Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

BUILDING AND OPERATING RESILIENT TRANSPORTATION YARDS USING SIMULATION

Hafsa Binte Mohsin
Jae Yong Lee
Vamshi Krishna Suvarna

Amazon Transportation Services
Amazon.com
11801 Domain Blvd
Austin, TX 78758, USA

ABSTRACT

Developing a comprehensive model is a practical approach for gaining insight into and analyzing complex systems such as transportation yards. Following this approach, we have developed a data-driven agent-based model for transportation yards at Amazon which captures the features and processes of yard operations. By simulating different scenarios and using simulation performance indicators such as yard/parking slip/dock door utilization, entry/exit gate queue, and late departure counts, the model helps to identify potential bottlenecks, inefficiencies, and risks in the system. Moreover, the model provides customers with recommendations for achieving maximum daily volume using what-if scenarios. The model's accuracy is evaluated using mean absolute error (MAE) and root mean squared error (RMSE), yielding promising results of 6 % and 7 % respectively. This paper presents an overview of the model, current use cases, and outlines future works to further improve the simulation model and enhance yard operations.

1 INTRODUCTION

The fulfillment of customer orders takes the highest priority among all Amazon networks requiring yards to process daily volumes as intended. refers to an area outside a facility where trailers arrive with trucks for processing. Trailers in the yard can either be waiting to be processed or waiting to be transported by a truck to their destination. The work process of yards can vary depending on the type of yards. One of the distinct characteristics of commercial retail business yards, including those operated by companies like Amazon, is their focus on reducing trailer dwell time (a couple of days) and achieving high turnover. The yard process starts with the entrance of a truck with or without a trailer by checking in at the guard shack (GS). The GS is a physical infrastructure where personnel confirm that the correct vehicle is entering/exiting the yard. During check-in, the associate updates the trailer location in the online tool, informing the truck driver where to drop/pick up the trailer in/from the yard. Post check-in, online tool reflects the truck's location. After entering the yard, trailers go to the dock door or parking slips, depending on the dock door availability. After a truck drops the trailer in the parking slip or dock door, an associate will operate a hostler to move trailers between a dock door and a parking slip and vice versa. Once the driver drops in/picks up a trailer, he checks out at the GS. GS personnel confirm the information of ending the truck's journey within Amazon's yard in online tool. Yards having high volume flow have offsite yards which help to carry extra trailers at peak volume season. Offsite shuttles move the trailers between onsite and offsite. Ontime customer order fulfillment requires smooth yard operation. Often yards face yard gridlock and long queues of trucks which can hamper the flow of volume. High yard utilization (gridlock) is an indication of inside yard congestion. Queuing up of trucks can impact public road traffic. Overall, high utilization of resources,

979-8-3503-6966-3/23/$31.00 ©2023 IEEE

yard, dock doors, and long queue of trucks are an indication of yard risk. The main objective of the simulation model for capacity analysis is to find one feasible solution in terms of maximum trailer load volume processing capacity avoiding yard risk.

The load volume process capacity of a yard depends on its physical attributes (parking capacity and process capacity in terms of dock doors), resources (hostlers and personnel) and the arrival schedule of trucks/trailers. The yard consists of several constraints that would limit the processing of additional volumes. Removing the yard constraints allows the site to not only increase the volume but also the on-time delivery to the customers' orders. Analytical models (Kourounioti et al. 2016; Lyu et al. 2022; Balster et al. 2020) are not capable of simulating random behavior of the system (Li et al. 2022). Simulation is often better than an analytical model for yards' capacity analysis because it can accurately represent yard operations' complex and dynamic nature (Amelia 2019; Vidal 2010). To address complex system behavior/interactions and provide data-driven insights to support decision-making and planning processes, we have developed an agent-based model for yard operations.

2 RELATED WORK

Yard simulation has become an area of interest for logistics and supply chain professionals to evaluate the performance of yards and identify strategies for improving efficiency and reducing costs. The performance and capacity of a yard depends on factors like yard configuration and handling system of loads (Zhang et. al 2017; Chu and Huang 2005). Arrival schedule of trucks/trailers is also an important input for yard capacity analysis and risk prediction. In their paper, Ramírez-Nafarrate et al (2017) discussed the impact of high variability of truck arrival on the yard at peak hour leading to gate congestion. Mizuyama (2020) also studied the dynamic behavior of arrival schedules and its effect on yard operation. To alleviate the congestion in the yard gate, Feng et al (2022) proposed a nonlinear mathematical model to allocate the space of incoming trailers considering the trucks waiting time in the yard. As the model's performance relies on the quality of inputs, studies of input improvement process are present in existing research. To improve the input of the model, integration of machine learning and simulation model is promising (Minbashi et al. 2023). Considering yard configuration, arrival schedule, trailer load processing times and the configuration of processing stations (docks) as inputs, we have developed an agent-based model for simulation. To develop a simulation model, modeler needs to understand the system. Demonstrating the process of simulating the yard in blocks instead of one single block, can capture the process in detail (Marinov and Viegas 2009). In this paper we described a model that utilizes input data and yard behavior specific to commercial retail yards like Amazon. The model aims to provide support for strategic decision-making processes within the commercial retail chain context. By developing and utilizing simulation models tailored to commercial retail chain yards, researchers and practitioners can gain valuable insights into yard performance, capacity analysis, and decision support. These models enable a deeper understanding of the complex dynamics of yard operations and facilitate the identification of strategies to improve the overall performance of commercial retail business yards.

