René (M.) B.M. de Koster is a professor of logistics and operations management at the Department of Technology and Operations Management, Rotterdam School of Management, Erasmus University (RSM), and Fellow of the Erasmus Research Institute of Management (ERIM) with the research programme Business Processes, Logistics and Information Systems (LIS). After completing his PhD at Eindhoven University of Technology (1988) he worked as a consultant. He joined Erasmus University in 1995. Professor De Koster’s research interests are warehousing, material handling, container terminal operations, behavioural operations and sustainable logistics. He is the author and editor of eight books and over 130 papers published in books and academic journals. He is in the editorial boards of several academic journals, including Operations Research, Transportation Science (SI), and Journal of Operations Management. He is member of several international research advisory boards (ELA: European Logistics Association, BVL: Germany (www.bvl.de), AIL: France, and university supervisory boards: University of Pisa and Aalto-Helsinky. He is chairman of Stichting Logistica, and founder of the Material Handling Forum (www.rsm.nl/mhf). He is involved in teaching at RSM at all levels: bachelor, master, post-experience, and executive development. He is also guest lecturer at several other universities in the Netherlands, Belgium, China, and South Africa.

This book marks the occasion of René de Koster’s 20 years full professorship at RSM and is particularly devoted to developments in material handling, a topic he has been teaching during these 20 years. Several authors take a retrospective look at the developments in the field of material handling both in practice and in research over the last twenty years, and make predictions for the future. The book focuses not only on technical developments in material handling, but also on insights obtained in the impact of behaviour and leadership on operational performance in combination with systems and procedures. Also included are some key representative papers used in courses on warehousing and material handling that highlight what we have learned and taught during this period.

ERIM

The Erasmus Research Institute of Management (ERIM) is the Research School (Onderzoekschool) in the field of management of the Erasmus University Rotterdam. The founding participants of ERIM are the Rotterdam School of Management (RSM), and the Erasmus School of Economics (ESE). ERIM was founded in 1999 and is officially accredited by the Royal Netherlands Academy of Arts and Sciences (KNAW). The research undertaken by ERIM is focused on the management of the firm in its environment, its intra- and interfirm relations, and its business processes in their interdependent connections.

The objective of ERIM is to carry out first rate research in management, and to offer an advanced doctoral programme in Research in Management. Within ERIM, over three hundred senior researchers and PhD candidates are active in the different research programmes. From a variety of academic backgrounds and expertises, the ERIM community is united in striving for excellence and working at the forefront of creating new business knowledge.

Inaugural Addresses Research in Management contain written texts of inaugural addresses by members of ERIM. The addresses are available in two ways, as printed hard-copy booklet and as digital fulltext file through the ERIM Electronic Series Portal.

Past and Future

Perspectives on Material Handling
Past and Future

Perspectives on Material Handling

On the occasion of René de Koster’s 20 years full professorship
at Rotterdam School of Management, Erasmus University
10 November 2015

René de Koster

Department of Technology and Operations Management
Rotterdam School of Management, Erasmus University
P.O. Box 1738
3000 DR Rotterdam
The Netherlands
E-mail: rkoster@rsm.nl
# Content

<table>
<thead>
<tr>
<th>Introduction by René de Koster</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreword by Jan Hommes</td>
<td>11</td>
</tr>
<tr>
<td><strong>Part 1: Developments in material handling. Different perspectives</strong></td>
<td>13</td>
</tr>
<tr>
<td><strong>1. What will the world of material handling look like in ten years?</strong></td>
<td>15</td>
</tr>
<tr>
<td>Jan van der Velden, Market Director Parts &amp; Components, Vanderlande</td>
<td></td>
</tr>
<tr>
<td><strong>2. Developments in material handling automation</strong></td>
<td>21</td>
</tr>
<tr>
<td>Tom Bonkenburg, St. Onge Company</td>
<td></td>
</tr>
<tr>
<td><strong>3. The role of warehousing / x-docking and material handling technologies for V&amp;D</strong></td>
<td>25</td>
</tr>
<tr>
<td>Theo Heemskerk, COO V&amp;D</td>
<td></td>
</tr>
<tr>
<td><strong>4. The emergence of material handling research in The Netherlands</strong></td>
<td>31</td>
</tr>
<tr>
<td>Kees-Jan Roodbergen, University of Groningen</td>
<td></td>
</tr>
<tr>
<td><strong>5. Developments in material handling research</strong></td>
<td>36</td>
</tr>
<tr>
<td>Ivo Adan, Manufacturing Networks Group, Department of Mechanical Engineering, Eindhoven University of Technology</td>
<td></td>
</tr>
<tr>
<td><strong>6. Compact Storage Systems: a new research area</strong></td>
<td>41</td>
</tr>
<tr>
<td>Yugang Yu, Professor of Logistics and Supply Chain Management, Executive Dean, University of Science and Technology of China, Hefei, PRC</td>
<td></td>
</tr>
</tbody>
</table>
Part 2: Selected academic contributions and publications
by René de Koster

