



Hybrid model for the design of a deep-lane multisatellite AVS/RS

Ilaria Battarra¹ · Riccardo Accorsi¹ · Riccardo Manzini¹ · Sara Rubini¹

Received: 8 February 2022 / Accepted: 16 May 2022
© The Author(s) 2022

Abstract

The autonomous vehicle storage and retrieval system (AVS/RS) significantly improves the responsiveness and throughput of the traditional automated storage and retrieval system (AS/RS) in regard to handling unit loads. The AVS/RS consists of multiple tiers connected to an elevator system and is equipped with at least two autonomous vehicles, that is, a shuttle and satellite. Other necessary equipment are the lifts and input/output buffer areas. This paper aims to present and apply an original hybrid analytical-simulative model for the design of a deep-lane and multisatellite AVS-RS by evaluating and controlling the system performance. This AVS-RS is equipped with multiple free and non-free satellites for each tier. As an original contribution, this study reviews the literature on AVS/RS according to the introduction of multiple features categorized into five homogeneous groups: (1) rack configuration, (2) vehicle kinematics and configuration, (3) dispatching rules, (4) modeling approach, and (5) validation. Two of the most critical issues in existing research studies are the random arrival time of storage and retrieval transactions and the random storage policy. The proposed modeling approach is data-driven and based on realistic assumptions, filling the gap between the literature and real applications. This hybrid model is applied to a case study of the beverage industry according to a what-if comparative and competitive multiscenario analysis. This data-driven assessment supports the decision-making process on the number of satellites for each tier, while simultaneously controlling the service and waiting times, system throughput, and vehicle utilization. The analysis based on the maximum system throughput estimation demonstrates that introducing more than two satellites does not increase the productivity of the system.

Keywords Automated warehouse · Automated vehicle storage and retrieval system (AVS/RS) · Analytical modeling · Simulation · Tier-captive system · Multisatellite AVS/RS

1 Introduction

Globalization is a growing trend that is driving companies to increase supply chain performance and warehouse efficiency. Just-in-time (JIT) production is essential to modern supply chains, requiring raw materials, WIPs (work in process), or assemblies in the right place at the right time. Simultaneously, customer demand is uncertain, being characterized by a complex inventory mix and short lead times. Moreover, the increase in e-commerce broadens the demand for customized products, requested in small volumes within a very short time [1]. Despite the costs, storage and warehousing systems play a pivotal role in achieving such goals throughout the

supply chain. In line with Industry 4.0, automated warehouses can manage and optimize goods storage, picking, and shipment processes, ensuring high flexibility and short response time.

The automated storage and retrieval system (AS/RS) is one of the most common automated warehouse solutions. It consists of one or more aisles, each equipped with storage racks on either side, a stacker crane (the storage and retrieval machine, S/R), and input and output stations (called I/O stations). This S/R machine can simultaneously move vertically and horizontally along the aisle. The FEM 9.851 standard is the most used criterion in industrial applications to evaluate the AS/RS performance by estimating the cycle time. Although these automated systems guarantee a high storage density, they are not flexible: at least one stacker crane in each aisle can move unit loads (UL) singly, and simultaneous operations on different tiers are impossible. In addition, this automatic storage solution can usually feed shelving to the maximum triple depth, limiting storage capacity and space efficiency.

✉ Riccardo Manzini
riccardo.manzini@unibo.it

¹ Department of Industrial Engineering, Alma Mater Studiorum – Bologna University, Viale Risorgimento 2, 40136 Bologna, Italy

Companies are moving towards autonomous vehicle storage and retrieval systems (AVS/RS) to overcome these weaknesses. These automated warehouses can improve flexibility and responsiveness in handling unit loads and increase productivity (i.e., throughput) but still require design-support tools and modeling techniques that can fully exploit their technological potential.

The system object of this study is a tier-captive, deep-lane AVS/RS equipped with lifts and buffer areas (as summarized in Fig. 1) that consists of n -independent tiers (system levels). The tiers are connected with an elevator system that can vary according to the number of lifts, their position within the storage area, and their function (e.g., lift for storage activities, lift for retrieval activities, or lift for both). Each tier has at least two autonomous vehicles: a shuttle and satellite. This is the so-called tier-captive configuration equipped with one or more satellites.

In a generic tier, an input conveyor moves unit loads from the lift to the buffer area (input bay), waiting for the shuttle. The conveyor and buffer area are the handling interfaces between the lift and shuttle. An output conveyor at each tier moves unit loads from the shuttle to the buffer area (output bay), waiting for the lift. In addition, there are input and output bays at the ground floor, where the fleet of man-on-board forklifts or flexible/autonomous vehicles, for example, laser-guided vehicles (LGV) or autonomous mobile robots (AMR), hold unit loads coming from the end-of-line area of a production/packaging system and directly catch unit loads to the shipment area. A typical application of such a handling and storage system configuration is the food and beverage industry equipped with automated filling and packaging machines.

Each tier has a main double-sided cross-aisle and multideep lanes (named channels). Depending on the lane depth, each lane contains one or more unit loads (ULs) of the same item. These are single stock keeping unit (SKU) channels in a last-in-first-out (LIFO) unit load handling. The shuttle moves back and forth along the main cross-aisle (x-direction in Fig. 1), transporting at least one satellite. The satellite departs from the shuttle and enters the storage lanes to store or retrieve the UL (z-direction in Fig. 1). A single satellite can transport one UL or two simultaneously. Therefore, the maximum handling capacity is four ULs if the shuttle has a double handling capacity along the x-direction, and each satellite has a double capacity along the z-direction.

During the storage activity, the UL is stored in a specific (x,y) lane owing to the lift and shuttle and in a specific z-location within the lane owing to the satellite. The satellite stops precisely in front of the last previously inserted UL, respecting the dynamic filling of the lane. The satellite performs the next retrieving or stocking action in this position, agreeing with the LIFO policy for put-away activities.

The shuttle and satellite can be two independent vehicles. Consequently, when the satellite enters a specific lane, the shuttle can perform other handling tasks (e.g., the shuttle can move to the input bay loading another incoming UL). We name this configuration “free satellite (FS),” which differs from the “not-free satellite (NFS).” The FS option allows an increase in the throughput of the system owing to the introduction of multiple satellites for each tier. This is called multisatellite AVS/RS.

Figure 1 summarizes and exemplifies the adopted notation for a unit load AVS/RS deep-lane of six ULs in the

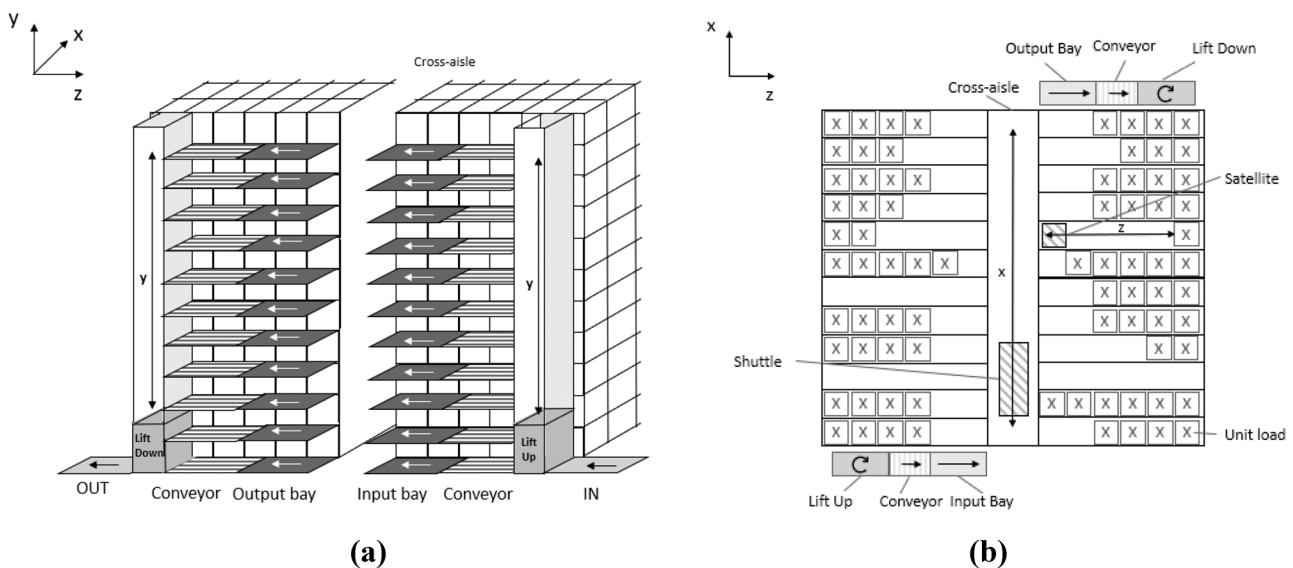


Fig. 1 a AVS/RS front view, b AVS/RS plan view

z-direction with one double-capacity shuttle and a single satellite.

For AVS/RS, there are no standards (see, for example, FEM 9.581 for AS/RS) for estimating cycle time. Therefore, most recent studies on AVS/RS focused on performance estimation.

The cycle time is one of the main factors affecting the overall performance of the system. The performance indicators that will be elucidated in this study are as follows.

- Average service time of vehicles (i.e., shuttle, satellite, and lift)
- Vehicle utilization
- Average waiting time of ULs
- Average number of ULs in the queues
- Throughput of the system

AVS/RS is an expensive material handling technology. Monitoring and optimizing these performance indicators can help to choose a suitable number of vehicles, which directly impacts the cost of the entire system [1]. Most life cycle costs and capabilities are established when selecting device technologies, throughput rate, storage capacity, and rack configuration.

Most of the studies on AVS/RS focused on developing analytical or simulative models to estimate some of these system performance indicators alone, for example, the cycle time. Moreover, the literature rarely discusses deep-lane and multisatellite systems that are widespread in the food and beverage industry. Finally, most AVS/RS performance models are validated through numerical examples and not by applying them to real case studies. These topics are discussed in detail in Sect. 2.

This paper presents an original hybrid analytical-simulative model to estimate and control the performance of a deep-lane and multisatellite AVS/RS, that is, one equipped with multiple FS and NFS for each tier. This model is applied to a case study of the beverage industry conducting a what-if comparative and competitive multiscenario analysis. The analysis is based on realistic assumptions, filling the gap between the literature and real applications. A real AVS/RS inspires the data entry. Therefore, the sequence of storage and retrieval transactions and storage policy is not random.

Another contribution of this paper is a review of the literature on AVS/RS according to the introduction of 14 features grouped into five homogeneous groups: (1) rack configuration, (2) vehicle kinematics and configuration, (3) dispatching rules, (4) modeling approach, and (5) validation.

The remainder of this paper is organized as follows. Section 2 presents a literature review of the AVS/RS systems. Sections 3.1 and 3.2 illustrate the hierarchical approach and the sequence of activities required to perform a storage or retrieval transaction. Section 4 describes the

analytical-simulative hybrid model for estimating the cycle time, system throughput, average queue length, average waiting time, and vehicle utilization (i.e., lift, shuttle, and satellite utilization). Section 5 shows the application of the proposed model to a case study that consists of a real AVS/RS in the beverage industry. Finally, Sect. 6 presents the conclusions and directions for further research.

2 Literature review

Research studies on AVS/RS are relatively recent. First, Malmberg [2] illustrated analytical conceptualizing tools for evaluating AVS/RS performance with a tier-to-tier configuration. Over the last few decades, many studies have focused on AVS/RS and SBS/RS (shuttle-based storage and retrieval systems). These two technologies differ in terms of the type of unit load handled. The AVS/RS manages the palletized unit loads, whereas the SBS/RS runs unit loads represented by a tote (i.e., a plastic container). Therefore, this literature review mentions studies on both technologies.

Most studies on AVS/RS and SBS/RS focus on measuring system performance. The literature usually adopts two main approaches to estimate system performance: (1) dynamic modeling and (2) analytical modeling.

The first is based on dynamic modeling via simulation applied to different system configurations that operate according to randomly generated storage and retrieval transactions. Existing studies illustrate two main methods to simulate a generic AVS/RS or SBS/RS, according to the design environment [3]: (1.1) library-based simulation (LIBS) and (1.2) language-based simulation (LABS). The LIBS provides a set of routines that can be adopted with the support of a programming language. LIBS modeling is designed for a specific problem instance, that is, application, according to a Taylor-made approach. This simulative model can replicate the system dynamics according to a specific configuration with high accuracy, providing accurate results. Therefore, developing a simulation model for a large and complex system is highly time consuming and labor intensive. One of the most challenging issues is to support the general results and guidelines for the design, management, and optimization of AVS/RS. The LIBS models the system using commercial digital twin platforms such as Automod [4–6], Arena [1, 7–9], and Simulink [10].

Ekren [1] presented a simulation model to evaluate the performance of an AVS/RS under various predefined design scenarios. The model simulated 55 types of experiments. Marchet et al. [9] investigated the compound effect of the kinematic behavior of vehicles and lifts and the influence of rack configuration on performance assessment. Both studies explored the system performance using the commercial simulation software Arena. Marolt et al.

[10] developed a simulation model in Simulink, focusing on the storage efficiency of a shuttle carrier and relocation strategies in a tier-captive multideep AVS/RS (without considering the lifts). The study compared nine different storage and relocation assignment strategy combinations, focusing on the assessment of dual-command cycle time.

The LABS provides a set of utilities to develop simulation models. This LABS modeling involves the use of a programming language such as VB.NET [11] or MATLAB [12–14]. The generic simulation model is easily parameterizable. Therefore, it fits with several system configurations supporting what-if, multiscenarios, and competitive analysis. This dynamic model is suitable for obtaining general results. A significant example of this modeling approach was presented by D’Antonio and Chiabert [12]. They illustrated original analytical models to evaluate AVS/RS cycle time and throughput according to different rack layouts, different fleet compositions, and different types of cycles. Based on a MATLAB routine, the simulation model runs a complete experimental plan of 1980 scenarios (replicated 10 times each). This study aimed to advocate both the design and deployment phases of AVS/RS by enabling performance estimation in a wide variety of scenarios.

The second approach to estimate system performance is based on analytical models and aims to evaluate the expected cycle times and throughput of the system. These performances can also be validated through simulations. Typically, there are two methods of investigating an analytical approach [15]: (1) cycle time modeling and (2) queuing network modeling.

Cycle time models are based on the mathematical relationships between inputs and outputs. The results are reasonably accurate. The researchers that follow this analytical approach aim to evaluate the cycle time, travel distance, and vehicle utilization (i.e., shuttle, satellite, and lift) in a reasonable time. Moreover, the development is typically not time consuming. However, changing the initial assumptions may render the analytical model invalid. In addition, developing an analytical model for a complex system such as an AVS/RS is difficult.

Lerher et al. [16] presented an analytical travel time model to estimate cycle time in an SBS/RS considering the real operating characteristics of the lift and the shuttle carrier (e.g., vehicle acceleration and deceleration). Thus, Manzini et al. [11] developed analytic models to determine the traveled distance and cycle time of shuttles, given alternative layout configurations and considering multideep systems. The results were validated through simulations using VB.NET programming language. Lerher et al. [17] presented analytical travel time models for computation cycle times and the throughput performance of AVS/RS with multiple-tier shuttle vehicles.

