Australasian Transport Research Forum 2016 Proceedings
16 – 18 November 2016, Melbourne, Australia
Publication website: http://www.atrf.info

Simulation-Optimization model applied on a closed loop railroad transportation system

Pedro Camargo[1], Claudio Cunha[2]

[1]Veitch Lister Consulting, Australia

[2]Politechnic School of the University of São Paulo, Brazil

Email for correspondence: c@margo.co

Abstract

Railroad transportation is one of the most efficient means used to move the Brazilian grains production to export ports. However, grain production has recently been growing at a faster rate than the railway infrastructure, leading to several operational restrictions for this mode of transportation. Since new investments take long to be completed, railroad operators seek to improve operational efficiency in order to ensure better utilization of existing resources.

In this context, this paper examines some operational strategies that could be used to achieve increased capacity for a grain rail transportation system operating in a closed loop, by means of a simulation-optimization model.

The proposed model consists of a traditional stochastic simulation model in which all decisions (assets distributions and queue management) were replaced by prioritization rules (very simple heuristics).

As there were tested as many as 10 possible prioritization rules for each one of the eight stochastic processes replaced, it became a combinatorial problem to find the best set of prioritization rules, which led us to apply a procedure analogous to the Hooke & Jeeves algorithm (Hooke & Jeeves 1961).

We provide the details of the whole modelling process from the developing of the stochastic simulation model used to characterize the system, including some very promising results. However, an especial attention is given to the prioritization rules that were incorporated into the model (as well as how they were matched together) and to the Hooke & Jeeves analogy used to search for the best set of prioritization rules for the system analysed.

1. Introduction

This paper describes a hybrid simulation-optimization model aimed at analysing the operational rules used in the transport of soy and corn in the FCA-EFVM rail system, South-East region of Brazil.

Moreover, the proposed model, implemented in a general programming language, is designed to allow for the analysis of different rules in treating queues in order to maximize the maximum system throughput from the unloading in the rail stations until the transfer to ships in the sea port.

Intermediary stages of the process such as train transit between stations, loading and unloading of rail cars and return of empty rail cars to the loading stations are also treated explicitly due to its stochastic nature, although no alternative operational rules are analysed for all of them.

A total of ten operations had their operational rules, usually First-in-First-Out (FIFO), replaced by prioritization rules that aim to reproduce the decisions that the system's operators are able to make when deciding how to treat each element in the queue. Due to the combinatorial nature of the problem, an heuristic procedure based on the Hooke & Jeeves algorithm is proposed to search for the best set of operational rules to be applied to each one of these ten decision points.

This paper is organized in five sections. The first section presents a brief literature review on research-optimization models applied to the problem in hand, while the following two sections provide an overview of the transport system being modelled and the proposed methodology. The last two sections present the computational experiments performed and the conclusions enabled by the model results.

2. Literature Review

Simulation-Optimization models are defined as models with elements of both mathematical programming (optimization), usually an objective function that needs to be maximized, and stochastic simulation, such as the presence of stochastic behaviour for part or all elements of the system. The definition, however, engulfs a large number of model types, as these two modelling frameworks can be articulated in a variety of fashions.

(Jacobson & Schruben 1989) classify simulation-optimization models as stochastic optimization, and state that this is a particularly difficult class of models to be solved. (Pflug 2008) provides an alternative interpretation to the problem, stating that this type of problem is characterized by the non-availability of complete information about the system's state at the time of each decision.

With this definition, (Pflug 2008) acknowledges that problems where variables have behaviour inherently stochastic become similar to problems where there is simple lack of information about the system's state.

One of the techniques used to integrate simulation and optimization models described by (Pflug 2008) is to replace trivial rules found in simulation systems such as FIFO and LIFO (last in, first out) to deal with queues and random selection of elements with heuristic procedures. Although not too often explored in the literature, this technique is perfectly adequate for analysing problems such as production sequencing.

This exact approach was adopted by (Everett 2001) in the problem of iron ore transport sequencing where the objective was to reach the exact quality measures in each pile formed in the port before loading. In his work, the author dealt with a problem where several mines, located at varied distances from the port were producing ores in with different chemical compositions, which allowed for a series of decisions to be made:

- Decision on which portion of each mine to be explored
- Decision on which train to dispatch first considering all possible origins
- Decision on which ore pile to unload the train at once in the port
- Decision on which pile to load in each ship

The similarities between the problem tackled by (Everett 2001) and the problem analysed in this research were many, but there was no particular focus on the closed-loop aspect of the problem, which imposes strict constrains on the delays in unloading rail cars that are needed elsewhere in the system.

