Transportation Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Robotized and Automated Warehouse Systems: Review and Recent Developments

Kaveh Azadeh, René De Koster, Debjit Roy

informs
http://pubsonline.informs.org/journal/trsc/

TRANSPORTATION SCIENCE
Vol. 53, No. 4, July–August 2019, pp. 917–945
ISSN 0041-1655 (print), ISSN 1526-5447 (online)

Robotized and Automated Warehouse Systems: Review and Recent Developments

Kaveh Azadeh,[a] René De Koster,[a] Debjit Roy[a,b]

[a] Rotterdam School of Management, Erasmus University, 3062 PA Rotterdam, Netherlands; [b] Indian Institute of Management Ahmedabad, Gujarat 380015, India

Contact: azadeh@rsm.nl, http://orcid.org/0000-0001-5267-1963 (KA); rkoster@rsm.nl, http://orcid.org/0000-0002-1740-7822 (RDK); debjit@iima.ac.in, http://orcid.org/0000-0001-5298-9664 (DR)

Received: July 18, 2017
Revised: January 12, 2018; May 4, 2018
Accepted: July 12, 2018
Published Online in Articles in Advance: June 28, 2019

https://doi.org/10.1287/trsc.2018.0873

Abstract. Robotic handling systems are increasingly applied in distribution centers. They require little space, provide flexibility in managing varying demand requirements, and are able to work 24/7. This makes them particularly fit for e-commerce operations. This paper reviews new categories of automated and robotic handling systems, such as shuttle-based storage and retrieval systems, shuttle-based compact storage systems, and robotic mobile fulfillment systems. For each system, we categorize the literature in three groups: system analysis, design optimization, and operations planning and control. Our focus is to identify the research issue and operations research modeling methodology adopted to analyze the problem. We find that many new robotic systems and applications have hardly been studied in academic literature, despite their increasing use in practice. Because of unique system features (such as autonomous control, flexible layout, networked and dynamic operation), new models and methods are needed to address the design and operational control challenges for such systems, in particular, for the integration of subsystems. Integrated robotic warehouse systems will form the next category of warehouses. All vital warehouse design, planning, and control logic, such as methods to design layout, storage and order-picking system selection, storage slotting, order batching, picker routing, and picker to order assignment, will have to be revisited for new robotized warehouses.

Keywords: warehousing • robotic systems • planning and control

1. Introduction

Warehouse operations tend to be labor intensive and require large spaces for facilities. Large buildings are needed to store item assortments in racks, to move stock, to unload and load trailers and containers, to inspect picked orders, to allow trucks to maneuver in the yard, and to dock the trucks. With the advent of e-commerce, companies store millions of unique items and handle large and variable daily order volumes. On the other hand, the most laborious and expensive process, order picking, is repetitive, often suffers from poor ergonomics, and requires high-quality labor willing to work in shifts, which is often difficult to get. It is therefore not surprising that warehousing systems and processes are key candidates for automation. In addition, the land available for warehouses (which should preferably be close to the demand points) has become scarce, and many warehouses have to operate 24/7. Together, these factors have given warehouse automation a big boost. Warehouse automation dates back to the 1960s, when the first high-bay (20–40 m high was quite standard) unit-load warehouses were established in Germany with aisle-captive cranes driving on rails, constructed as silo buildings (Industrie Forum 2004). These so-called automated storage and retrieval (AS/R) systems were able to store bulk stock on unit loads (pallets, or totes [miniload system]). They could also work in conjunction with manual pick stations as a parts-to-picker system, where the retrieved unit load was restored after picking units from it. Since then, AS/R systems have become very popular in practice, and research has gained momentum with the papers by Hausman, Schwarz, and Graves (1976) and Bozer and White (1984). Hundreds of papers have been published on these systems. An overview of AS/R system classification and research studies is given by Roodbergen and Vis (2009).

During the last decade, warehouse automation has developed rapidly. A big boost to this development has been given by autonomous vehicle–based storage and retrieval (AVS/R), or shuttle-based storage and retrieval, systems. These systems use racks with aisles and deploy autonomous shuttles that operate at each level in each aisle. Vertical transport is enabled by lifts. Another important development has been automated pallet stacking and destacking technologies, in particular, also by mixed-case palletizing technology developed in the early 2000s. A new generation of automated guided vehicles (AGVs), supporting the order-picking process, has recently been introduced. These systems will gradually result in automated picking processes. Pioneered by Witron, combining multiple technologies has led to the advent of completely automated warehouses, particularly in the store-based retail industry (mostly grocery). Based on the authors' experience, in Western Europe alone, about 40 fully automated warehouses are in operation, and many are under development. Although these warehouses are large, they are much smaller (and supposedly more cost-efficient) than their conventional, manual counterparts. Figure 1 shows a flow diagram of such a warehouse with typical storage and handling systems.

In such an automated retail warehouse, selected suppliers unload their own trucks and feed the preannounced single–stock-keeping unit (SKU) pallets to a check-in conveyor (Step 1). The pallets are then stored in an AS/R system (Step 2). When a certain product is requested, the pallet is off loaded and automatically destacked (Step 3). The loose cases are then often put on trays to ease manipulation and are stored in a miniload AS/R or in an AVS/R system (Step 4). When the store order arrives, the cases are retrieved and sequenced (Step 5), and mixed-case palletizers build the pallets or roll cages in a store-specific sequence that allows rapid shelving in the store (Step 6). These roll cages then wait in an order consolidation buffer (OCB), usually an AS/R system (Step 7), until the departure truck arrives, after which they are retrieved and loaded in the sequence determined by the stop sequence in the truck route.

Apart from the (many) technicians needed to keep the system alive, no manual handling is involved. In addition to these fully automated warehouses, many partially robotized warehouses have been built. According to Buck Consultants International (2017), in the Netherlands alone, 63 large new warehouses were constructed in the period 2012–2016 using robot technologies. However, the majority of warehouse research still focuses on conventional storage and order-picking methods. The overview by De Koster, Le-Duc, and Roodbergen (2007) provides some avenues for research into (semi)automated picking methods. Because of rapid system developments, it is time for an update, as the new technologies have provided new and interesting research opportunities. This paper structures the new automated technologies and provides an overview of these technologies and the research carried out already. It also reviews the modeling techniques used and the research opportunities they provide. We focus on the design and control aspects of order-picking systems because they form the heart and soul of any warehouse. In doing so, we include the corresponding automated product storage and handling techniques. Figure 2 categorizes the automated picking systems, both the classical as well as the newly developed automated picking systems.

In this study, we focus on recent robotic automated picking systems, in particular systems that use free-roaming retrieval robots such as shuttles and free-roaming AGVs (the grey-shaded systems in Figure 2). The more conventional systems, such as cranes, automated forklifts, carousels, and automated dispensers have been reviewed in other papers (Roodbergen and Vis 2009; Litvak and Vlasiou 2010; Gagliardi, Renaud, and Ruiz 2012; Boysen and Stephan 2016); we highlight only a few key articles. To find articles, we used the following search terms in Scopus: "autonomous vehicle/shuttle storage and retrieval systems," robotic mobile fulfillment system," "puzzle-based storage system," "compact warehouse storage systems," and "robotic warehouse storage and retrieval systems," as well

Figure 1. (Color online) Material Flow in a Typical Automated Warehouse

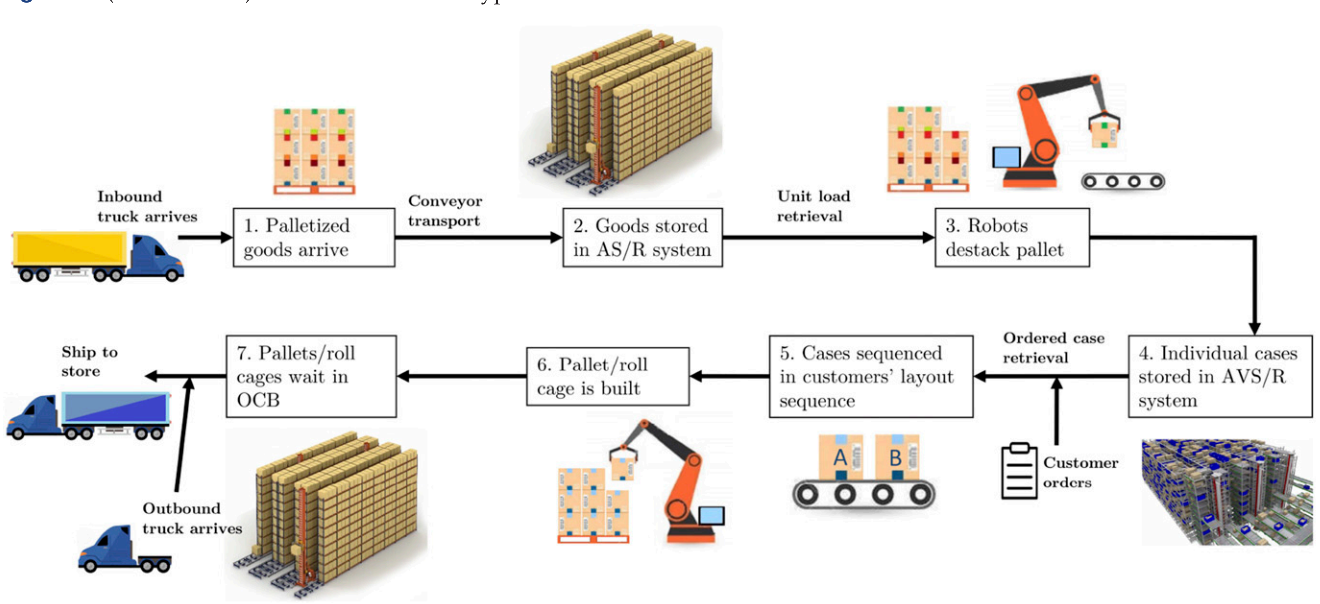
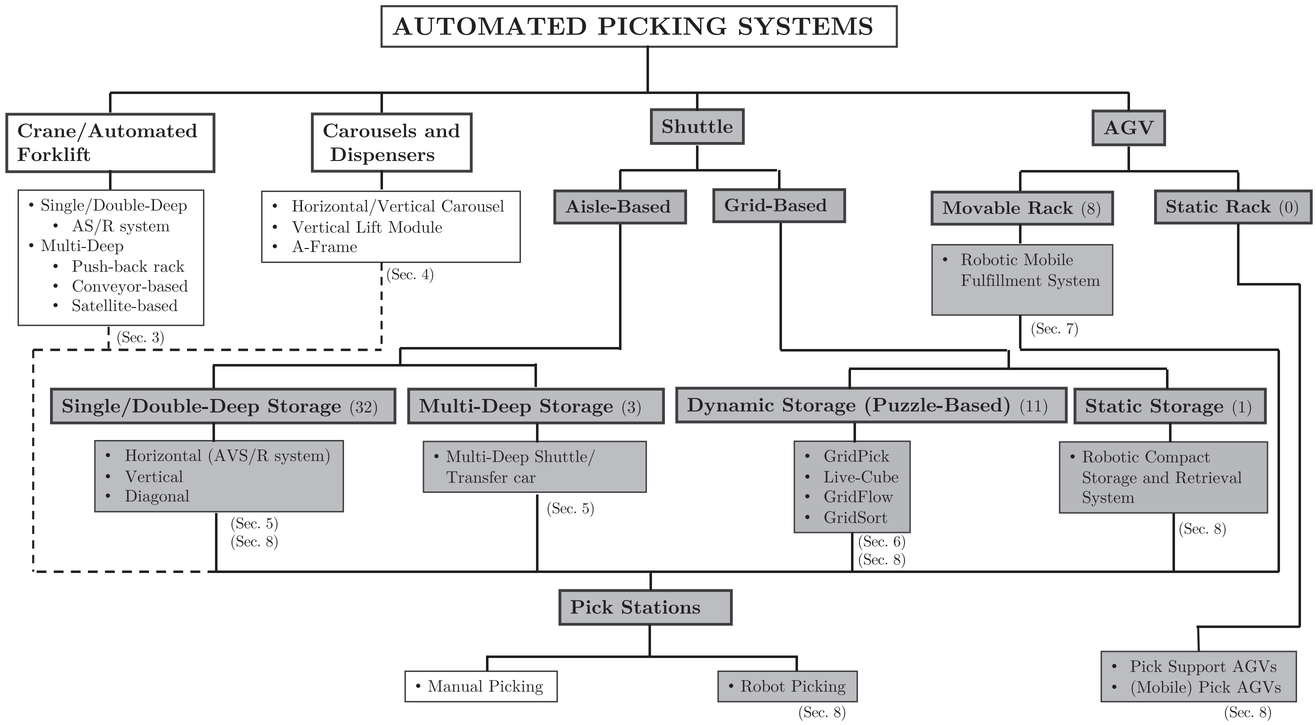


Figure 2. Classification of Automated Picking Systems



Notes. The literature of the gray-shaded systems is reviewed. The numbers placed next to the systems indicate the numbers of reviewed papers.

as variants of these search terms. We review papers published in high-quality journals, complemented by some working papers and proceedings for prominent systems that have not received much attention yet. We review 55 papers on the core systems indicated in the gray-shaded boxes in Figure 2.

We first describe various modeling methods used in the design and operation of the systems and the associated objectives (Section 2). Section 3 deals with the "conventional" AS/R systems, that have been researched intensively and then continues with less conventional crane and automated forklift-based systems, such as multideep racks operated by cranes and satellites. Section 4 discusses different types of carousels, vertical lift modules (VLMs), and automated dispenser systems. Section 5 discusses various types of aisle-based AVS/R systems, and Section 6 considers grid-based storage and retrieval systems. Section 7 continues with robotic movable rack systems. Section 8 discusses directions for future research and includes emerging technologies, in particular, humans picking in collaboration with AGVs. We conclude in Section 9.

2. Modeling Methods and Objectives in Storage, Transport, and the Order-Picking Process

Two types of approaches exist to model the systems: *analytical* and *simulation based*. Simulation-based models can mimic reality accurately and produce the least error. However, conceptualizing and designing a detailed and accurate simulation model is time intensive. Optimizing the entire design space may require the development of multiple models. Therefore, at an early stage, analytical models are preferred, to reduce the design search space and to identify a limited number of promising configurations. Compared with simulation modeling, analytical models run faster and can obtain the optimal configuration either directly or with a quick enumeration over a large number of design parameters. The error made in the estimated performance measures using analytical models is usually acceptable for the conceptualization phase. Section 2.1 explains analytical models. Section 2.2 discusses what the different objectives and decisions are in evaluating automated warehouses and how the analytical models are used to optimize those objectives. We also present the classification scheme that we use for reviewing articles.

2.1. Analytical Models

The most common analytical models for storage and retrieval are classified into three categories: linear programming (LP) and mixed-integer programming (MIP) models, travel time models, and queuing network (QN) models.

2.1.1. Linear and Mixed-Integer Programming Models.

Many of the design and operational decisions in automated systems can be optimized using LP or nonlinear and MIP models. For instance, LP and MIP models can be used for optimizing the shape of the system, obtaining the right choice of storage policy, scheduling and sequencing order transactions, and establishing order batching rules. LP and MIP models are usually used in a deterministic setting. To capture the stochasticity, travel time and queueing network models are preferred.

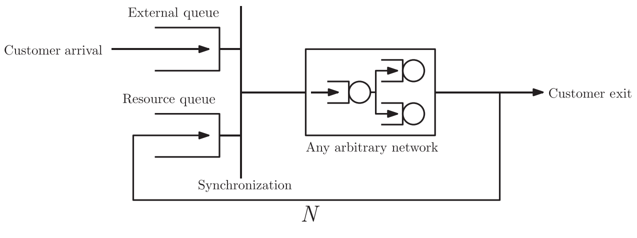
Solution Methods for Linear and Mixed-Integer Programming Models. LP models can be solved exactly in polynomial time. However, the exact solutions for the majority of the MIP models are intractable. As a result, metaheuristic algorithms are developed, which provide near-optimal solutions in a short time. The notion behind metaheuristic algorithms is to find the best solution out of all possible feasible solutions. Some notable examples of metaheuristic algorithms include genetic algorithms, tabu search, simulated anealing, and adaptive large neighborhood search. See Glover and Kochenberger (2006) for a more detailed overview of the different metaheuristic algorithms. Recent developments in exact and heuristic algorithms have resulted in an integrated technique using matheuristics. In a matheuristic, the problem is decomposed into several small subproblems that can be solved using exact algorithms. Later, the results of subproblems are used in the heuristic algorithm (see Puchinger and Raidl 2005).

2.1.2. Travel Time Models. Using travel time models, the design engineer can obtain the amount of time that it takes for a resource to move from one location to another. For instance, in an automated parts-to-picker picking context, travel time models can be used to obtain closed-form expressions for the expected load storage and retrieval time. The closed-form travel time expressions are usually simple and computationally friendly. Therefore, they can be used to limit the search space before adopting a detailed simulation, or for optimizing the design choices. They can also be used to estimate the expected service time of a server in a network of queues. Despite the simplicity of the travel time models, they are not capable of capturing several factors such as interaction between multiple resources, parallel processing by multiple resources, or queueing within the system. In these scenarios, QN models are preferred.