3 METHODOLOGY

3.1 Agent-Based Model

An agent-based model (ABM) is a computational modeling technique for complex systems. Three main steps involving ABM are defining the agents, environment, and rules. The agents of this model are trucks, trailers, packing slips, and dock doors and have their own characteristics and behavior. Simulation environment is the yard represented by process flow. The agents' interactions in a simulation environment happen based on standard yard-specific rules. We used AnyLogic software, version 8.8.2, for simulation modeling. The agents' behavior changes with the change of yard configuration or rules, leading to different simulation results. Without the need for expensive and time-consuming real-world experimentation, experimenting the scenario in simulation can help to take strategic decisions.

3.2 Model Assumption and Input Data

The model assumes the operations would abide by the standard operating procedure when executing tasks. The inputs that populate resources, equipment, and physical attributes come from five tables among yard management system. First, the dock specific processing capabilities generated based on the historical load types ranging five priorities. Second, the physical attributes such as the number of check in/out lanes, hostlers, offsite capacity are set based on the space management tool. Third, the task duration to complete a specific task is measured based on the employees' interaction with the onboard devices such as moving a trailer from dock door to a parking spot or processing time for checking in a trailer. Fourth, the initialization table populates the yards with the historical presiding trailer loads to avoid starting simulation with empty yard. As this model runs every day to predict the next three day's yard status, we used actual yard data instead of warming up the model. Fifth, the arrival schedule enters and removes trailers from and to the yard executing the move events consisting of trailer load type, equipment type, and volume. These inputs are used to generate agents. Table 1 represents some details of input data across the five databases.

Table 1: Input Data for Simulation Model.

Input	Description	Data type
Dock Processing Time	Time to process a trailer load at the dock door	Continuous
Resource Processing Time	Time to process a truck/trailer by resource	Continuous
Dock Door	No dock doors	Discrete
Parking Slips	No parking slips.	Discrete
Guard Shack	No guard shacks.	Discrete
Hostler	No of Hostlers	Discrete
Arrival Time	Trailer arrival time in the yard	Timestamp
Departure time	Trailer departure time from the yard	Timestamp
Trailer Load type	Process type of loads	String
Offsite yard	No parking slips outside of the main yard	Discrete

The yard inputs in terms of physical and resource counts are mostly the same over time since it requires renovation or equipment addition to the yard. The count for physical and resource remains constant since construction and adding equipment takes time. In contrast, the dock processing and the resource processing time continues to change. The model uses triangular distribution to address the processing time variability. A study (Fairchild et al. 2016) supports the use of triangular distribution, where data distribution is vaguely known. From the minimum, maximum, and mode values present in the database for processing time data, the distribution has been created using Equation (1).

$$P(x) = \begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & \text{for } a \leq x < c \\ \frac{2(b-x)}{(b-a)(b-c)} & \text{for } c \leq x \leq b \end{cases} \tag{1}$$

In the above equation, a is minimum, b maximum, and c is the mode value of the distribution. Diverse loads take various times to process based on their distribution. Figure 1 represents the dock processing time of a trailer specific load.

3.3 Model Development

The model logics are compartmentalized to avoid logic overlaps among yard activities. The model logic was then developed based on the standard operating procedure assuming the operation to take place by the regulation. Trailer agent was developed based on twelve parameters represented as $[L_{l1}, L_{l2}, \ldots, l_{ln}]$ where n = {1, 2...,12}, l is the trailer agent, and L is the parameter value. For example, an equipment type is a parameter value of an agent distinguishing whether the agent is a box truck or a linehaul with different processes. Similarly, the dock door agent also has five parameters from the highest to the lowest priorities of load types determining which trailer carrying load would end up on a particular dock door as [D_{d1}, D_{d2}, …, D_{dm}] where m = {1, 2...,5}. The parking slip agent has a single parameter since being a single entity for stationing a load. The trailer agent's entry and exit to/from the yard takes place with a truck agent enabling the movement of the loads. The physical attributes such as the guard shack and the equipment attributes like a hostler are resource referring the historical attributes. Figure 2 illustrates the process flow of an incoming trailer to the yard.