These analytical cycle time models focus on service times, ignoring the interactions between vehicles (e.g., lift shuttle and shuttle satellite). Therefore, waiting times are usually ignored.

The second analytical approach focuses on queuing network (QN) models to evaluate the waiting time between lifts and shuttles (and vice versa) and the cycle time to perform storage and retrieval transactions. A QN is a collection of service stations organized in the sequence that customers visit to satisfy their service requirements [18]. There are three main types of QN: the open queuing network (OQN), closed queueing network (CQN), and semi-open queuing network (SOQN). An OQN is a network in which customers arrive at a specified rate from outside to one or more queues, waiting to be processed at a finite number of servers. A CQN differs from OQN because the number of customers circulating in the system is fixed. An SOQN is a variation of the first two queueing network systems: customers arrive at a specified rate, and the maximum number of customers that can be processed at any time is fixed. A synchronization station (the so-called semaphore) decouples the arriving customers from the network resources using a semaphore queue.

Many studies on AVS/RS and SBS/RS have adopted the SOQN. In particular, Tappia et al. [14] modeled each tier of the SBS/RS as a multiclass SOQN, whereas an open queue modeled the vertical transfer. According to Jia and Heragu [19], using an SOQN rather than an OQN or a CQN allows capturing the pairing between transactions and shuttles. This enables a better estimation of the shuttle utilization and transaction waiting time for an available shuttle. Deng et al. [7] presented a SOQN model with class switching to analyze the throughput time of a single-tier shuttle-based compact storage system by varying the number of shuttles, arrival rate, and depth/width ratio. Moreover, they modeled the parallel movement of shuttles and “transfer car” using a fork-join queuing network.

Thus, Marchet et al. [8] modeled an AVS/RS using a QN approach to quantify the waiting time component of the transaction cycle time. In this case, the selected QN was an OQN because it reduces the effort in the modeling procedure to be implemented and to fulfill the need to minimize the time required in the “conceptualization” phase of an AVS/RS design [20]. The approach followed by Marchet et al. [8] did not impose restrictions on the queue length. This is a strong initial assumption. Eder overcame this assumption in their recent studies [15, 21, 22] by modeling the SBS/RS with a continuous open queuing system with limited capacity. Recently, Kumawat and Roy [23] illustrated an original method for solving multistage SOQNs and validated it using a case study on a multitier shuttle-based compact storage system. Finally, they demonstrated the model accuracy by benchmarking their results with an existing approach in the literature.

QN models for AVS/RS start with the following assumptions, unlike real industrial applications.

1. The arrival and service times follow a Poisson distribution and an exponential distribution (Markov chain basic hypotheses). These parametric functions roughly simplify the real dynamics of arrival and service times. In industrial applications, incoming and outgoing ULs are frequently subject to seasonal trends because of the throughput of production systems producing ULs and demand requests generating retrieval transactions.
2. QN models assume that the queue length is infinite. Therefore, unit loads always wait at the input/output stations for the lift or shuttle. Consequently, the estimation of the maximum throughput of the system can be unrealistic. The shuttle and satellite at each tier are always involved in storage or retrieval transactions to achieve maximum throughput. In the long term, the output buffer or input buffer may be full. Therefore, the shuttle and satellite are forced to stop, and the real throughput of the system decreases.

The literature illustrates several tools and methods for calculating the performance of AVS/RS. Table 1 presents a complete review of the literature by the classification of the most significant contributions according to the features and related options previously mentioned.

A list of the features and related options identified and compared for each group of factors is as follows.

1. Rack configuration. The features that primarily affect the performance of the system are the number of tiers, lane depth, and location and capacity of I/O stations.

In the food and beverage industry, AVS/RS exploits the relevant vertical development. The performance of the system changes according to the number of tiers. This correlation is not linear. In particular, Marchet et al. [9] developed a simulation model using Arena to analyze the system performance, demonstrating that the maximum throughput does not necessarily correspond to the maximum possible rack height configuration. The lift was the bottleneck of the system in which the configuration had high racks. Therefore, vehicles created a bottleneck for a given storage capacity as the number of tiers was reduced. Consequently, a model that estimates system performance must consider all tiers and not a single-tier configuration. Most studies related to AVS/RS estimate system performance considering a single-tier system [7] or simply by multiplying the performance of a single tier by the number of tiers in the system [15, 21, 22].

Moreover, the AVS/RS is typically a deep-lane. However, studies on AVS/RS primarily discuss single-deep

lanes [2, 8, 16] or double-deep lanes [26, 27]. Few studies have assumed multideep lanes [11, 13]. The lane depth significantly affects the system performance, particularly the storage capacity, space area utilization, vehicle utilization, and system throughput. For instance, Lerher [26] proposed a model to calculate the performance of a double-deep SBS/RS, focusing on cycle time estimation. This study introduced the advantage of the double-deep system, which primarily concerns the reduction in the number of aisles, given the same storage capacity. This is a more efficient use of floor space than that in a single-deep lane system. This result can be extended to multideep lane systems to extrapolate the relationship between lane depth, number of aisles, and system performance. Finally, Manzini et al. [11] studied a multideep AVS/RS. They proposed alternative analytical models for different layout configurations, distinguishing the location of the I/O points and the shape ratio (which varies according to the number of aisles and lane depth).

As presented by Manzini et al. [11], the location of I/O stations influences the system performance. Most studies on AVS/RS allocate the I/O stations at the first tiers, usually one in front of the other to decongest the access and exit area to the warehouse. Krenczyk et al. [28] compared two simulation environments platforms (FlexSim and GPenSIM) and presented a case study involving an elevator as a part of an SBS/RS. Each tier had input and output locations split in front of the lifting table. The location of I/O stations in the first tier is commonly spread in the real AVS/RS.

I/O stations are also characterized by the loading capacity, which is the number of ULs in the buffer area waiting for handling devices. The station loading capacity is a physical limit of the system that indirectly affects the system performance. Ignoring it can result in an overestimation of the performance. Typically, studies on AVS/RS do not consider the buffer loading capacity. For instance, Fukunari and Malmborg [4] illustrated a network queuing approach for estimating performance measures. This requires assumptions, including the existence of a steady state and unlimited buffer capacity. Generally, all models based on the QN consider infinite capacities. Few studies have introduced the buffer capacity of I/O stations. In particular, Marchet et al. [9] developed a simulation model to evaluate SBS/RS performance. In the definition of the primary assumption, they limited the maximum number of totes to one UL in the queue at buffer out. This assumption made the retrieval cycle more realistic because each time a vehicle retrieves a tote and goes to the buffer out, it must wait till the buffer is full.

Table 1 An overview of the literature on AVS/RS and SBS/RS

Authors	Rack configuration			Vehicle kinematics			Vehicles configuration			Dispatching rules			Modeling approach			Validation	
	System	Depth	I/O area	Speed*	Cycles	Free satellite	Satellites per shuttle	Loading capacity	S/R policy	SKU	Analytical	Simulative	Tool	Objective	Application		
Malmberg (2002) [2]	AVS/RS TT	SD	FT	Yes	SC/DC	No	1 Sat*4	1	Random	No	No	/	State equation	No	CT, TH, U	Numerical example	
Kuo et al. (2007) [24]	AVS/RS TT	SD	FT, 1UL	Yes	SC	No	1 Sat*4	1	Random	No	No	/	QN	No	CT, U	Numerical example	
Fukunari and Malmberg (2009) [4]	AVS/RS TT	SD	FT, +UL	No	SC/DC	No	1 Sat*4	1	Random	No	Yes	AutoMod	QN	Yes	CT, U	Numerical example	
Ekren (2011) [1]	AVS/RS TT	SD	FT	Yes	SC/DC	No	1 Sat*4	1	Random	No	Yes	Arena	/	Yes	Costs	Case study	
Marchet et al. (2012) [8]	AVS/RS TC	SD	FT, 1UL	Yes	SC (R)	No	1 Sat*4	1	Random	No	Yes	Arena	OQN	Yes	CT, TH	Experimental results	
Marchet et al. (2013) [9]	AVS/RS TC	SD	FT, 1UL	Yes	SC (R)	No	1 Sat*4	1	Random	No	Yes	Arena	/	Yes	CT, TH	Case study	
Roy et al. (2015) [5]	AVS/RS TT	SD	FT	No	SC	No	1 Sat*4	1	Random	No	Yes	AutoMod	SOQN	Yes	CT, TH, U	Numerical example	
Lerher et al. (2015) [16]	SBS/RS TC	SD	FT, 1UL	Yes	SC/DC	No	1 Sat*4	1	Random	No	Yes	/	Travel time model	Yes	CT	Numerical example	
Roy (2016) [25]	AVS/RS TT	SD	FT	No	SC	No	1 Sat*4	1	Random	No	No	/	SOQN	No	CT, TH, U	/	
Manzini et al. (2016) [11]	AVS/RS TC	MD	FT	No	SC/DC	No	1 Sat	1	Random	No	Yes	VB.NET	Travel time models	Yes	CT	Numerical example	
Tappia et al. (2017) [14]	SBS/RS TC	MD	FT	No	SC	No	1 Sat	1	Random	No	Yes	MATLAB	SOQN	Yes	CT	/	
Roy et al. (2017) [6]	AVS/RS TT	SD	FT	No	SC	No	1 Sat*4	1	Random	No	Yes	AutoMod	SOQN	Yes	CT, TH, U	Numerical example	
D'Antonio et al. (2018) [13]	AVS/RS TT	MD	FT	Yes	SC/DC	Yes	1 Sat	1	Random, closest floor, closest channel	Yes	Yes	MATLAB	Analytical model	Yes	CT, TH	Numerical example	
D'Antonio and Chiabert (2019) [12]	AVS/RS TT	MD	FT	Yes	SC/DC	Yes	1 Sat	1	Random, closest floor, closest channel	Yes	Yes	MATLAB	Analytical model	Yes	CT, TH	Numerical example	
Eder (2020) [15]	SBS/RS TC	MD	FT, +UL	No	SC/DC*1	No	1 Sat*4	1	Random	No	Yes	/	Continuous-time OQN	Yes	TH	Optimization example	
Eder (2020) [21]	SBS/RS TC	SD	FT, +UL	No	SC/DC*1	No	1 Sat*4	1	Class-based	No	No	/	Continuous-time OQN	No	TH	Optimization example	
Eder (2020) [22]	SBS/RS TC	MD	FT, +UL	No	SC/DC*1	No	1 Sat*4	1	Random	No	Yes	/	Continuous-time OQN	Yes	CT, TH	Optimization example	
Lerher et al. (2021) [17]	AVS/RS TT	SD	FT, 1/2UL	Yes	SC/DC	No	1 Sat*4	1	Random	No	Yes	/	Analytical model	Yes	CT	Numerical example	

Table 1 (continued)

Authors	Rack configuration		Vehicle kinematics		Vehicles configuration			Dispatching rules		Modeling approach		Validation			
	System	Depth	I/O area	Speed*	Cycles	Free satellite	Satellites per shuttle	Loading capacity	S/R policy	SKU	Analytical	Simulative	Tool	Objective	Application
Marolt et al. (2022) [10]	AVS/RS TC	MD	FT, +UL	Yes #2	DC	No	1 Sat	1	Random, depth-first, nearest neighbour	No	/	Yes	Simulink	CT	Numerical example
Deng et al. (2021) [7]	SBS/RS TC	MD	FT, +UL	No	SC	Yes	N Sat	1	Random	No	SOQN+JFQN	Yes	Arena	TH, U	Case study
Present work TC	AVS/RS TC	MD	FT, 2/3UL	Yes	SC/DC	Yes	1/2/3 Sat	2	Case study*³	Yes	Analytical-simulative model	Yes	MATLAB	CT, U, TH	Case study

QN queueing network, SOQN semi-open QN, OQN open QN, JFQN join fork QN, FT first tier, TT tier-to-tier, TC tier-captive, SC single command, DC double command, CT cycle time, TH throughput, U vehicle utilization, SD single deep, MD multi-deep

*speed: Yes = the model considers velocity, acceleration, and deceleration; No = the model considers a constant velocity and ignores acceleration

¹lift operates only in single command; ²load and unload fix time are not realistic (1 s); ³storage and retrieval policy follows the real case study policy; it is aligned with the dynamic filling of the AVS/RS; *shuttle and satellite are the same vehicle; FT, first tier; UL* = buffer area's size, in terms of unit loads (JUL, single capacity; 2UL, double capacity; + UL, infinite capacity)

- Vehicle configuration. The features that influence the system performance are the number of satellites per shuttle and FS option. In AVS/RS, at least one satellite is assigned for each shuttle. Most studies on AVS/RS simplify the system transactions by defining a single vehicle per tier, representing both the shuttle and the satellite. In particular, [15, 21, 22] Eder recently considered the shuttle and satellite as a single vehicle and linked the number of vehicles to the number of tiers. Thus, Lerher [26] studied double-deep AVS/RS considering a single vehicle, called a shuttle carrier, to perform the task along the cross-aisle and within the lanes. Few studies have analyzed different vehicle configurations. For instance, D'Antonio et al. [12, 13] illustrated original analytical models capable of assessing the performance of a tier-to-tier AVS/RS considering an arbitrary number of shuttles and assigning only one satellite per shuttle. One of the contributions of the present study is the comparative analysis between scenarios with an arbitrary number of satellites per shuttle. An AVS/RS with a single vehicle that performs both operations is indirectly an AVS/RS with the NFS. Therefore, the shuttle must wait for the satellite while performing a task within a lane (and vice versa). Similar to the number of satellites, the FS/NFS configuration directly affects the system productivity. Existing studies on AVS/RS rarely consider this option. Some studies considered two vehicles assigned to each tier, one for the handlings along the cross-aisle and one within the lanes. However, they usually work together and expect each other. For instance, Tappia et al. [14] recognized two independent vehicles, the shuttle and the transfer car (i.e., the satellite), but the assumption was that both movements were sequential. Therefore, the two vehicles moved as one. No simultaneous operations were performed. Thus, Marolt et al. [20] considered two vehicles (shuttle and satellite), although they could not move independently from one another. In both cases, this is the NFS option. Recently, Deng et al. [7] modeled a single-tier specialized shuttle-based compact storage system considering a parallel processing policy. The parallel movement of vehicles was modeled using a fork-join queuing network.
- Vehicle kinematics. The most critical factors in vehicle kinematics are the velocity profile of the lifts, shuttles, and satellites; the type of command cycles that each vehicle can execute; and the vehicle loading capacity. Furthermore, many studies on AVS/RS ignore acceleration and deceleration [4, 5, 7] or assume a constant vehicle velocity [15, 21, 22]. Acceleration and deceleration significantly affect system performance, especially in short-distance routes.

There are two main types of vehicle routing to perform storage and retrieval transactions: the single com-

mand cycle (SCC) and the double command cycle (DCC). In the AVS/RS literature, some studies consider both SCC and DCC, defining a priori a percentage of transactions to execute as SCC and a percentage as DCC. For example, Lerher et al. [17] developed an analytical travel time model to compute cycle times in an AVS/RS system with multiple-tier shuttle vehicles. Other contributions have focused solely on SCC [7, 8, 25]. In particular, Marchet et al. [8] developed an analytical model to estimate AVS/RS performance considering retrieval SCC because retrieval transactions represent the most critical activities from an organizational viewpoint. Some studies have only considered the DCC. In particular, Marolt et al. [20] developed a discrete event simulation of multiple-deep AVS/RS with a tier-captive shuttle carrier capable of performing DCC because the vehicle performed more efficiently: it operated with fewer empty travels than an SCC.