2.1.3. Queueing Network Models. Automated picking systems can be modeled as multistage service systems using QNs. In a QN, a customer arrives in the system, undergoes several stages of service, and leaves the system. Several types of queueing networks have been studied: *open queuing networks* (*OQNs*), *closed queuing networks* (*CQNs*), and *semiopen queuing networks* (*SOQNs*). In an OQN, customers, such as orders to be picked, arrive from an external source and, after receiving service in different nodes, leave the system. An OQN is particularly useful to estimate expected order throughput time. However, in many systems, resources accompany orders during the whole or a part of the process, for example, a transport vehicle, a transport roll container, or a pallet. Often, the number and the capacity of the resources that affect the performance of the system are limited. For instance, orders might be transported by expensive robots in the system. In this scenario, an OQN is not capable of accurately estimating the performance of the system, as it assumes an infinite supply of robots. One way to overcome this challenge is to model the system as a CQN. In a CQN, a limited number of resources are paired with the incoming orders. Once an order is completed, the resource becomes available to serve another order. The limited number of resources enforces a population constraint in the CQN. However, it is implicitly assumed that an infinite number of orders are waiting outside the system (Heragu et al. 2011). CQNs are useful to estimate the maximum throughput capacity of the system. Using a CQN to model the systems in which the incoming customers and the resources are paired together throughout the process leads to an underestimation of the true customer waiting time. The reason lies in the assumption (infinite number of customers waiting externally in a CQN). However, in reality, there are times when a customer needs to wait for a resource or vice versa. In this situation, an SOQN is a suitable model because it can accurately capture the external transaction waiting time. As it illustrated in Figure 3, an SOQN (in the literature sometimes called an open queueing network with limited capacity) possesses a synchronization station in which incoming customers waiting at an external queue are paired with available resources in the resource queue. Then, each customer is processed using the resource that carries the customer to prespecified different nodes (Cai, Heragu, and Liu 2013; Roy et al. 2015b; Roy 2016).

Solution Methods for Evaluating Queueing Networks. One of the most important methods for calculating performance measures of product-form queueing

Figure 3. A General Semiopen Queueing Network with N Circulating Resources



networks (Baskett et al. 1975) is *mean value analysis* (*MVA*) (Reiser and Lavenberg 1980). The MVA algorithm is based on Little's law and the arrival theorem. However, networks used in analyzing automated picking systems usually do not have product-form solutions for a number of reasons, such as nonexponentially distributed service times, customer blocking, or non-Markov routing. Therefore, approximation algorithms are used to estimate the performance measures of the system. Several approximation techniques such as approximate mean value analysis and the parametric decomposition approach proposed by Whitt (1983) have been developed based on the characteristics of the network. Bolch et al. (2006) provide a detailed overview of exact and approximate algorithms to evaluate the performance of open and closed queueing networks. The SOQN does not have a product-form solution, even for Poisson arrivals and exponential servers. The matrix-geometric method, aggregation, network decomposition, parametric decomposition, and performance bounds are the most common solution approaches for approximating the performance of an SOQN. Detailed overviews of solution techniques to evaluate an SOQN are presented by Jia and Heragu (2009) and Roy (2016). When it is not possible to analytically solve a queueing network, it is always possible to obtain its performance measures by simulation.

2.2. Decision Variables and Performance Objectives

Two levels of decision making can be distinguished in warehouse planning and design: long term (tactical) and short term (operational).

In long-term planning, decisions revolve around the hardware design selection and optimization of the system. At this level, the prime objective is to maximize the throughput and the storage capacity of the system. The objectives are affected by several decision variables, such as the physical layout configuration (e.g., the number of aisles, the depth of each aisle, the number of cross aisles, and the number of tiers), the numbers of robots and lifts, and the numbers and locations of load/unload (L/U) points and workstations. At this stage, the focus is on the decisions that are hard to alter once the system is in place.

Short-term decision making focuses on operational planning and control. The prime objectives are to minimize lead time, waiting time, response time, and resource idleness, and so forth. Decisions include vehicle assignment policies; blocking prevention protocols; dwell point use of the vehicles, that is, selecting the location where a vehicle without a job (idle vehicle) is parked; storage slotting; and workstation assignment rules.

Analytical models can address both long-term and short-term decision making. LP models are used to optimize any objective function (e.g., cost) while satisfying multiple constraints. With a (usually nonlinear) travel time model, it is (sometimes) possible to obtain a closed-form expression of the performance measures, such as the average processing time. By taking derivatives with respect to the desired decision variables, one can optimize the system with regard to the performance measure. However, deriving a closed-form expression of system measures such as transaction time (including waiting) is often not possible. For this purpose, queueing network and simulation-based models are used. Design performance optimization is then done by enumerating the decision variables. Sometimes, combinations of decision variables have a joint effect on the performance of the system. As a result, some authors, such as Ekren and Heragu (2010b), suggest using regression models with interaction variables to evaluate the combined effect of decision variables on the performance of the system. Then, the enumeration is done over the variables and their combinations to examine the effect on the desired performance measure.

Table 1 presents a framework of different objectives and decision variables and the suitable modeling approaches to address them. When reviewing the articles in Sections 5, 6, and 7, we leverage the framework presented in Table 1 and group the articles based on the prime objective being investigated. The categories include: system analysis, design optimization, and operations planning and control. System analysis articles focus on modeling techniques to estimate the performance of the system without focusing on any optimization. Design optimization articles focus on hardware optimization of the system (e.g., system layout), and operations planning and control articles focus on the software optimization of the system (e.g., block prevention policies).

3. Automated Storage and Retrieval Systems with Cranes or Automated Forklifts

Crane-based automated storage and retrieval systems were introduced in the 1960s. Initially, their main application was in pallet warehouses storing bulk inventories. Later, miniload warehouses and more compact multideep order-picking warehouses were also automated. In this section, we discuss the different types of crane/automated forklift-based automated storage and retrieval systems, as mentioned in Figure 2.

3.1. Single/Double-Deep Storage

Such a system consists of racks and automated handling systems such as cranes or automated forklifts. These handling systems can be aisle captive (typically cranes) or aisle roaming (typically high-bay automated forklifts). To perform a storage operation, a crane picks up a load, usually from a conveyor, and stores it in the 30- to 40-m-high racks. Driving and lifting in the aisle take place simultaneously. The process sequence is

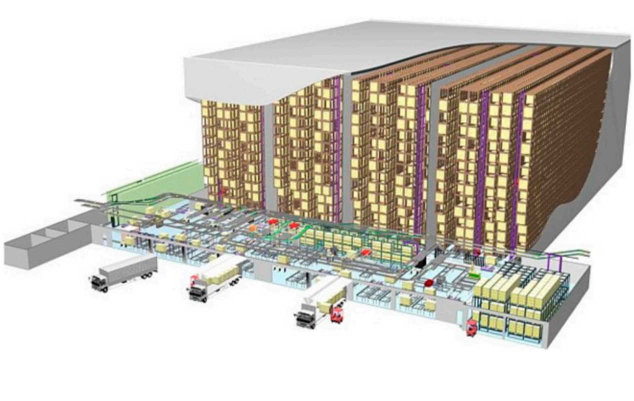
Table 1. Decision-Making Framework and Appropriate Modeling Methods

Decision level	Prime objectives	Decision variables	Modeling approach
Long-term decisions (design optimization)	Maximize: Throughput capacity Storage capacity	Physical layout: Number of aisles Number of cross aisles Depth of the aisle Number of tiers Number of robots Number of lifts L/U point(s) Workstation(s) location	Simulation Travel time model Closed queueing network Semi-open queueing network Deterministic optimization (LP, IP, MIP)
Short-term decisions (operational planning and control)	Minimize: Lead time Waiting time Response time Resource idleness	Vehicle assignment policy Block prevention policy Dwell point policy Storage policy Resource scheduling Sequencing transactions	Simulation Travel time model Closed queueing network Semi-open queueing network Deterministic optimization (LP, IP, MIP)

reversed for a retrieval operation. It is also possible to carry out a dual command cycle, in which a storage and a retrieval job are combined. This would save one movement per dual command cycle; however, there may be an additional wait for pairing a storage transaction with a retrieval. If totes instead of pallets are stored, the system is referred to as a miniload system. Figure 4 shows an example of such a warehouse.

Unit-load and miniload aisle-captive single-deep AS/R systems have been studied extensively. One of the first scientific articles is by Bozer and White (1984). They calculate the average cycle time of the crane for single command cycles and assume that crane travel to any location within the rack has the same probability (random storage policy). Their expected cycle time is $E[T] = \left(1 + \frac{(t_y/t_x)^2}{3}\right).t_x$, in which t_x is the travel time to the farthest location in the rack, and t_y is the lifting time to the highest location in the rack. The formula assumes that the crane drives and lifts at the same time and that the travel time to the farthest location is longer than the lifting time. Using this formula, the optimal ratio between the length and height of an aisle can be obtained,

Figure 4. (Color online) Automated High-Bay Warehouse for Pallets with Aisle-Captive Cranes (De Koster 2015)



which proves to be square in time, meaning that the travel time to the farthest location and the lifting time to the highest location are identical. Assuming that a crane travels approximately four times faster than it lifts, the length of the aisle should therefore be four times its height to minimize the cycle time. Later on, this formula was adjusted to include other aspects of the warehouse, such as different storage strategies (such as ABC storage), dual command cycles, and different locations of the load and unload points. (The above formula assumes one such point, at the lower corner of the rack.) We refer to Roodbergen and Vis (2009) for an extensive overview of the literature on AS/R systems. Furthermore, Gagliardi, Renaud, and Ruiz (2012) provide an overview of the simulation-based models for AS/R systems. Boysen and Stephan (2016) present a novel classification scheme for defining various crane-scheduling problems in AS/R systems. Later, they applied the scheme to review the literature.

In the case of ABC (or product turnover-based) storage, the items are divided into classes (e.g., classes A, B, and C) based on item turnover rate. The locations are also divided into groups based on travel time to the L/U point. This ensures that the items from the class with the highest turnover rate are located closest to that point. Hausman, Schwarz, and Graves (1976) investigated the cycle time calculations with ABC storage and economic order quantity–based replenishment. Later, their results were extended to N product classes by Rosenblatt and Eynan (1989). Hausman, Schwarz, and Graves (1976) calculated the optimal class boundaries for known ABC demand curves, for example, 20/70 demand curves, whereby 20% of the items (or unit loads) are responsible for 70% of the demand. In the calculation, they considered product restocking according to a continuous review $<s, Q>$ policy, with the stocking quantity Q being equal to the optimal order quantity.

However, they did not take into account that the more storage classes there are, the fewer items per class are stored. This requires more space per item stored in the class, because the space within the classes cannot be shared by the items, which lengthens crane travel time. In the extreme case of one item per class, the space required is $\sum\lceil Q_i + SS_i\rceil$ whereas in the extreme case of one class containing all items (i.e., random storage), the space required is $\sum\lceil\frac{Q_i}{2} + SS_i\rceil$. This means that an optimum number of storage classes can be distinguished. In practice, the optimal number of classes is small (about three to five), but the cycle time is relatively insensitive to the exact number. At such a limited number of classes, products can perfectly share the space available in a class. However, the required number of locations on top of the average stock level quickly amounts to an additional 40% (Yu, De Koster, and Guo 2015).

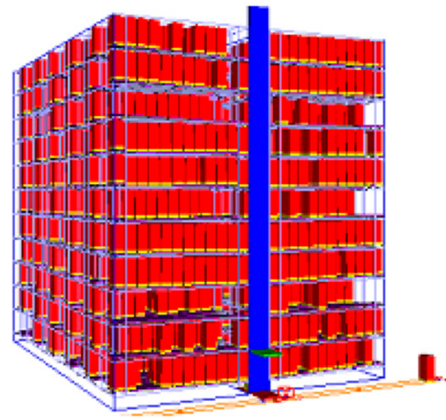
3.2. Multideep (Compact) Storage

AS/R systems can also be used to store loads double deep in the racks. To this end, the cranes can be equipped with double-deep telescopic forks. Deep lane, or compact, multideep (three-dimensional [3D]) AS/R systems can store loads even more deep in storage lanes (see Figure 5). The storage depth (e.g., 5–15 loads) depends on the type of product and the technology. These systems are particularly popular for storing products when storage space minimization is a primary concern, for example, fresh produce and cold storage warehouses. In a typical crane-based compact storage system, a storage and retrieval (S/R) crane takes care of movements in the horizontal and vertical directions of the rack, and an orthogonal conveying mechanism takes care of the depth movement. Multideep lane crane-based compact storage systems can be further classified into three categories based on the mechanism of the depth movement. These three categories are push-back rack, conveyor based, and satellite-based systems (see Figure 2) which are described below.

Push-back rack. In this variant, the crane (or automated forklift) stores the loads by mechanically pushing them into the storage lanes. The system works according to the last-in-first-out (LIFO) principle. A slight slope on the storage lane utilizes gravity to ensure that a load is always available in front of the storage lane. The depth of the lane in a push-back pallet rack is up to about five loads.

Conveyor-based system. The racks in these systems are equipped with conveyors (see Figure 6). If the conveyor can move in two directions, the operation is LIFO, similar to the push-back racks. The conveyors can also operate in pairs (either by gravity or powered). On the inbound conveyor, unit loads flow to the rear end of the rack. The outbound conveyor is located next to the inbound conveyor. On the outbound conveyor, unit

Figure 5. (Color online) A Crane-Based Multideep Compact Storage System (De Koster, Le-Duc, and Yugang 2008)

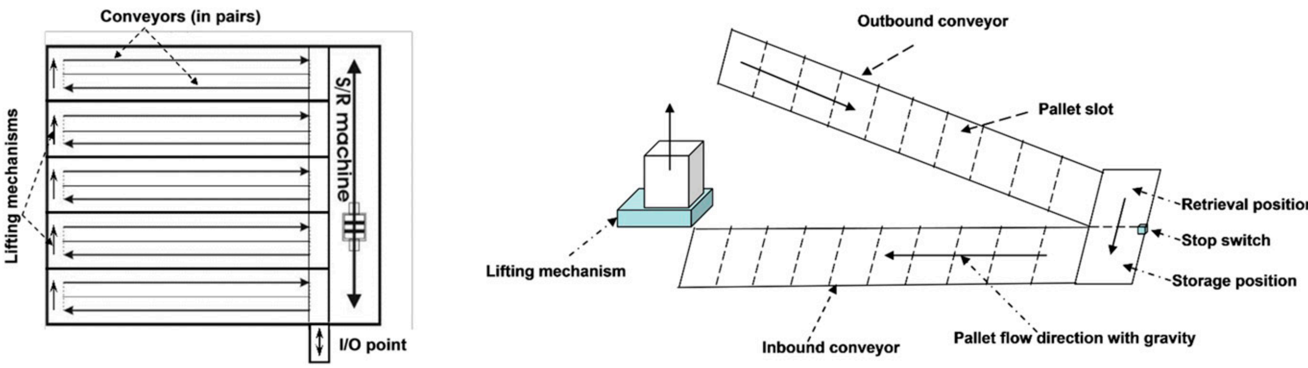


loads flow to the rack's front end and stop at the retrieval position of the crane. In the case of a gravity conveyor, the rack is equipped with a simple elevating mechanism at the back of the rack to lift unit loads from the down inbound conveyor to the upper outbound conveyor (see Figure 6). A stop switch located at the front side of the outbound conveyor stops a unit load when it is needed for retrieval. The lift drives the rotation of unit loads and, as it is the slowest element, it determines the effective rotation speed. To retrieve a pallet, the two neighboring gravity conveyors should have at least one empty slot (De Koster, Le-Duc, and Yugang 2008). The system with powered conveyors does not need lifts, but uses more expensive powered conveyors (that are not so easy to fix in the case of a malfunction). However, powered conveyors allow more dense storage because racks with powered conveyors can be constructed deeper than racks with gravity conveyors.

Satellite-based system. In this variant, a satellite (connected to the crane) or a shuttle (freely roaming) is used to perform the depth movement. The crane with a shuttle picks up a storage pallet and travels to the storage lane. Then the crane releases the shuttle in the rack and the shuttle travels along the storage lane to store the load. Likewise, to retrieve a load, the shuttle travels underneath the load to retrieve the pallet and completes the remaining operations in a reverse sequence. In some cases, the shuttles can also be dedicated to lanes. If a system has fewer shuttles than storage lanes, the crane moves the shuttles between the lanes (Stadtler 1996).

Unlike single-deep AS/R systems, the number of papers on multideep AS/R systems is limited. Sari, Saygin, and Ghouali (2005) develop closed-form travel time expressions for a flow rack AS/R system. The expressions, which rely on a continuous storage rack approximation, are validated using discrete-even simulations. The simulations use a discrete rack dimensional

Figure 6. (Color online) Working Mechanism of Gravity Conveyor (De Koster, Le-Duc, and Yugang 2008)



approach. They find that the percentage errors are quite reasonable (varying between 11% and 14%). Hence, such models can be used to estimate system throughput capacity.