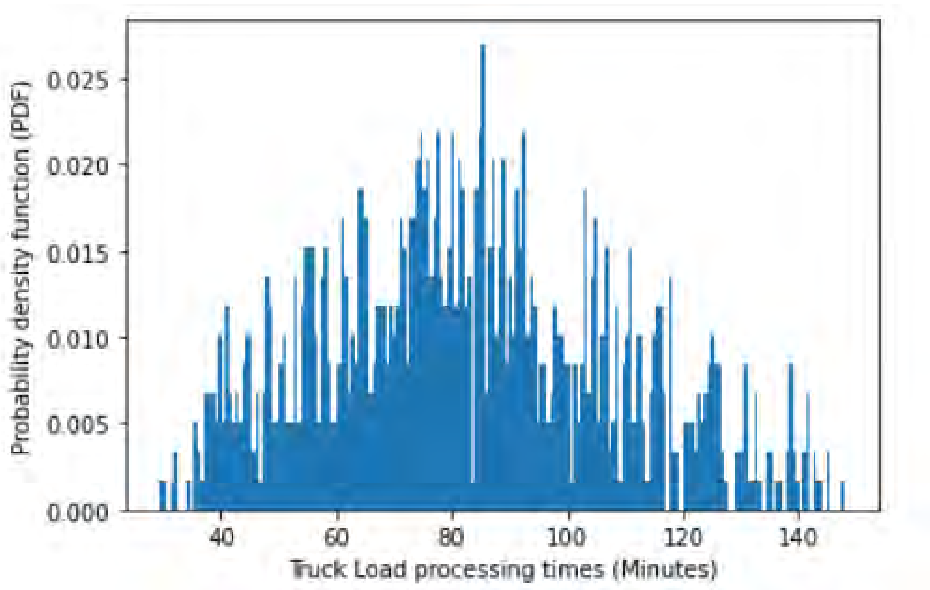


Figure 1: Dock processing time distribution for a specific type of truck (trailer) load.

The model validation process starts after model development. After validation, the model is ready for capacity analysis. The complete process of running and analyzing the simulation model follows the following steps depicted in Figure 3.

3.4 Model Performance Measure

The risk signal metrics per scenario are the utilization (yard, dock door, packing slip, hostler, offsite shuttle), check-in queue, late departure count, and trailer check-in threshold. The ratio of occupied to total number of assets represents the utilization of respective yard assets. The queues are the number of trucks accumulating beyond the guard shack signaling public road interference. The late departure is the number of trailers departing late for 30 minutes or more from their scheduled departure time. Finally, the trailer check-in time is the count of trailers entering the yard per hour. These risk signals are measured for all analysis since any signals violated could lead to volume flow limitation. Equation 2 to 6 are used to calculate yard metric values. d_{bt}, d_a, p_{bt}, p_a, N_{lt} represents the number of occupied dock doors, number of total dock doors, number of occupied parking slips, number of total parking slips, and number of vehicles at time t. Here the maximum value of t is sixty minutes.

$$\text{Mean Yard Utilization / Hour} = \sum_1^t \frac{d_{bt}+p_{bt}}{d_a+p_a} / t \tag{2}$$

$$\text{Mean Dock Door Utilization / Hour} = \sum_1^t \frac{d_{bt}}{d_a} / t \tag{3}$$

$$\text{Mean Gate Queue Length / Hour} = \sum_1^t [N_{\text{lt(Entry)}} - N_{\text{lt(Exit)}}]/\,t \tag{4}$$

$$\text{Mean Hostler Utilization / Hour} = \sum_1^t \frac{h_{bt}}{h_a}/t \tag{5}$$