Finally, the loading capacity of each vehicle is rarely discussed in the literature. Lerher et al. [17] illustrated a multiple-tier AVS/RS equipped with a lift that provided vertical movements. They considered two main lift installations: one and two lifting tables. Deng et al. [7] concluded their study by declaring that future research should investigate the system performance with transactions requiring more than one unit load and considering both single- and dual-command cycles.

4. Dispatching rules. The dispatching rules primarily concern vehicle allocation to S/R transactions, the management of different SKUs, and the storage policy.

Vehicle allocation to S/R transactions is a feature of tier-to-tier systems (i.e., systems where vehicles can move between tiers) or tier-captive systems with multiple shuttles or multiple satellites.

Most studies on AVS/RS focused on tier-captive systems, with one vehicle and one satellite per tier. In these systems, S/R transactions are typically performed according to the first-in-first-out (FIFO) policy. Therefore, an assignment rule is not required. Instead, in tier-to-tier systems or tier-captive systems with multishuttles or multisatellites, S/R transactions can be performed by several vehicles. Therefore, an assignment rule is not introduced. This paper presents a rule inspired by a real application, both for the assignment of the satellite to the shuttle and the UL to the shuttle.

In addition, S/R transactions involve one or more ULs, each of which is characterized by a specific SKU type (i.e., product type). Real AVS/RS manages hundreds of SKU types, and each lane is a mono-SKU. Consequently, the sequence of S/R transactions that handle a specific SKU type can involve the lanes assigned to the SKU type (single SKU lane assignment). Currently, few studies have addressed these topics. D'Antonio and

Chiabert [12] and D'Antonio et al. [13] considered four different SKUs in the illustrated case study, and Deng et al. [7] introduced the assumption that each storage lane holds one product. Furthermore, there are no standards available for storage assignment policy. Most studies on AVS/RS adopted a random storage policy [7, 15, 21, 22, 24, 25]. Models based on this storage assignment policy do not consider the real filling of a warehouse. Few studies have focused on this topic. In particular, D'Antonio and Chiabert [12] and D'Antonio et al. [13] introduced three storage criteria: random, closest floor, and closest channel. They compared the performance of these different storage policies by adopting a simulation model that assumed a starting time filling ratio of 50%.

5. Validation. Many studies have validated their models through numerical examples [4, 20] or experimental results [1]. Few studies have translated their numerical applications in real case studies [9]. The application of the proposed models to real AVS/RS or SBS/RS allows the estimation of system performance as realistic as possible.

The analytical-simulative hybrid model proposed in this study overcomes the limits of both analytical and simulative approaches. The analytical approach allows the evaluation of travel time to perform storage or retrieval transactions. The simulative approach aims to estimate the waiting times in agreement with the real trends of incoming ULs and real demand for retrieval transactions (outgoing ULs). This dynamic step, based on simulation, allows the parameterization of the variables of the system. It enables a what-if comparative and competitive multiscenario analysis owing to the parameterization of the variables of the system (e.g., the number of tiers, kinematics of vehicles, and number of satellites). Moreover, this language-based simulation overcomes the assumptions of QN models. In particular, the proposed model follows a data entry where a real AVS/RS inspires the “arrival time” of storage and retrieval transactions. Moreover, the buffer capacity (i.e., queue length) is finite and depends on the rack configuration. The obtained performance, for example, the system throughput, vehicle utilization, and WIP unit loads, is robust and realistic.

3 Methodology

3.1 Data-driven hierarchical approach

This section illustrates the novel adopted approach to model the selected deep-lane and multisatellite AVS/RS according to the evaluation of some critical performance.

This approach is hybrid, that is, mixed analytic and simulative. The aim is to measure multiple performance indicators, such as the throughput and productivity of the entire

system and the utilization of each material handling resource (shuttle, satellite, lift, and bay). These performance levels were significantly affected by the cycle waiting times and service times quantified for each vehicle. Many studies have focused on service times without interest in evaluating buffering areas corresponding to the queues of the handling system [11, 16]. In real AVS/RS, the queue time can represent a considerable percentage of the total time spent by each UL within the system [5]. Therefore, queue dynamics should not be neglected. Figure 2 summarizes the proposed hierarchical analytical-simulative hybrid approach to quantify the system performance and compare different system configurations.

As illustrated in Fig. 2, the starting point, which is one of the original contributions of this study, relates to the data entry definition (detailed in Sect. 4), inspired by real AVS/RS applications. The second step is the analytical travel time modeling for data entry service time determination. The third is the simulative time modeling, which focuses on the waiting time determination according to system capacity constraints at the buffering areas. The performance dashboard section details the evaluation of the key performance indicators (KPIs) in a what-if multiscenario environment.

This hybrid approach considers four time contributions to model storage and retrieval transactions:

1. Input lift (output lift) waiting time in storage (retrieval) transactions
2. Idle/waiting of shuttle and satellites
3. Input lift (output lift) service time in storage (retrieval) transactions
4. Service time of the shuttle and satellites to perform the storage or retrieval transaction in the location (x,y,z)

According to the storage and retrieval transaction data entry, the analytical model quantifies the minimum service times for shuttles, satellites, and lifts to perform each task. This time is obtained from applying a heuristic procedure that minimizes the total service time transaction. Subsequently, storage and retrieval transactions are processed by the simulative part of the proposed model, which estimates the waiting times for shuttles, satellites, and lifts. Finally, the model provides a performance dashboard that reports the system's average performance.

3.2 System configuration and transactions

In an AVS/RS, the cycle time of a generic transaction, throughput of the system, and vehicle utilization depend on the system configuration and the sequence of storage and retrieval tasks. In addition to the previously introduced rack

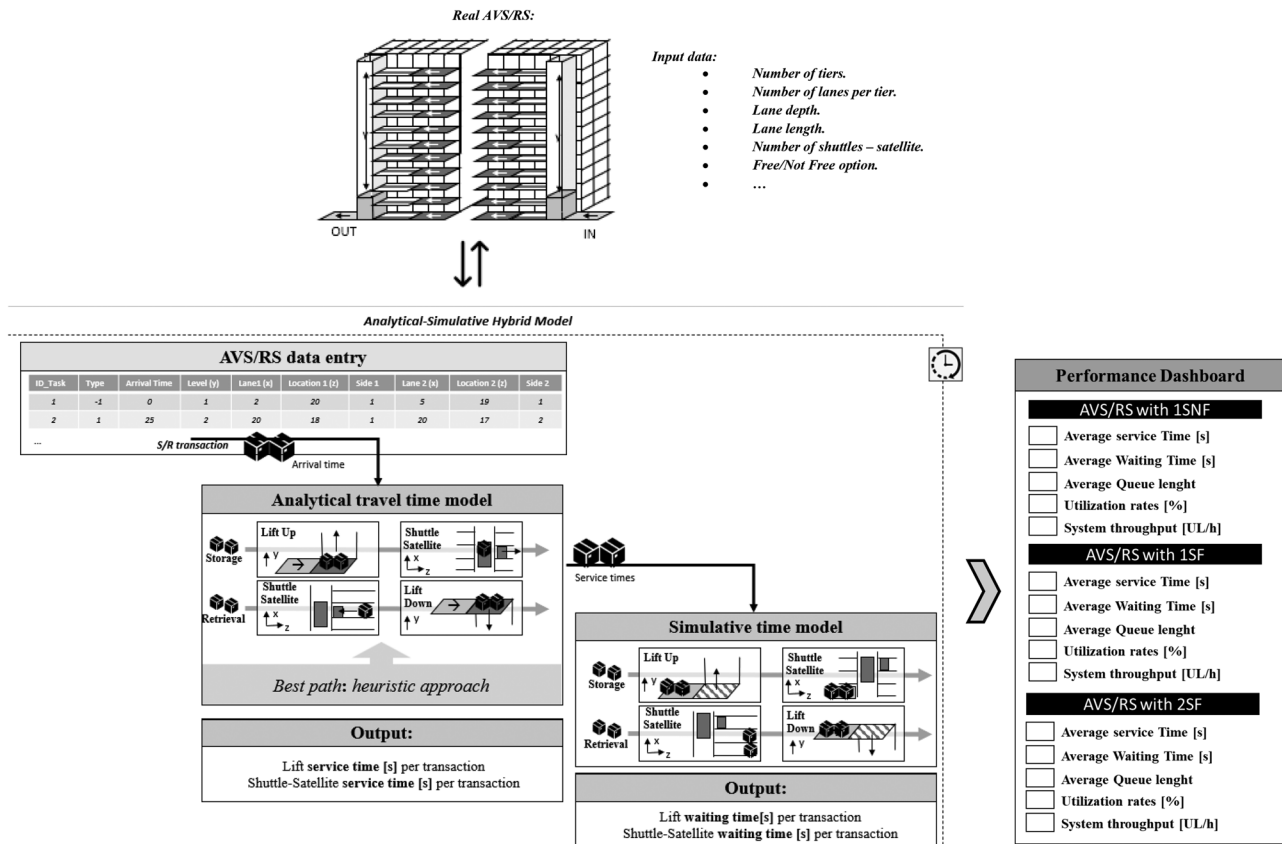


Fig. 2 The analytical-simulative hybrid approach

configuration features (see Table 1), the following issues critically influence the system performance, as demonstrated by industrial applications:

1. Number and loading capacity of lifts. This study assumes two lifts: one for storage (LiftS) and one for retrieval (LiftR) transactions. Both simultaneously can move two unit loads (i.e., a double-capacity lift).
2. Number and loading capacity of the bays. This study assumes two bays at each tier, one for input (BayI) and one for output (BayO). Moreover, there are two bays on the ground floor, corresponding to LiftS and LiftR. Each bay can accommodate up to three ULs.

Similarly, concerning the vehicle assumptions, these are the additional features:

1. Shuttle loading capacity. This study assumes one shuttle per tier that can handle up to two ULs simultaneously (i.e., a double-capacity shuttle). Therefore, each storage and retrieval transaction involves two ULs.
2. Satellite loading capacity. This study assumes that at least one satellite is assigned to each shuttle. The satellite can move one UL (i.e., a single-capacity satellite).

The number of free and non-free satellites has rarely been discussed in the literature. Therefore, this study aims

to study the impact of the number of satellites on the cycle time, measuring the throughput of the system and vehicle utilization. In particular, the previously introduced FS option allows the shuttle to release the satellite within a lane and perform other tasks, that is, movements, while decreasing idle waiting time. When a satellite goes back and forth within a lane, the shuttle can simultaneously perform operations, minimizing the total time needed to complete the transaction. The possibility of working with free satellites and their number generates the following three alternative scenarios:

1. 1NFS, i.e., AVS/RS TC with one NFS per shuttle
2. 1FS, i.e., AVS/RS TC with one FS per shuttle
3. 2FS, i.e., AVS/RS TC with two FS per shuttle

According to these scenarios and the other assumptions, the sequence of tasks to perform a storage or retrieval transaction can differ significantly. Hence, Table 2 illustrates the tasks to visit two storage locations (SL1 and SL2) at the assigned tier (AT). The ULs come from a production system, for example, a filling and packaging line producing single SKU ULs in the beverage industry.

Figure 3 exemplifies the storage activity, illustrated from the explicit task ID reported in Table 2. The (x, z) location of SL1 and SL2 is also represented. The lane is assumed to be a single SKU subject to LIFO unit load handling. The lane depth is equal to 6.

Table 2 The sequence of tasks to visit the storage locations in the three scenarios

<i>Storage task ID</i>	1NFS	1FS	2FS
(1)	The two ULs move from the production to the storage system	The two ULs move from the production to the storage system	The two ULs move from the production to the storage system
(2)	The ULs wait for LiftS	The ULs wait for LiftS	The ULs wait for LiftS
(3)	The ULs are loaded by LiftS and conducted to the AT	The ULs are loaded by LiftS and conducted to the AT	The two ULs are loaded by LiftS and conducted to the AT
(4)	The ULs move through the input conveyor to BayI	The ULs move through the input conveyor to BayI	The two ULs move through the input conveyor to BayI
(5)	The ULs wait for the shuttle	The ULs wait for the shuttle	The two ULs wait for the shuttle
(6)	The shuttle retrieves the satellite and moves to BayI	The shuttle decides whether to retrieve the satellite; then it moves to BayI	The shuttle decides whether to retrieve satellites; then it moves to BayI
(7)	The shuttle loads the ULs and moves to the nearest SL	The shuttle loads the ULs and moves to the nearest SL if the satellite is on board; otherwise, it moves to recover it and then to the nearest SL	The shuttle loads the ULs and, if at least one satellite is on board, it moves to the nearest SL; otherwise, the shuttle recovers it and then moves to the nearest SL
(8)	The shuttle releases the satellite with the first UL inside the lane and waits for it (NFS)	The shuttle releases the satellite with the first UL inside the lane and waits for it	The shuttle releases the satellite with the first UL inside the lane and decides whether to wait for it or retrieve another. If another satellite is already on board, it moves to the second SL
(9)	The shuttle moves to the other SL and releases the satellite (with the second UL) inside the lane	The shuttle moves to the second SL and releases the satellite (with the second UL) inside the lane	The shuttle moves to the second SL and releases the satellite (with the second UL) inside the lane



Fig. 3 Exemplifying a sequence of tasks for a storage cycle in the three scenarios

Similarly, Table 3 illustrates the tasks to visit the two retrieval locations (RL1, RL2) at the assigned tier (AT) to collect two unit loads (ULs).

Figure 4 exemplifies the retrieval process for the three alternative scenarios. Each task is related to the explicit task ID reported in Table 3.

Table 3 The sequence of tasks to visit the retrieval locations in the three scenarios

Retrieval task ID	INFS	IFS	2FS
(1)	The shuttle retrieves the satellite and moves to the nearest RL	The shuttle retrieves the satellite and moves to the nearest RL	The shuttle retrieves one or two satellites and moves to the nearest RL
(2)	The shuttle releases the satellite inside the lane and waits for it	The shuttle releases the satellite inside the lane and waits for it	The shuttle releases the satellite inside the lane
(3)	The shuttle retrieves the satellite with the first UL and moves to the second RL	The shuttle retrieves the satellite with the first UL and moves to the second RL	(a) The shuttle waits for the satellite, loads the first UL, and moves to the second RL (b) The shuttle retrieves another satellite and moves to the second RL
(4)	The shuttle releases the satellite inside the lane and waits for it	The shuttle releases the satellite inside the lane and waits for it	(a) The shuttle releases the satellite and waits for it (b) The shuttle releases the satellite and waits for it
(5)	The shuttle retrieves the satellite (with the second UL) and moves to the BayO	The shuttle retrieves the satellite (with the second UL) and moves to the BayO	(a) The shuttle loads the second UL and moves to the BayO (b) The shuttle loads the second UL, and if the first one is not on board, it moves to collect it and finally to the BayO
(6)	The shuttle downloads the two ULs at the BayO	The shuttle downloads the two ULs at the BayO	The shuttle downloads the two ULs at the BayO
(7)	The ULs move through the output conveyor and wait for the LiftR	The ULs move through the output conveyor and wait for the LiftR	The ULs move through the output conveyor and wait for the LiftR
(8)	The LiftR loads the ULs and brings them to the BayO on the ground floor	The LiftR loads the ULs and brings them to the BayO on the ground floor	The LiftR loads the ULs and brings them to the BayO on the ground floor
(9)	The ULs exit the LiftR	The ULs exit the LiftR	The ULs exit the LiftR

3.3 Free satellite option

The previous subsection illustrates how the sequence of operations changes according to the number of satellites adopted by each shuttle and the FS option. In particular, the examples in Figs. 3 and 4 show that the FS configuration has several implications. The most critical issues are as follows.