De Koster, Le-Duc, and Yugang (2008) develop closed-form travel time expressions for a crane-based compact storage system with rotating conveyors, using a single-command cycle and random storage policy. The crane's expected retrieval travel times are identical for both gravity and powered conveyors. Using the expected travel time expressions, they calculate the optimal ratio between the three dimensions that minimizes the travel time. They also provide an approximate travel time expression for dual command cycles and use it to optimize the system dimensions. They find a counterintuitive result that the cube-in-time dimensions for the rack is not the optimal choice. The performance for a cube-in-time rack is still fairly good and deviates from the optimal rack configuration (optimal ratio along the three dimensions, 0.72:0.72:1) by about 3%. Yu and De Koster (2009a) extend the analysis of De Koster, Le-Duc, and Yugang (2008) for a turnover-based storage policy and determine the optimal rack dimensions that minimize the expected cycle time. They analytically determine the optimal rack dimensions for any given rack capacity and ABC curve skewness. They find that with greater skewness of the ABC curve, savings in the expected time increase compared with the random storage policy. Yang et al. (2015) further extend the analysis of De Koster, Le-Duc, and Yugang (2008) by optimizing the shape of the system and by considering the acceleration and deceleration of the S/R machine, which has a direct impact on the optimal shape of the system. For the special case of constant speed of the S/R machine, their findings are in line with the results of De Koster, Le-Duc, and Yugang (2008). Hao, Yu, and Zhang (2015) also develop expected travel time expressions and optimize the rack layout for a random storage policy. However, they choose an input/output (I/O) point located in the middle of the rack (which, in reality, is difficult to construct for aisle-captive cranes). Under the same operating conditions, they obtain lower expected travel time and higher throughput.

One of the biggest disadvantages of dense storage is that the pallets are accessible from only one side. Therefore, pallets are either retrieved based on the LIFO principle or they undergo multiple relocations/reshuffles to allow access to the right pallet. Stadtler (1996) uses the retrieval time estimate of each pallet and proposes a storage and retrieval assignment planning tool considering this issue. The decision models are formulated as mixed-integer programs and are solved using a tabu search heuristic. The results show that the compact storage systems can operate at heavy workload and high storage rack utilization with a small number of pallet relocations (6% relocations at 78% rack utilization over a 42-day period of operation). Yu and De Koster (2012) develop heuristic approaches to sequence a block of storage and retrieval transactions for a compact conveyor-based storage system operating in a dual-command cycle. They compare the makespan performance for five sequencing heuristics: (1) first come, first served; (2) nearest neighbor, in which the sequence is based on the minimum travel distance between storage and retrieval locations; (3) shortest leg, in which the open storage location lies on the Tchebychev path leading to the retrieval location; (4) shortest dual cycle, in which sequencing is done in a way to minimize the dual cycle time in every step; and (5) percentage priority to retrievals with shortest leg (PPR-SL), in which a certain percentage of retrievals are given a higher priority for prepositioning than the storage open locations. Numerical results suggest that the PPR-SL strategy outperforms all sequencing strategies by 20% or more. For a compact AS/R system with shuttles or satellites, one of the biggest challenges is the additional time required to reshuffle unit loads and retrieve the right unit. Many companies, therefore, use a dedicated storage policy per lane, which reduces the reshuffle time, but decreases lane utilization (and requires a larger system). To overcome this shortcoming, Zaerpour, De Koster, and Yu (2013) propose a mathematical model for a shared storage policy that minimizes the total retrieval time in a cross-dock/temporary storage environment. They solve the model using a construction and improvement heuristic. They show that for most real cases, shared storage outperforms dedicated storage, with a shorter response time and better lane utilization. Yu and De Koster (2009b) focus on identifying the optimal class zone boundaries for a compact 3D crane-based systems with two storage classes (a high turnover class and a low turnover class). They formulate the problem as a nonlinear integer program and obtain a solution using a decomposition technique and a one-dimensional search scheme. They show that the crane travel time is significantly influenced by zone dimensions, zone boundaries, and the ABC curve skewness.

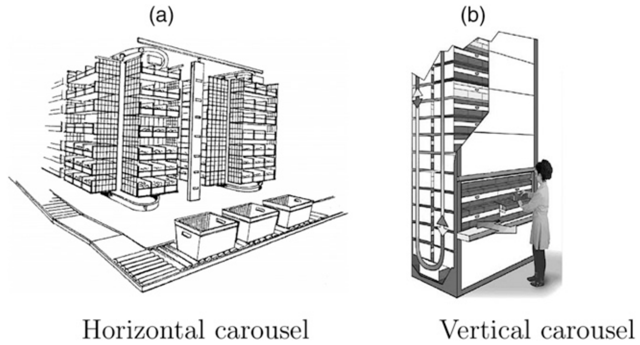
4. Carousels, Vertical Lift Modules, and Automated Dispensing Systems

Carousels are automated storage and retrieval systems in which shelves are linked together and rotate in a closed loop. The rotation is either horizontal or vertical (see Figure 7, (a) and (b)). In this system, the picker has a fixed location in front of the system, and the system transports the items to the picker. Carousels are especially suitable for small and midsize items such as books and health and beauty products (Litvak and Vlasiou 2010).

A VLM is similar to a carousel, but operates differently. It consists of two columns of trays with a lift-mounted inserter/extractor in the center (see Figure 8(a)). When an item is needed, the inserter/extractor locates the trays in which the item is stored and brings the tray to the picker, who is located in front of the system, like in a carousel (MHI 2017c). The static location of the picker in these systems eliminates pickers walking (Meller and Klote 2004), which can improve picking productivity. The pickers can also perform other tasks such as packing and labeling or even serving another carousel or VLM while waiting for the carousel to retrieve items.

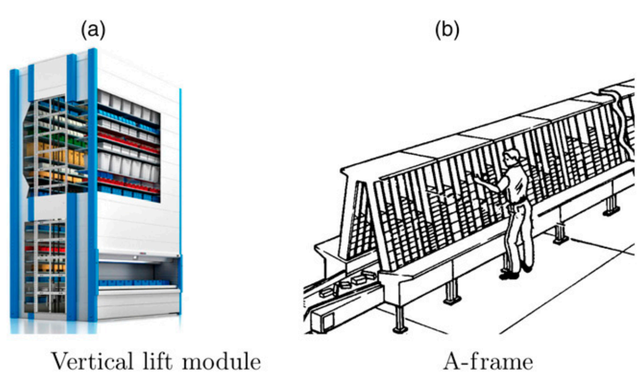
In an automated dispensing system, products are dispensed automatically. The replenishment is still carried out manually, but it can be done without interrupting the picking process. A common automated dispensing system is the A-frame. This system consists of product channels positioned in an "A" shape layout creating a tunnel in which the collection belt is located. The orders are filled by automatically dispensing the corresponding products in a virtual window on the conveyor belt (see Figure 8(b)). A-frames are suitable for large orders of small-sized items. The systems are mainly used in pharmaceutical, cosmetic, and mail-order industries (Pazour and Meller 2011, MHI 2017a).

Figure 7. Carousels (Meller and Klote 2004, Litvak and Vlasiou 2010)



Horizontal carousel models have been extensively studied in the literature dating back to the 1980s, when the basic foundation for studying carousals was laid out by Bartholdi and Platzman (1986). Different aspects have been studied, such as storage arrangement, response time, and design issues. Litvak and Vlasiou (2010) give an extensive literature overview on performance evaluation and design of carousel systems. Pazour and Meller (2013) investigate the effect of batch retrieval on the performance of the horizontal carousel system. They show that batching retrievals reduces the cycle time in the carousel by 20% compared with sequential processing. The number of studies on horizontal carousels have declined, and the only recent study is by Pazour and Meller (2013). The reason could be that more and more companies are replacing their horizontal carousels with shuttle-based storage and retrieval systems, which we discuss in Section 5. VLMs, on the other hand, have been studied in only a handful of articles. Meller and Klote (2004) develop a throughput model for a single VLM pod. Dukic, Opetuk, and Lerher (2015) extend the research to model the throughput of a dual-tray VLM. Rosi et al. (2016) use simulation to analyze the throughput performance of the single-tray VLM for different design profiles (height and width of

Figure 8. (Color online) Vertical Lift Module and A-Frame (MHI 2017b, c)



VLM) and the lift velocity. Similar to VLMs, A-frames and automatic dispensers have been studied in few articles. Caputo and Pelagagge (2006) develop a decision support system for an A-frame system. They use a heuristic approach to determine the number of channels in the system, reorder level, and maximum quantity to be dispensed based on recorded performance of the last period and the forecasted demand. Meller and Pazour (2008) investigate a SKU assignment problem for an A-frame and use a knapsack heuristic approach to solve it. Pazour and Meller (2011) develop a mixed-integer linear program to determine the infrastructure investment of an A-frame as well as SKU allocation to the A-frame. They develop a heuristic solution to a solve real-size SKU allocation problem. They also propose a closed-form equation to calculate the system throughput of an A-frame. Imahori and Hase (2016) investigate the SKU assignment of an A-frame as well as the optimal sequencing of the order retrievals to minimize the total retrieval time. They analyze the problem for computational complexity and develop a graph-based heuristic to obtain the best sequence for retrieving orders and SKU allocations. Kim et al. (2016) study the effects of different ejecting zone (EZ) methods on the performance of an A-frame system. An EZ is a segment of the conveyor belt that is dedicated to an order on which the required SKUs for that specific order are ejected while that zone passes through the A-frame. They investigate three EZ methods: unequal, equal, and combined. They use simulation to show which EZ method is suitable for the system, depending on the order throughput time and energy usage.

5. Aisle-Based Shuttle Systems

Throughput capacity of AS/R systems is constrained because only one crane is responsible for handling loads at all vertical levels within a given storage aisle. This led to a new generation of automated order-picking systems, autonomous vehicle–based storage and retrieval systems, which were first introduced by Savoye Logistics in the 1990s. Such systems are increasingly popular because the required investment is similar to that of AS/R systems, although they offer a much higher retrieval capacity and are also significantly more flexible in capacity. By using additional shuttles, system capacity can be increased, and by removing shuttles, capacity can be decreased. Typical AVS/R systems use shuttles, which can drive in the x direction and the y direction on any level in the aisle, and lifts move shuttles (or unit loads) between the levels. In this variant, shuttles can move only horizontally, and rely on lifts for vertical movements. Recently, several robotic solutions have emerged, in which the shuttles (called robots) have the ability to not only move horizontally but also elevate up to different tiers by either moving diagonally or vertically (Azadeh, Roy, and De Koster 2018). Therefore, the AVS/R system can be classified based on their shuttles' movement capability into three categories: horizontal, vertical, and diagonal systems (see Figure 2). In this section, we discuss different types of horizontal systems and leave the discussion on vertical and diagonal systems for Section 8.

5.1. System Description

5.1.1. Single/Double-Deep Storage. The storage area in an AVS/R system consists of aisles with multitier storage racks on both sides and a cross aisle that runs orthogonal to the aisles (see Figure 9). To perform storage and retrieval actions, a lift is used for vertical movements between tiers and autonomous vehicles or shuttles are used for the horizontal movements within the tier (Roy 2011). To retrieve a tote, a shuttle moves to the tote's storage location and picks up the tote, pulls it on board and moves toward the lift for vertical travel. Then the shuttle either hands the tote to the lift (tier-captive system; Heragu et al. 2008), or uses the lift to move the load to a lower level (tier-to-tier system; Heragu et al. 2008) where it is transferred to the pick station by conveyor belt. After picking, the tote again uses the lift and a shuttle to be stored in the system.

5.1.2. Multideep (Compact) Storage. Crane-based compact storage systems lack flexibility in the volumes they can handle. Shuttle-based multideep storage systems, using lifts instead of cranes, have more throughput flexibility by adding or removing shuttles. They are adapted for the safe and secure handling of a variety of products such as textiles, automobile spare parts, and fresh produce.

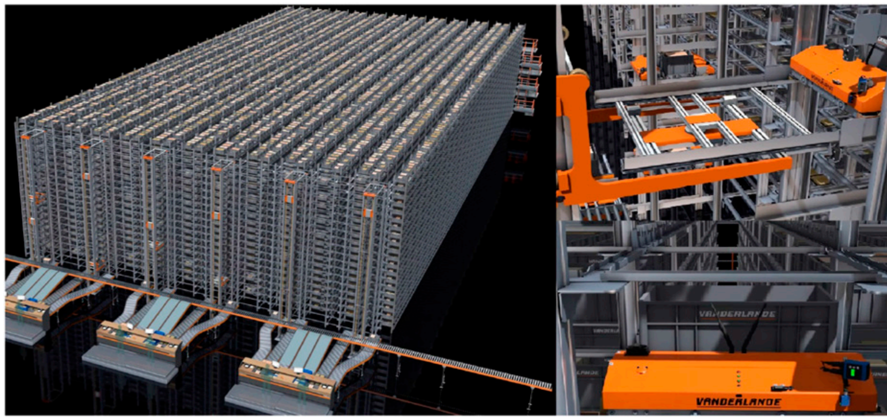
These systems consist of multiple tiers of multideep storage lanes, each of which holds one type of product (see Figure 10). The loads in a lane are managed using a LIFO policy unless the retrieval is possible from opposite sides. In such a system, the vertical transfer of loads (usually pallets) across multiple tiers is carried out using lifts, whereas the horizontal transfer of loads within a tier is carried out using shuttles. These shuttles move underneath the loads within each storage lane to store or retrieval the loads.

The horizontal movements of shuttles and loads in the system can be carried out either by "specialized" shuttles and a transfer car, or by "generic" shuttles that can move in both horizontal directions without the transfer car.

5.2. Literature

5.2.1. Single/Double-Deep Storage. Using the framework we discuss in Section 2.2, the literature on the single/double-deep horizontal AVS/R systems is categorized in three categories: system analysis, design optimization, and operations planning and control.

Figure 9. (Color online) Adapto AVS/R System



Source. Vanderlande.

System Analysis. Malmborg (2002) was the first to analyze the AVS/R system. He developed a state equation model to estimate the vehicle utilization and cycle time of the unit-load AVS/R system. He estimates the vehicle cycle time to be $(1-\alpha)t_{SC} + \alpha t_{DC}/2$, in which t_{SC} and t_{DC} denote the single-command and dual-command cycle times, respectively, and α is the proportion of all cycles that are dual-command cycles. Malmborg (2003a) emphasizes the design advantage of an AVS/R system relative to an AS/R system, which is the ability to adapt the vehicle fleet size in response to the transaction demand. Malmborg (2003b) extends the state equation model by including the number of pending transactions in the state space description, to estimate α, in a system with opportunistic interleaving; that is, dual-command cycles are used only if storage and retrieval requests are pending in the transaction queue at the time when the cycle is initiated. However, the state equation approach is computationally inefficient for solving large-scale problems. Therefore, Kuo, Krishnamurthy, and Malmborg (2007) and Fukunari and Malmborg (2008) propose a computationally efficient model to overcome this problem. In this approach, the lift system is modeled as a closed queueing network that is nested within a separate vehicle closed queueing network. They model the queuing dynamics between vehicles and transactions using an $M/G/V$ queue (with V vehicles), and the dynamics between transactions/vehicles and lift using a $G/G/L$ queue (with L lifts). The two systems are analyzed iteratively until the performance measures converge. Although the nested queueing approach is computationally efficient, it is not able to model a scenario in which the cycle starts outside of the storage rack, that is, when loads are received from outside the storage rack. Fukunari and Malmborg (2009) propose a queueing network model as an alternative to address this drawback. They propose a closed queueing network for estimating resource utilization in AVS/R systems. Although the earlier models are effective in estimating vehicle utilization with reasonable accuracy, they are ineffective in estimating transaction waiting times. Using a series of queuing approximations, Zhang et al. (2009) address this problem by dynamically choosing among three different queueing approximations, based on the variability of transaction interarrival times. This procedure significantly improves the accuracy of transaction waiting time estimates.

Recent studies use a semiopen queueing network to analyze the performance of the AVS/R system and estimate the external transaction waiting time with better accuracy. Roy et al. (2012) build a multiclass SOQN with class switching for a single-tier AVS/R system and design a decomposition method to estimate system performance. Ekren et al. (2013) model a tier-to-tier AVS/R system as an SOQN and present an analytical approximation by extending the algorithm of Ekren and Heragu (2010a) to estimate the performance measures. Ekren et al. (2014) improve the estimation of the number of transactions waiting in the vehicle queue by developing a matrix-geometric method for the SOQN model. Cai, Heragu, and Liu (2014) model a tier-to-tier system as a multiclass multistage SOQN

Figure 10. (Color online) Multideep Shuttle-Based Compact Storage System (Tappia et al. 2016)

and use matrix-geometric methods to analyze it. Ekren (2011) performs a case study by simulating the performance of a real AVS/R system under predefined design scenarios (number of aisles, bays, tiers, and vehicles). He also includes the total cost of the system in his analysis. The number of studies on tier-captive configurations is limited. Heragu et al. (2011), Marchet et al. (2012), and Epp, Wiedemann, and Furmans (2017) use the open-queueing network approach to estimate the transaction cycle time of the AVS/R system with tier-captive vehicles. Heragu et al. (2011) use an existing tool called the manufacturing performance analyzer to compare the performance of AVS/R systems and traditional AS/R systems. Ekren (2017) uses simulation to model the system and provide a graph-based solution for performance evaluation of the system (utilization of lifts and the cycle time) under various design configurations. Roy et al. (2017) model the system as an integrated queueing network and estimate the cycle time and resource utilization. They model each tier as a semiopen queueing network and the vertical transfer unit as a multiclass queueing network with $G/G/1$ queues corresponding to each vertical transfer segment. They replace each tier subsystem with a single load-dependent queue and approximate the first and second moments of interdeparture times using embedded Markov chain analysis. Then they solve the integrated model by capturing the linkage between arrivals and departures in the tier subsystem and the vertical transfer unit. Lerher et al. (2015) and Lerher (2016) develop travel time models for single-deep and double-deep AVS/R systems, respectively. They develop closed-form expressions for the cycle time and consider the effects of shuttle acceleration and deceleration.