$$\text{Number of Vehicles Checked in / hour} = \sum_1^t N_t \tag{6}$$

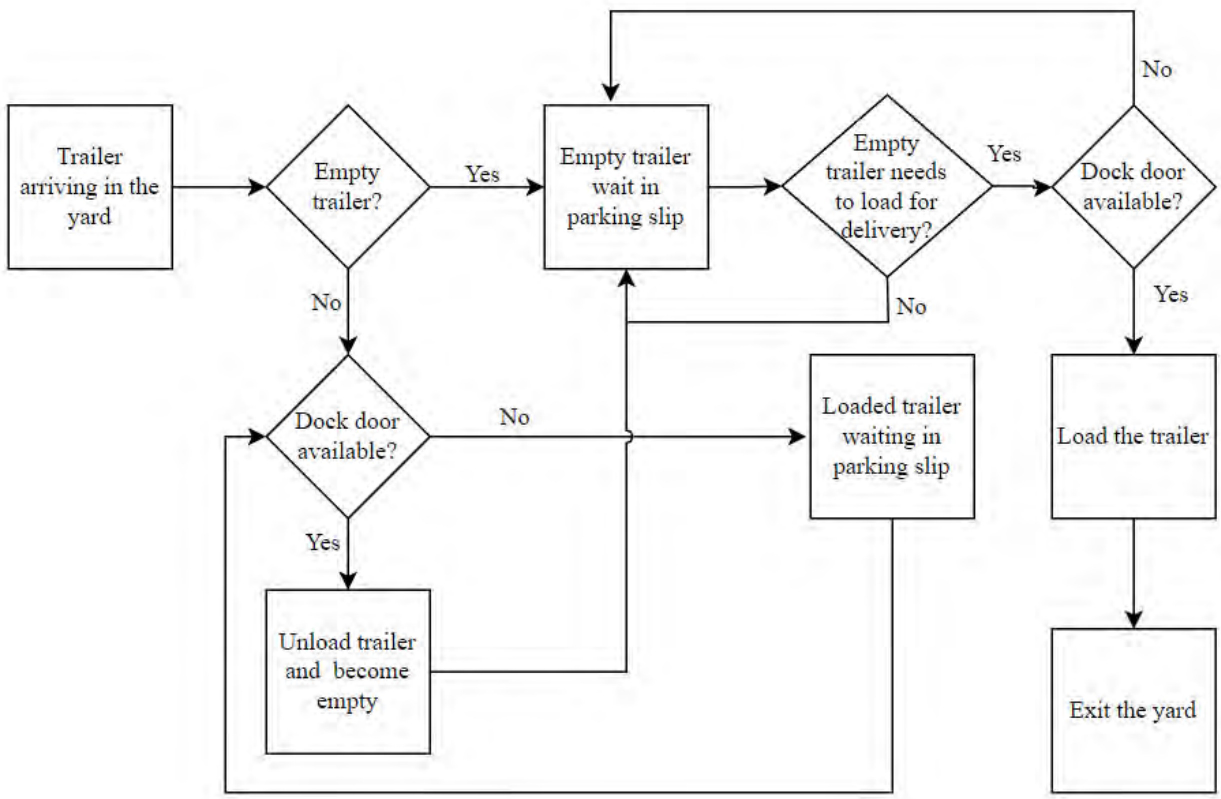


Figure 2: Process flow diagram of incoming trailer.

3.5 Model Validation and Sensitivity

More than a hundred yards have been selected to validate the simulation model. We used yard utilization for validation as it represents the total status of the yard. With actual input data, we ran the simulation model for two days for all the yards and collected the simulated yard utilization data. Using two sample t-tests, we checked the model's performance. With a 95 % confidence level, we observed that there is no significant difference between the values of simulated and actual yard utilization [p-value (1.21)>α (0.05)]. It implies that the model is statistically valid. We have also used mean absolute error (MAE), root mean square error (RMSE), and dynamic time wrapping (DTW) methods (Table 2) to measure the performance of the model. MAE and RMSE provide the indication of deviation of simulated value from actual value where DTW looks for similarity in pattern between two time series data. Based on both MAE and RMSE, the model's error percentage is less than 8 %. DTW constructs a cost matrix that specifies the distance between every pair of points in the two-time series. Lower cost represents the higher similarity of two data sequences. From forty-eight hours of time series data, we calculated DTW to check the similarity in patterns between simulated and actual yard utilization. The observed DTW value for our model is 2.13.

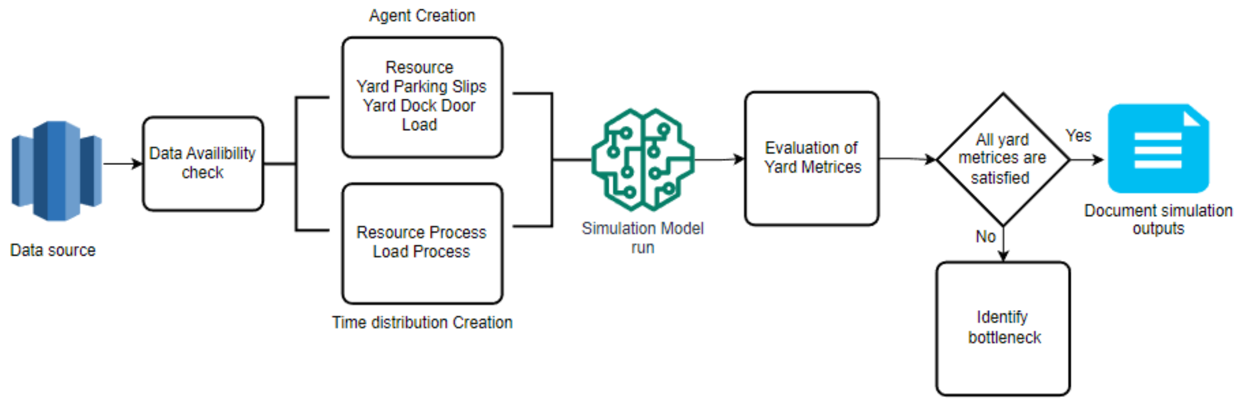


Figure 3: Conceptual framework of Simulation model analysis for Amazon Yard.

Table 2: Performance calculation for the Simulation model.

Model Performance Measure	**Method**	**Average Value**
Accuracy Metrics	MAE	6 %
	RMSE	7 %
Distance Measure	DTW	2.123

Consistency of model output can provide an idea of the model's robustness. To check the robustness of the model, we ran it for four consecutive days for all the yards assessed before. The main difference in this setup is the different schedule for different days. For MAE and RMSE, the error % varied by ~1 % across the days and DTW varied from 0.05 to 0.24 (Table 3). Less variation of performance indicates the robustness of the model.