1. A generic free satellite can be on its shuttle or within a lane when it does not perform any transaction. If the satellite is not free, it is always on the shuttle when performing any task. Consequently, in the presence of the free option, the shuttle must decide whether to recover a satellite and which satellite recovers in the presence of more than one.
2. The FS option allows shuttles and satellites to perform different operations simultaneously. For example, the shuttle, after releasing the satellite inside a lane to carry out a storage transaction, can go to the input bay and load two new ULs for storage. While performing this task, the satellite completes the previous storage transaction within the lane. The proposed heuristic procedure explores this option to best perform generic transactions.

In practical applications, satellites are not typically used to load the ULs in the input bay. This assumption is maintained in this study. Moreover, the management of simultaneous operations is more complex for retrieval than storage transactions because the shuttle must recover the ULs collected by satellites before going to the output bay. When the shuttle waits for the satellite, it simply rejoins it. However, if it decides to execute another task while the satellite performs the retrieval, it has to postpone the collection of the satellite with the ULs on board.

3.3.1 Next task selection

This subsection illustrates a heuristic procedure to define the sequence of tasks executed by vehicles of a multisatellite AVS/RS according to throughput maximization. The approach is heuristic because of the hard computational complexity of such a sequencing and scheduling decisional problem that involves multiple tasks and multiple vehicles operating in a stochastic environment owing to the proposed data-driven approach inspired by real applications. In particular, the number of alternative options to schedule the following available tasks significantly increases in real applications.

With respect to the shuttle, the choice of whether to wait for the satellite relies on two factors: *travel time* (1) and *availability of the satellite* (2). These measures must be continuously evaluated whenever the shuttle has to decide to wait for the satellite because this choice directly impacts the global handling time of generic transactions. Therefore, it is

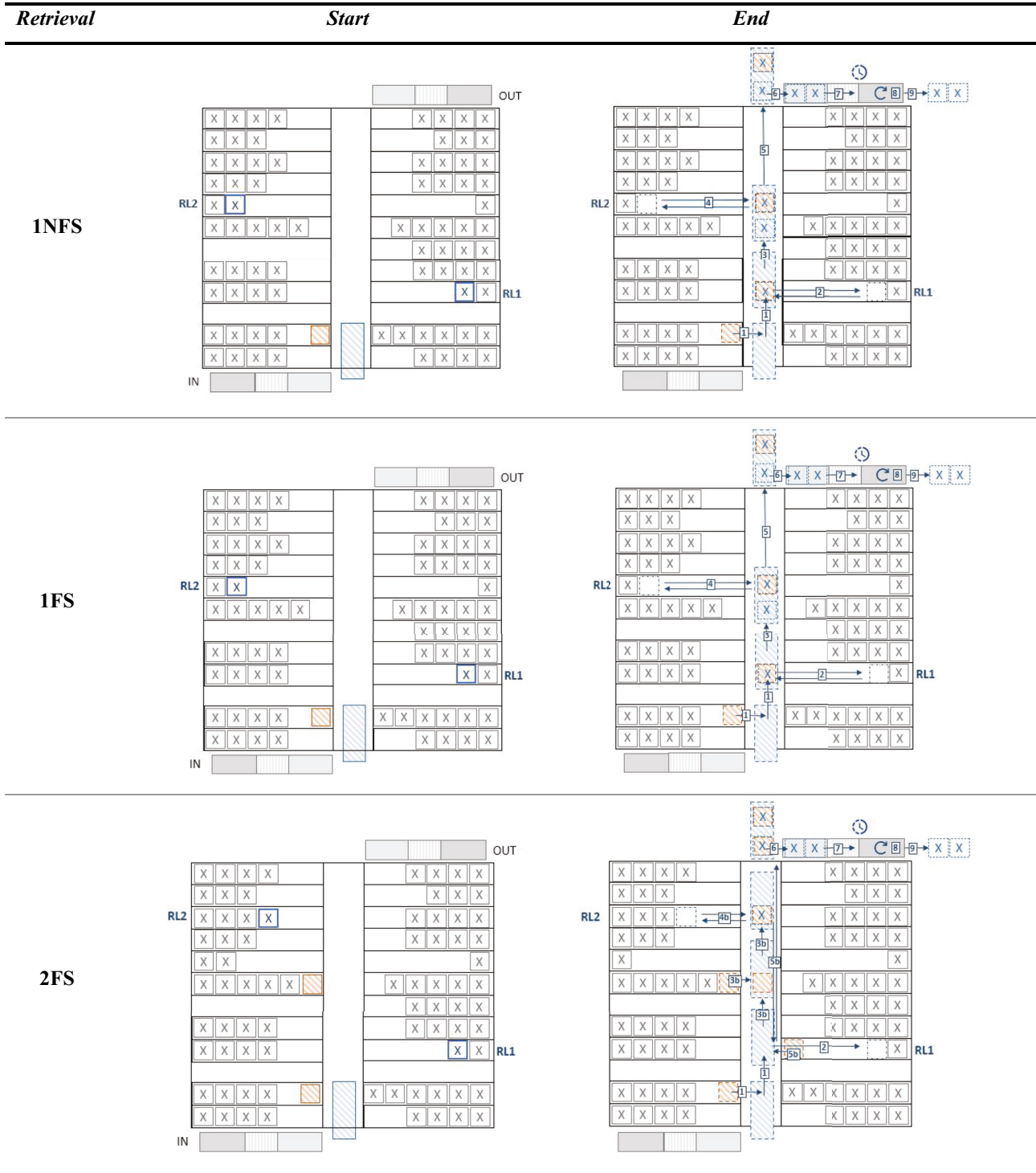


Fig. 4 Exemplifying a sequence of tasks for a retrieval cycle in the three scenarios

necessary to introduce three analytical time-based equations to support the measurement and monitoring of these two factors to select the best shuttle task to be executed.

Table 4 reports the notation adopted to quantify the following:

1. *Travel time.* Every time the shuttle releases a satellite into a lane, it can choose whether to perform another operation (e.g., recover another satellite) while the satellite is running a storage (or retrieval) transaction or is in idle time waiting for the satellite. The choice is the result of com-

Table 4 Analytical time models' notation, Eqs. (1)–(3)

Notation	Description
i	Operation i
nSh	Number of shuttles (equal to the number of tiers)
nSat	Number of satellites
$A_{sat(n)}$	Availability of the satellite n , with $n = 1, \dots, nSat$
$T_{free_sat(n)}$	The time the satellite will be free, at the interface with the cross-aisle, with $n = 1, \dots, nSat$
$T_{busy_sat(n)}$	The time the satellite is busy, with $n = 1, \dots, nSat$
y_i, x_i, z_i	y (tier), x (lane), z (location) coordinates of the operation i
$y_{sat(n)}, x_{sat(n)}, z_{sat(n)}$	y (tier), x (lane), z (location) coordinates of the satellite n , with $n = 1, \dots, nSat$
$y_{sh(k)}, x_{sh(k)}, z_{sh(k)}$	y (tier), x (lane), z (location) coordinates of the shuttle, with $k = 1, \dots, nSh$
t_{wait}	Time to wait for the satellite into the lane $x(i)$ [s]
t_{move}	Time to perform the following operation $(i + 1)$ [s]
v_{sat_e}, v_{sh_e}	Horizontal velocity of the satellite/shuttle when it is empty [m/s]
a_{sat_e}, a_{sh_e}	Horizontal acceleration/deceleration of the satellite/shuttle when it is empty [m/s ²]
v_{sat}, v_{sh}	Horizontal velocity of the satellite/shuttle when it is not empty [m/s]
a_{sat}, a_{sh}	horizontal acceleration/deceleration of the satellite/shuttle when it is not empty [m/s ²]
$t_{coupling}$	Time required to couple/uncouple the satellite (i.e., go up/down the shuttle)[s]
$t_{drop}/t_{collect}$	Time to drop/collect the UL from the satellite [s]

paring the satellite waiting time, named t_{wait} , with the time the shuttle would take to perform the subsequent operation, t_{move} . If t_{move} is lower than t_{wait} , the shuttle executes the next transaction. Otherwise, the shuttle waits for the satellite in front of the lane where the satellite is visiting. For example, in the AVS/RS with 2FS described in Fig. 5, shuttle k can wait for satellite n to retrieve the first UL or recover satellite j to perform the second retrieval. The two periods can be evaluated as the sum of different contributions: fixed times (e.g., $t_{coupling}$) and variable times affected by the traveling distance according to (x, z) locations, vehicle speed (v), and acceleration (a):

$$t_{move} = t \left[\left| x_{sh(k)} - x_{sat(j)} \right|, v_{sh_e}, a_{sh_e} \right] + t_{coupling} + t \left[\left| x_{sat(j)} - x_{i+1} \right|, v_{sh_e}, a_{sh_e} \right] \tag{1}$$

$$t_{wait} = t \left[z_i, v_{sat_e}, a_{sat_e} \right] + t_{collect} + t \left[z_i, v_{sat}, a_{sat} \right] + t_{coupling} \tag{2}$$

where $T[\Delta x, v, a]$ is the shuttle travel time to achieve the x -location along the cross-aisle, and $T[z, v, a]$ is the satellite travel time to achieve the z -location within the lane.

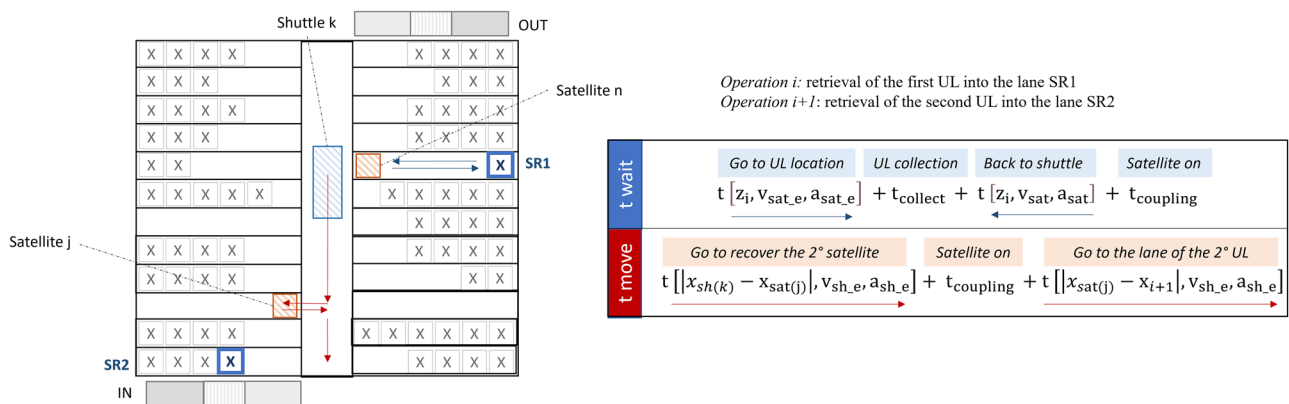


Fig. 5 Comparison between the time awaiting the first satellite (t_{wait}) and the time to recover another one (t_{move})

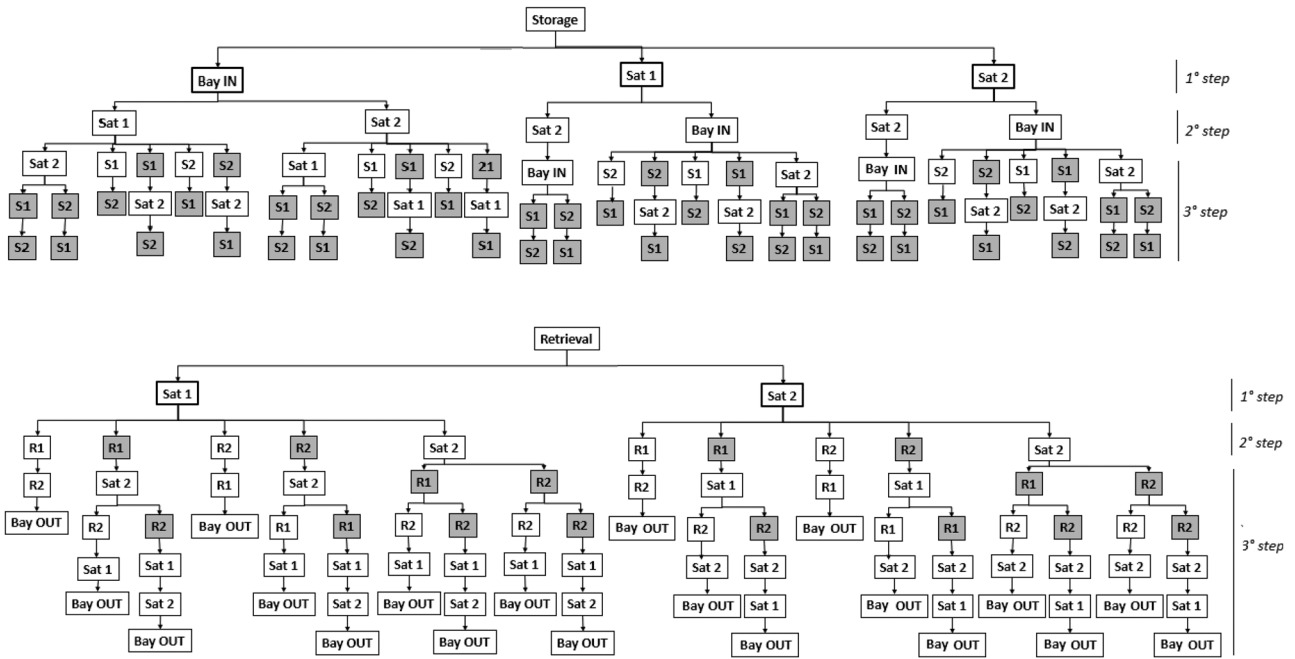


Fig. 6 Storage and retrieval decision tree diagram (note: blocks with colored background represent operations that can be performed at the same time)

2. *Availability of satellites.* The shuttle can retrieve a satellite if it is not involved in another operation. The binary variable $A_{sat(n)}$ measures whether satellite n is available. Satellite n is defined as available if it is free to be loaded by shuttle k when the shuttle k arrives in front of the lane of satellite n .

$$A_{sat(n)} = \begin{cases} 1 & \text{if } t \left[|x_{sh(k)} - x_{sat(n)}|, v_{sh_e}, a_{sh_e} \right] \geq T_{free_{sat(n)}} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

For example, in the AVS/RS with 2FS described in Fig. 5, $A_{sat(j)}$ is 1 if satellite j will be available to be loaded on the shuttle when shuttle k is ready to retrieve it.