This aspect of the problem is well analysed, however, by (Fioroni 2008), who modelled a close-loop rail system in a purely stochastic framework and attached the model to an optimization engine charged with altering model parameters to maximize the system's throughput. (Fioroni 2008), however, concluded that the utilization of random draw of decision parameters and queue treatments were sub-optimal, which is the exact problem tackled in this research.

3. The FCA-EFVM-Tubarão grain transportation system

The system being modelled is formed by three major components, two railroads and one port, as illustrated on Figure 1.

Figure 1: FCA-EFVM-Tubarão system

The first railroad, the *Ferrovia Centro-Atlântica* (FCA) is a low performance railway system with mostly single tracks that extends for over 6,500 km, mostly in the South-East portion of Brazil, mostly built in the 19th and 20th centuries and capable only of supporting low weight railcars and low powered engines.

Although fairly antiquated, the portion of FCA that transports grains was upgraded to support railcars with up to 25 tons/axle, although the size of the compositions remain highly constrained by short crossing yards and an unfavourable power/weight ratio varying from 1.03 ton/HP to 0.42 ton/HP

The second railroad in the system, the *Estrada de Ferro Vitória a Minas* (EFVM) extends for 570km in the states of *Minas Gerais* and *Espírito Santo* and is composed mostly by double tracks and a power/weight ratio of 1.96 ton/HP. As this railroad was built to transport in the excess of 100 million tons of Iron Ore per year, its design and operational standards are much higher than those for FCA, and the compositions are as long as 240 railcars.

It is important to highlight that in the junction of the two railroads there is a large classification railyard commonly referred to as intermediary station, where trains are reclassified in order to form trains adequate for transit in each of the railroads and where final train dispatch is formalized.

The third element in this system is the Port of *Tubarão*, which was the Brazilian port with highest throughput in 2007 in terms of mass, with over 104 million tons of total movements.

The port has two terminals dedicated to the iron ore operation and a third terminal that concentrates the operations of agricultural grains (exported) and fertilizers (imported).

The unloading of grains is done in two separate structures capable of unloading 19 cars at a time, taking around 190 minutes per set of cars, of which only 98 minutes are effectively unloading and the remaining 92 minutes are taken by issues im placing the railcars in place and operational problems (~52% utilization rate).

These operational parameters result in a capacity of 520 thousand[1] tons per month, according to the operator of the system.

The port also has 9 different silos with capacities varying from 43 to 51 thousand tons, and a total of 444 thousand tons of storage. It is important to highlight that not all silos are available to the regular operations, as they are owned by third parties. Due to the confidential nature of the commercial agreements regarding these silos, all silos will be considered available to the operation.

The port of Tubarão follows operational rules consistent with other ports in the world, which are fairly complex and in the interest of brevity they will not be described in this document, and it suffices to say that all these rules were properly included in the model developed for this system.

3.1 System's current planning framework

The current planning framework adopted for this system is divided in several different horizons, varying from the strategic, multiyear plan to the service order for loading and unloading issued on a daily basis.

The strategic plan, which targets the proper alignment of the future transport volumes with the strategic objectives of the company, and the annual plan, which includes non-product-specific volumes with a monthly resolution, are not relevant for the work presented in this document.

The monthly plan is the first stage of the planning system where all the relevant information on the demand for grain transport and schedule of ships is available down to a weekly resolution, and it is in this stage that the annual plan is either confirmed or amended to what is likely to final transportation numbers for the system.

The weekly and daily plans are purely operational and are tasked with solving conflicts in railcar loading, train dispatch and sequencing of loading and unloading operations throughout the system.

It is in the realm of the decisions performed in the weekly and daily plans that the model presented in this paper finds itself, and defining this correspondence was crucial for model validation, as it was necessary to establish correspondence between current practice and the model in order to obtain a sensible feedback from the operation personnel.

3.2 Product transportation demand

The transportation demand considered in the model corresponds to a total of 5.12 million tons unequally distributed as 4 different products. This product classification does not account for different clients and smaller product variations, which would drive the number of products to 20 different SKU.