Design Optimization. Roy et al. (2012) develop a semiopen queuing network model and optimize the shape of the system. Their results suggest that the layout configuration with depth-to-width ratio $D/W = 2$ for a system with the lift in the middle provides the best system performance. Roy et al. (2015a) extend the model, and show that the end of the aisle is the optimal cross-aisle location for the system. Ekren and Heragu (2010b) provide a simulation-based regression analysis for the rack configuration of the system. In their regression model, the average cycle time is chosen as the output variable, and the input variables are the numbers of tiers (T), aisles (A), and bays (B). The regression function demonstrates that the cycle time is positively related to T and B, but is negatively related to $T*A$ as well as to $T*B$. Marchet et al. (2013) simulate an AVS/R systems with a tier-captive configuration and illustrate the effect of rack configurations on the throughput performance. By varying the rack configuration and observing the performance impact, they optimize the shape of the system.

Operations Planning and Control. Ekren et al. (2010) develop a simulation-based experimental design to identify the effect of a combination of several input factors (dwell point policy, scheduling rule, I/O location, and interleaving rule) on the performance of the system (average cycle time, average vehicle, and lift utilization). They investigate the effect of up to four-way interactions of input variables on the performance of the system. Kuo, Krishnamurthy, and Malmborg (2008) use the closed queueing network approach to investigate the effect of a class-based storage policy on the cycle time of an AVS/R system. They conclude that class-based storage policies can mitigate the cycle time inflation effect of vertical storage, while keeping the space efficiency of the random storage intact. Kumar, Roy, and Tiwari (2014) simulate an AVS/R system in which the vehicles are captive in vertical zones rather than in tiers. They show that the optimal partitioning of vertical zones can reduce the transaction cycle times by up to 12% compared with the tier-captive configuration. Roy et al. (2012) develop a semiopen queuing network model and analyze the effect of vehicle location, the number of storage zones, and vehicle assignment policies on the performance measures. They show that using multiple zones reduces travel time along the cross aisle, which improves the performance of the system. However, increasing the number of zones beyond a threshold results in longer transaction waiting time and worsens the system performance. Finally, they observe that the most efficient vehicle assignment policy is the random policy. Roy et al. (2015a) extend the model to analyze different dwell point policies. They shows that the best dwell policy is the L/U point dwell policy. He and Luo (2009) use colored time Petri nets to dynamically model AVS/R systems and established the necessary conditions to have a deadlock-free system. Roy et al. (2014) use a semiopen queueing network to investigate the effect of vehicle blocking within a single tier of the AVS/R system. Their results show that the blocking delays could contribute significantly (up to 20%) to the transaction cycle time. They also show that the percentage of blocking delays goes up as the number of vehicles increases. However, the effect of blocking decreases as the utilization of vehicles increases, because the waiting time to obtain a free vehicle dominates in a system with high vehicle utilization. Roy et al. (2016) arrive at a similar conclusion using a simulation model. Roy et al. (2015b) evaluate congestion effects in a multitier AVS/R system. They develop a semiopen queuing network and use a decomposition-based approach to solve it. Their model provides the steady-state distribution of the vehicles at the cross aisles and aisles of each tier, conveyor loops, at the L/U point. The model also captures the blocking delays at the cross aisle and aisle nodes. Zou et al. (2019) investigate a scenario in which the lift and vehicles in the tier-captive AVS/R system are requested to move a load simultaneously rather than sequentially. They model the system with a fork–join queueing network. They show that the parallel processing policy improves the response time of the system by at least 5.5% compared with the sequential processing policy, for small-sized systems (system with fewer than 10 tiers). In large systems with more than 10 tiers and a ratio of aisle length to rack height of more than seven, they find a critical point for the retrieval transaction arrival rate. Before that rate, the parallel processing policy performs better. For arrival rates more than the critical point, the sequential processing policy should be used.

5.2.2. Multideep (Compact) Storage. The number of research articles on multideep AVS/R systems is limited. The literature is categorized in two categories: system analysis and design optimization.

System Analysis. Manzini et al. (2016) develop an analytical model to determine the travel time and travel distance for single and dual-command cycles for a layout configuration. D'Antonio et al. (2018) present an analytical model to calculate the cycle time and its standard deviation for a system.

Design Optimization. Tappia et al. (2016) model each tier and the vertical transfer mechanism using a multiclass semiopen queuing network and an open queue, respectively. They suggest that generic shuttles may reduce the total travel distance for storage and retrieval operations because additional shuttle movements in the cross aisle without a load are not required. However, they argue that a specialized shuttle might be attractive from an economic perspective, because a generic shuttle is about twice as expensive as a specialized one. They also show that a single-tier system with a depth-to-width ratio of around 1.25 minimizes the expected throughput time. Manzini et al. (2016) calculate the optimal location of the L/U point and the optimal shape of the system. They also calculate the optimal number and depth of the lanes depending on the demand pattern by minimizing the operative costs and maximizing the storage space efficiency.

An overview of the literature on shuttle-based storage and retrieval systems with aisles is presented in Table A.1.

6. Grid-Based Shuttle Systems

In this section, we discuss a variant of the shuttle-based automated storage and retrieval systems in which shuttles move on a grid. In a grid-based system, the storage locations are either dynamic or static (see Figure 2). In a dynamic storage system (or puzzle-based system), the stored SKUs need to move (on a shuttle) to store or retrieve an item. We discuss the static storage systems in Section 8.

6.1. System Description

Gue (2006) shows that the storage density of a k-deep aisle-based system is less than or equal to $2k/(2k+1)$, that is, 2/3 for a single-deep and 4/5 for a double-deep system. To achieve an absolute maximum storage density, a new concept based on Sam Loyd's famous puzzle game has been developed; the 15-slide puzzle (Loyd 1914). The 15-slide puzzle is a game in which 15 numbered tiles slide within a 4×4 grid, and the objective of the game is to arrange the tiles in the correct numerical sequence, starting from a random initial arrangement.

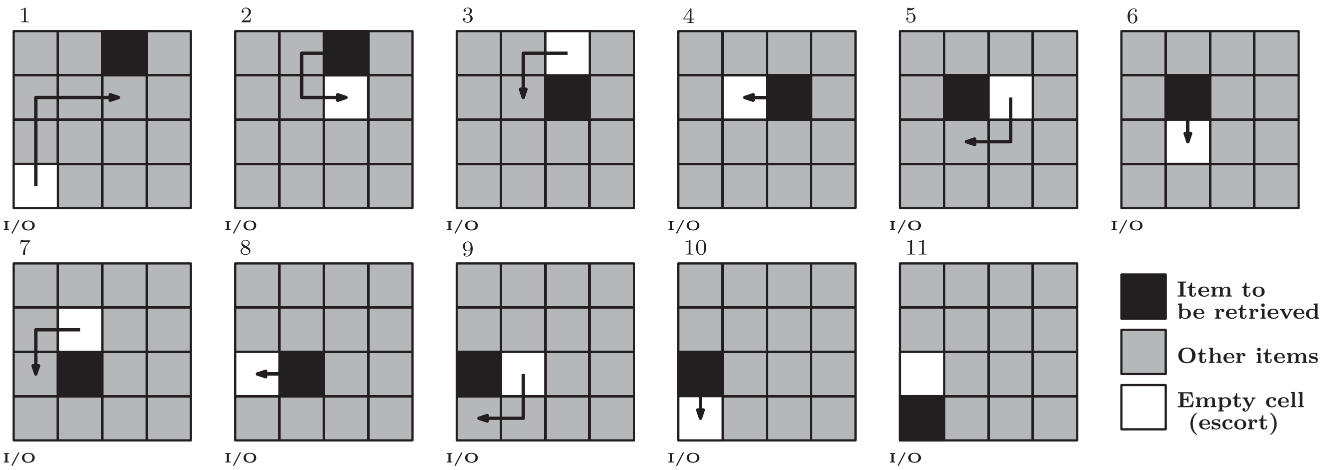
The puzzle-based storage and retrieval concept (Gue and Kim 2007), follows a similar idea. A tile represents a tote, a pallet, or even a container that is stored in a grid with only one open spot on the grid, which allows a $(n-1)/n$ storage density, where n is the number of cells in the grid. To retrieve a requested unit load, the system repeatedly moves the open locations, which ultimately brings the load to the I/O point. This is illustrated in Figure 11.

To retrieve a load, an open location first needs to be moved next to the requested item. Then the open location should be used to move the item to the I/O point. In other words, the open location "escorts" the requested item to the I/O point. An open location is called an *escort* (Gue and Kim 2007). Several compact storage system variants have emerged from the puzzle-based concept in practice and in the literature.

6.1.1. GridStore. Building on the puzzle-based storage system concept, Gue et al. (2014) propose a high-density storage system for physical goods called GridStore. The system consists of a rectangular grid of square conveyor modules with the capability to move items in the four cardinal directions. The modules can communicate with their neighboring modules as well as with the item they carry. At the south side of the grid, the retrieval conveyor moves products away from the grid. At the north, a replenishment conveyor moves products that need to be stored in the grid. Figure 12 illustrates the movement of the items toward the retrieval conveyor.

6.1.2. GridPick. Based on the GridStore architecture, Uludag (2014) introduces an order-picking systems called GridPick. The system is filled with high-density storage containers, without any fixed lanes or aisles; only a few open spots on the grid allow items to move during the retrieval process. The objective of the system is to provide a high order-picking rate while minimizing any congestion effects. Unlike the GridStore, items do not leave the grid in the GridPick system. Only

Figure 11. Maneuvering a Load (Item) to the I/O Point



containers holding the requested item move to the edge of the system, called the pick face. The picker picks the items and accumulates the order in a picking cart. There is also a backward movement, away from the pick face, to balance the empty cells in each row. This balancing rule helps to avoid deadlocks in the system.

Figure 13 illustrates an instance of GridPick. The gray items are not-requested stored items, and the black items are the requested items that are moving toward the pick face. Black circles on top of some gray items are balancing items moving in the opposite direction from the pick station. The numbers on top of the items display the order number for the requested item. The next order for picking is released when all the items from an order have arrived in the pick face of the system.

When comparing the GridPick with its equivalent gravity flow rack counterpart, Gue and Uludag (2012) show that the gravity flow rack results in a larger system. Therefore, the average productivity (measured in picks per hour) is higher for smaller orders in the GridPick system because it reduces travel. However, as order size increases, walking time of both systems converges to the same number.

6.1.3. Live-Cube Compact Storage. A multilevel system in which each floor is based on a puzzle-based storage architecture is called a live-cube storage system (Zaerpour, Yu, and De Koster 2015). As illustrated in Figure 14, the essential parts of the system are multiple levels of storage grids, shuttles, lifts, and the I/O points. Each level of the system forms a grid-based storage system where shuttles move in x and y directions with the load on top of them. With at least one escort available in each level, the shuttles maneuver the requested item to the lift, which transports the load to the I/O point. The I/O point is usually located at the lower left corner of the system.

6.1.4. GridFlow System. A major drawback of the puzzle-based system is that the physical layout cannot be changed easily. Therefore, the concept of GridFlow is proposed by Furmans, Nobbe, and Schwab (2011) to offer a cheaper and a more flexible system.

Figure 12. Items Movements Toward the Retrieval Conveyor in GridStore

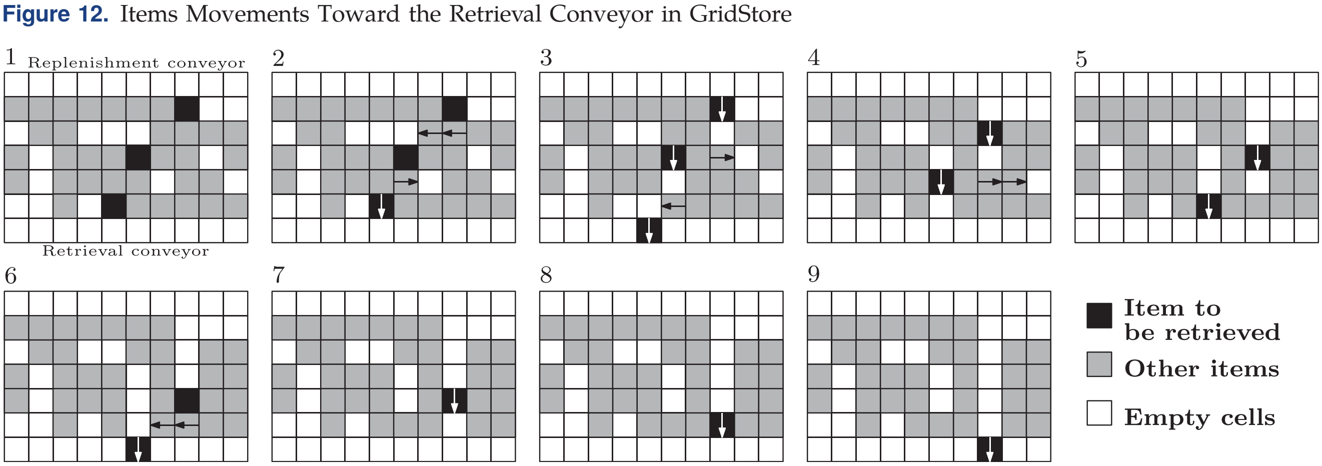
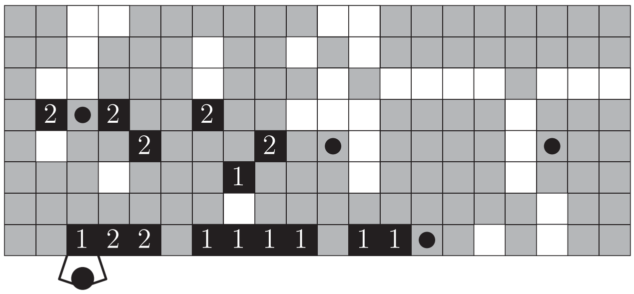


Figure 13. An Instance of the GridPick System



In this system, instead of conveyors, AGVs are used to move the pallets. The use of AGVs instead of conveyors makes the system more flexible with respect to design and throughput changes. Vehicles can form grids of any shape without any additional investment. Figure 15 illustrates the GridFlow system and the vehicle movements in the system.

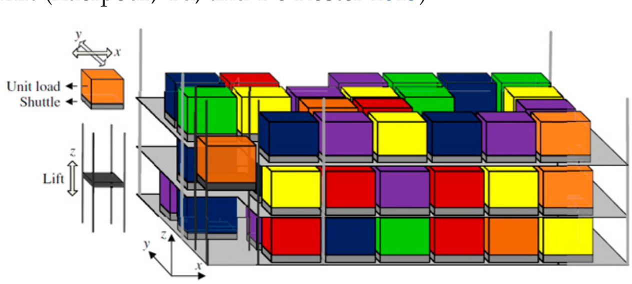
Many system manufacturers are developing puzzle-based systems in different variants. The number of actual implementations and prototypes based on this concept is growing in many different fields, especially in automated parking systems (e.g., "park, swipe, leave" parking system, Automation Parking Systems 2018; Hyundai integrated parking system, Hyundai Elevator Co. 2016; Wöhr Parksafe, Wöhr 2016).

6.2. Literature

The literature on the puzzle-based storage and retrieval systems is categorized in three categories: system analysis, design optimization, and operations planning and control.

6.2.1. System Analysis. Gue and Kim (2007) develop an algorithm to find an optimal path to retrieve an item in the puzzle-based system with a single escort positioned at the I/O point. They propose a dynamic programming approach for multiple escorts and a heuristic for larger instances. Their results confirm the intuition that having more escorts shortens the retrieval time. The only exception occurs for smaller systems with many escorts at the I/O point. They also compare the performance of the puzzle-based system with its aisle-based counterpart. They find that aisle-based systems have shorter retrieval times than puzzle-based systems, unless the desired storage density is more than 90%. Kota, Taylor, and Gue (2015) develop a closed-form expression for the retrieval time in the puzzle-based storage system with a single or two randomly scattered escorts within the grid. They propose a heuristic solution for more than two escorts in the system. Their heuristic gives a near optimal solution, except for the time when free escorts are congested near the edge of the grid. Zaerpour, Yu, and De Koster (2015) investigate a multitier puzzle-based (live-cube) storage system. They assume that there are sufficient escorts available at each level so that a virtual aisle can be created (minimum number of escorts is the maximum of the rows and columns in the system). They use traditional methods for the aisle-based system and derive a closed-form formula for expected retrieval time. Zaerpour, Yu, and De Koster (2017b) propose a two-class-based storage policy for a live-cube system. They derive closed-form formulas to calculate the expected retrieval time of the system. They conclude that their proposed storage policy can improve the average response time of the system up to 55% compared with the random storage policy, and up to 22% compared with the cuboid two-class-based storage policy.