The proposed decision-making process for “next task selection” is based on the following steps for storage transactions:

1. Shuttle moved to the input bay to load the two ULs. Meanwhile, it decides whether to retrieve one or two satellites.
2. Shuttle loads the two ULs and retrieves the nearest one if there is no satellite onboard. If it only has one satellite onboard, it decides whether to retrieve another satellite. Finally, it moved to the nearest SL.
3. After releasing satellite into the lane, shuttle decides whether to wait or retrieve another for the subsequent storage transaction.

Similarly, the following is for retrieval transactions:

1. Depending on the number of satellites on board, shuttle can choose to retrieve the nearest one. Finally, it moved to the nearest RL.
2. Shuttle releases satellite in the destination lane (x) and decides whether to wait in RL or move to retrieve another UL.
3. Before going to the output bay, shuttle retrieves satellites with the recovered ULs (unless they are already on board).

These steps reduce the complexity of multisatellite management for both storage and retrieval transactions. The number of alternative options is significantly increased in real applications. For example, in a tier-captive AVS/RS with two free satellites per shuttle, the alternative paths for storage transactions in SCC are 28. For retrieval transactions in SCC, they amount to 20, whereas for transactions in DCC, there are 640 alternative decisional paths. Figure 6 illustrates the decision tree diagram for an SCC of storage and an SCC of retrieval in an AVS/RS with 2FS. The previously introduced three steps are explicitly reported as the 1st, 2nd, and 3rd steps.

The high level of complexity of this sequencing and scheduling problem requires an original heuristic approach, whose aim is to control the system throughput. Figure 7 illustrates this heuristic procedure to dynamically select the next task for the shuttle and its satellites at the generic tier of an AVS/RS.

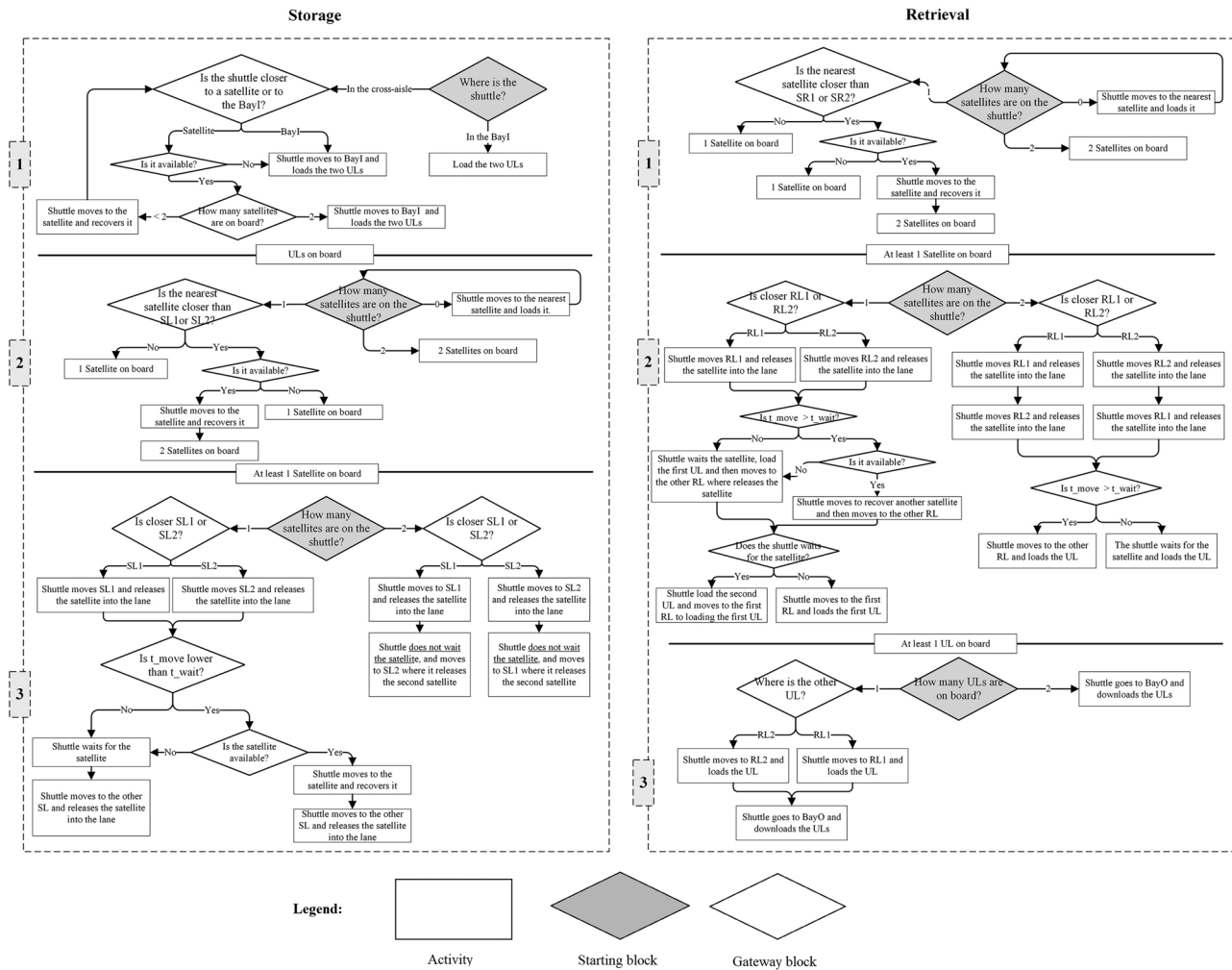


Fig. 7 A heuristic approach for storage and retrieval transactions with free satellite

4 Analytical-simulative hybrid model

The proposed data-driven hybrid model for designing and controlling a multishuttle AVS/RS is analytical and simulative. In particular, data entry is made of a chronological sequence of storage and retrieval transactions. The generic transaction record refers to the (x,y,z) system location and is expressed as follows:

1. Date and time (e.g., 29/07/2021, 12:00:03)
2. SKU name (for each mission, the type of the two ULs is the same)
3. Tier (yS1/yR1 and yS2/yR2)
4. Lane (xS1/xR1 and xS2/xR2)
5. Storage location (zS1/zR1 and zS2/zR2)

The input file can be generated following a random storage policy or by extrapolating data from a real AVS/RS. Section 6 illustrates how the system performance changes in line with the origin of the data entry.

According to the transaction arrival time, the model processes the storage and retrieval missions by considering both single and double command cycles. The type of cycle depends on the sequence of the incoming transactions.

Table 5 introduces the notation of the proposed hybrid model.

Section 4.1 illustrates the travel time analytical equations for cycle time determination. Another contribution to cycle time is the result of the application of a dynamic model measuring the waiting times (see Subject. 4.2).

4.1 Cycle time analytical model

The following equations of motion assume constant acceleration and deceleration for the generic vehicle (i.e., shuttle, satellite, and lift). Two types of velocity profiles can be distinguished depending on whether the obtained velocity peak $v(t_p)$ at time t_p is less than v_{max} (type 1) or equal to v_{max} (type II). For clarity,

Table 5 Analytical-simulative hybrid model notation

Parameters	Description
nTier	Number of tiers in the warehouse
hTier	Height of each tier (y) [m]
nLane	Number of lanes at each side of the main cross-aisle
wLane	Width of each lane (x) [m]
nPP	Number of locations for palletized unit loads (UL) within a lane
lPP	Length of each location (z) [m]
lConv	Length of the input/output conveyor (z) [m]
nSh	Number of shuttle (equals to the number of tiers; the system is tier captive)
$x_{sh(k)}$	x (lane) coordinate of the shuttle k, with $k = 1, \dots, nSh$
nSat	Number of satellites
$x_{sat(n)}, z_{sat(n)}$	x (lane), z (location) coordinates of the satellite n, with $n = 1, \dots, nSat$
nLiftS	Number of Lifts up (dedicated to storage transactions)
$y_{ls(j)}$	y (tier) coordinate of the lift-up j, with $j = 1, \dots, nLiftS$
nLiftR	Number of Lifts down (dedicated to retrieval transactions)
$y_{lr(p)}$	y (tier) coordinate of the lift-down p, with $p = 1, \dots, nLiftR$
cSh	Shuttle's loading capacity: single, double capacity [#UL]
cSat	Satellite's loading capacity: single, double capacity [#UL]
cBay	Bay's loading capacity: single, double capacity [#UL]
$y_{BayI}, x_{BayI}, z_{BayI}$	y(tier), x (lane), z (location) coordinates of input bay
$y_{BayO}, x_{BayO}, z_{BayO}$	y(tier), x (lane), z (location) coordinates of output bay
v_{lift_e}, v_{lift}	Maximum Lift speed without/with load [m/s]
v_{sh_e}, v_{sh}	Maximum Shuttle speed without/with load [m/s]
v_{sat_e}, v_{sat}	Maximum Satellite speed without/with load [m/s]
v_{conv}	Maximum Conveyor speed [m/s]
a_{lift_e}, a_{lift}	Lift acceleration/deceleration without/with load [m/s ²]
a_{sh_e}, a_{sh}	Shuttle acceleration/deceleration without/with load [m/s ²]
a_{sat_e}, a_{sat}	Satellite acceleration/deceleration without/with load [m/s ²]
$t_{coupling}$	Fixed coupling/uncoupling time [s]
t_{load}/t_{unload}	Fixed load/unload time of shuttles and lifts [s]
$t_{collect}/t_{drop}$	Fixed load/unload time of satellite [s]
y_{Sj}, x_{Sj}, z_{Sj}	Tier (y), lane (x), location (z) for the storage of the j UL, with $j = 1, 2$
y_{Rj}, x_{Rj}, z_{Rj}	Tier (y), lane (x), location (z) for the retrieval of the j UL, with $j = 1, 2$

v_{max} of the shuttle equals v_{sh} if the shuttle is full; otherwise, it is v_{sh_e} . The v_{max} of the satellite equals v_{sat} if the satellite is full; otherwise, it is v_{sat_e} . The v_{max} of the lift equals v_{lift} if the lift is full; otherwise, it is v_{lift_e} . The two velocity profiles are as follows:

1. Type 1: if the obtained peak velocity $v(t'_p)$ is less than v_{max} , the time to reach $v(t'_p)$ is $t'_p = \frac{v}{a}$ [s], and the distance traveled after the time t_p is equal to $x_p = \frac{1}{2}at_p'^2$ [m]. Consequently, the time spent traveling for the total distance $d = 2x_p$ is equal to $t_{total} = 2\sqrt{\frac{d}{a}}$ [s].
2. Type 2: if the obtained peak velocity $v(t_p)$ is equal to v_{max} , the time to reach $v(t_p)$ is equal to $t_p = \frac{v}{a}$ [s], and the distance traveled after time t_p is equal to $x_p = \frac{1}{2}at_p^2$ [m]. Consequently, the time spent traveling for the total distance $d = 2x_p + \Delta x$ is equal to $t_{total} = 2t_p + \frac{\Delta x}{v}$ [s].

The velocity profiles are illustrated in Fig. 8:

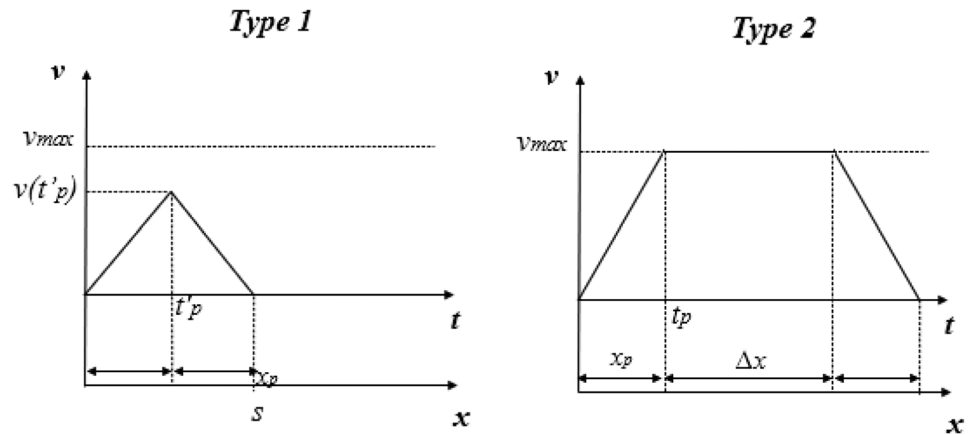
The motion equations depend on velocity (v), acceleration (a), and distance (d) traveled. Therefore, the time contribution (t) to perform each task is expressed considering these three elements: $t[d, v, a]$.

The time contributions shown for storage transactions in Table 6 and for retrieval in Table 7 are effective for the AVS/RS system with 1NFS. In scenarios with one or two free satellites, the sequence of operations can change. We decided not to illustrate these scenarios.

4.2 Time-dependent simulative model

The proposed hybrid approach requires a simulation model to analyze the interactions between the different entities of the system (i.e., shuttles, satellites, and lifts) and quantify

Fig. 8 Velocity profiles



waiting times and the number of items in the queues. The simulation model is a discrete event simulation based on a MATLAB routine. Figure 9 illustrates the simulation model adopting a notation from the queueing network theory.

The simulation is powered by the data entry inspired by the real AVS/RS system. Storage and retrieval transactions are ordered chronologically and dispatched among the tiers (in number). The simulation is time-driven. Therefore, the transactions arrive at the system according to the arrival time declared in the data entry.

Storage transactions (S) arrive at the interface with the climb elevator (node lift-up in Fig. 9), waiting for the resource in queue Qs. Each storage transaction is composed of one or two items that need to be brought to the assigned tier (AT); if the items have two different ATs, the elevator stops at the nearest one and then moves to the other AT. The lift-up waits

for the following storage request in the Q_{LU_p} queue, whereas the items wait for the shuttle in the queue Q_{SR_i} at tier i .

The retrieval transactions (R) arrive directly to tier i (e.g., R1 at tier 1), waiting for the resource in queue Q_{SR_i} . Each retrieval transaction is composed of one or two items that need to be collected. At each tier, the storage and retrieval transactions are processed by shuttles and satellites. Every time the resource turns free, another transaction can be performed. For retrieval transactions, after being processed by the shuttle and satellite, the items wait for the elevator (node lift-down in Fig. 9) in queue Q_R . When the lift-down is free, it moves to the tier of the retrieval request, in agreement with the FIFO policy, and simultaneously loads one or two unit loads. Then, it moves to the ground floor and downloads the items.