Table 3: Simulation Model's performance for four consecutive days.

Day	**MAE (%)**	**RMSE (%)**	**DTW**
1	6.62	7.19	1.31
2	5.78	6.35	1.07
3	5.40	6.06	1.02
4	5.63	6.20	1.08

As discussed in Section 3.2, several inputs are required to develop this model. Yard capacity analysis is one of the use cases of this model. For yards, capacity is translated to number of trailers processing ability. To check the sensitivity of the model, we ran it for a single yard varying the total number of trailers flowing to the yard. We used gate queue metric to observe the impact. From Figure 4, for four different truck volume scenarios (Volume_Sc1 to Volume_Sc2), the number of trucks waiting outside the gate (Gate queue) is different which confirms the claim of model's sensitivity is true. For 'Volume_Sc2', the impact is significant from rest of the scenarios.

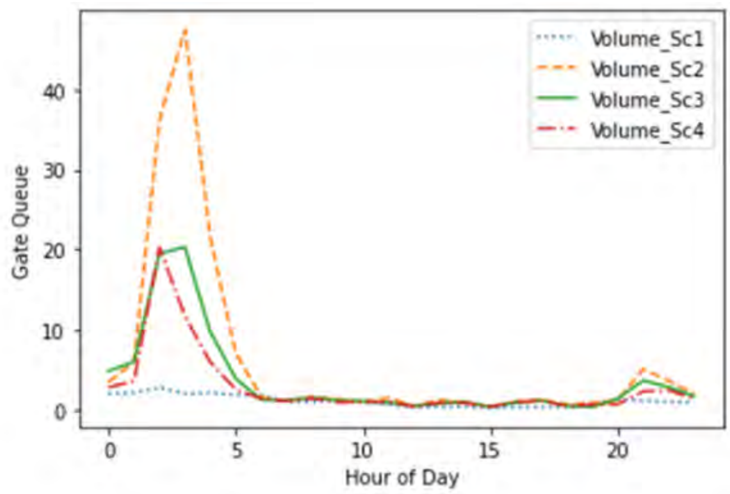


Figure 4: Sensitivity check of Simulation Model.

4 MODEL USE CASE

The two primary use cases of this model are yard capacity analysis and yard risk prediction. For capacity analysis, desired trailers volume is the primary input. Total trailer volume flows to the yard following a schedule. Maintaining historical arrival and departure pattern, total number of trailers are distributed over twenty-four hours to generate a schedule. Keeping other inputs fixed, we run the simulation model for forecasted schedule. Evaluation of simulated yard metric provides insights into the processing capacity of the yard. For the use case of yard risk prediction, the model consumes the forecasted trailer volume schedule and predicts the next three days status of the yard. Risk prediction helps the risky yards to take advance precautions.

4.1 Capacity Analysis

The main objective for capacity analysis is to achieve maximum trailer volume flow without setting the yard at risk. We set the boundary constraints before running the simulation model. As an example, for capacity analysis of Yard ‘A,’ boundary thresholds for desirable output metrics are set. Yard utilization, check-in, gate queue, dock utilization, and onsite hostlers are the targeted metrics, and thresholds set to these metrics are 95 %, 109, 6, 95 %, and 95 %, respectively. As this simulation model is stochastic, using random seeds for different simulation replications will make the analysis more robust. The number of replications required for simulation relies on the acceptable rate of error from that analysis.

For this study, we have considered the error rate as 10 %. We used equation 8 to get the number of replications. Here, d_n is an error as a fraction of the sample mean, n is the number of simulation replications, tn-1, α/2 is the critical value of the student's t-distribution, α is the level of significance, S_n is the sample standard deviation, and X_n is the sample mean.

$$d_n = \frac{t_{n-1, \alpha/2} * (\frac{S_n}{\sqrt{n}})}{X_n} \qquad (7)$$

The experiment presented in this section aims to identify the maximum trailer volume processing capacity of the yard without breaching any thresholds. As a starting point, we run the model for specific trailer volume scenario and check for target metrics. Depending on the metrics output, the next volume scenario run takes place. For this case, after starting Volume Scenario 1, the volume was increased twice as the boundary thresholds were not breaching for any metrics (Table 4). At Volume Scenario 3, gate queue threshold breached (>6) and required the decrease in volume. As Volume Scenario 5 was not violating any thresholds, we identified it as a feasible volume for the yard.

Table 4: Capacity analysis scenarios.