Figure 14. (Color online) A Live-Cube Storage System with Lift (Zaerpour, Yu, and De Koster 2015)



6.2.2. Design Optimization. Gue et al. (2014) analyze the optimal shape of the GridStore system. They find that a system with more columns has a higher throughput with the same number of stored items. Zaerpour, Yu, and De Koster (2015) propose and solve a mixed-integer nonlinear model to optimize the dimensions of a live-cube system by minimizing the retrieval time assuming a random storage policy. Zaerpour, Yu, and De Koster (2017b) extend this work by considering a two class-based storage policy. Their results show that the optimal dimensions of the system are identical for two class-based and for a random storage policy. Zaerpour, Yu, and De Koster (2017a) propose a mixed-integer nonlinear model to optimize the dimensions and zone boundaries of the two-class live-cube storage system by minimizing the response time. Furmans, Nobbe, and Schwab (2011) investigate the design choices for the GridFlow system with one vehicle and one escort. They conclude that putting the I/O point in the middle of the longer side of the grid produces the best performance. Furthermore, they show that the 2:1 aspect ratio results in the lowest retrieval time when the number of storage locations is less than 2,000. Their results are not conclusive for larger storage capacities.

6.2.3. Operations Planning and Control. Taylor and Gue (2008) investigate the effect of the distribution of escorts in the puzzle-based system. They examine three choices for the initial location of escorts: (1) near the

I/O (located at a lower left corner of the grid), (2) along the diagonal from lower left to upper right, and (3) randomly on the grid. They show that when the number of escorts is above 25%, having the escorts along the diagonal always outperforms the other strategies. The only exception occurs when the storage is based on an ABC policy, in which random placement for the escorts is the best option. Yu et al. (2017) consider a puzzle-based storage system with multiple escorts, in which multiple loads and escorts are allowed to move simultaneously and in blocks (simultaneous movement of loads in a line). Using integer programming, they obtain the optimal retrieval time of a single item in the system. Their results show that allowing loads and escorts to move simultaneously and in blocks can save up to 70% in the total number of needed moves to retrieve an item. Mirzaei, De Koster, and Zaerpour (2017) propose an approach for simultaneous retrieval of multiple items. They derive the optimal retrieval time for double-item and triple-item retrieval using enumeration. They propose a heuristic algorithm for more than three simultaneous item retrievals. They show that double-item retrieval policy reduces the storage/retrieval time by an average of 17% compared with a sequential retrieval policy. Cycle time savings can be further increased by performing multi-item retrievals. Gue et al. (2014) propose a decentralized assess–negotiate–convey control scheme for the GridStore system, in which each conveying cell can execute the same set of instructions based on its local condition. They also investigate the effects of work in process (WIP) and the number and distributions of escorts per row on the throughput. First, they assume that escorts are uniformly distributed in the rows. They show that for medium and low levels of WIP in the system, the throughput increases with an increasing rate with an additional request. Next, they investigate two additional distribution of escorts: more escorts in the southern row (increasing k) and fewer escorts in the southern row (decreasing k). They show that the distribution of escorts has no effect on the throughput for a low level of WIP. The increasing k performs better at low to moderate WIP levels, and all distributions perform equally well at a high level of WIP. Alfieri et al. (2012) investigate the GridFlow system with a limited number of vehicles. They propose a heuristic algorithm to optimize the movement of shelves and to dispatch the AGVs optimally.

An overview of the literature on the puzzle-based storage and retrieval systems is presented in Table A.2.

7. Robotic Mobile Fulfillment Systems

Internet retailers typically have warehouses with large assortments of small products. Their demands usually consists of multiline small-quantity orders. In manual picking systems, much nonvalue added time is needed by the pickers to travel along the aisles. The robotic mobile fulfillment (RMF) system is a system in which robots capable of lifting and carrying movable shelves retrieve the storage pods (i.e., movable shelf racks) and transport them to the pickers, who work in ergonomically designed workstations. Bringing the inventory to the picker instead of the picker traveling to the inventory can double the picker productivity (Wurman, D'Andrea, and Mountz 2008). The system is also very flexible in throughput capacity, as more robots and pods can be added to the warehouse. This is particularly important for internet retailers who face volatile demand. The RMF system was conceptualized by Jünemann (1989) and was patented in the Unites States by KIVA Systems Inc. (Mountz et al. 2008), which then was acquired by Amazon and rebranded to Amazon Robotics. Today the system is operational in many Amazon facilities. Meanwhile, other providers have also entered the market with mobile racks in combination with robots, such as CarryPick™ by Swisslog, Butler™ by GreyOrange, the Scallog System™, and Racrew™ by Hitachi (Banker 2016).

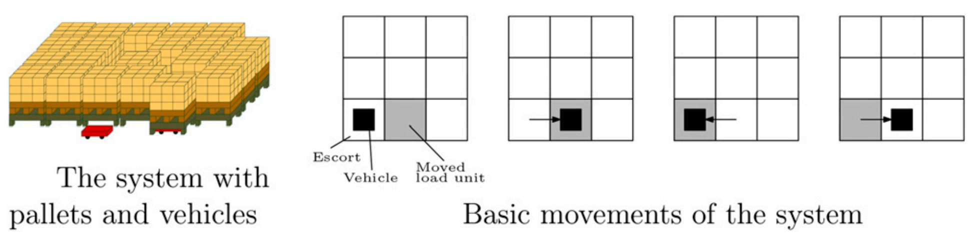
7.1. System Description

The RMF system consists of three major components:

1. *Robotic drive units*. These robots are instructed by the central computer to transport inventory pods to the workstation for restocking or for picking. Nowadays also decentrally (or locally) controlled systems exist.

2. *Inventory pods.* Pods are movable shelf racks that contain the stored products. Pods come in two standard sizes. Smaller pods are used for weights up to 450 kg, and large pods are used for weights up to 1300 kg.

Figure 15. (Color online) The GridFlow System (Furmans, Nobbe, and Schwab 2011)



3. *Workstations.* Workstations are ergonomically designed areas where human workers perform pod replenishment, picking and packing functions (MWPVL International 2012). Figure 16 presents an RMF system workstation and its components.

To pick an ordered item with the RMF system, the order is first assigned to one of the workstations. Then the item is assigned to a pod and one robot. The robot then moves from its dwell location to retrieve the pod. At this point, the robot moves without a load and can therefore move underneath the pods, without using the designated travel aisles. Once the robot reaches the desired pod, it moves underneath it, lifts the pod, and transports it to the workstation via the travel aisles. The robot enters the workstation buffer and waits for its turn (see Figure 16). The picker takes the requested products and adds them to the customer order bin placed in a different rack. The robot then returns the item pod to a storage location that accounts for the frequency of the requests for the pod. The storage locations are therefore fully dynamic (Wurman, D'Andrea, and Mountz 2008; Enright and Wurman 2011). The layout can be fully adapted both dynamically and automatically to the product and order characteristics.

7.2. Literature

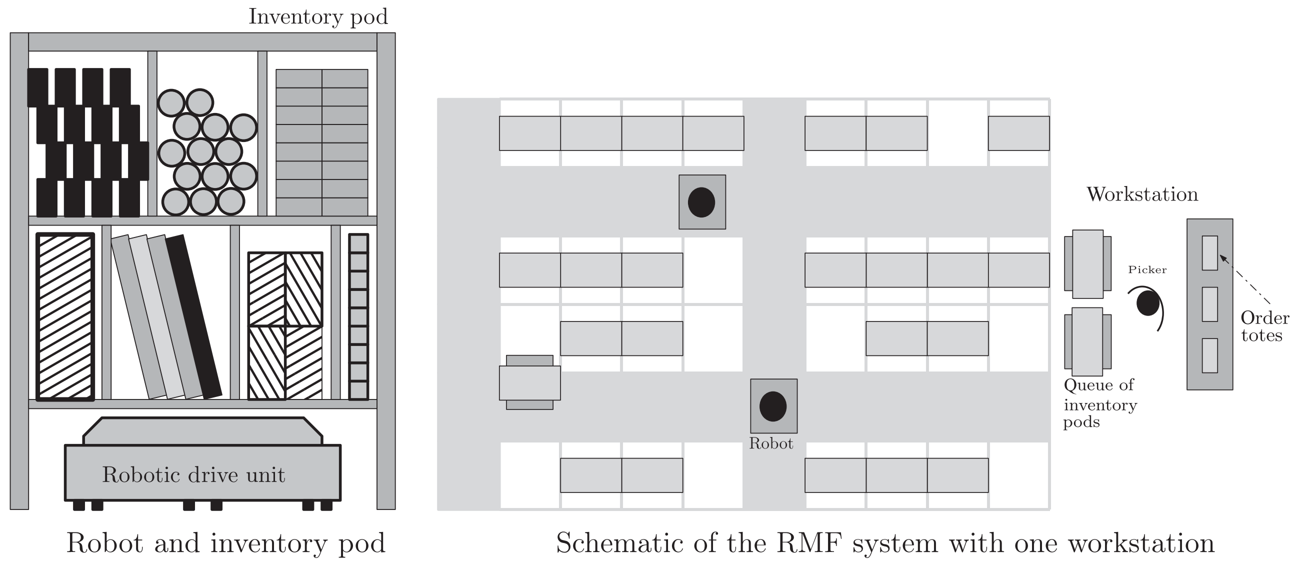
The performance of RMF systems has hardly been studied scientifically. The literature on the RMF systems is categorized in three categories: system analysis, design optimization, and operations planning and control.

7.2.1. System Analysis. Nigam et al. (2014) develop a closed queueing network model for an RMF system. They estimate order throughput time for single-line orders in an RMF system with a turnover class-based storage policy. Lamballais, Roy, and De Koster (2017a) extend the work of Nigam et al. (2014) by deriving travel time expressions for multiline as well as single-line orders in an RMF system with storage zones. They develop a SOQN to estimate the average order cycle time and the utilization of robots and workstations.

7.2.2. Design Optimization. Lamballais, Roy, and De Koster (2017a) show that the maximum throughput capacity in an RMF system with storage zones is insensitive to the length-to-width ratio of the storage area (unless the ratio is strongly skewed). However, they show that the positions of the workstations around the storage area directly affects the throughput capacity. Within their settings, the workstations should be located west and east of the storage area when turnover-based zoned storage is used, and north and south of the storage area when zoned storage is not used, to maximize the throughput. Yuan and Gong (2017) develop an OQN to estimate the total throughput time of the RMF system. Using the developed model, they calculate the optimal number of the robots and their required average speed to achieve a certain throughput time. Zou et al. (2019) optimize the shape of the system using an SOQN.

7.2.3. Operations Planning and Control. Nigam et al. (2014) show that the closest-open-location pod storage strategy does not use the storage space efficiently compared with the random location pod storage policy. However, the closest-open-location policy achieves a slightly higher throughput capacity. Yuan, Cezik, and Graves (2018) investigate the performance of a velocity-based storage assignment for an RMF system. A velocity-based policy is a policy in which the popular items are stored closer the pick stations. By using a fluid model, they show that a two- or three-class velocity-based storage policy reduces the travel distance by 8% or 10%, respectively, compared with a random storage policy. Lamballais, Roy, and De Koster (2017a) show that the maximum throughput in an RMF system with storage zones can be increased by almost 50% by using pod turnover–based storage zones. One of the drawbacks of the analysis by Lamballais, Roy, and De Koster (2017a) is that they assume items on one pod are all the same; for a multiline order, multiple pods are required. However, in reality, each pod contains multiple products. Therefore, it might be possible that a single pod can fulfill multiple requests of an order. Lamballais, Roy, and De Koster (2017b) address this issue by investigating how the inventory of products should be spread across storage pods. They develop an SOQN to estimate the throughput time. They then optimize the number of pods per product, the ratio of the number of workstations to replenishment stations, and the replenishment level for each pod to minimize the throughput time. The results show that the inventory should be spread across as many pods as possible to minimize the throughput time. Furthermore, they find that the optimal ratio of pick stations to replenishment stations is two to one, and that the optimal replenishment level is about 50%. Boysen, Briskorn, and Emde (2017) investigate sequencing picking orders at the work stations of an RMF system. They formulate the problem as a mixed-integer program. Their results show that by optimally sequencing the picking orders, the order fulfillment process can be done with half of the fleet size of the robots compared with the first-come-first-serve order sequencing rule. Furthermore, they show that the robot fleet can be further reduced by using the shared storage policy, in which the same SKUs are spread over multiple pods. Zou et al. (2019) investigate different battery recovery strategies for the RMF systems. They develop an SOQN to model the charging process. They conclude that inductive charging provides the best system throughput time. They also realized that the battery swapping

Figure 16. Elements and Layout of the RMF Systems



strategy outperforms plug-in charging strategy, but it is more expensive (Roy et al. 2019).

An overview of the literature on RMF systems is presented in Table A.3.

8. Directions for Future Research

Our review shows that four major topics require investigation for all systems:

1. *System analysis:* How does the system perform with respect to important performance measures (such as throughput and throughput time) for a given system configuration?
2. *Design optimization:* How can the system be designed to optimize certain performance measures, including the optimal shape of the system, and optimal number and locations of workstations?
3. *Operational policies:* What is the impact of different operational policies on the system performance, such as the effect of storage policies, efficient robot blocking prevention, and dwell point policies?
4. *System comparison:* How do different systems compare on performance, space and resource utilization, and operational costs?

Not all these questions have been addressed yet for all systems reviewed. In this section, we first discuss important generic research topics for established automated systems, which are not system specific and require further scientific investigation (Section 8.1). Next, we provide several research topics that are specific to the automated systems reviewed in this paper, that is, shuttle systems and RMF systems (Section 8.2). Then we identify promising new emerging technologies that have hardly received any research attention yet (Section 8.3). We finally discuss other areas related to automated warehouses that require further scientific research.

8.1. Generic Research Topics for Established Systems

8.1.1. Integrated Models. Almost all existing studies on automated or robotized warehouses analyze storage and pick systems in isolation. For instance, the literature on shuttle systems focuses primarily on storage systems; optimal policies are derived without considering the effect of the storage system configuration on downstream pick performance. Likewise, the literature concerning RMF systems focuses mostly on design issues, rather than operational policies that integrate the picking, storage, and replenishment processes.

To design optimal system configurations, integral consideration of interactions between both upstream processes (such as receiving and reserve storage) and downstream processes (such as picking and packing) in the warehouse is crucial. Such integrated models can capture the variations in the receiving and the picking throughput requirements, which may vary across days and weeks. In particular, the replenishment rates may be lumpier compared with the pick rates. Researchers can take inspiration from integrated models developed in the recent past for container terminal operations (see Meisel and Bierwirth 2013). Although the berth allocation and quay crane allocation problem have been dealt with separately in literature, new algorithms with an integrated focus can improve the joint performance.

8.1.2. Nonstationary Demand Profiles. Existing research on automated systems focuses largely on performance analysis using stationary inputs. However, because of the ever-changing demand profile, especially in e-commerce environments, it is crucial to take into account nonstationary demand profiles to create a robust design with dynamic operational policies.

Sample research questions may be the following: How can we develop dynamic operational policies, such as selecting dwell point locations and shuttle blocking prevention policies, when the demand profile is nonstationary? What is the optimal shape of the storage area in an RMF system when the demand is nonstationary? Or, what is the optimal pod repositioning policy in an RMF system when the demand changes over time?

8.1.3. New Storage Policies. Large amounts of e-commerce data provide new insights on customer shopping behavior. In particular, it is possible to estimate, with a very high accuracy, which items will be ordered together. As a result, new storage policies can be introduced that incorporate the product affinity (Mirzaei, Zaerpour, and De Koster 2018). The associated research question may be, how can product affinity be exploited in a storage policy, and how does that compare with other storage policies, such as class-based or random policies? In decision support systems and marketing literature, market basket analysis (also known as association-rule mining) has been used to discover customer purchase patterns by extracting associations or co-occurrences from transactional databases (see Chen et al. 2005). Using real-time customer behavior data, dynamic storage policies may be developed, which can improve pick cost and responsiveness.

8.1.4. Product and Order Sequencing. Some of the systems and processes shown in Figure 1 (the steps and systems that can be found in fully automated warehouses) have not yet received much research attention. For example, in Step 4, totes with products have to be retrieved for multiple orders, from, from an AVS/R system, to arrive at the stacking robots in the proper stacking sequence (Steps 5 and 6). Usually these robots have some freedom in item selection. However, it is still very important to have a correct retrieval sequence to improve the performance. So the question is, how can the retrieval shuttles in the AVS/R system be scheduled, with precedence constraints, to improve the stacking process? Currently, heuristics are used, and much slack is built into the systems. Order sequencing can also improve the efficiency of picking operations for RMF systems. Specifically, by sequencing the orders, it is possible to improve pod coverage; that is, more items can be picked per pod (see Boysen, Briskorn, and Emde 2017, who recently published a first paper on this topic).