Despite reducing the complexity of the problem, this simplification preserves the multi-product aspect of the problem while making it compatible with the level of detail applied to other portions of the model.

4. Methodology

As mentioned in the introduction and literature review, the appropriate modelling framework for defining operational rules that maximize system throughput has to include elements of operations research and stochastic simulation systems, which can come together on a simulation-optimization system as previously described.

---

[1] Figure of 2008

As touched briefly in the literature review, the model proposed in this paper can be summarized as an event-based stochastic simulation model with complex rules for queue treatment rules and system resources distribution, in line with the approaches of (Everett 2001) and (Fioroni 2008).

The biggest challenge to the implementation of such model, however, is that stochastic simulation packages are not designed to have complex rules for queue treatment and resource distribution that are based on multiple system parameters. For this reason, the entire model was implemented in scripting language, which allowed greater control over the elements being simulated, but also eliminated the possibility of sophisticated visualization outputs for the model.

4.1 The events-oriented simulation model

One of the most challenging aspects of implementing of an event oriented simulation system outside a traditional simulation package is that clear relationships between each and every event and process triggered by the start and completion of each other event and process in the system need to be established in a logical model. In turn, such a model is the guide for the entire model development and corresponds to the formulation of the model itself. The final logical model developed for this model is presented on Figure 2.

Figure 2: Stochastic simulation's logical model

Event \ Effect on the System	Rail engine is queued	Empty railcars are queued	Loaded railcar is queued	Truck is queued	Estimates ship arrival	Ship is queued	Updates Silo status	Free rail unloading structure	Free ship loading equipment	Atempt to start truck unloading	Starts railcar loading	Train with loaded cars is dispatched	Train with empty cars is dispatched	Starts railcar unloading	Start ship loading	Draw new Truck arrival at station	Ship is put in queue for loading
Truck arrives at train station				■						■						■	
Truck finishes unloading							■	■		■	■						
Empty railcars arrive at loading station	■	■									■	■					
Railcar finishes loading process			■				■	■		■	■	■					
Loaded railcars arrive at intermediary station	■		■									■	■				
Loaded railcars arrive at the port	■		■										■	■			
Empty railcars arrive at intermediary station	■	■										■	■				
Railcar finishes unloading		■					■	■					■	■	■		
Ship finishes loading							■		■					■	■		
Ship arrives at the port						■									■		
Ship is put in queue for loading					■												■

In order to establish the exact behaviour of each element in the model, the following subsections provide details in all the elements of the model and the processes and events associated to each one of them. It should be noted, however, that all the assumptions described in the following subsections were subjected to sensitivity analysis tests in order to assess their impact in the model results.

4.1.1 Rail loading stations

There are for possible processes associated with the rail loading stations: Truck unloading, railcar loading, train formation and train splitting (after arrival). Due to lack of data to estimate some parameters of the model, a few assumptions were necessary.

The first assumption necessary regarded the truck unloading structures, as there was not data on their capacity and operation times. For this reason, it was arbitrated a single unloading structure per station with a capacity that would result in utilization rate of 75% if all the demand is served.

It was also assumed that a truck would not wait in line for more than 36h, time after which it would be considered “lost demand”, which reflects the need for a minimum level of service required by the truck operators.

Further, the storage capacity in the origin stations is considered fixed per product throughout the entire simulation period, and the storage capacity in all the stations to add to the same storage capacity in the port.

Lastly, the railcar loading structure is also assumed to have a capacity that would result in utilization rate of 75% if all the demand is served.

4.1.2 Rail Transport system

The rail transport itself, including train crossings and breakdowns were not modelled explicitly, but rather as a set of probabilistic distributions for transit times for each pair of stations (loading stations, intermediary station and the port).

All such probability distributions were provided by the system’s operator, and constitute a level of detail above and beyond the information currently used to realize the monthly and weekly plan, all done in deterministic fashion with no intervals of confidence.

Operational times for train formation, disassembly and reclassification were also modelled with probability distributions provided by the operator, which allowed for a representative reproduction of the overall transport times in the system.

4.1.3 Grains and the port system

The port system includes the arrival of loaded trains, unlading of railcars, grain storage ans loading of ships.

As considered in the rail transport system, the rail operations occurring inside the port such as train formation and dispatch were modelled with probability distributions, but the capacity of the rail yard was considered unlimited for these operations, as they represent a small part of the overall port operations and capacity is currently sufficient for all rail operations inside the port.