Volume Scenario	Change in Volume	Yard Utilization (%)	Number of Check-in	Gate Queue	Dock utilization (%)	Hostler Utilization (%)
Scenario 1	_	68.0	48	1	63.00	44.00
Scenario 2	30 % increase	83.0	61	4	87.83	57.80
Scenario 3	8 % increase	90.0	62	7	93.00	64.00
Scenario 4	2 % decrease	86.0	63	7	93.04	61.46
Scenario 5	**2 % decrease**	**83.0**	**61**	**6**	**93.00**	**60.21**

The yards always want to process more volume. Though Volume Scenario 5 met all the safety thresholds, we wanted to test whether the two bottom scenarios difference is significant or not. As the main factor is gate queue for this case, we conducted a statistical test on gate queue for both the scenarios. To perform paired t-test, the data should satisfy normality test. The two-dataset showed skewness and required Wilcoxon signed-rank test based on equation (8).

$$W = \sum_{i=1}^{N_r} \left[sgn(x_{2,i} - x_{1,i}). R_i \right] \tag{8}$$

Here,
W = Test statistics
N_r = Sample size, excluding pairs where x_1=x_2
sign = Sign function
$x_{1,\,i}$, $x_{2,\,I}$ = Corresponding ranked pairs from two distributions
R_i = rank i

We have conducted this test considering the null hypothesis as the median gate queue for Volume Scenario 4 and 5 are the same. The significance value considered is α=0.05. Wilcoxon's result showed that the (p-value=0.0440) p-value is less than α (0.05) which implies the rejection of null hypothesis. With rejection of null hypothesis, we accept the fact that two volume scenarios are different, and the volume of Scenario 5 is the maximum capacity for the yard.

4.2 Risk Prediction

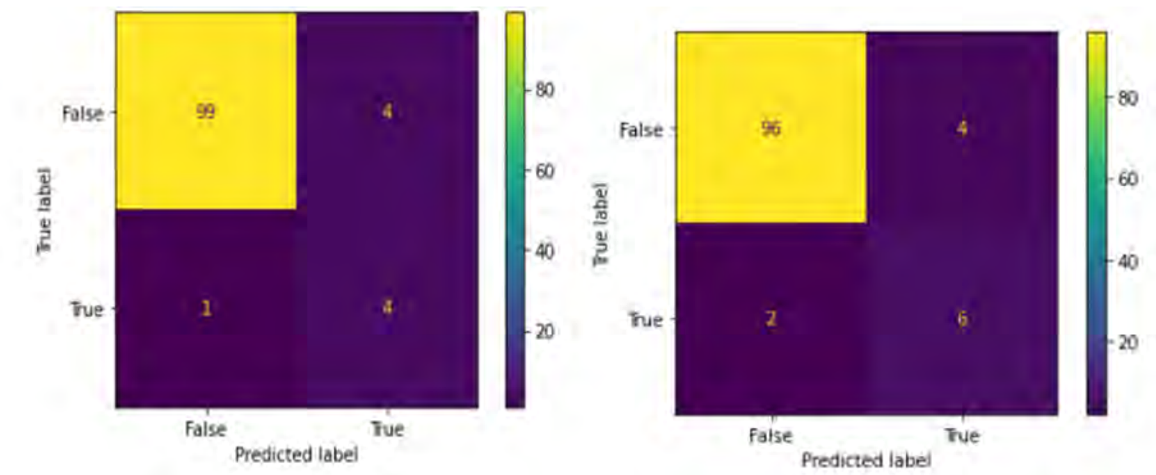
Yard risk prediction helps a yard to take necessary steps beforehand to avoid congestion/gridlocks and maintain smooth volume flow. This simulation model predicts risk for yards based on the forecasted trailer volume later converted to the arrival schedule.

4.2.1 Risk Score

The risk score depends on the value of yard utilization metric. To calculate the yard risk score, hourly yard utilization data are collected. The yard utilization reaching more than 90 % is the critical point for the yard. For hours having yard utilization more than 90 % is marked as risk score ‘1’. The risk score is ‘0’ for the rest of the yard utilization value. Based on the collected data, we categorize the data in two classes - “No risk" with a risk score of zero and "Has risk" with a risk score of 1. Depending on the presence of risk hours in a yard’s hourly data, yards are categorized as ‘Has Risk’ and ‘No Risk’ class.