The service times of lift, shuttle, and satellite for each storage and retrieval transaction are provided by the analytical model

Table 6 Storage travel time in an AVS/RS with INFS

	Description	Equation
Lift-up	Time to come back to the input bay on the ground floor	$t \left[\left y_{ls(j)} - y_{Bay1} \right , v_{lift_e}, a_{lift_e} \right]_+$
	Time to load the two UL and go to the AT	$t_{load} + t \left[\left y_{Bay1} - y_{AT} \right , v_{lift}, a_{lift} \right]_+$
	Time to go to the BayI through the conveyor	$t \left[Conv, v_{Conv} \right]$
Shuttle satellite	Time to join with the satellite and go to BayI	$t_{coupling} + t \left[\left x_{sh(k)} - x_{Bay1} \right , v_{sh_e}, a_{sh_e} \right]_+$
	Time to load the first UL, move along x-direction of one location, and load the second UL	$t_{load} + t \left[w_{lane}, v_{sh}, a_{sh} \right]_+ t_{load} +$
	Time to reach the nearest Storage Location: S1	$t \left[\left x_{Bay1} - x_{S1} \right , v_{sh}, a_{sh} \right]_+$
	Time to release the satellite, to achieve the right location, drop the UL and come back	$t_{coupling} + t \left[z_{S1}, v_{sat}, a_{sh} \right] + t_{drop} + t \left[z_{S1}, v_{sat_e}, a_{sh_e} \right]_+$
	Time to join with the satellite and go to the second Storage Location: S2	$t_{coupling} + t \left[\left x_{S1} - x_{S2} \right , v_{sh}, a_{sh} \right]$
Lift-down	Time to release the satellite, to achieve the right location, and drop the UL	$t_{coupling} + t \left[z_{S2}, v_{sat}, a_{sh} \right] + t_{drop}$

Table 7 Retrieval travel time in an AVS/RS with 1NFS

Description		Equation
Lift-up		
Shuttle satellite	Time to join with the satellite and go to the nearest Retrieval Location: R1	$t_{\text{coupling}} + t \left[\left x_{\text{sh}(k)} - x_{R1} \right , v_{\text{sh}_e}, a_{\text{sh}_e} \right] +$
	Time to release the satellite, to achieve the right location, collect the UL and come back	$t_{\text{coupling}} + t \left[z_{R1}, v_{\text{sat}_e}, a_{\text{sh}_e} \right] + t_{\text{collect}} + t \left[z_{R1}, v_{\text{sat}}, a_{\text{sh}} \right] +$
	Time to join with the satellite and go to the second Retrieval Location: R2	$t_{\text{coupling}} + t \left[\left x_{R1} - x_{R2} \right , v_{\text{sh}}, a_{\text{sh}} \right] +$
	Time to release the satellite, to achieve the right location, collect the UL and come back	$t_{\text{coupling}} + t \left[z_{R2}, v_{\text{sat}_e}, a_{\text{sh}_e} \right] + t_{\text{collect}} + t \left[z_{R2}, v_{\text{sat}}, a_{\text{sh}} \right] +$
	Time to join with the satellite and go to the output	$t_{\text{coupling}} + t \left[\left x_{R2} - x_{\text{BayO}} \right , v_{\text{sh}}, a_{\text{sh}} \right]$
Lift-down		
Shuttle satellite	Time to come back to the AT	$t \left[\left y_{\text{lr}(p)} - y_{\text{AT}} \right , v_{\text{lif}_e}, a_{\text{lif}_e} \right]$
	Time to go to the LiftR through the conveyor	$t \left[l_{\text{Conv}}, v_{\text{Conv}} \right] +$
	Time to load the UL at BayO and go to the ground floor	$t_{\text{load}} + t \left[\left y_{\text{AT}} - y_{\text{BayI}} \right , v_{\text{lif}_t}, a_{\text{lif}_t} \right] + t_{\text{unload}}$

previously described according to the proposed mixed analytical and simulative approach. Therefore, the results generated by the simulation model vary according to the configuration selected for the analytical model (i.e., 1NFS, 1FS, and 2FS).

5 Case study

This section presents the results of the application of the proposed analytical-simulative hybrid model to aid the design of a new AVS/RS system. This case study was inspired by a real application in the beverage industry.

What is the number of satellites and which free satellite option best performs the system throughput? The proposed approach is suitable for addressing this challenge in a comparative and competitive multiscenario analysis. Table 8 summarizes the features and technological constraints in 1NFS, 1FS, and 2FS. The resulting storage capacity of the system is 6080 unit loads (i.e., $\{32 \text{ [lane/dx]} * 22 \text{ [UL/lane]} + 32 \text{ [lane/sx]} * 16 \text{ [UL/lane]}\} * 5 \text{ [tier]}$).

For clarity, the layout of the multideep AVS/RS proposed in this case study is reported in Fig. 10.

Fig. 9 The time-dependent simulation model

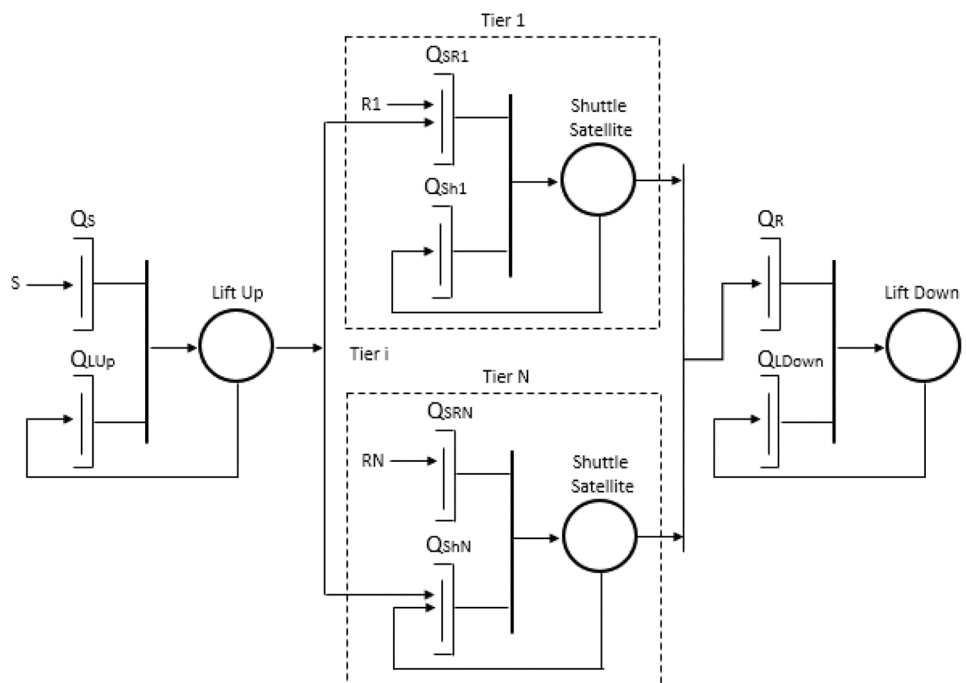


Table 8 Features and technological constraints

Parameter	Value
nTier	5 [tier]
hTier	2.25 [m]
nLane	32 [lane]
wLane	1.57 [m]
nPP	16 left [UL] to 22 right [UL]
lPP	1.90 [m]
lConv	2.25 [m]
nSh	5 [1 per tier]
$x_{sh(k)}$	$x_{sh(k)} = x_{BayI} \forall k$
nSat	5 [1 per shuttle] in 1NSF, 1FS 10 [2 per shuttle] in 2FS
$x_{sat(n)}, z_{sat(n)}$	$x_{sat(n)}, z_{sat(n)} = x_{BayI}, z_{BayI} \forall n$
nLiftS	1
$y_{ls(j)}$	$y_{ls(j)} = y_{BayI} \forall j$
nLiftR	1
$y_{lr(p)}$	$y_{lr(p)} = y_{BayO} \forall p$
cSh	2 [PP]
cBay	2 [PP]
cSat	1 [PP]
$y_{BayI}, x_{BayI}, z_{BayI}$	$(y_{BayI}, x_{BayI}, z_{BayI}) = (0;0;0)$
$y_{BayO}, x_{BayO}, z_{BayO}$	$(y_{BayO}, x_{BayO}, z_{BayO}) = (0;0;0)$
v_{lift_e}	1 [m/s]
v_{lift}	1 [m/s]
v_{sh_e}	3 [m/s]
v_{sh}	3 [m/s]
v_{sat_e}	1.2 [m/s]
v_{sat}	1 [m/s]
v_{conv}	0.30 [m/s]
a_{lift_e}	0.35 [m/s ²]
a_{lift}	0.35 [m/s ²]
a_{sh_e}	1.2 [m/s ²]
a_{sh}	0.8 [m/s ²]
a_{sat_e}	1.2 [m/s ²]
a_{sat}	1 [m/s ²]
$t_{coupling}$	5 [s]
t_{load}/t_{unload}	1 [s]
$t_{collect}$	6.5 [s]
t_{drop}	7.5 [s]

5.1 Analysis 1: real data entry

The data entry selected for this analysis is a list of chronologically ordered storage and retrieval historical transactions for an existing unit load storage system made of five tiers. The following five features characterize generic transactions:

1. Arrival time that defines the sequence of transactions
2. Transaction type (storage or retrieval)
3. y-level assigned to the transaction

4. x-lane assigned to the transaction
5. z-storage location (z) where the UL are collected or dropped

The real data entry counts 23,662 storages and 25,976 retrievals corresponding to a horizon time of 1 month. The number of SKUs was 140.

The proposed hybrid model processes the storage and retrieval transactions following this data entry and quantifies several KPIs. Table 9 summarizes some relevant KPIs such as the average service time (service), average waiting time (waiting), and impact of waiting time on total time (% wait) for a storage or retrieval transaction. Values are reported for each tier of the AVS/RS and each configuration (1NFS, 1FS, and 2FS). The temporal contribution of the lift was not reported because it did not change according to the selected system configuration. Therefore, it does not directly influence the system performance at tier.

The comparison between the three scenarios shows that introducing the “free option” and increasing the number of satellites minimize service and waiting times. In 1FS configuration, the average service time required by the shuttle and the satellite to perform the operations is approximately 10% lower for storage and retrieval tasks than the 1NFS configuration. In addition, if the number of FS increases to two (2FS), the average service time decreases by 15% compared to 1NFS and by 6% compared to 1FS.

Thus, the waiting time and impact on the total time also decreased from the 1NFS to the 2FS configuration. The average waiting time of configuration 2FS decreases by 26% compared to the 1NFS configuration and by 21% compared to the 1FS configuration. In general, the impact of waiting time on total time is significant, depending on the data entry. This demonstrates the effectiveness of the proposed hybrid approach that does not assume random arrival times and random storage locations.

The previous observations were related to the average values. The proposed model allows data collection on the distribution of service and waiting times for all processed storage and retrieval transactions. Tables 10 and 11 report the distributions of the following:

1. Service time: the time required by the shuttle and satellite to perform a storage or retrieval transaction.
2. Waiting time: the waiting time for a storage (retrieval) transaction in the input buffer at the tier (in the location within the lane).
3. Throughput time: the total time required to perform a storage (retrieval) transaction. For storage transactions, it is the time interval between the UL arriving at the input buffer of the lift-up and the UL being dropped into the storage location. For retrieval transactions, it is the time interval between the UL being

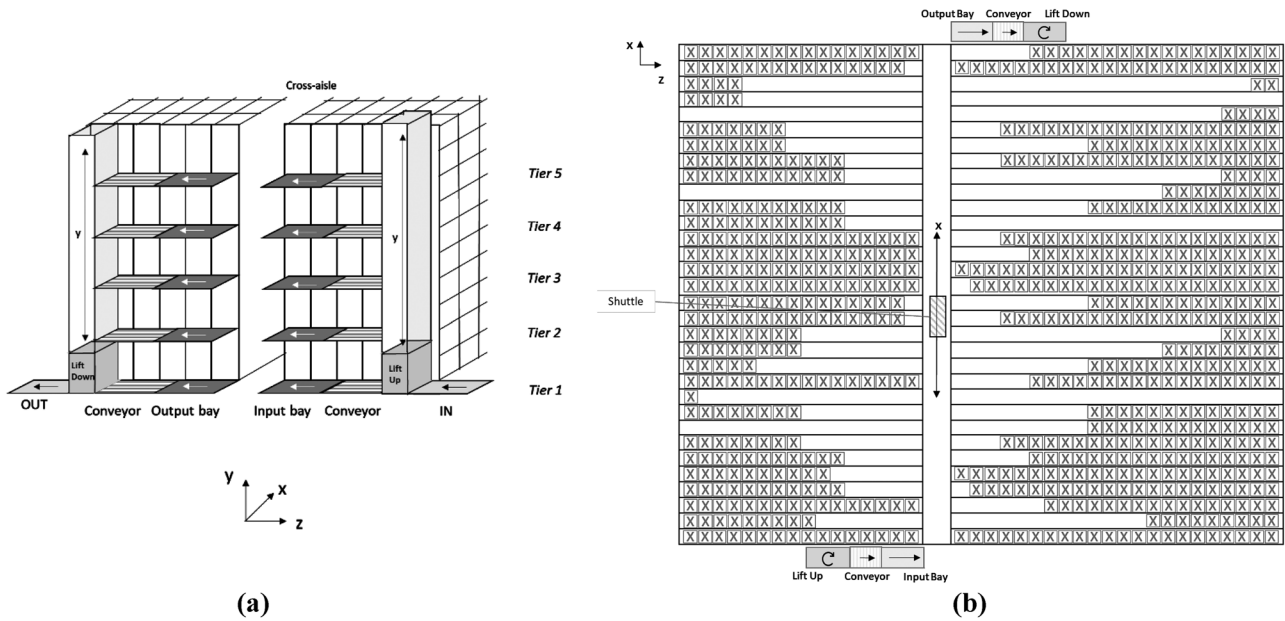


Fig. 10 a AVS/RS front view, b AVS/RS plan view

collected and downloaded in the output bay on the ground floor.

The proposed model quantifies the average number of items in the generic queue. Table 12 presents the distribution of ULs at the interface with the input lift (LiftS), at the input bay of each tier (Tier), and at the interface with the output lift (LiftR). The number of items varies in a range of 0–3 ULs, and it is reported for the three configurations selected (1NFS, 1FS, and 2FS).

Table 12 demonstrates that the bottleneck in the system is the capacity of the buffer located at the point of output in each tier, which is typical of such a storage system. The column “LiftR” illustrates how the percentage related to the probability of having 0 UL in the output buffer decreases by 15% (from 72 to 57%) in the configuration with 2FS compared with 1NFS. By increasing shuttle and satellite performance on the tier, the output buffer fills up faster. The output buffer has a finite capacity due to the physical limits of the warehouse. Consequently, a performance

improvement of shuttle and satellite handling positively affects the system throughput on one hand and generates bottlenecks on the other. It is essential to control bottlenecks and throughput simultaneously as in real operating conditions. Typically, production systems produce large batches of single SKU ULs, located in dedicated and homogeneous lanes. Retrieval transactions have to be executed as quickly as possible to guarantee high levels of customer service. The storage and retrieval transactions sequence is variable with a day, a week, a month, and a year.

The proposed hybrid and data-driven approach allows the time spent by each vehicle moving back and forth within the warehouse and the idle times to be tracked. Table 13 illustrates the utilization of the input lift (LiftS), shuttle and satellite (tier 1–tier 5), and the output lift (LiftR) according to the free/not-free option and number of satellites.

The utilization of LiftS and LiftR is lower than those of the shuttle and satellite. This result supports the previous statement regarding the output buffer, which is declared as the critical queue in this system.