The biggest assumption necessary for this subsystem regarded the rail unloading structure capacity, which had to be set at 754 thousand tons per month (compared to the 520 thousand informed by the operator), since the actual past operations and current system demand far surpassed the capacity informed initially.

The capacity assumed would also result in a utilization rate of 75% if all demand is served, and the block for unloading was assumed to be equal to 10 railcars, which normalizes the railcar blocks to the considered throughout the entire system.

4.1.4 Ship loading

Ship loading is modelled with a single berth and a single loading equipment, and the ship size was normalized to 42 thousand tons, which represent all the ship sizes used in this system, which present very little capacity variation around the 42-thousand-ton average.

The system's capacity corresponds to the total capacity of the dedicated berth and the partial use of a second berth in the port, corresponding to a monthly capacity of 57 ships per month and 2.4 million tons, resulting on a 72% berth occupation for the monthly peak period.

It is practice in the port to not accept the programming of ships for future loading if it is expected that the delays in docking or lack of available of products for loading would generate large payments in demurrage. However, the model has been constructed with assumptions similar to those used for the arrival of trucks, by defining a limit of 72 hours for a ship to be queued, after which it will be considered "lost capacity".

4.2 Prioritization rules

As mentioned before, the proposed hybrid model consists of replacing the queueing and resource distribution treatments with heuristic procedures. The prioritization rules, as they are described in the model, are specified by a number of system variables, described as follows:

p – Product index;

E – Loading station;

CN – Ship capacity;

CC – Truck capacity;

CV – Railcar capacity;

VT_i^p – List of railcar blocks in transit;

CAP_p^E – Storage capacity of product *p* at station *E;*

S_p^E – Amount of product *p* stored at station *E;*

${}_iFC_p^E$ – Truck *i* on queue at station E loaded with product *p; and*

${}_iFN_p$ – Ship *i* on queue at the port to be loaded with product *p;*

4.2.1 Proportion of silo occupation on loading stations

For the case when one needs to decide which product will be loaded at a certain railcar, it is possible to use the measure formulated on (1) to decide which product needs to be loaded, which aims to ensuring that there will always be storage capacity at stations for arriving trucks to unload, configuring a " pushed" system, where the supply of product takes precedence over the demand of it in the port.

$$\max \frac{S_p^E}{CAP_p^E} \tag{1}$$

4.2.2 Proportion of silo occupation on loading stations, alternative

For the case when one needs to choose which truck will be unloaded at a station, one can choose to unload the product for which the silo is the emptiest, as formulated on (2), in order to avoid that one product will become unavailable in the system, likely resulting in larger ship queues.

$$\min \frac{S_p^E}{CAP_p^E} \tag{2}$$

4.2.3 Port status: Silo occupation

Considering now that priority must be given for the product that has the smallest occupation ratio (formulated on (3)) at the port when deciding for which truck to unload, which railcar block will be transported or unloaded at the port, one will be prioritizing the port demand, and not the products at are more available in the system (Port is “pulling” the system). It is also possible to use the absolute value version of this criteria, as described on (4).

$$\min \frac{Silo_p^{port}}{CAP_p^{port}}. \tag{3}$$

$$\max\left(Silo_p^{port} - CAP_p^{port}\right) \tag{4}$$

4.2.4 Port status: Silo occupation Vs. Ships in queue

A natural derivation of the previous criteria is to consider not only the proportion of silo occupation at the port, but also the demand from ships that are already queued for docking, as formulated in (5) for the relative measure and in (6) for the absolute alternative.

$$\min \frac{S_p^{port}}{CAP_p^{port} + CN * \left(\sum_{i=1}^{n} {}_i FP_p\right)} \tag{5}$$

$$\max\left\{\left[CAP_p^{port} + CN * \left(\sum_{i=1}^{n} {}_i FP_p\right)\right] - S_p^{port}\right\} \tag{6}$$

4.2.5 Port status: Silo occupation and stock in transit Vs. Ships in queue

A second derivation of the port status criteria is to consider not only the product stored in the port and the ships queued for docking, but also the stock in transit, which should be arriving in the port for unloading within a short amount of time, as formulated in (5) for the relative measure and in (6) for the absolute alternative.