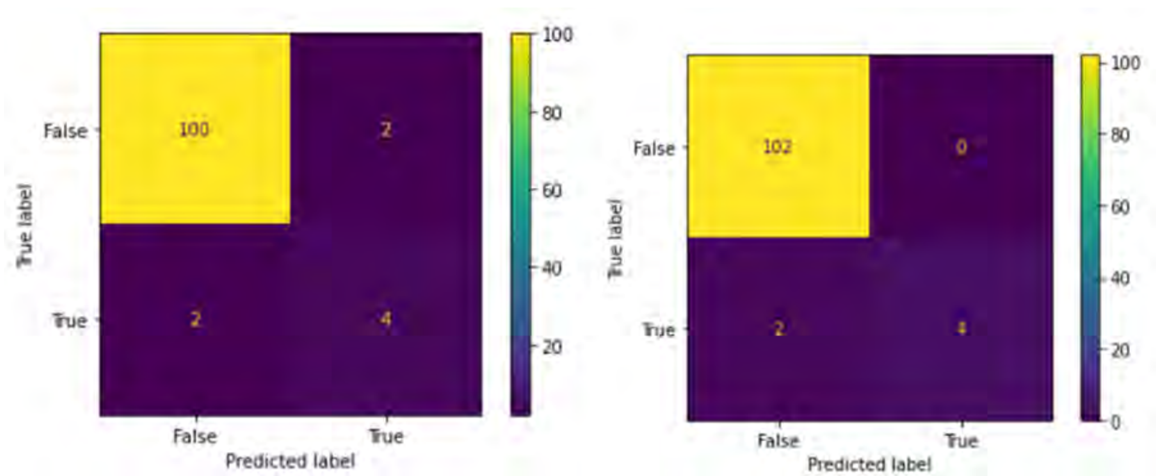
4.2.2 Prediction Performance

For hundred and eight yards, we have measured prediction performance for four consecutive days. Here, the performance measurement is in yard level. The prediction performance analysis will provide an idea of model's accuracy to predict yards status in terms of safety. Figure 5 and Figure 6 represent the confusion matrix of four days. Here the dataset has more data for the 'No Risk' class (False) than the 'Has Risk' class (True). As the data set is imbalanced, we have calculated precision, recall, and f1 score along with accuracy to measure the prediction performance.



(Day 1) (Day 2)

Figure 5: Confusion matrix for first and second day.



(Day 3) (Day 4)

Figure 6: Confusion matrix for third and fourth day.

Tables 5–8 show the prediction performance of the simulation model. The classification score for each day has accuracy, precision, recall, and f1-score data. Accuracy tells the percentage of total data predicted correctly. As the dataset here has more data for the 'No Risk' (>100 out of 108) class, the model's accuracy will mostly depend on this class's prediction performance. We can see from Table 5 to Table 8 that the total prediction accuracy is more than 90 % for all four days. To measure the prediction performance for the 'No Risk' and 'High Risk' classes separately, we need to check the precision, recall, and F1-score values. Precision for a class is the ratio of the number of correct predictions to the total number of predictions that the data belongs to that class. A precision of 99 % for class 0 represents that ninety-nine yards have been classified correctly as a 'No Risk' class out of one hundred yards predicted as a 'No Risk' class. For the 'High Risk' class, the precision is 50 %, meaning the model correctly identified this class 50 % of the time. Recall is the ratio of correct predictions to the total number of actual data that the data belongs to that class. F1-score is the average of precision and recall. For class-based performance, the f1-score for the 'High Risk' class varies from 62 % to 80 % from day 1 to day 4 (Tables 5 to 8). Macro avg provides the class average for all the metrics, whereas weighted average provides weight to the respective class while calculating the weighted average. As the class 'No Risk' has a higher percentage of data, the weight for this class is high. The weighted average value for precision, recall, and F1-score inclined towards 'No Risk' class values for these metrics. The overall performance of the model from the f1-score perspective for the 'High risk' class is >60 %, and for the 'No Risk' class,>95 % across four days. Among the four days, f1-score for fourth day is better than the rest of the days (Table 8).

Table 5: Classification scores for Day 1.

		Class	**Precision**	**Recall**	**F1-score**	**Support**
Label	No Risk	0	0.99	0.96	0.98	103
	Has Risk	1	0.5	0.8	0.62	5
	Accuracy				0.95	108
	Macro avg		0.74	0.88	0.8	108
	Weighted avg		0.97	0.95	0.96	108

Table 6: Classification scores for Day 2.

		Class	**Precision**	**Recall**	**F1-score**	**Support**
Label	No Risk	0	0.98	0.96	0.97	100
	Has Risk	1	0.60	0.75	0.67	8
	Accuracy				0.94	108
	Macro avg		0.79	0.85	0.82	108
	Weighted avg		0.95	0.94	0.95	108

Table 7: Classification scores for Day 3.

		Class	**Precision**	**Recall**	**F1-score**	**Support**
Label	No Risk	0	0.98	0.98	0.98	102
	Has Risk	1	0.67	0.67	0.67	6
	Accuracy				0.96	108
	Macro avg		0.82	0.82	0.82	108
	Weighted avg		0.96	0.96	0.96	108

Table 8: Classification scores for Day 4.

		Class	Precision	Recall	F1-score	Support
Label	No Risk	0	0.98	1.0	0.99	102
	Has Risk	1	1.00	0.67	0.80	6
	Accuracy				0.98	108
	Macro avg		0.99	0.83	0.90	108
	Weighted avg		0.98	0.98	0.98	108

5 CONCLUSION AND FUTURE WORK

This paper describes the utilization of a simulation model to evaluate processing capacity and assess risks in commercial retail business yards. By identifying bottlenecks and analyzing the system's performance against capacity goals, the simulation model allows for improvements to be made in yard features. The yard risk prediction data generated by the model is then used by field teams to make tactical decisions and mitigate yard gridlocks. This model enables cost-effective strategic decision-making for Amazon yards which can be applied other commercial retail business yards.