8.1.5. Multiline Order Picking. While single-line orders form the majority of e-commerce order volumes, we expect that the share of multiline orders will increase, thereby improving packaging efficiencies and reducing carbon footprints. Many retailers offer free shipping for minimum buy quantities. However, most analytical models consider only single-line order picking. Obtaining design insights with multiline order picking is crucial.

8.1.6. Analytical Model Accuracy. Models inherently have to make assumptions to make them tractable. Also, the majority of the existing (stochastic) analytical models have been validated using discrete-event simulation. The credibility of the obtained insights will increase if the models are validated using output measures from real system implementations. In particular, the external queue length measure, which reflects the number of customer transactions waiting to be served, is an important measure for practical decisions. Existing analytical models validated with discrete event simulation reflect errors of up to 40% under various input scenarios (see Roy 2016). It is not clear how these errors compare with real output data. Other issues, such as nonstationarity of demand, discussed earlier, also play a role. We think more effort must be put into verifying the validity of the output.

8.2. Research Topics for Shuttle Systems and RMF Systems

8.2.1. Shuttle Systems.

Multiple Input/Output Points. The majority of the literature on shuttle systems provides design and operational choices assuming a single I/O point for the system. However, many systems have multiple I/O points. New studies are required to investigate questions such as, what is the effect of having multiple I/O points on the design and operational choices, such as depth-to-width ratio and storage policies?

Automated Replenishment. Some systems combine automated storage and replenishment of the pick system (like steps two and three shown in Figure 1) with manual picking. Particularly, if the number of pick slots is smaller than the number of products, scheduling the retrievals so that the picker does not have to wait, is challenging. Product bins from which units have been picked already have to be returned to the bulk storage system (Step 2). This problem has been studied to a limited extent by some researchers, but only in combination with manual pick processes (Yu and De Koster 2010; Ramtin and Pazour 2014, 2015; Schwerdfeger and Boysen 2017; Füßler and Boysen 2017, 2019). Further research is needed on systems with automated picking and for different storage and retrieval configurations.

8.2.2. Robotic Mobile Fulfillment Systems.

Storage Decisions. For RMF systems, two storage decisions have to be taken: first, how the pods should be stored in the storage area, and second, how the SKUs should be divided over the pods. Artificial intelligence and deep learning can be very helpful to understand order patterns, which can be then used to match the right SKUs to pods as well as to decide where to store the pods dynamically every time they have been retrieved for picking.

Replenishment Policy. The pod replenishment policy for RMF systems differs from that for other systems because multiple SKUs are stored in each pod. Therefore, deciding when to retrieve the pod for replenishment is a challenging question. So the research question is, what is the optimal inventory threshold for replenishment of the pods?

8.3. Description and Research Topics for Emerging Technologies

8.3.1. System Description.

Vertical and Diagonal AVS/R Systems. In these systems, a single robot can independently roam the storage rack to perform storage and retrieval operations (no lift is required). In a diagonal system, robots also move diagonally, and in a vertical system, robots also move vertically inside the rack structure to elevate to the upper levels. The RackRacer (see Figure 17(a)) developed by Fraunhofer Institute for Material Flow and Logistics (2014) is an example of a diagonal system. Perfect Pick® developed by OPEX Corporation (2013) and the Skypod™ (see Figure 17(c)) developed by Exotec Solutions (2017) are two examples of vertical systems. The Perfect Pick system uses robots, called iBots® (see Figure 17(b)), to perform storage and retrieval actions (Azadeh, Roy, and De Koster 2018).

The diagonal system has not yet been studied, whereas the vertical system has been studied in only one paper. Azadeh, Roy, and De Koster (2018) model a single aisle of the vertical system using a closed queueing network to optimize the shape of the system. They also investigate the effect of different robot blocking policies on the performance of the system. Finally, they also compare the operational performance and costs of vertical and horizontal systems.

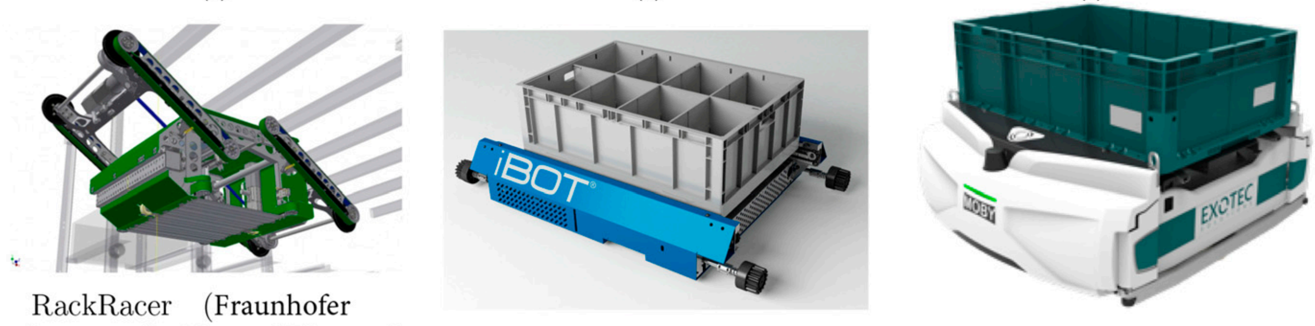
Robot-Based Compact Storage and Retrieval Systems. The robot-based compact storage and retrieval (RCSR) system is another type of grid-based system (see Section 6) in which items are stored in a very dense storage stack with a grid on top (see Figure 18). In each cell of the grid, bins that contain the items are stacked on top of each other and form the storage stacks. The workstations are located at the lowest level next to the storage stacks. Robots roam on top of the storage block on the grid. The robots have lifting capabilities and can extract bins from the storage frames and transport them to the workstations (Zou, De Koster, and Xu 2016). AutoStore™, developed by Hatteland is the first implementation of an RCSR system (see AutoStore 2018). Recently the British retailer Ocado developed a similar system.

Zou, De Koster, and Xu (2016) are the only ones to investigate the RCSR system. They model the system as a semiopen queueing network and compare two storage policies, namely, dedicated and shared storage. They show that the dedicated policy results in a shorter throughput time, whereas the shared policy has financial benefits because of substantial cost savings in the total storage space. They also optimize the shape of the system and show that the width-to-length ratio is around 2/3 when using random storage stacks, and slightly larger when using zoned storage racks. They also show that immediate reshuffling can improve the dual command throughput time compared with delayed reshuffling.

GridSort. The GridSort system is based on the GridFlow system discussed in Section 6. It uses modular four-directional conveyors, called FlexConveyors (Furmans, Schönung, and Gue 2010), or AGVs to transport and sort the loads. Recently, Libiao Robotics developed a different type of GridSort system, used by several parcel carriers in China, that uses a fleet of hundreds of autonomous AGVs on a grid to sort parcels by destination.

Pick Support AGVs. Most retail warehouses still use manual order-picking systems. Retail stores usually place large replenishment orders at the distribution center. The distribution center then ships the orders in multiple roll cages or on pallets. Therefore, a single order requires multiple pick tours (trips between pick locations and the depot). Recently, AGV-based pick systems, called pick support AGV systems, have been developed to minimize the picker travel time to fill large orders (see Figure 19). In these systems, an AGV automatically follows the picker closely and transports the roll cages, so that the picker can drop off the retrieved items. Once the roll cage is full, the AGV is automatically swapped with a new AGV carrying an empty roll cage. The picker can continue the picking route without returning to the depot, and the AGV automatically transports the full roll cage to the depot. AVGPick™ developed by Swisslog and Pick-n-Go™ developed by Kollmorgen (see Kollmorgen 2010) are two examples of such systems. Locus Robotics (2018) has developed another variant of this system. Instead of following the picker, their AGV (called LocusBots™) automatically goes to the pick location and waits for the picker to arrive. Once the picker puts the item into a customer tote carried by the AGV, the AGV goes to

Figure 17. (Color online) Robots in Single-Touch Systems



(a) RackRacer (Fraunhofer Institute for Material Flow and Logistics)

(b) iBot (OPEX Corporation)

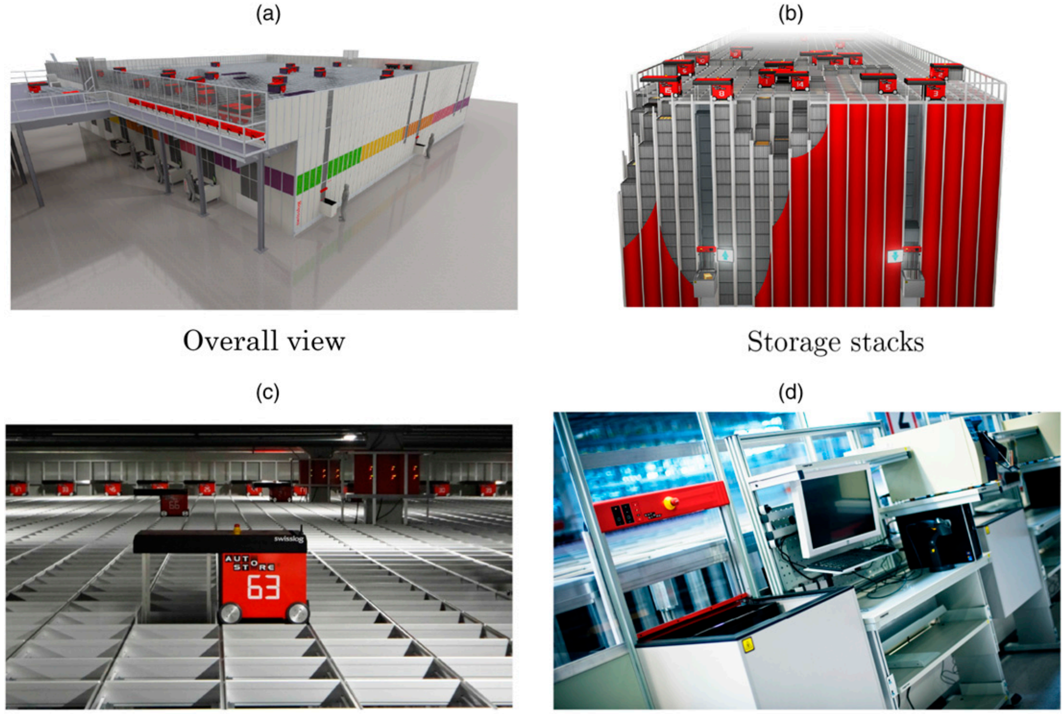
(c) Skypod (Exotec Solutions)

the next location. When the order is complete, the AGV transports the tote to the depot. Some systems automate the whole picking process. An example is the TORU™ picking robot. In this variant, the AGV automatically goes to the picking location and picks up the item without any help from the picker. Similar to the previous variants, once the order is complete, the AGV transports the picked items to the depot (see Magazino 2017).

8.3.2. Research Topics. All the four general research questions mentioned at the beginning of this section should be investigated for these new technologies (i.e., system analysis, design optimization, operational policy, and system comparison). Furthermore, there are unique characteristics that lead to some system specific research questions.

The distinguishing characteristic of vertical/diagonal systems compared with the horizontal systems is the roaming flexibility of the robots. This provides several routing trajectories to perform storage and retrieval actions. Therefore, a key research question is, what is the appropriate routing trajectory for the robots, considering performance, blocking delays, and operational costs?

Figure 18. (Color online) Robot-Based Compact Storage and Retrieval System



(a) Overall view

(b) Storage stacks

(c) Robot

(d) Workstation

Source. Hatteland.

Figure 19. (Color online) Pick Support AGVs

(a) (b)

A picker following AGVs — TORU

Source. The photo of the TURO is from Magazino.

The differentiating factor of RCSR systems compared with other systems is the fact that items are stacked on top of each other. Therefore, it is crucial to take into account reshuffling and congestion effects when analyzing the system. A particular research question could be, what is a good storage policy to minimize reshuffling and maximize throughput of the system?

GridSort differs fundamentally from conventional conveyor-based sorters and new models will be required to evaluate its performance. For example, in GridSort the movements of the shuttles (that carry the loads) depend on the empty spaces on the grid. So a research question may be, how can the empty spaces on the grid be exploited to simultaneously move multiple loads in the system efficiently?

The differentiating characteristic of pick support AGVs is the collaboration between human pickers and AGVs. The parallel movement of pickers and AGVs makes the modeling, analysis, and optimization of this system completely different from manual picking systems or the robotic systems mentioned earlier in this paper. Evaluating the performance of these systems is an interesting stream for future research. The interesting research question is, how can we coordinate the parallel movement of the pickers and AGVs to maximize the throughput?

Note that the different systems and research questions will require different methods, as suggested in Table 1.

8.4. Other Research Areas

8.4.1. Dynamic Optimization Using Data Analytics. Robotic warehouse systems are much more flexible in capacity, portability, and extension options than conventional automated storage and retrieval systems. These new capabilities make it possible to dynamically optimize warehouse decisions using real-time data. However, so far, no research has been published. Questions that justify attention include when to scale up or down retrieval capacity, when to open or close pick stations, when to reallocate tasks to workers, when to change picking methods (e.g., in a system with pick support AGVs), and when to modify the system layout or storage strategies (in RMF systems or pick support AGVs). These decisions can be taken at any point in time and some can be executed rapidly.

8.4.2. Comprehensive Evaluation of System Variants. With the rapid introduction of new automation technologies, distribution center managers are confronted with multiple technology options. Therefore, it is crucial to create a framework to help distribution center managers in evaluating different automated systems. In other words, how can we develop a framework for the comprehensive evaluation of technologies against cost and time? The only (published) paper in this area is by Pazour and Meller (2014). However, more research is needed particularly to include more recent automated systems.

8.4.3. Modular System Analysis Architectures for Proposing High-Performing New Technologies. To date, academic research on system design is reactive; that is, researchers evaluate designs after implementation. Is it possible to develop a modular system analysis architecture to analyze different design elements and propose a new system configurations? Solving this problem presents significant challenges because of the large design space, interaction between design elements, and diverse need of warehouse designers. So the question is, how can we develop a modeling framework which assists in analyzing design options and provides quick insights for developing new system designs?

8.4.4. Analyzing Robot Pick Stations. Most automated systems described in Figure 2 and the previous sections retrieve unit loads and bring them to an order pick station. Even with new technologies, like RCSR systems, puzzle-based systems, or RMF systems, piece picking at the station is done manually (De Koster and Yu 2008; Füßler and Boysen 2017, 2019). However, new technologies, like deep learning and rapid image processing, are developing that make automated product recognition, selection of the appropriate gripper, and rapid automated picking with robots possible. For the coming decade or so, robot picking stations alone cannot do the job cost-efficiently and at sufficient speed, which means humans may have to collaborate with robots. This calls for further study of human–machine interaction.

8.4.5. Human–Machine Interaction. Although warehouse processes are becoming increasingly automated, humans will still be required to do a part of the work. They will have to cooperate and interact with machines. The interaction between man and machine has received little attention in operations management literature. Researchers may focus on addressing the following questions:

- What are the types of human jobs that should remain to maximize joint performance during tasks in cooperation with machines?
- How can we minimize the discomfort of order pickers (see also Larco et al. 2017)?

Or researchers may focus on addressing more behavioral questions:

- How do we incentivize people, or what type of personality should a person have to maximize the performance of joint work?

A recent study showed that the organization of the pick process, work incentives, and personalities of the pickers strongly interact and can have a major effect on picking performance (de Vries, De Koster, and Stam 2016).

8.4.6. Warehouse Sustainability. Increased social awareness, together with governmental regulations for carbon emissions and waste management, has transformed sustainability from an idealistic idea to an absolute necessity for companies (Chaabane, Ramudhin, and Paquet 2011). While increasing attention has focused on supply chain sustainability (e.g., Seuring and Müller 2008; Ballot and Fontane 2010; Barjis et al. 2010; Barber, Beach, and Zolkiewski 2012), the environmental impact of automated warehouses has not received much attention. Colicchia, Melacini, and Perotti (2011) offer several approaches for more sustainable warehouses, such as using green energy sources, optimizing travel distance and storage assignment policies, and adopting energy-efficient material handling equipment. Tappia et al. (2015) propose a mathematical model to evaluate the energy consumption and environmental impact of AS/R and AVS/R systems. Zaerpour, Yu, and De Koster (2015) do a similar analysis for a live-cube storage system. Hahn-Woernle and Günthner (2018) propose a power-load management policy for multi-aisle miniload AS/R systems to reduce energy consumption peaks, thereby lowering the energy cost of the automated warehouse, with only a slight decrease in throughput. However, more studies are needed to incorporate environmental aspects into the decision models revolving around new material handling technologies.

8.4.7. New Methods. In addition to the methods that are discussed in Section 2, new techniques might need to be developed, or other existing tools can be used to evaluate the performance of the automated systems. For instance, data-driven techniques such as data envelopment analysis can be used to benchmark automated systems. New modeling methods need to be developed, for example, for performance evaluation with nonstationary transaction arrivals (see Dhingra et al. 2018). Many of the recent robotic solutions have flexible capacities (e.g., vertical AVS/R, picking AGVs, RMF systems). In these systems, the number of robots can be adjusted and workstations can be opened or closed depending on the needed capacity. Thus, new methods are needed to estimate how the capacity of the systems can be dynamically adjusted and allocated to different activities. New data-driven techniques, such as deep learning, may find an application in answering these questions.