The advantage of this formulation is that there is less lag in measuring the effect of prioritizations already made in previous railcar loading and dispatching events.

$$\min \frac{S_p^{port} + CV * \left(\sum_{i=1}^{n} {}_i VT_i^{\,p}\right)}{CAP_p^{port} + CN * \left(\sum_{i=1}^{n} {}_i FP_p\right)} \tag{1}$$

$$\max CAP_p^{port} + CN * \left(\sum_{i=1}^{n} {}_i FP_p\right) - \left(S_p^{port} + CV * \left(\sum_{i=1}^{n} {}_i VT_i^{\,p}\right)\right) \tag{2}$$

4.2.6 Ships in Queue

Another alternative criteria is prioritizing movement of products throughout the system that correspond to the maximum number of ships queued at the port, as formulated on (9). Although simplistic, this heuristic represents well a “pulled system” setting.

$$\max\left(\sum_{i=1}^{n} {}_i FN_p\right) \tag{3}$$

4.3 Model decisions

As previously mentioned, the 9 prioritization rules just established were used as replacement for the queue management in resource distribution tasks in the simulation model developed for this problem.

In order to analyse the impact of these prioritization rules when applied to all possible decisions existing in the simulation model, it was necessary to identify and describe all these decisions, which are presented below.

4.3.1 Railcar loading: Decision at the hub station

Deciding which product will be loaded in each railcar block at the time they are sent from the intermediary station to loading stations is a fairly similar assumption to the real world operation, and at this stage it is possible to decide which product will be loaded using the criteria listed above.

4.3.2 Railcar loading: Decision at the loading station

An important addition to the previous decision is the possibility of re-assigning the product to be loaded in a railcar block at the time of loading. This alternative does not reflect precisely the most common real system operation procedure for this system, but its ability of reducing line of railcars waiting for specific products suggests that it is important to have this choice as a real alternative both in the model and in practice.

4.3.3 Ship loading

Since ships arrive at the port with a designated product for loading, the only decision possible regards the order of ships that will dock whenever there is more than one ship waiting in queue.

It should be highlighted that it is common practice for ports to operate under a strict FIFO queue. However, it is necessary to quantify what is the impact in total system productivity that this rigidity in operation and to measure the possible gains to be obtained if that rule where to be relaxed.

4.3.4 Train formation at the loading station

In loading stations with high throughput of grains, it is common to exist a rather long list of loaded railcars waiting for an engine to be dispatched towards the intermediary station. In these cases, there is the possibility of choosing which railcar blocks will be in the next train that departs that station.

4.3.5 Train formation at the intermediary station (loaded railcars)

Once the loaded railcar blocks are in the intermediary station, it is necessary to reclassify all the blocks and form the train that will be moving through the EFVM. As it was the case with train formation in the loading stations, in the intermediary station there is also the possibility of choosing which railcar blocks will be in the next train departing towards the port.

4.3.6 Truck unloading

Unloading trucks at the rail loading station is conditioned by the availability of two resources: The truck unloading equipment and empty space in the station silos. Since the truck unloading equipment is unique for each station, one can choose the next truck to unload among all of the trucks loaded with products for which there is available space in the silos.

4.3.7 Railcar unloading

Railcar unloading at the port is analogous to the truck unloading at the rail stations. However, it is necessary to also assume that there is enough space and resources to manoeuvre the railcar blocks in queue in order to unloaded the chosen one. In the other hand, there is no risk of “losing demand” in case one element is kept in the queue for long periods of time.

4.3.8 Railcar distribution

Resource distribution in this system, namely the choice of which loading stations will receive the next train of empty railcars is probably the most complex activity in the real life operation of the system, and thus the decision that resembles the least the assumptions of regular stochastic simulation models, which generally assumes random distribution proportional to the demand of each station.

By looking at the availability of products in silos and in trucks being unloaded, it is possible to ensure that railcars will seat empty while waiting for product to be loaded the least amount of time, hence better utilizing that resource.

5. Computational experiments

The computational experiments made with the proposed model were undertaken in two phases. In the first phase, the system was simulated[2] as a purely stochastic model with no prioritization rule being used in any of the model decisions, which resulted in 4.02 million tons of product transported and 1.1 million tons of "lost demand".

In the second phase, the objective was to determine the best combination of priority rules to be used in the system in order to maximize throughput. However, the combinatorial nature of the problem results in over 10 million decision-rule combinations, rendering unfeasible the analysis of the entire solution space. To address this issue, a heuristic loosely based on the Hooke & Jeeves method (Hooke & Jeeves 1961).