However, the paper also acknowledges certain drawbacks of the simulation model. One limitation is that simulation models are built based on standard operating procedures and may not accurately replicate all sites or facilities to the same degree. For instance, the use of an offsite yard is required when the main yard lacks space. Historical data on trailer movement between yards may not fully justify this standard as trailers may be moved between yards without following a set rule. Another challenge is the accuracy of input data, which can impact the precision of risk predictions.

To address these limitations, future work involves integrating machine learning techniques with the simulation model. By mining rules from historical data instead of relying solely on predefined standards, the model's accuracy and prediction performance can be improved. Additionally, efforts will be made to improve data quality. These improvements aim to enhance the model's capacity evaluation and yard risk prediction capabilities with greater confidence.

REFERENCES

Amelia, P., and A. Latifah, 2019. “Dynamics Analysis of Container Needs and Availability in Surabaya Container Terminal with Agent-based Modeling and Simulation”. *Procedia Computer Science* 161:910–918.

Balster, A., O. Hansen, H. Friedrich, and A. Ludwig,. 2020. “An ETA Prediction Model for Intermodal Transport Networks Based on Machine Learning”. *Business & Information Systems Engineering* 62:403–416.

Fairchild, K.W., L. Misra, and Y. Shi. 2016. “Using Triangular Distribution for Business and Finance Simulations in Excel”. *Journal of Financial Education* 42(3/4):313–336.

Feng, X., Y. He, and K.H. Kim. 2022. “Space Planning Considering Congestion in Container Terminal Yards”. *Transportation Research Part B: Methodological* 158:52–77.

Kourounioti, I., A. Polydoropoulou, and C, Tsiklidis. 2016. “Development of Models Predicting Dwell Time of Import Containers in Port Container Terminals – An Artificial Neural Networks Application”. *Transportation Research Procedia* 14:243–252.

Li, N., H. Haralambides, H. Sheng, and Z. Jin. 2022. “A New Vocation Queuing Model to Optimize Truck Appointments and Yard Handling-Equipment Use in Dual Transactions Systems of Container Terminals”. *Computers & Industrial Engineering* 169:108216.

Lyu, Z., D. Pons, Y. Zhang, and Z. Ji. 2022. “Minimum Viable Model (MVM) Methodology for Integration of Agile Methods into Operational Simulation of Logistics”. *Logistics* 6(2):37.

Marinov, M., and J. Viegas. 2009. “A Simulation Modeling Methodology for Evaluating Flat-shunted Yard Operations”. *Simulation Modelling Practice and Theory* 17(6):1106–1129.

Matsuyama, H. 2020. “Agent-based Modeling and Analysis of Dynamic Slab Yard Management in a Steel Factory”. *IFIP Advances in Information and Communication Technology* 591:37–44.

Minbashi, N., H. Sipilä, C.W. Palmqvist, M. Bohlin,, and B. Kordnejad. 2023. “Machine Learning-assisted Macro Simulation for Yard Arrival Prediction”. *Journal of Rail Transport Planning & Management* 25:100368.

Misra, J. 1986. “Distributed Discrete-event Simulation”. *ACM Computing Surveys* 18(1):39–65.

Petering, M.E.H. 2011. “Decision Support for Yard Capacity, Fleet Composition, Truck Substitutability, and Scalability Issues at Seaport Container Terminals”. *Transportation Research Part E: Logistics and Transportation Review*, 47(1):85–103.

Ramírez-Nafarrate, A., R.G. González-Ramírez, N.R. Smith, R. Guerra-Olivares, and S. Voß. 2016. “Impact on Yard Efficiency of a Truck Appointment System for a Port Terminal”. *Annals of Operations Research* 258(2)195–216.

Vidal, J. M., N. and Huynh. 2010. “Building Agent-based Models of Seaport Container Terminals”. In *Proceedings of the 6th Workshop on Agents in Traffic and Transportation, May 11th*, Toronto, Canada, 49–58.

Zhang, L., M. Jin, Z. Ye, H. Li, D.B. Clarke, and Y. Wang. 2017. “Macrolevel Classification Yard Capacity Modeling”. *Transportation Research Record: Journal of the Transportation Research Board* 2608(1):125–133.

AUTHOR BIOGRAPHIES

HAFSA BINTE MOHSIN is a Data Scientist at Amazon. She earned her Ph.D. in Industrial Engineering from the University of Texas at Arlington. Her research interest is in agent-based, discrete event, and Hybrid simulation. Her email address is hbmohsin@amazon.com.

JAE YONG LEE is a Data Scientist at Amazon with a research interest in agent-based and discrete event simulation focusing on constraint analysis. His email is jaeylee@amazon.com, and his LinkedIn is www.linkedin.com/in/leejaey/.

VAMSHI KRIHSNA SUVARNA is a Data Science Manager part of Amazon Transportation Service at Amazon. His primary responsibility is analyzing the risk of the Amazon Transportation yards leveraging Simulation and Machine Learning based models and supporting strategic and tactical decision-making processes. His email address is vssuvarn@amazon.com.