Warehouse assessment in a single tour

Design and control of order picking

Warehouse automation
De Koster, R. (2013), Magazijnautomatisering, in: R. Jansen and A. van Goor (Eds.), 40 jaar logistiek, 40 jaar VLM, pp. 41-44

Warehouse math

Self-storage warehousing
De Koster, R. (2013), Boosting revenues in the self-storage warehouse industry, RSM Insight 16, 8-10

Accidents will happen
De Koster, R., Stam, D., Balk, B. (2011), Accidents will happen: do hazard-reducing systems help?, RSM Insight 5, 8-11

Container terminal operations: overview
Optimal design of container terminal layout

Publication list René de Koster

PhD candidates supervised by René de Koster
Introduction by René de Koster

Two decades of Material Handling

The world of material handling has changed enormously over the past two decades. These changes have occurred primarily in the areas of information technology and social networks, automation and robotization, ergonomics, safety, sustainability, and efficiency – but also in the place of material handling in the world of academia.

Twenty years ago very little attention was paid to material handling both within academic research and higher education. This was true in the Netherlands but certainly also elsewhere, with the exception of some reputed research institutes like Georgia Tech, Virginia Tech, and Fraunhofer IML.

For me, with a background in material handling consulting, this was odd, since storage and transhipment are key to the Dutch economy, providing employment to many people and responsible for a considerable share of the gross domestic product. I wanted to change this, and the best way I could see how was to simply start conducting my own research and teaching in this area.

In teaching, I started a course in Warehouse Management, one that I am still teaching. Many of the supply chain master’s students I have taught in this course found their first job during their warehousing project. Some of my former PhD students continued to conduct research in the field, became professors, and are now teaching similar courses at their respective universities.

In research, I had a number of subjects in mind: the control and design of automated guided vehicle systems, order picking methods, storage strategies, facility layouts, container terminal operations, and compact storage systems. Since then, thanks to several gifted PhD students, we have been able to explore all these areas and move our research forward considerably. New research areas have emerged: sustainability in material handling and, in particular, the human factor in material handling.

Operational performance in organizations is largely determined by system design, IT control systems and procedures. But it appears to also depend, sometimes substantially, on how managers lead the operation, take decisions, and the behaviour of employees. This area is studied in behavioral operations management. Leadership styles, incentives, the personality of the leader and the
personalities of employees interact with the system design and procedures and have a strong impact on the productivity, quality, rate of accidents, and job satisfaction of the employees.

In this booklet, several authors take a retrospective look at the developments in the field of material handling both in practice and in research over the last twenty years, and make some predictions for the future. Also included are some key representative papers used in courses on warehousing and material handling that highlight what we have learned and taught during this period.

I thank all the authors for their valuable contributions to this book, as well as ERIM, Stichting Logistica, and the Material Handling Forum (MHF) for making the publication of this book possible.

*René de Koster*
Foreword by Jan Hommes

For someone to work twenty years in a line of business is not so remarkable. For someone to radically change that line of business over the course of twenty years, certainly is. And that is precisely what René de Koster has done.

René de Koster has given a face to Material Handling in the Netherlands, and facilitated its emergence as a unique and valuable area within the broader field of logistics.

The contributions in this celebratory book are largely a tribute to this achievement. They showcase the developments that have taken place in the field of Material Handling in the Netherlands from an academic perspective, and from the point of view of practitioners.

That such developments have occurred across both sectors is an appropriate reflection of the way in which René has conducted his work as a professor: an approach that has united academia and practice. A great example of this can be seen in his work as chairman of the Prize Safest Warehouse, and even more so in the Material Handling Forum he founded, a platform that regularly draws together academics, practitioners and suppliers.

