A (2000) A Simulation Support for Railway Infrastructure Design and Planning Processes, 7th International Conference on Computers in Railways 2000, Bologna, Italy

Klima V (1997) Virtual Railway Marshalling Yard, Transportation Systems 97, Vol. 1-3

Klima V (2001) Software Tool VIRTUOS - Simulation of railway station operation, 15th European Simulation Multiconference (ESM 2001) June 6-9, Prague, Czech Republic

Konstantinov D (1969) Manual for Technology and Projection of Rail Stations and Junctions, VMEI, Sofia, Bulgaria (written in Bulgarian)

Kraft E (2002) The Yard: railroading's hidden half. In: Trains, Vol. 62, Part I: VI/02, and Part II: VII/02

Kumar S (2004) Improvement of Railroad Yard Operations, Handbook of transportation engineering, Chapter 25, Myer Kutz – Editor, President, Myer Kutz Associates, Inc. ISBN 0-070139122-3, Copyright by The McGraw-Hill Companies, Inc.

Lee A (1968) Applied Queuing Theory, Operational Research, Air Canada

Marinov M (2007) Analysis and Evaluation of Formation Yard Performances, PhD Thesis, Instituto Superior Tecnico at Technical University of Lisbon, Portugal

Nadel R and Rover E (1967) The Use of a Computer Simulation Model for Classification Yard Design, PROC, New York

Pachl J and White T (2003) Efficiency through Integrated Planning and Operation, International Heavy Haul Association, Virginia Beach

Petersen E (1977a) Railyard Modeling: Part I. Prediction of Put-through Time, Transportation Science, Vol. 11

Petersen E (1977b) Railyard Modeling: Part II. The Effect of Yard Facilities on Congestion, Transportation Science, Vol. 11

Raikov R (1986) Organization of Shunting and Train Processes in Railway Stations, VTU "Todor Kableshkov", Sofia, Bulgaria (written in Bulgarian)

Razmov T (2004) “Forecasting and Managing the Processing of Car Flows over the Railway Network”, Ph.D. Thesis, approved by Scientific Council on “Mechanical Technologies and Transport”, Sofia, Bulgaria (written in Bulgarian)

Rhodes M (1988) *The Illustrated History of British Marshalling Yards*, Sparkford: Haynes Oxford Publishing & Co

Rhodes M (2003) *North American Railyards*, St. Paul (USA): Motorbooks International, MBI Publishing Company

Shalliker J and Ricketts Ch (1997) *An Introduction to SIMUL 8*, School of Mathematics and Statistics, University of Plymouth

Shore H (1988a) Simple Approximations for GI/G/c Queue-I: The steady state probabilities, *J. Operational Research Society*, Vol. 39, No. 3

Shore H (1988b) Simple Approximations for GI/G/c Queue-II: The moments, the inverse distribution function and the loss function of the number in system and queue delay, *J. Operational Research Society*, Vol. 39, No. 4

Takus D and Profozich D (1997) *ARENA® Software Tutorial*, Proceedings of the 1997 Winter Simulation Conference, ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson

Tasev Y (1969) Technical Indices for Projection of Track Development, *Journal of Railway Transport*, Vol. 11, Sofia, Bulgaria (written in Bulgarian)

Tasev Y and Karagyosov K (1983) *Manual for Technology and Projection of Railway Stations and Junctions*, VNVTU “T. Kableshkov”, Sofia, Bulgaria (written in Bulgarian)

Wegner R (2003) Classification Yards, Map of the Month, In: *Trains IV/2003*

White T (2003a) *Elements of Train Dispatching, Volume 1 – History and Overview of Train Dispatching and Railroad Operation*, VTD Rail Publishing

Whitt W (1983) The Queuing Network Analyzer, *Bell System Technical Journal*, Vol. 62, No. 9