Table 9 Average service times and waiting times of shuttle and satellite/s for storage and retrieval tasks

	1NFS			1FS			2FS		
	Service	Waiting	% wait	Service	Waiting	% wait	Service	Waiting	% wait
Tier 1	146.93 s	84.02 s	36%	133.10 s	72.01 s	35%	124.16 s	59.92 s	33%
Tier 2	146.90 s	75.17 s	34%	132.39 s	64.32 s	33%	124.63 s	47.19 s	27%
Tier 3	147.35 s	75.77 s	34%	133.57 s	64.23 s	32%	125.86 s	35.57 s	22%
Tier 4	147.76 s	77.50 s	34%	133.41 s	66.19 s	33%	125.87 s	36.83 s	23%
Tier 5	148.06 s	78.53 s	35%	133.98 s	66.84 s	33%	126.01 s	36.83 s	23%
Average	147.42 s	78.19 s	35%	133.31 s	66.72 s	33%	125.30 s	43.27 s	26%

Table 10 Distribution of times for storage transactions

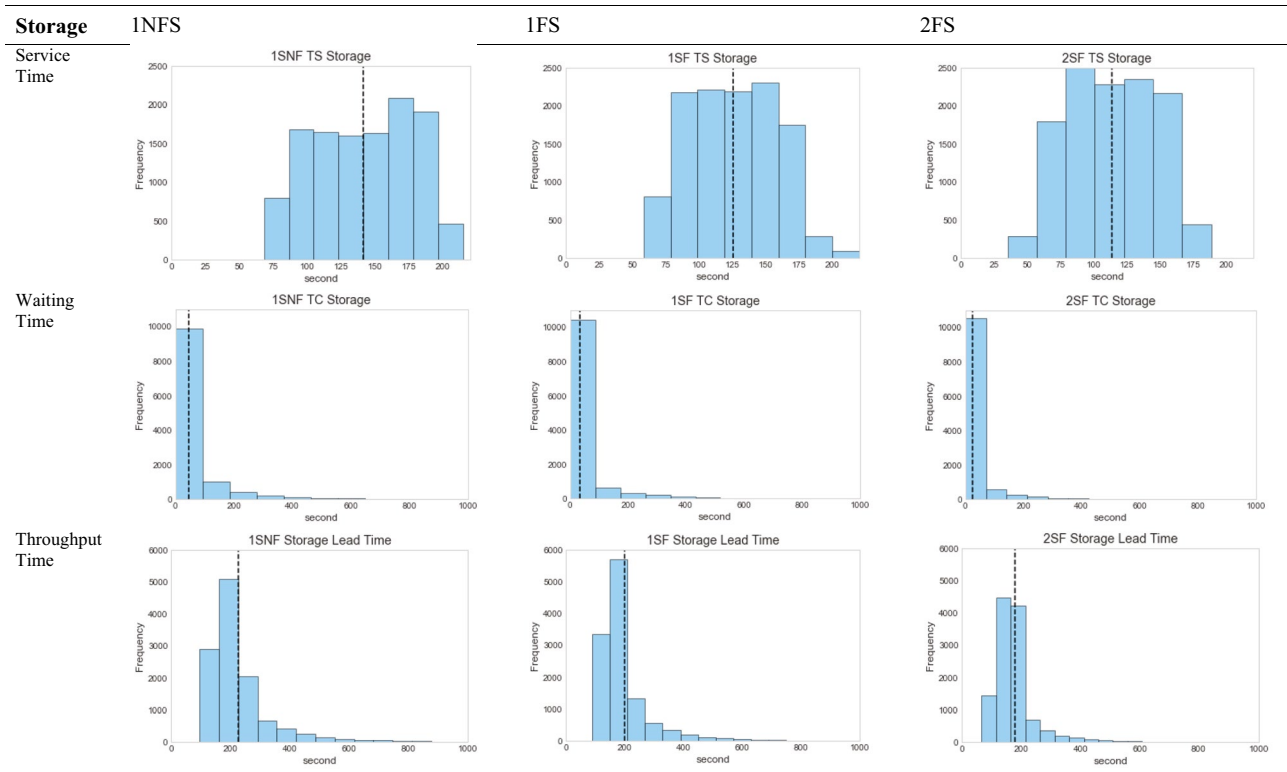


Table 11 Distribution of times for retrieval transactions

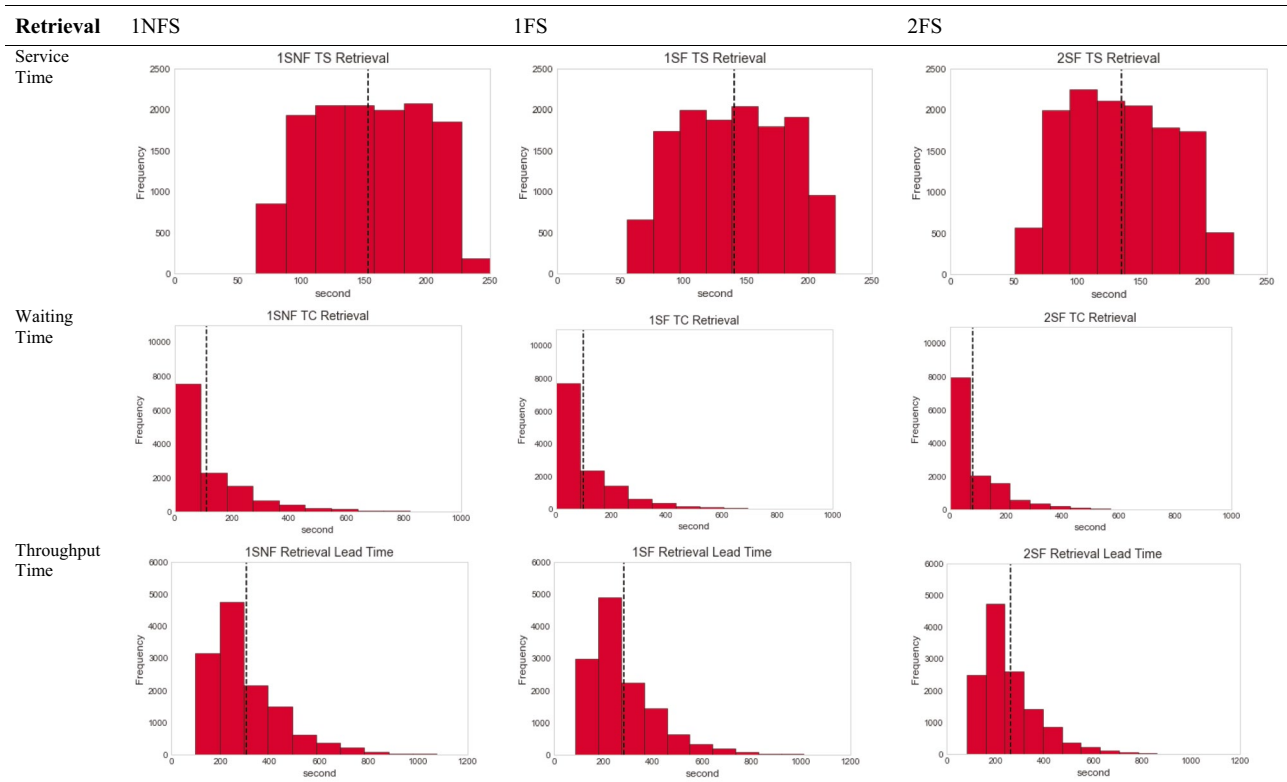


Table 12 Distribution of ULs at lift-up, tier, and lift-down, according to the three scenarios

	LiftS		Tier		LiftR	
	#UL	% distribution	#UL	% distribution	# UL	% distribution
1NFS	0	82%	0	68%	0	72%
	1	17%	1	28%	1	28%
	2	1%	2	4%	2	0%
	3	0%	3	0%	3	0%
1FS	0	82%	0	74%	0	69%
	1	17%	1	23%	1	31%
	2	1%	2	3%	2	0%
	3	0%	3	0%	3	0%
2FS	0	82%	0	77%	0	57%
	1	17%	1	21%	1	37%
	2	1%	2	2%	2	5%
	3	0%	3	0%	3	1%

Table 13 Vehicle utilization

	LiftS	Tier 1			Tier 2			Tier 3			Tier 4			Tier 5			LiftR
		Sh	Sat1	Sat2	Sh	Sat1	Sat2	Sh	Sat1	Sat2	Sh	Sat1	Sat2	Sh	Sat1	Sat2	
1NFS	13%	29%	/	/	25%	/	/	27%	/	/	30%	/	/	30%	/	/	15%
1FS	13%	25%	23%	/	21%	20%	/	23%	22%	/	25%	24%	/	25%	24%	/	15%
2FS	13%	22%	17%	5%	19%	15%	5%	21%	16%	5%	23%	18%	5%	23%	18%	5%	15%

In conclusion, the model suggests that the 2FS configuration minimizes the service and waiting times and the number of items in the queue. However, this solution also generates the worst utilization for shuttles and satellites. Finally, the system throughput amounts to 33 ULs per hour, and it depends on the arrival rate of the data entry. In particular, the system throughput is the same for the three scenarios (1NFS, 1FS, and 2FS) because of the incoming sequence and timing of transactions.

An original contribution of this study is the data entry inspired by a real AVS/RS. Therefore, storage and retrieval tasks follow the trends of production lines and shipments. The following subsections illustrate the impact of different data entries on the system performance. In particular, Subsection 5.2 illustrates the adoption of a random storage assignment policy (Analysis 2). Subsection 5.3 quantifies the maximum throughput of the system (Analysis 3). Both

analyses assumed the same sequence of transactions adopted in Analysis 1 and came from real applications.

5.2 Analysis 2: random storage assignment

This analysis reports some relevant KPIs obtained starting from a new data entry where the level and lane assigned to each transaction are randomly generated. The storage locations (z) follow the dynamic filling of the lane according to the LIFO policy. The sequence and number of storage and retrieval transactions are the same of the data entry previously discussed (real case study, named “Seq_Time_xyz”). This new data entry is named “Seq_Time_Random.” Table 14 reports the results in terms of the average service times and average waiting times for scenarios 1NFS, 1FS, and 2FS.

Figure 11 illustrates the percentage deviation between the average service time and the average waiting time to

Table 14 Comparison between Seq_Time_Random and Seq_Time_xyz data entry

	Storage				Retrieval			
	Seq_Time_Random		Seq_Time_xyz		Seq_Time_Random		Seq_Time_xyz	
	Service	Waiting	Service	Waiting	Service	Waiting	Service	Waiting
1NFS	147.68 s	55.01 s	139.45 s	49.46 s	159.70 s	111.30 s	154.64 s	104.43
1FS	132.87 s	39.71 s	125.67 s	35.94 s	145.21 s	100.78 s	140.23 s	94.98 s
2FS	86.25 s	23.44 s	114.55 s	27.22 s	133.36 s	59.81 s	135.10 s	57.96 s

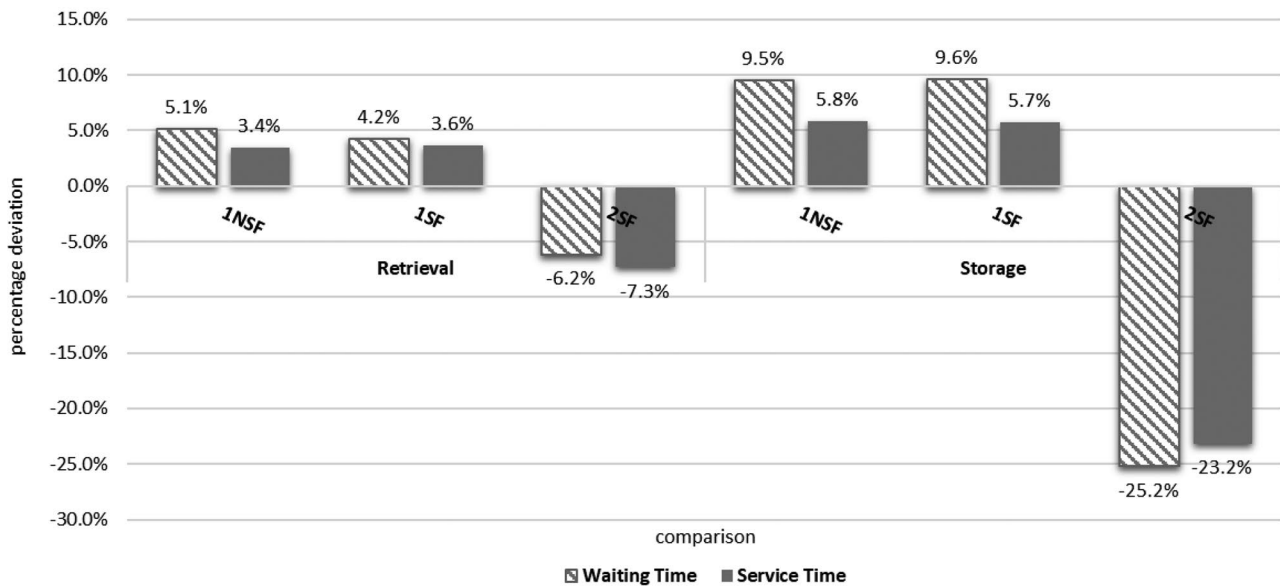


Fig. 11 Percentage deviation of average values between Seq_Time_Random and Seq_Time_xyz scenarios

perform a storage and a retrieval transaction according to the two data entries. The values are referred to Seq_Time_xyz data entry.

In the 2FS configuration, the service time generated by the Seq_Time_xyz data entry for storage transactions is 23.2% lower than that obtained by the Seq_Time_Random data entry. Thus, the waiting times are also 25.2% lower. The main reason is that Seq_Time_xyz data entry follows a storage policy that favors locations close to each other for unit loads that belong to the same production batch. While the first satellite is within a lane (z-direction) to drop a UL, the shuttle moves to the lane of the second UL (x-direction) and releases the second satellite to perform another task (z-direction). The total service time includes all the temporal contributions along the z-direction and the ones along the x-direction. The storage policy directly influences the contributions along the x-direction, whereas the contributions along the z-direction depend on the dynamic filling of the lanes. Therefore, the two data entries generate comparable contributions along the z-direction, whereas the contributions along the x-direction are different. Consequently, if the Seq_Time_xyz data entry favors locations close to each other, the service and waiting times are significantly lower than those obtained with the Seq_Time_Random data entry due to the temporal contributions along the x-direction.

In retrieval transactions, the difference between service times amounts to 7%. These transactions follow customer demand (characterized by a variable product variety), requiring different unit loads located in random lanes. Seq_Time_Random data entry is similar to the Seq_Time_xyz data entry for retrieval transactions. Therefore, the difference between the two service times is lower than for storage.

Finally, Fig. 11 shows that the gap between service times is more relevant for the 2FS configuration than for 1NSF and 1FS. In the 1NSF and 1FS configurations, when the satellite performs a task inside the first lane, the shuttle must wait before moving to the second lane. In the 2FS configuration, the temporal contributions along the z-direction can be reduced because different tasks can be performed simultaneously. Consequently, the effect of the temporal contribution along the x-direction on the total time is lower for 1NSF and 1FS than for 2FS. In conclusion, even if the Seq_Time_xyz data entry favors lanes close to each other for storage transactions, this temporal reduction has a minor effect on the total time.

Existing literature studies did not address these important issues, further demonstrating the effectiveness of the proposed hybrid approach.