René never misses an opportunity to promote and advance the field of Material Handling via this platform: to develop new research proposals time and time again and to involve business and industry in his endeavours. The most compelling example may be seen in a comprehensive scientific research project that studied different order picking methods, with goods made available by suppliers and with the research jointly executed with real order pickers.

René knows how to inspire and motivate his students to continuously push and develop themselves as scientists. It is for this reason that, in both 2014 and 2015, the winners of the Dutch Master’s Thesis prize were students of René.

Erasmus University Rotterdam can take pride in having René among its professors for as long as twenty years.
Past and Future
Perspectives on Material Handling

Part 1
Developments in material handling.
Different perspectives
1. What will the world of material handling look like in ten years?

Jan van der Velden,
*Market Director Parts & Components, Vanderlande
President of FEM*

The material handling industry supports many logistical processes. To understand how this industry will evolve we need to look at the trends that drive logistics and thus the key market drivers of the developments within material handling. What can we say about how the industry will be ten years from now? In the next pages we hope to give you a glimpse into this new world from the perspective of our organisation.

**Global trends will drive logistical complexity and flexibility**

A great many changes are currently taking place in the world. The major drivers for these are shown in the figure below.

**Impact on logistics**

- **Robotics and automation**
  - Capabilities increase and costs decrease
  - Autonomous control and distributed intelligence
  - Self driving / flying
  - Miniaturisation

- **Sensors / internet of things**
  - Real time and remote control
  - Wireless
  - (predictive) Warnings of problems
  - Optimising routing and delivery decisions

- **Big data, predictive analytics**
  - Analysis of massive quantities of data
  - Data mining
  - Data visualisation
  - Better decisions in logistics and operations

- **The changing workforce**
  - Attracting and keeping an adequate workforce
  - Different workforce and skills than today
  - Innovation of collaboration and workplace
### Impact on logistics (continued)

<table>
<thead>
<tr>
<th>Circular economy</th>
<th>The growth of e-commerce</th>
<th>Relentless competition</th>
<th>Mass personalisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Local production</td>
<td>• Fast order fulfilment</td>
<td>• Third-party logistics</td>
<td>• Any time, anywhere, any way</td>
</tr>
<tr>
<td>• Usership versus ownership</td>
<td>• Individual items</td>
<td>• Price</td>
<td>• Diverse set of order and distribution channels</td>
</tr>
<tr>
<td>• Waste management</td>
<td>• Products are in stock, without excessive safety stock</td>
<td>• Service</td>
<td>• Rising customer expectations</td>
</tr>
<tr>
<td>• Returns</td>
<td>• Returns</td>
<td>– Shorter lead times</td>
<td>• Solution as a service</td>
</tr>
<tr>
<td>• Standardisation for smaller units</td>
<td>• Push to pull</td>
<td>– End-to-end visibility</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• From planned to instant delivery</td>
<td>– Respond to changes in delivery location and time</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Urbanisation</th>
<th>Humanising technology</th>
<th>Value chain</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Challenges for last-mile distribution facilities</td>
<td>• Virtual / augmented reality</td>
<td>• Integration of the value chain</td>
</tr>
<tr>
<td>• Home delivery</td>
<td>• Ergonomic design</td>
<td>– Who manages the grid?</td>
</tr>
<tr>
<td>• Smaller quantities</td>
<td>• Location-based services</td>
<td>• Client might be potential future competitor</td>
</tr>
<tr>
<td>• Higher variety of products</td>
<td>• Wearable computing</td>
<td>• Responsiveness and flexibility</td>
</tr>
<tr>
<td>• Infrastructure and legislation restrictions</td>
<td>• Push to pull</td>
<td></td>
</tr>
</tbody>
</table>
**IT and new technologies will become key in future solutions for all market segments**

Material handling solutions are being developed in two market segments: transport and storage. Solutions in both of these areas are currently developing independently of the other.

The integrated solution starts where these two worlds meet, and it is here that IT starts to play an important role. Relevant technologies will, in the future, enable us to create new products and solutions that can be used in all market segments, while IT will continue to play a key role in the optimisation of these market segments.

**What will the future look like in ten years?**

It is evident that the need for automation remains and will continue to grow. The other certainty is that, as business continues to undergo rapid change, the demand for complete and on-time delivery is becoming greater. To cope with these fast-changing requirements, automation will require two main characteristics: adaptability and scalability. Current material handling solutions have very limited scalability and almost no adaptability.
ADAPTO: experiencing the future today

Vanderlande has developed ADAPTO as a highly flexible and scalable concept and a viable option to the challenge of adaptability and scalability.