The resulting algorithm was based on an initialization step, tasked with finding a good starting solution, and an iterative phase, tasked with improving on the starting solution.

The starting solution was built by running the simulation model with only one decision being guided by a prioritization rule at a time, which resulted in 56 model evaluations (9 prioritization rules and 8 decisions, with not all combinations being applicable). Once it has been established what prioritization rule yielded the best model output (maximum system throughput), the initial solution was defined as the set of all one-on-one best prioritization rules applied at the same time.

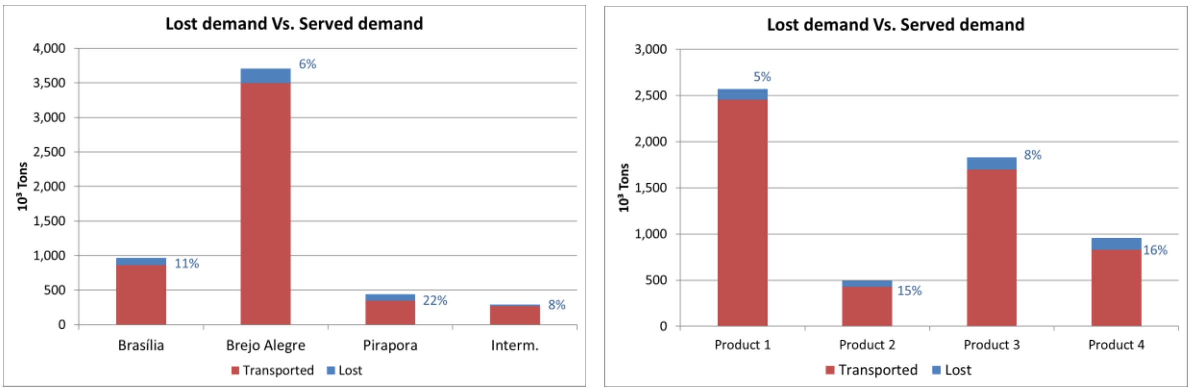
The iterative phase was conducted by looping through one model decision at a time and searching the best prioritization rule for each decision, given a fixed set of prioritization rules for the other decision. Once the best prioritization rule has been found for that particular decision model decision, the algorithm would move to the next decision and repeat the process.

The algorithm was stopped once a series of iterations through all model decisions had not yielded any improvements to the objective function.

The final result yielded a 4.57 million tons of throughput, which corresponds to 13.7% increase in total throughout and 50% reduction in "lost demand". The performance of the model was not uniform when analysing one loading station at a time and each product at a time, however, as clearly depicted on Figure 3.

[2] Each model run was composed of 100 repetitions

Figure 3: Model performance per Loading station and product



Despite the differences found among different stations and different products, the chart on Figure 4 shows that relative prioritization rules (1, 2 and 9) resulted in a more homogenous treatment of the products, as it would be expected.