5.3 Analysis 3: maximum system throughput

The data entry named Seq_Time_xyz generates a system throughput of 33 ULs per hour in all scenarios (1NSF, 1FS, and 2FS). This section illustrates how system performance changes according to different arrival times, given the same number and type of storage and retrieval transactions (i.e., 23,662 storages and 25,976 retrievals). In particular, we assumed the same arrival time for each UL, strictly respecting the sequence of transactions and incoming/outgoing SKU. Therefore, the transactions executed in Analyses 1 and 2 are now processed by the system according to a push-pull policy, i.e., incoming ULs are pushed and outgoing ULs are pulled respecting the original sequence of transactions. The first application of this push-pull policy estimates the so-called minimum arrival time for each transaction and

Table 15 Comparison between the distribution of ULs at lift-up, tier, and lift-down of the Seq_Time_xyz and Seq_Time_Random data entry

	LiftS			Tier			LiftR		
	#UL	%Distribution		#UL	%Distribution		#UL	%Distribution	
		A1	A3		A1	A3		A1	A3
1NFS	0	82%	29%	0	68%	19%	0	72%	64%
	1	17%	24%	1	28%	29%	1	28%	36%
	2	1%	35%	2	4%	39%	2	0%	0%
	3	0%	12%	3	0%	13%	3	0%	0%
1FS	0	82%	25%	0	74%	21%	0	69%	61%
	1	17%	22%	1	23%	30%	1	31%	39%
	2	1%	38%	2	3%	37%	2	0%	0%
	3	0%	15%	3	0%	12%	3	0%	0%
2FS	0	82%	25%	0	77%	41%	0	57%	48%
	1	17%	20%	1	21%	41%	1	37%	44%
	2	1%	39%	2	2%	17%	2	5%	7%
	3	0%	16%	3	0%	2%	3	1%	1%

releases a list of new arrival times. The “minimum arrival time” is the time when the transaction can arrive and be hosted by the system without saturating the input and output buffers. The maximum capacity of each buffer is a parametric value. In this analysis, it was set to three (inspired by a real case study). The new Seq_Time_xyz data entry reports all chronologically sorted transactions according to new arrival times.

The system performance can be estimated by applying the proposed model to new data entry. The obtained time to complete all transactions is now 284 h for the 1NSF configuration, 259 h for the 1FS configuration, and 227 h for the 2FS configuration. The time to complete all transactions decreases by increasing the number of satellites and in the presence of the free option. The average service times of the shuttle, satellite, and lift did not change.

Another observation from Analysis 3 is that arrival times have no impact on service times. Waiting times and UL distribution in the buffers change according to the new arrival times, which are assumed to be in line with the push-pull policy.

Table 15 illustrates the distribution of ULs at the interface with the input lift (LiftS), at the input bay of each tier (Tier), and at the interface with the output lift (LiftR). Both

the number of items in the queues (with the related probabilities) of Analysis 1 (named A1) and Analysis 3 (named A3) are reported for each buffer.

Table 15 illustrates how the buffer saturation increases when the time interval between one transaction and the next decreases. Moreover, Table 16 shows how also vehicle utilization increases.

The highest level of utilization is for shuttle one. Nevertheless, shuttles never reach 100% because the transaction arrival rates are set to never exceed buffer capacities. The utilization reaches 100% if buffer capacity is ignored, confirming that the shuttle is the bottleneck of the system. Thus, the system throughput also changes according to the configuration and the possibility of considering buffers capacity (named BC). Table 17 illustrates how the system throughput changes according to these two hypotheses (BC Yes/No).

This analysis demonstrates that buffer capacity significantly affects system performance. In 1NSF and 1FS configurations, the throughputs are 24% lower than those that do not consider buffer capacity (BC: No). Moreover, the throughput of the 2FS configuration decreases by 30% compared with the same configuration that does not consider buffers capacity.

Table 16 The utilization of the Seq_Time_Random data entry

	LiftS	Tier 1			Tier 2			Tier 3			Tier 4			Tier 5			LiftR	
		Sh	Sat1	Sat2	Sh	Sat1	Sat2	Sh	Sat1	Sat2	Sh	Sat1	Sat2	Sh	Sat1	Sat2		
1NFS	A1	13%	29%	/	/	25%	/	/	27%	/	/	30%	/	/	30%	/	/	15%
	A3	34%	74%	/	/	64%	/	/	69%	/	/	76%	/	/	76%	/	/	37%
1FS	A1	13%	25%	23%	/	21%	20%	/	23%	22%	/	25%	24%	/	25%	24%	/	15%
	A3	37%	72%	65%	/	61%	55%	/	67%	66%	/	74%	66%	/	74%	67%	/	40%
2FS	A1	13%	22%	17%	5%	19%	15%	5%	21%	16%	5%	23%	18%	5%	23%	18%	5%	15%
	A3	42%	70%	56%	15%	63%	49%	13%	70%	54%	13%	76%	59%	14%	77%	59%	14%	46%

Table 17 Throughput variation according to buffer capacities, following the Seq_Time_Random data entry

	Throughput [BC: Yes]	Throughput [BC: No]
1NSF	174 ULs/hour	229 ULs/hour
1FS	190 ULs/hour	252 ULs/hour
2FS	217 ULs/hour	313 ULs/hour

In conclusion, scenario 2FS guarantees the maximum throughput, considering both the buffer capacity and the lower utilization. A fourth scenario, with three free satellites (3SF), was tested to identify the maximum throughput of the system. The throughput of the system does not change compared to that of the 2FS configuration and is equal to 217 ULs per hour. Therefore, introducing an additional satellite (i.e., more than two per shuttle) does not increase the system performance. This result refers to a specific case study. However, the proposed hybrid approach of modeling gives the decision maker the opportunity to best design and configure any AVS/RS control and best performing the performance in coherence with real operating conditions.

6 Conclusions and further research

This study introduces an original data-driven hybrid approach for the design and control of an AVS/RS based on analytical and simulative models, multiple performance indicators (e.g., vehicle utilization, system throughput, service and waiting times of S/R transactions), real data entry (with realistic sequence and timing of S/R transactions, dispatching rules, storage policies, etc.), dynamic filling of the lane with homogeneous unit loads (single SKU of the same production batch), and free and not-free multiple satellites working with shuttles operating at different tiers. This approach attempts to solve the limits of traditional models and assumptions, producing useful results for real applications.

The proposed hybrid model is applied to a real case study involving three analyses based on different assumptions:

1. Analysis 1: number and typology (free and not free) of satellites. It assumes the real arrival time and sequence of storage and retrieval transactions according to the production rate of the filling beverage machine and customer demand. This analysis demonstrates the effectiveness of the proposed hybrid model that quantifies the service time, waiting time, and throughput time for each transaction (1); the time-varying number of ULs at the input/output lift interfaces and tier input bay (2); and vehicle utilization (3). Moreover, it identifies the system bottlenecks, demonstrating the importance of buffer capacity. This data-driven approach is necessary to configure an AVS/RS to control system performance.

2. Analysis 2: random storage assignment. The literature frequently adopts a random storage policy, i.e., given a storage transaction, the selection of the lane in the x-direction is random. In real applications, it is necessary to fill the generic lane according to the single SKU and production batch constraints that are rarely discussed by existing studies, especially for many SKUs. This analysis demonstrates that storage assignment significantly affects the system performance, and that the proposed data-driven and hybrid approach is effective.
3. Analysis 3: maximum system throughput. The proposed hybrid approach quantifies the maximum system throughput. The time to complete all transactions decreases by increasing the number of satellites and adopting the free option. The average service times of the shuttle, satellite, and lift did not change. The level of generic buffer capacity significantly affects system throughput. Finally, introducing more than two satellites did not increase the productivity of the system.

The sequence and timing of storage and retrieval transactions significantly affect the system performance. Any theoretical assumption on storage policy (e.g., random storage policy), arrival times (e.g., Poisson distribution of arrival times), and service times (e.g., exponential distribution of service times) is not acceptable. This general result makes it necessary to combine the proposed data-driven approach with analytical and dynamic models to realistically support the design and control of deep-lane multisatellite AVS/RS.

Further research on the application of this hybrid model to different system configurations (e.g., layout, lane depths, number and location of lifts, and vehicle UL capacity) and applications (i.e., case studies coming from different sectors) is required to support the technological development of new vehicles. New studies concerning the space efficiency control and maximization combined with system throughput analysis and according with different system configuration, equipment, and S/R policies are achieved.

Acknowledgements The authors warmly thank the Italian company Elettric80 SpA, part of the E80 Group SpA, deeply involved in this study, in particular, Eng. Alberto Lodini, R&D Director in E80 Group, for his priceless work and cooperation to this research project.

Funding Open access funding provided by Alma Mater Studiorum - Università di Bologna within the CRUI-CARE Agreement.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ekren BY (2011) Performance evaluation of AVS/RS under various design scenarios: a case study. *Int J Adv Manuf Technol* 55(9–12):1253–1261. <https://doi.org/10.1007/s00170-010-3137-x>
- Malmberg CJ (2002) Conceptualizing tools for autonomous vehicle storage and retrieval systems. *Int J Prod Res* 40(8):1807–1822. <https://doi.org/10.1080/00207540110118668>
- Sulistio A, Yeo CS, Buyya R (2004) A taxonomy of computer-based simulations and its mapping to parallel and distributed systems simulation tools. *Softw Pract Exper* 2004(34):653–673. <https://doi.org/10.1002/spe.585>
- Fukunari M, Malmberg CJ (2009) A network queuing approach for evaluation of performance measures in autonomous vehicle storage and retrieval systems. *Eur J Oper Res* 193(1):152–167. <https://doi.org/10.1016/j.ejor.2007.10.049>
- Roy D, Krishnamurthy A, Heragu S, Malmberg C (2015) Stochastic models for unit-load operations in warehouse systems with autonomous vehicles. *Ann Oper Res* 231(1):129–155. <https://doi.org/10.1007/s10479-014-1665-8>
- Roy D, Krishnamurthy A, Heragu S, Malmberg C (2017) A multi-tier linking approach to analyze performance of autonomous vehicle-based storage and retrieval systems. *Comput Oper Res* 83:173–188. <https://doi.org/10.1016/j.cor.2017.02.012>
- Deng L, Chen L, Zhao J, Wang R (2021) Modeling and performance analysis of shuttle-based compact storage systems under parallel processing policy. *PLoS One* 16(11):e0259773. <https://doi.org/10.1371/journal.pone.0259773>
- Marchet G, Melacini M, Perotti S, Tappia E (2012) Analytical model to estimate performances of autonomous vehicle storage and retrieval systems for product totes. *Int J Prod Res* 50(24):7134–7148. <https://doi.org/10.1080/00207543.2011.639815>
- Marchet G, Melacini M, Perotti S, Tappia E (2013) Development of a framework for the design of autonomous vehicle storage and retrieval systems. *Int J Prod Res* 51(14):4365–4387. <https://doi.org/10.1080/00207543.2013.778430>
- Marolt J, Kosanić N, Lerher T (2022) Relocation and storage assignment strategy evaluation in a multiple-deep tier captive automated vehicle storage and retrieval system with undetermined retrieval sequence. *Int J Adv Manuf Technol* 118:3403–3420. <https://doi.org/10.1007/s00170-021-08169-x>
- Manzini R, Accorsi R, Baruffaldi G, Cennerazzo T, Gamberi M (2016) Travel time models for deep-lane unit-load autonomous vehicle storage and retrieval system (AVS/RS). *Int J Prod Res* 54(14):4286–4304. <https://doi.org/10.1080/00207543.2016.1144241>
- D'Antonio G, Chiabert P (2019) Analytical models for cycle time and throughput evaluation of multi-shuttle deep-lane AVS/RS. *Int J Adv Manuf Technol* 104(5–8):1919–1936. <https://doi.org/10.1007/s00170-019-03985-8>
- D'Antonio G, de Maddis M, Bedolla JS, Chiabert P, Lombardi F (2018) Analytical models for the evaluation of deep-lane autonomous vehicle storage and retrieval system performance. *Int J Adv Manuf Technol* 94(5–8):1811–1824. <https://doi.org/10.1007/s00170-017-0313-2>
- Tappia E, Roy D, de Koster R, Melacini M (2017) Modelling, analysis, and design insights for shuttle-based compact storage systems. *Transp Sci* 51(1):269–295. <https://doi.org/10.1287/trsc.2016.0699>
- Eder M (2020) An approach for a performance calculation of shuttle-based storage and retrieval systems with multiple-deep storage. *Int J Adv Manuf Technol* 107(1–2):859–873. <https://doi.org/10.1007/s00170-019-04831-7>
- Lerher T, Ekren BY, Dukic G, Rosi B (2015) Travel time model for shuttle-based storage and retrieval systems. *Int J Adv Manuf Technol* 78(9–12):1705–1725. <https://doi.org/10.1007/s00170-014-6726-2>
- Lerher T, Ficko M, Palčič I (2021) Throughput performance analysis of Automated Vehicle Storage and Retrieval Systems with multiple-tier shuttle vehicles. *Appl Math Model* 91:1004–1022. <https://doi.org/10.1016/j.apm.2020.10.032>
- Ekren BY, Heragu SS (2012) Performance comparison of two material handling systems: AVS/RS and CBAS/RS. *Int J Prod Res*. <https://doi.org/10.1080/00207543.2011.588627>
- Jia J, Heragu SS (2009) Analysis of semi-open queueing networks via analytical matrix geometric methods. *Oper Res* 57:391401. <https://doi.org/10.1287/opre.1080.0627>
- Heragu SS, Cai X, Krishnamurthy A, Malmberg CJ (2009) Analysis of autonomous vehicle storage and retrieval system by open queueing network. *IEEE Int Conf Autom Sci* 2009:455–459. <https://doi.org/10.1109/COASE.2009.5234100>
- Eder M (2020) Analytical model to estimate the performance of shuttle-based storage and retrieval systems with class-based storage policy. *Int J Adv Manuf Technol* 107(5–6):2091–2106. <https://doi.org/10.1007/s00170-020-04990-y>
- Eder M (2020) An approach for performance evaluation of SBS/RS with shuttle vehicles serving multiple tiers of multiple-deep storage rack. *Int J Adv Manuf Technol* 110(11–12):3241–3256. <https://doi.org/10.1007/s00170-020-06033-y>
- Kumawat GL, Roy D (2021) A new solution approach for multi-stage semi-open queueing networks: An application in shuttle-based compact storage systems. *Comput Oper Res* (125). ISSN 0305–0548. <https://doi.org/10.1016/j.cor.2020.105086>
- Kuo P, Krishnamurthy A, Malmberg CJ (2007) Design models for unit load storage and retrieval systems using autonomous vehicle technology and resource conserving storage and dwell point policies. *Appl Math Model* 31(10):2332–2346. <https://doi.org/10.1016/j.apm.2006.09.011>
- Roy D (2016) Semi-open queueing networks: a review of stochastic models, solution methods and new research areas. *Int J Prod Res* 54(6):1735–1752. <https://doi.org/10.1080/00207543.2015.1056316>
- Lerher T (2016) Travel time model for double-deep shuttle-based storage and retrieval systems. *Int J Prod Res* 54(9):2519–2540. <https://doi.org/10.1080/00207543.2015.1061717>
- Lin Y, Wang Y, Zhu J, Wang L (2021) A model and a task scheduling method for double-deep tier-captive SBS/RS with alternative elevator-patterns. *IEEE Access* 9:146378–146391. <https://doi.org/10.1109/ACCESS.2021.3120418>
- Krenczyk D, Davidrajuh R, Skolud B (2019) Comparing two methodologies for modelling and simulation of discrete-event based automated warehouses systems. In *Lecture Notes in Mechanical Engineering* (pp. 161–175). Pleiades Publishing. https://doi.org/10.1007/978-3-030-18789-7_15

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.