ADAPTO is a 3D concept, built on a modular configuration of racking, shuttles and lifts. It consists of the following components:

- A racking structure with an integrated shuttle track system.
- Multidirectional microshuttles that transport product carriers between rack locations and system exits / entries.
- Lifts that allow the shuttles to move between rack levels and system exits / entries.
- Traffic control software that uses unique strategies to maximise system throughput. This meets the highest standards of quality and usability. The user interface is very user friendly and offers a real-time overview of the entire system and its components.

The modular design of racking, shuttles and lifts makes it possible to adapt the system to changing business needs. Storage capacity and throughput can be scaled independently.

ADAPTO requires a low initial investment and allows for a step-by-step extension as a business grows.
Autonomous or centralised control?

Current solutions are being built with centralised control. An example of a screening area is shown in the figure below. In the bag screening area, two loop sorters facilitate the flow to and from the screening machines. The only reason for the double sorter is redundancy.

This system has been squeezed into the available footprint within the building. At Vanderlande, we have investigated the use of AGVs to eliminate everything in this space – except the screening machines. The green dots in the alternative illustration represent the AGVs.

In this particular example, 130 AGVs could deliver the load throughout the day, offering clear benefits such as facilitating failure recovery, reducing redundancy and creating more free space. Prototype AGVs have been built to verify these design choices, the current level of technology, and the cost breakdown.
What can we learn from this exercise?

- Autonomous vehicles...
  - Provide superior robustness and scalability.
  - The current state-of-the-art trend is for low vehicle density.
  - High vehicle density applications are not yet fit for use.
- Until then...
  - Traffic control for warehouse automation (ADAPTO) is applicable.
- Typically 30-40% of the cost is directly related to vehicles...
  - Starting at the lowest capacity possible.
  - There is more reusable infrastructure to come.

Fully scalable solutions will build from a finite set of reusable components

The figure below indicates that the merging of these three worlds will result in workable integrated equipment.
2. Developments in material handling automation

Tom Bonkenburg,
St. Onge Company

René de Koster must have heard many strange stories from his students during his twenty years at Erasmus University. But I think this story might be new to him.

My opportunity to attend university came about because of a forklift.

Like most young students in the United States I needed to pass my high school exams before I could move on to higher education. I had an extra challenge, however: I was one quarter of my father’s warehouse workforce. As a savvy businessman he knew that losing me to college would hurt his bottom line and this was not ideal.

For much of my early life I spent many evenings and weekends working in our family business, which consisted of my father, my mother, my brother, and I. My father ran the business side and did the real bulk of the work. My role was loading and unloading trucks, picking orders, packing boxes, and generally lifting heavy things and putting them where my mother told me to. My future university career depended on my finding a replacement for this manual material-handling grunt work. Salvation came in the form of a used forklift that arrived the day before I was leaving for university.

My father had faced a challenge that many modern companies are facing – how to continue to get orders out the warehouse door while the workforce keeps disappearing. A recent study by BCG shows that over the next fifteen years Germany alone could see labour shortages grow up to 10 million workers. My father took the course that most companies do today. He added material handling equipment to improve his remaining workforce: when using our new forklift my mother could now do her job, my job, and much more with greater speed and far fewer complaints (while in the warehouse I was also our complaint department, i.e. the person who generated the most complaints).

Research shows that 80% of current warehouses are manual. These warehouses have dealt with the demands for increased productivity and throughput by supporting existing workers with good layout design, mobile material handling equipment, and constantly improving IT.
15% of warehouses are mechanized. In addition to the technology used in manual warehouses, these distribution centres also use some type of material handling automation such as conveyors, sorters, goods to picker solutions, or other mechanized equipment to further improve the existing workforce.

This has been the trend in material handling automation over the last three decades; most centres with very large order fulfilment requirements follow this approach. These technologies supplement the manual workforce, multiplying their efforts. Today a well-designed e-commerce warehouse using a goods-to-picker concept can achieve a pick rate of 300+ lines per hour per picker.

While this is impressive, modern mechanized equipment still requires an accomplished worker to make it run, just as our little forklift required my mother. Traditional mechanization is not flexible or easily adaptable. A well-designed system pushes the need for flexibility to the people while removing tasks that are more mundane and standardized. If you remove the people from the system it will completely shut down in a matter of minutes.