9. Conclusion

This paper presents an overview of the recent trends in automated warehousing, especially the use of robotic technologies to fulfill orders. The advantages of automation are mainly savings in space, savings on labor costs, 24/7 availability (it is not always easy to find unskilled personnel willing to do warehouse work), and savings on other operational costs, such as heating and lighting. Furthermore, robotic technologies provide scalability and throughput flexibility, which is essential in e-commerce environments where the demand variability is high. Automation of storage and order picking requires considerable scale and a long-term vision, as the investments can be earned back only in the medium and longer term. Therefore, it is crucial to develop tools to help decision makers find the correct solutions for their warehouses. As a result, studies have been carried out to model and optimize the performance of the various automated systems. We present modeling techniques as well as corresponding solution approaches in evaluating the performance of the automated systems. We also illustrate how the models are used in long-term and short-term decision-making processes (design, operational control, and planning). We describe well-established automated technologies (AS/R, shuttle-based, and AGV-based systems) as well as the literature related to the various design and control problems in these systems, such as optimally shaping the system, the impact of dwell point policies, block prevention protocols, and storage assignment. These systems differ in terms of infrastructural requirements, operational protocols, and equipment movement, and although the frameworks are common, models need to be customized to each system's unique characteristics. We also discuss emerging technologies and aspects that have not received enough (or any) attention in the literature. We summarize the unaddressed research questions in established systems and pose research questions for emerging technologies. Human picking in collaboration with AGVs is one of the most recent technologies that is becoming popular in practice because of its simplicity and flexibility, but it has not yet been adequately studied. Also, automated replenishment and sequencing, integrated systems, human–machine interaction, and warehouse sustainability are areas that require more attention from researchers.

Appendix. Overview of the Literature

Table A1. Overview of the Literature (34 Papers) on Shuttle-Based Storage and Retrieval System (Aisle-Based)

Research category	System	Article	Research issue	Methodology
System analysis	Single/Double Deep (tier-to-tier)	Malmborg (2002; 2003a, b)	Estimate vehicle utilization and cycle time	State equation model
		Kuo et al. (2007), Fukunari and Malmborg (2008)	Estimate vehicle utilization and cycle time	Nested queueing model
		Fukunari and Malmborg (2009)	Estimate vehicle utilization and cycle time interfacing material flow system	Closed-queueing network
		Zhang et al. (2009)	Estimate transaction waiting time	Variance-based nested queueing model
		Ekren (2011)	Evaluate performance of a real system under predefined design scenarios	Simulation
		Ekren et al. (2013, 2014) Cai et al. (2014)	Model the system	Semi-open queueing network
	Single/Double Deep (tier-captive)	Marchet et al. (2012), Epp et al. (2017)	Estimate transaction cycle time	Open-queueing network
		Heragu et al. (2011)	Estimate transaction cycle time, Compare with AS/RS	Open-queueing network
		Lerher et al. (2015), Lerher (2016)	Estimate mean travel time	Closed-form solution
		Ekren (2017)	Graph-based solution for performance evaluation of the system	Simulation
		Roy et al. (2017)	Estimate transaction cycle time and resource utilization	Multi-stage semi-open queueing network
	Multi-Deep	Manzini et al. (2016), D'Antonio et al. (2018)	Estimate cycle time	Travel time model
Design optimization	Single/Double Deep (single tier)	Roy et al. (2012)	Optimal rack configuration	Semi-open queueing network
		Roy et al. (2015a)	Optimal cross-aisle location	Semi-open queueing network
	Single/Double Deep (tier-to-teir)	Ekren and Heragu (2010b)	Optimal rack configuration	Simulation-based regression
	Single/Double Deep (tier-captive)	Marchet et al. (2013)	Optimal rack configuration of the system	Simulation
	Multi-Deep	Manzini et al. (2016)	Optimal L/U point location, Optimal layout configuration, Optimal number and depth of the lanes	Semi-open queueing network
		Tappia et al. (2016)	Optimal layout configuration, choice of shuttle and vertical transfer	Semi-open queueing network

Table A1. (Continued)

Research category	System	Article	Research issue	Methodology
Operations planning and control	Single/Double Deep (single tier)	Roy et al. (2012)	Effect of design choices on cycle time and vehicle utilization	Semi-open queueing network
		Roy et al. (2014)	Effect of vehicle blocking on performance	Semi-open queueing network
		Roy et al. (2015a)	Optimal dwell-point policy	Semi-open queueing network
		Roy et al. (2016)	Effect of vehicle blocking on performance	Simulation model
	Single/Double Deep (tier-to-tier)	Kuo et al. (2008)	Effect of class-based storage on cycle time	Closed-queueing network
		He and Luo (2009)	Deadlock-free control policy	Colored time Petri nets
		Ekren et al. (2010)	Effect of combination of dwell-point, I/O location, scheduling and interleaving rule on performance	Simulation, ANOVA
		Kumar et al. (2014)	Optimal partitioning of vertical zones in the system	Simulation
		Roy et al. (2015b)	Congestion effect on the performance of the system	Semi-open queueing network
	Single/Double Deep (tier-captive)	Zou, Xu, Gong, and De Koster (2016)	Simultaneously vs. sequentially requesting vehicles and lifts	Fork-join queueing network

Table A2. Overview of the Literature (11 Papers) on Puzzle-Based Storage and Retrieval Systems

Research category	System	Article	Research issue	Methodology
System analysis	Puzzle-Based	Gue and Kim (2007)	Optimal retrieval path with fixed escort positions, performance comparison with aisle-based	Dynamic programming, heuristics
		Kota et al. (2015)	Retrieval time estimation with randomly located escorts	Closed-form expression, heuristics
	Live-Cube	Zaerpour et al. (2015)	Retrieval time expression with random storage policy	Closed-form expression
		Zaerpour et al. (2017b)	Retrieval time expression with two class-based storage policy	Closed-form expression
Design optimization	Puzzle-Based	Taylor and Gue (2008)	Effect of escort locations	Discrete time simulation
	Live-Cube	Zaerpour et al. (2015)	Optimal shape of the system with random storage policy	Mixed-integer nonlinear model
		Zaerpour et al. (2017b)	Optimal shape of the system with two-class-based storage policy	Close-form expression
		Zaerpour et al. (2017a)	Optimal zone boundary in two class-based storage policy	Mixed-integer nonlinear model
	GridFlow	Furmans et al. (2011)	Optimal shape of the system, choice of I/O point	Discrete time simulation
	GridStore	Gue et al. (2014)	Optimal shape of the system, effect of WIP and escorts on the throughput rate	Discrete time simulation
Operations planning and control	Puzzle-Based	Yu et al. (2017)	Effect of simultaneous and block movement of items and escorts	Integer-programming
		Mirzaei et al. (2017)	Simultaneous multi-load retrieval	Monte Carlo simulation, heuristics
	GridStore	Gue et al. (2014)	Deadlock free decentralized control scheme, effect of WIP and escorts on the throughput rate	Discrete time simulation
	GridFlow	Alfieri et al. (2012)	Gridflow with limited number of vehicles, Optimally dispatch AGVs, Optimize the shelves' movement	Heuristics

Table A3. Overview of the Literature (Eight Papers) on Robotic Mobile Fulfillment Systems

Research category	Article	Research issue	Methodology
System analysis	Nigam et al. (2014)	Estimate order throughput time for single-line orders	Closed-queueing network
	Lamballais et al. (2017a)	Estimate average order cycle time for multi-line orders	Semi-open queueing network
Design optimization	Lamballais et al. (2017a)	Optimal length-to-width ratio of storage area, optimal location of workstations	Semi-open queueing network
	Yuan and Gong (2017)	Optimal number of robots and their required average speed	Open queueing network
	Zou et al. (2019)	Optimal shape of the system	Semi-open queueing network
Operations planning and control	Nigam et al. (2014)	Efficient pod storage policy	Closed queueing network
	Yuan et al. (2019)	Velocity-based storage assignment	Fluid model
	Lamballais et al. (2017a)	Efficient zoning policy	Semi-open queueing network
	Lamballais et al. (2017b)	Efficient replenishment policy	Semi-open queueing network
	Boysen et al. (2017)	Sequencing picking orders at the workstation	Mixed-integer nonlinear
	Zou et al. (2019)	Battery charging and swapping strategies	Semi-open queueing network
	?	Analyze tradeoffs between dedicated and pooled robot assignment, Analyze tradeoffs between random vs. shortest queue allocation of robots	Closed queueing network

References

Alfieri A, Cantamessa M, Monchiero A, Montagna F (2012) Heuristics for puzzle-based storage systems driven by a limited set of automated guided vehicles. *J. Intelligent Manufacturing* 23(5):1695–1705.

Automation Parking Systems (2018) Park, swipe, leave … it is that simple. Accessed December 11, 2018, http://automotionparking.com/.

AutoStore (2018) AutoStore: Introduction. Accessed September 3, 2018, https://www.youtube.com/watch?v=iHC9ec591lI.

Azadeh K, Roy D, De Koster R (2018) Design, modeling, and analysis of vertical robotic storage and retrieval systems. Working paper, Erasmus University, Rotterdam, Netherlands.

Ballot E, Fontane F (2010) Reducing transportation CO2 emissions through pooling of supply networks: Perspectives from a case study in French retail chains. *Production Planning Control* 21(6):640–650.

Banker S (2016) Robots in the warehouse: It's not just Amazon. *Forbes* (January 11), https://www.forbes.com/sites/stevebanker/2016/01/11/robots-in-the-warehouse-its-not-just-amazon/#3acec8a740b8.

Barber KD, Beach R, Zolkiewski J (2012) Environmental sustainability: A value cycle research agenda. *Production Planning Control* 23(2–3):105–119.

Barjis J, Verbraeck A, Tan K-S, Daud Ahmed M, Sundaram D (2010) Sustainable enterprise modelling and simulation in a warehousing context. *Bus. Process Management J.* 16(5):871–886.

Bartholdi JJ III, Platzman LK (1986) Retrieval strategies for a carousel conveyor. *IIE Trans.* 18(2):166–173.

Baskett F, Chandy KM, Muntz RR, Palacios FG (1975) Open, closed, and mixed networks of queues with different classes of customers. *J. ACM* 22(2):248–260.

Bolch G, Greiner S, de Meer H, Trivedi KS (2006) *Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications* (John Wiley & Sons, Hoboken, NJ).

Boysen N, Stephan K (2016) A survey on single crane scheduling in automated storage/retrieval systems. *Eur. J. Oper. Res.* 254(3):691–704.

Boysen N, Briskorn D, Emde S (2017) Parts-to-picker based order processing in a rack-moving mobile robots environment. *Eur. J. Oper. Res.* 262(2):550–562.

Bozer YA, White JA (1984) Travel-time models for automated storage/retrieval systems. *IIE Trans.* 16(4):329–338.

Buck Consultants International (2017) Robotisering bespaart de komende 15 jaren 35.000 arbeidsplaatsen in nederlandse distributiecentra. Accessed May 30, 2017, http://www.bciglobal.com/nieuws_detail.asp?cat=5006&dc=26663.

Cai X, Heragu SS, Liu Y (2013) Modeling automated warehouses using semi-open queueing networks. Smith J, Tan B, eds. *Handbook of Stochastic Models and Analysis of Manufacturing System Operations*, International Series in Operations Research & Management Science, vol. 192 (Springer, New York), 29–71.

Cai X, Heragu SS, Liu Y (2014) Modeling and evaluating the AVS/Rs with tier-to-tier vehicles using a semi-open queueing network. *IIE Trans.* 46(9):905–927.

Caputo AC, Pelagagge PM (2006) Management criteria of automated order picking systems in high-rotation high-volume distribution centers. *Indust. Management Data Systems* 106(9):1359–1383.

Chaabane A, Ramudhin A, Paquet M (2011) Designing supply chains with sustainability considerations. *Production Planning Control* 22(8):727–741.

Chen YL, Tang K, Shen RJ, Hu YH (2005) Market basket analysis in a multiple store environment. *Decision Support Systems* 40(2):339–354.

Colicchia C, MelaciniM, Perotti S (2011) Benchmarking supply chain sustainability: Insights from a field study. *Benchmarking Internat. J.* 18(5):705–732.

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. *Internat. J. Advanced Manufacturing Tech.* 94(5–8):1811–1824.

De Koster R (2015) *Past and Future: Perspectives on Material Handling* (Erasmus Research Institute of Management, Rotterdam, Netherlands).

De Koster R, Yu M (2008) Minimizing makespan and throughput times at Aalsmeer flower auction. *J. Oper. Res. Soc.* 59(9):1182–1190.

De Koster R, Le-Duc T, Roodbergen KJ (2007) Design and control of warehouse order picking: A literature review. *Eur. J. Oper. Res.* 182(2):481–501.

De Koster R, Le-Duc T, Yugang Y (2008) Optimal storage rack design for a 3-dimensional compact AS/RS. *Internat. J. Production Res.* 46(6):1495–1514.

De Vries J, De Koster R, Stam D (2016) Aligning order picking methods, incentive systems, and regulatory focus to increase performance. *Production Oper. Management* 25(8):1363–1376.

Dhingra V, Kumawat GL, Roy D, De Koster R (2018) Solving semi-open queuing networks with time-varying arrivals: An application in container terminal landside operations. *Eur. J. Oper. Res.* 267(3):855–876.

Dukic G, Opetuk T, Lerher T (2015) A throughput model for a dual-tray vertical lift module with a human order-picker. *Internat. J. Production Econom.* 170(December):874–881.

Ekren BY (2011) Performance evaluation of AVS/Rs under various design scenarios: A case study. *Internat. J. Advanced Manufacturing Tech.* 55(9):1253–1261.

Ekren BY (2017) Graph-based solution for performance evaluation of shuttle-based storage and retrieval system. *Internat. J. Production Res.* 55(21):6516–6526.

Ekren BY, Heragu SS (2010a) Approximate analysis of load-dependent generally distributed queuing networks with low service time variability. *Eur. J. Oper. Res.* 205(2):381–389.

Ekren BY, Heragu SS (2010b) Simulation-based regression analysis for the rack configuration of an autonomous vehicle storage and retrieval system. *Internat. J. Production Res.* 48(21):6257–6274.

Ekren BY, Heragu SS, Krishnamurthy A, Malmborg CJ (2010) Simulation based experimental design to identify factors affecting performance of AVS/RS. *Comput. Indust. Engrg.* 58(1):175–185.

Ekren BY, Heragu SS, Krishnamurthy A, Malmborg CJ (2013) An approximate solution for semi-open queueing network model of an autonomous vehicle storage and retrieval system. *IEEE Trans. Automation Sci. Engrg.* 10(1):205–215.

Ekren BY, Heragu SS, Krishnamurthy A, Malmborg CJ (2014) Matrix-geometric solution for semi-open queuing network model of autonomous vehicle storage and retrieval system. *Comput. Indust. Engrg.* 68(February):78–86.

Enright JJ, Wurman PR (2011) Optimization and coordinated autonomy in mobile fulfillment systems. *Proc. 9th AAAI Conf. Automated Action Planning Autonomous Mobile Robots* (Association for the Advancement of Artificial Intelligence, Palo Alto, CA), 33–38.

Epp M, Wiedemann S, Furmans K (2017) A discrete-time queueing network approach to performance evaluation of autonomous vehicle storage and retrieval systems. *Internat. J. Production Res.* 55(4):960–978.

Exotec Solutions (2017) Exotec Solutions realizes order picking systems based on fleets of robots evolving in three dimensions. Accessed May 30, 2017, https://www.exotecsolutions.com/?lang=en.

Fraunhofer Institute for Material Flow and Logistics (2014) RackRacer—Versatile and flexible rack feeder for container racks. Technical report, Fraunhofer Institute for Material Flow and Logistics, Dortmund, Germany.

Fukunari M, Malmborg CJ (2008) An efficient cycle time model for autonomous vehicle storage and retrieval systems. *Internat. J. Production Res.* 46(12):3167–3184.

Fukunari M, Malmborg 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.

Furmans K, Nobbe C, Schwab M (2011) Future of material handling–modular, flexible and efficient. Amato NM, Chen IM, De Luca A, Jenkins C, Kragic D, Papanikolopoulos N, Park F, Parker L, Sugano S, van der Stappen F, eds. *IEEE/RSJ Internat. Conf. Intelligent Robots Systems* (Institute of Electrical and Electronics Engineers, Piscataway, NJ).

Furmans K, Schönung F, Gue KR (2010) Plug and work material handling systems. *Progress Material Handling Res.* 132–142.

Füßler D, Boysen N (2017) Efficient order processing in an inverse order picking system. *Comput. Oper. Res.* 88(December):150–160.

Füßler D, Boysen N (2019) High-performance order processing in picking workstations. *EURO J. Transportation Logist.* 8(1):65–90.

Gagliardi JP, Renaud J, Ruiz A (2012) Models for automated storage and retrieval systems: A literature review. *Internat. J. Production Res.* 50(24):7110–7125.

Glover FW, Kochenberger GA (2006) *Handbook of Metaheuristics*, vol. 57 (Springer Science & Business Media, Berlin).

Gue KR (2006) Very high density storage systems. *IIE Trans.* 38(1):79–90.