Figure 4: Comparison of relative and absolute prioritization rules

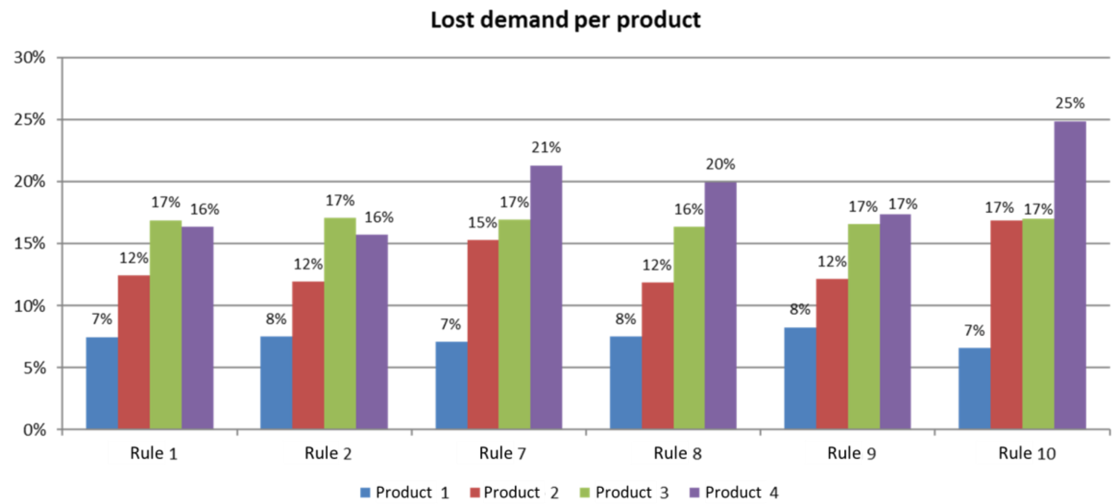
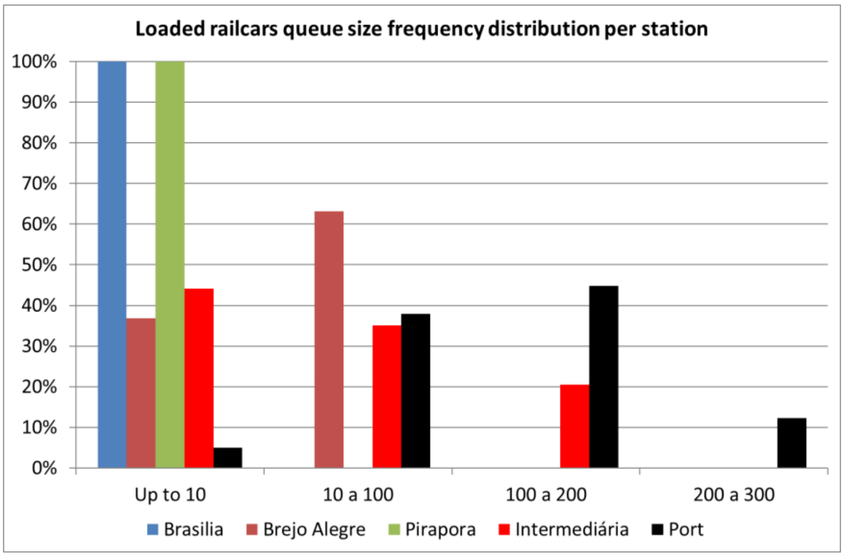


Figure 5: Queue size frequency distribution per station (loaded railcars)



Another interesting result was that the bottleneck of the system seems to be the port, as the it can be verified in the queue size frequency distribution per station for loaded railcars chart depicted on Figure 5.

Lastly, the optimal set of rules found after model convergence suggests that the system has its optimum throughput when operating as a pulled system (when priority is given for products needed at the port), which is in line with the expectation of the system operators.

6. Conclusions

A few conclusions can be made after developing this research. The first and most important conclusion, is that the substitution of standard queue treatment rules with more sophisticated and "system-aware" rules was extremely successful and indicates a more sensible path for the simulation this type of systems.

The second conclusion was that a simulation-optimization approach yielded a more accurate representation of the system, as the final model throughput was much closer to the real system throughput than a purely stochastic model.

The third important conclusion is that the model failed to consider strict correlation between truck arrival and ship arrival, resulting in consistently different "losses of demand" due to ship turn down and truck turn down.

7. References

Everett, J.E., 2001. Iron ore production scheduling to improve product quality. *European Journal of Operational Research*, 129(2), pp.355–361. Available at: http://www.sciencedirect.com/science/article/pii/S0377221700002332 [Accessed June 1, 2016].

Fioroni, M.M., 2008. Railroad simulation on closed loop and it's applications in Brazil evaluation of alternatives on choosing train destination. Available at: http://www.teses.usp.br/teses/disponiveis/3/3135/tde-03062008-180002/ [Accessed June 1, 2016].

Hooke, R. & Jeeves, T.A., 1961. `` Direct Search'' Solution of Numerical and Statistical Problems. *Journal of the ACM*, 8(2), pp.212–229. Available at: http://dl.acm.org/citation.cfm?id=321062.321069 [Accessed May 9, 2016].

Jacobson, S.H. & Schruben, L.W., 1989. Techniques for simulation response optimization. *Operations Research Letters*, 8(1), pp.1–9. Available at: http://dl.acm.org/citation.cfm?id=2309687.2309828 [Accessed June 1, 2016].

Pflug, G.C., 2008. *Optimization of Stochastic Models: The Interface between Simulation and Optimization*, Kluwer Academic Publishers. Available at: http://www.springer.com/us/book/9780792397809 [Accessed June 1, 2016].