The last 5% of warehouses are automated. The reality today is that these ‘automated’ warehouses are typically just highly mechanized environments. True, fully automated distribution operations are hard to find and history shows that business has not been kind to them. Every year you read in trade magazines of one or two more highly automated fulfilment centres being built. What you do not read however is that, every year, one or two highly automated fulfilment centres are torn down or heavily modified because they did not meet expectations or easily flex with changing business needs.

The material handling industry has tried to take the techniques of the ‘mechanized’ world and extend them to create ‘automated’ systems, but with limited success. This approach makes sense from the point of view of the equipment vendor. Developing new technology is risky and expensive. Vendors are therefore on much safer ground taking small risks with their current equipment and doing so only when a customer has agreed to pay for it on a project-by-project basis.

But overall this philosophy has, unfortunately, failed, because typical material handling mechanization does not have the flexibility required to cover the needs of the current dynamic business climate. Vendors and customers alike see these failures and react accordingly: the risks taken get smaller and they happen less often. True material handling innovation stops.
The only way to break this cycle is to develop an entirely new type of technology with more inherent flexibility. Vendors in the material handling world are stuck. They do not have the tools, knowledge, or capital to develop this next level of automation.

But there is hope on the horizon. The next trend in material handling automation does not seem to be coming from the traditional equipment vendors but, rather, from a combination of three groups: universities, start-up companies, and e-commerce retailers. Many of these groups are focusing on using robotic technology as a way to bring more flexibility into the material handling world.

In the context of material handling, one of the main advantages of humans is our ability to identify, pick up, sort, count, position, and otherwise manipulate an endless variety of items. If we can see it on a shelf or in a box we generally have the ability to ‘pick it and pack it’.

Traditional manufacturing robots have no such ability – they are blind and dumb. The good news is that recent advances are just starting to change this. While it is still early days, universities around the world are conducting research in computer science that is advancing the perceptive capability of robots to identify an object, locate its position, and plan a path of motion that will allow the item to be picked up.

This is an extremely challenging technical problem to solve that, even a few years ago, was close to impossible. Yet with the advent of low cost sensors from the consumer electronics industry, much faster computers, and smarter software algorithms, we are starting to see this area of research, one that I will call ‘robotics with advanced perception’, bear fruit.

Several startup companies are using the knowledge generated by these university studies to develop the next generation of robots. The most talked about robot in the media today is named Baxter and comes from a company called Rethink Robotics. Rethink was founded by an MIT professor who took his lab research and brought it into the industrial world. Baxter has two arms with three built-in cameras that it uses to identify and pick up objects.

The system used by Baxter is a great leap forward in robotic design on many fronts, but it has yet to be a commercial success. The reason for this is simple: it is just not good enough. The technology is promising but still falls far short of a
skilled human worker. Yet while Baxter’s current performance is limited, it has helped illustrate to the market the great potential of this technology. Rethink Robotics has since raised $113 Million USD in investment capital to develop and advance this technology further. How many material handling vendors could afford to invest this much money into new research?

Amazon.com also sees the potential of robots. In 2013 they spent $750 million USD to buy Kiva, another robotics startup company with a focus on warehouse logistics. In 2014, Google bought six robotic startup companies including one focused on automatic trailer unloading using advanced perception. Overall, Google has spent a rumoured $100+ million USD to get into the high-tech robotics game.

A new startup company that is focused on the distribution market is Fetch Robotics. Fetch has developed a robotic arm that drives around on a mobile base to pick items off of a standard warehouse shelf and put them into an order tote. Amazon is extremely interested in this concept and, earlier this year, sponsored a contest with a cash prize for anyone who could develop a robotic system to pick items off of a shelf. Twenty-eight groups entered the contest. The prize went to a group from the Technical University in Berlin for a robot that picked ten items.

This is only the beginning. While no one knows what the future will bring, there are some clear trends emerging. Labour shortages continue to grow and the large e-commerce retailers are trying to create their own path towards advanced automation. Universities, startup companies, and capital investors with deep pockets are all showing interest in developing robotics with advanced perception. Traditional material handling companies are taking notice and trying to find ways to integrate these new technologies into their mechanized solutions.