Gue KR, Kim BS (2007) Puzzle-based storage systems. *Naval Res. Logist.* 54(5):556–567.

Gue KR, Uludag O (2012) A high-density, puzzle-based order picking system. Carrano AL, Gue K, De Koster R, Ogle M, Montreuil B, Smith J, eds. *12th Internat. Material Handling Res. Colloquium* (College Industry Council on Material Handling Education, Gardanne, France), 209–220.

Gue KR, Furmans K, Seibold Z, Uludağ O (2014) GridStore: A puzzle-based storage system with decentralized control. *IEEE Trans. Automation Sci. Engrg.* 11(2):429–438.

Hahn-Woernle P, Günthner WA (2018) Power-load management reduces energy-dependent costs of multi-aisle mini-load automated storage and retrieval systems. *Internat. J. Production Res.* 56(3):1269–1285.

Hao J, Yu Y, Zhang LL (2015) Optimal design of a 3D compact storage system with the I/O port at the lower mid-point of the storage rack. *Internat. J. Production Res.* 53(17):5153–5173.

Hausman WH, Schwarz LB, Graves SC (1976) Optimal storage assignment in automatic warehousing systems. *Management Sci.* 22(6):629–638.

He S, Luo J (2009) Deadlock control of autonomous vehicle storage and retrieval systems via coloured timed petri nets and digraph tools. *Internat. J. Production Res.* 47(12):3253–3263.

Heragu SS, Cai X, Krishnamurthy A, Malmborg CJ (2008) Striving for warehouse excellence. *Indust. Engineer* 40(12):43–47.

Heragu SS, Cai X, Krishnamurthy A, Malmborg CJ (2011) Analytical models for analysis of automated warehouse material handling systems. *Internat. J. Production Res.* 49(22):6833–6861.

Hyundai Elevator Co. (2016) HIP (Hyundai integrated parking system). Accessed May 30, 2016, http://hyundaielevator.co.kr/eng/parking/car/automobile.jsp.

Imahori S, Hase Y (2016) Graph-based heuristics for operational planning and scheduling problem in automatic picking system. *J. Adv. Mech. Design Systems Manufacturing* 10(3):JAMDSM0039.

Industrie Forum (2004) Vom gelochten winkelprofil zum vollautomatischen hochregallager. Accessed March 28, 2017, http://www.industrie-forum.net/de/if/online/2811670/jsessionid=ahrBozpKA35dkBIwL7.

Jia J, Heragu SS (2009) Solving semi-open queuing networks. *Oper. Res.* 57(2):391–401.

Jünemann R (1989) *Materialfluß und Logistik* (Springer, Berlin).

Kim WR, Kim WR, Kim N, Kim N, Chang YS, Chang YS (2016) A study on the control of an automatic picking system (APS) in a warehouse. *Assembly Automation* 36(4):388–397.

Kollmorgen (2010) Coop Pick-n-Go. Accessed September 3, 2018, https://www.youtube.com/watch?v=3d6fo9TlBV0&t=3s.

Kota VR, Taylor D, Gue KR (2015) Retrieval time performance in puzzle-based storage systems. *J. Manufacturing Tech. Management* 26(4):582–602.

Kumar A, Roy D, Tiwari M (2014) Optimal partitioning of vertical zones in vehicle-based warehouse systems. *Internat. J. Production Res.* 52(5):1285–1305.

Kuo PH, Krishnamurthy A, Malmborg 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.

Kuo PH, Krishnamurthy A, Malmborg CJ (2008) Performance modelling of autonomous vehicle storage and retrieval systems using class-based storage policies. *Internat. J. Comput. Appl. Tech.* 31(3–4):238–248.

Lamballais T, Roy D, De Koster R (2017a) Estimating performance in a robotic mobile fulfillment system. *Eur. J. Oper. Res.* 256(3):976–990.

Lamballais T, Roy D, De Koster R (2017b) Inventory allocation in robotic mobile fulfillment systems. Report, Erasmus Research Institute of Management, Erasmus University, Rotterdam, Netherlands.

Larco JA, De Koster R, Roodbergen KJ, Dul J (2017) Managing warehouse efficiency and worker discomfort through enhanced storage assignment decisions. *Internat. J. Production Res.* 55(21): 6407–6422.

Lerher T (2016) Travel time model for double-deep shuttle-based storage and retrieval systems. *Internat. J. Production Res.* 54(9): 2519–2540.

Lerher T, Ekren BY, Dukic G, Rosi B (2015) Travel time model for shuttle-based storage and retrieval systems. *Internat. J. Advanced Manufacturing Tech.* 78(9–12):1705–1725.

Litvak N, Vlasiou M (2010) A survey on performance analysis of warehouse carousel systems. *Statistica Neerlandica* 64(4):401–447.

Locus Robotics (2018) Simply pick faster—Locus Robotics overview. Accessed September 3, 2018, https://www.youtube.com/watch?v=2_I9j-1QJKU.

Loyd S (1914) *Cyclopedia of Puzzles* (Lamb Publishing Company, New York).

Loyd S, Gardner M (1959) *Mathematical Puzzles*, vol. 1 (Courier Corporation, North Chelmsford, MA).

Magazino (2017) Pick-by-robot: Warehouse automation with Magazino at FIEGE Logistik. Accessed September 3, 2018, https://www.youtube.com/watch?v=kj8NaHAoLJw.

Malmborg CJ (2002) Conceptualizing tools for autonomous vehicle storage and retrieval systems. *Internat. J. Production Res.* 40(8): 1807–1822.

Malmborg CJ (2003a) Design optimization models for storage and retrieval systems using rail guided vehicles. *Appl. Math. Model.* 27(12):929–941.

Malmborg CJ (2003b) Interleaving dynamics in autonomous vehicle storage and retrieval systems. *Internat. J. Production Res.* 41(5): 1057–1069.

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). *Internat. J. Production Res.* 54(14):4286–4304.

Marchet G, Melacini M, Perotti S, Tappia E (2012) Analytical model to estimate performances of autonomous vehicle storage and retrieval systems for product totes. *Internat. J. Production Res.* 50(24): 7134–7148.

Marchet G, Melacini M, Perotti S, Tappia E (2013) Development of a framework for the design of autonomous vehicle storage and retrieval systems. *Internat. J. Production Res.* 51(14):4365–4387.

Meisel F, Bierwirth C (2013) A framework for integrated berth allocation and crane operations planning in seaport container terminals. *Transportation Sci.* 47(2):131–147.

Meller RD, Klote JF (2004) A throughput model for carousel/VLM pods. *IIE Trans.* 36(8):725–741.

Meller RD, Pazour JA (2008) A heuristic for SKU assignment and allocation in an a-frame system. Fowler J, Mason S, eds. *Proc. 2008 Indust. Engrg. Res. Conf.* (Institute of Industrial and Systems Engineers, Peachtree Corners, GA), 770–775.

MHI (2017a) A-frame. Accessed December 06, 2017, http://www.mhi.org/solutions-community/solutions-guide/aframe.

MHI (2017b) Storage equipment. Accessed January 25, 2018, http://www.mhi.org/learning/cicmhe/resources/taxonomy/StorEq/index.htm.

MHI (2017c) Vertical lift modules. Accessed January 20, 2017, http://www.mhi.org/ofs/solutions-guide/vertical-lift.

Mirzaei M, De Koster RB, Zaerpour N (2017) Modelling load retrievals in puzzle-based storage systems. *Internat. J. Production Res.* 55(21):6423–6435.

Mirzaei M, Zaerpour N, De Koster R (2018) An integrated cluster-based storage assignment policy. Working paper, Rotterdam School of Management, Erasmus University, Rotterdam, Netherlands.

Mountz MC, D'andrea R, LaPlante JA, David PLI, Mansfield PK, Amsbury BW (2008) Inventory system with mobile drive unit and inventory holder. U.S. Patent 7,402,018.

MWPVL International (2012) Is Kiva Systems a good fit for your distribution center? An unbiased distribution consultant evaluation. Accessed May 30, 2016, http://www.mwpvl.com/html/kiva_systems.html.

Nigam S, Roy D, De Koster R, Adan I (2014) Analysis of class-based storage strategies for the mobile shelf-based order pick system. Smith J, Ellis K, De Koster R, Lavender S, Montreuil B, Ogle M, eds. *13th IMHRC Proc.* (CICMHE, Charlotte, NC), 19.

OPEX Corporation (2013) A unique approach to automated order fulfillment for small- to mid-sized ecommerce businesses. Technical report, OPEX Corporation, Moorestown, NJ.

Pazour JA, Meller RD (2011) An analytical model for a-frame system design. *IIE Trans.* 43(10):739–752.

Pazour JA, Meller RD (2013) The impact of batch retrievals on throughput performance of a carousel system serviced by a storage and retrieval machine. *Internat. J. Production Econom.* 142(2): 332–342.

Pazour JA, Meller RD (2014) A framework and analysis to inform the selection of piece-level order-fulfillment technologies. Smith J, Ellis K, De Koster R, Lavender S, Montreuil B, Ogle M, eds. *13th IMHRC Proc.* (CICMHE, Charlotte, NC), 22.

Puchinger J, Raidl GR (2005) Combining metaheuristics and exact algorithms in combinatorial optimization: A survey and classification. Mira J, Álvarez JR, eds. *Artificial Intelligence and Knowledge Engineering Applications: A Bioinspired Approach,* Lecture Notes in Computer Science, vol. 3562 (Springer, Berlin), 41–53.

Ramtin F, Pazour JA (2014) Analytical models for an automated storage and retrieval system with multiple in-the-aisle pick positions. *IIE Trans.* 46(9):968–986.

Ramtin F, Pazour JA (2015) Product allocation problem for an AS/RS with multiple in-the-aisle pick positions. *IIE Trans.* 47(12): 1379–1396.

Reiser M, Lavenberg SS (1980) Mean-value analysis of closed multichain queuing networks. *J. ACM* 27(2):313–322.

Roodbergen KJ, Vis IF (2009) A survey of literature on automated storage and retrieval systems. *Eur. J. Oper. Res.* 194(2):343–362.

Rosenblatt MJ, Eynan A (1989) Deriving the optimal boundaries for class-based automatic storage/retrieval systems. *Management Sci.* 35(12):1519–1524.

Rosi B, Grasic L, Dukic G, Opetuk T, Lerher T (2016) Simulation-based performance analysis of automated single-tray vertical lift module. *Internat. J. Simulation Model.* 15(1):97–108.

Roy D (2011) Design and analysis of unit-load warehouse operations using autonomous vehicles. Unpublished PhD thesis, University of Wisconsin, Madison.

Roy D (2016) Semi-open queuing networks: A review of stochastic models, solution methods and new research areas. *Internat. J. Production Res.* 54(6):1735–1752.

Roy D, Krishnamurthy A, Heragu SS, Malmborg CJ (2012) Performance analysis and design trade-offs in warehouses with autonomous vehicle technology. *IIE Trans.* 44(12):1045–1060.

Roy D, Krishnamurthy A, Heragu SS, Malmborg CJ (2014) Blocking effects in warehouse systems with autonomous vehicles. *IEEE Trans. Automation Sci. Engrg.* 11(2):439–451.

Roy D, Krishnamurthy A, Heragu SS, Malmborg CJ (2015a) Queuing models to analyze dwell-point and cross-aisle location in autonomous vehicle-based warehouse systems. *Eur. J. Oper. Res.* 242(1):72–87.

Roy D, Krishnamurthy A, Heragu SS, Malmborg CJ (2015b) Stochastic models for unit-load operations in warehouse systems with autonomous vehicles. *Ann. Oper. Res.* 231(1):129–155.

Roy D, Krishnamurthy A, Heragu SS, Malmborg CJ (2016) A simulation framework for studying blocking effects in warehouse systems with autonomous vehicles. *Eur. J. Indust. Engrg.* 10(1): 51–80.

Roy D, Krishnamurthy A, Heragu SS, Malmborg C (2017) A multi-tier linking approach to analyze performance of autonomous vehicle-based storage and retrieval systems. *Comput. Oper. Res.* 83(July): 173–188.

Roy D, Nigam S, De Koster R, Adan I, Resinf J (2019) Robot-storage zone assignment strategies in mobile fulfillment systems. *Transportation Res. Part E: Logist. Transportation Rev.* 122: 119–142.

Sari Z, Saygin C, Ghouali N (2005) Travel-time models for flow-rack automated storage and retrieval systems. *Internat. J. Advanced Manufacturing Tech.* 25(9–10):979–987.

Schwerdfeger S, Boysen N (2017) Order picking along a crane-supplied pick face: The SKU switching problem. *Eur. J. Oper. Res.* 260(2):534–545.

Seuring S, Müller M (2008) From a literature review to a conceptual framework for sustainable supply chain management. *J. Cleaner Production* 16(15):1699–1710.

Stadtler H (1996) An operational planning concept for deep lane storage systems. *Production Oper. Management* 5(3):266–282.

Tappia E, Marchet G, Melacini M, Perotti S (2015) Incorporating the environmental dimension in the assessment of automated warehouses. *Production Planning Control* 26(10):824–838.

Tappia E, Roy D, De Koster R, Melacini M (2016) Modeling, analysis, and design insights for shuttle-based compact storage systems. *Transportation Sci.* 51(1):269–295.

Taylor GD, Gue KR (2008) The effects of empty storage locations in puzzle-based storage systems. *IIE Annual Conf. Proc.* (Institute of Industrial Engineers, Peachtree Corners, GA), 519–523.

UludagO (2014) GridPick: A high density puzzle based order picking system with decentralized control. Unpublished doctoral dissertation, Auburn University, Auburn, AL.

Whitt W (1983) The queueing network analyzer. *Bell Labs Tech. J.* 62(9):2779–2815.

Wöhr (2016) Parksafe for the Luxury Albany Apartments in Liverpool providing 84 car park places. Accessed June 20, 2019, http://www.woehr.de/en/project/liverpool-parksafe-583.html.

Wurman PR, D'Andrea R, Mountz M (2008) Coordinating hundreds of cooperative, autonomous vehicles in warehouses. *AI Magazine* 29(1):9–19.

Yang P, Miao L, Xue Z, Qin L (2015) Optimal storage rack design for a multi-deep compact AS/RS considering the acceleration/deceleration of the storage and retrieval machine. *Internat. J. Production Res.* 53(3):929–943.

Yu Y, Yu H, De Koster R, Zaerpour N (2017) Optimal algorithm for minimizing retrieval time in puzzle-based storage system with multiple simultaneously movable empty cells. Working paper, School of Management, University of Science and Technology of China, Hefei.

Yu Y, De Koster R (2009a) Designing an optimal turnover-based storage rack for a 3D compact automated storage and retrieval system. *Internat. J. Production Res.* 47(6):1551–1571.

Yu Y, De Koster R (2009b) Optimal zone boundaries for two-class-based compact three-dimensional automated storage and retrieval systems. *IIE Trans.* 41(3):194–208.

Yu M, De Koster R (2010) Enhancing performance in order picking processes by dynamic storage systems. *Internat. J. Production Res.* 48(16):4785–4806.

Yu Y, De Koster R (2012) Sequencing heuristics for storing and retrieving unit loads in 3D compact automated warehousing systems. *IIE Trans.* 44(2):69–87.

Yu Y, De Koster R, Guo X (2015) Class-based storage with a finite number of items: Using more classes is not always better. *Production Oper. Management* 24(8):1235–1247.

Yuan R, Graves SC, Cezik T (2019) Velocity-based storage assignment in semi-automated storage systems. *Production Oper. Management* 28:354–373.

Yuan Z, Gong YY (2017) Bot-in-time delivery for robotic mobile fulfillment systems. *IEEE Trans. Engrg. Management* 64(1):83–93.

Zaerpour N, De Koster R, Yu Y (2013) Storage policies and optimal shape of a storage system. *Internat. J. Production Res.* 51(23–24): 6891–6899.

Zaerpour N, Yu Y, De Koster R (2015) Small is beautiful: A framework for evaluating and optimizing live-cube compact storage systems. *Transportation Sci.* 51(1):34–51.

Zaerpour N, Yu Y, De Koster R (2017a) Optimal two-class-based storage in a live-cube compact storage system. *IISE Trans.* 49(7): 653–668.

Zaerpour N, Yu Y, De Koster R (2017b) Response time analysis of a live-cube compact storage system with two storage classes. *IISE Trans.* 49(5):461–480.

Zhang L, Krishnamurthy A, Malmborg CJ, Heragu SS (2009) Variance-based approximations of transaction waiting times in autonomous vehicle storage and retrieval systems. *Eur. J. Indust. Engrg.* 3(2): 146–169.

Zou B, De Koster R, Xu X (2019) Evaluating dedicated and shared storage policies in robot-based compact storage and retrieval systems. *Transportation Sci.* Forthcoming.

Zou B, Xu X, Gong Y, De Koster R (2016) Modeling parallel movement of lifts and vehicles in tier-captive vehicle-based warehousing systems. *Eur. J. Oper. Res.* 254(1):51–67.

Zou B, Xu X, Gong YY, De Koster R (2018) Evaluating battery charging and swapping strategies in a robotic mobile fulfillment system. *Eur. J. Oper. Res.* 267(2):733–753.