In recent years, true material handling innovation has stalled. Our best hope is to find a new technology that has the potential to break paradigms and kick-start the next wave of logistics development. Will robotics with advanced perception be this technology? The potential is enormous and, I think we all agree, will be an interesting trend to watch during René de Koster’s next twenty years at Erasmus University.

My engineering career was unlocked by a forklift that allowed me to sidestep a life of manual warehouse labour and go to university. My hope is that many more careers will be unlocked and lives improved by future developments in material handling automation and advanced robotics.
3. The role of warehousing / x-docking and material handling technologies for V&D

Theo Heemskerk,
COO V&D

History of V&D Logistics

Since about eleven years, René and I are jury members for the Logistics Manager of the Year Award. Many candidates we have evaluated over time have contributed to their company success through innovation in material handling. Also V&D has realized remarkable improvements by restructuring its warehouses.

In 2006, V&D had four warehouses located in Oedenrode, Aduard, Amsterdam and Utrecht. These four warehouses were a heritage of an earlier consolidation round of a maximum of fourteen regional warehouses.

Utrecht was the fashion warehouse, and the others each processed one or more “hard”-goods departments, such as luggage, stationary or bedlinen. The fashion warehouse was equipped with an extensive Equinox overhead hanging rail system for stock keeping storage and sorting of including a sorter for hanging garments. For the so-called “flat” goods, a space saver for receiving and labelling goods together with two Equinox bomb-bay sorters with six infeed stations were operational.

At the warehouse in Aduard, a Beumer tilt-tray sorter was used. All warehouses were also using manual sortation systems supported by screens or scanners. SAP provided the WMS and the WCSs were from Equinox and Beumer with interfaces to and from with the WMS.

From these warehouses, the (regional) stores received deliveries from of at least one truck per day from Utrecht, supplying fashion stock, and one truck per day supplying staple / hard goods.

In 2008, V&D opened its web shop. The warehousing and fulfilment requirements for the web shop are currently done by a third party in Waalwijk. The site in Waalwijk receives deliveries from the new Nieuwegein warehouse in the same way as a regular store.
The new warehouse in Nieuwegein was opened in 2012, and the warehouses in Utrecht and Amsterdam were closed. In 2014, Aduard closed and all processes were integrated into the Nieuwegein warehouse. Transportation was combined, resulting in one delivery per day per store.
Our main goal was to improve the level of service we provide to our stores by improving goods reliability and goods availability.

One-stop store delivery, which combines the delivery of fashion and staple items, has been made possible by combining hanging and flat goods transportation within the same truck. Facilitating cross-dock processes at the warehouse and consolidating transportation has also minimized direct store deliveries by suppliers (i.e. A-brands); for the bigger stores this means a reduction of up to fifty deliveries a day.

Another very important development has been the shift from push flows to more controlled pull flows, which allows for holding back part of the stock in the warehouse. This helps to prevent overloading stores with unwanted quantities at unwanted moments. Goods are also shelf-ready delivered, consolidated on product group level and store floor level.

This helps to control the logistics handling and costs within stores.

The Nieuwegein warehouse was initially designed to handle V&D fashion processes and some hard good departments, like luggage. E-commerce fulfilment was also in scope. However in Q2 2011 the decision was made not to integrate Adward and e-commerce, based on the required short implementation time for the new site, including new systems and material handling equipment. The decisions to push forward these targets made it necessary to enforce some capacity cuts in material handling solutions.

The total footprint of the Nieuwegein warehouse is 25,000 m², with mezzanines of 56,000 m² of useable floors.
Total footprint of the warehouse is 25,000m², with mezzanines 56,000m² useable floors

Material handling equipment

The warehouse handles three main flows: hanging, flat goods and cross dock. The equipment supplier for all flows is SSI-Schaefer.

For hanging garments, Schaefer (Meiko) installed a rail system with chain driven trolleys. The trolleys are provided with RFID-chips for tracking purposes.
The system extends from the receiving dock into a dynamic stock keeping area with space for 100,000 pieces of hanging garments and static storage for another 285,000 pieces. Both areas are built on dedicated mezzanines, each 6,000 m².

The dynamic storage area feeds the sorter (ladder-type sorter). The WCS for the whole system including the sorter is supplied by HSP (part of Schaefer). After sortation, goods are manually moved into the loading area on rolling racks and the goods are manually loaded onto bars in the trucks.

For flat goods, the material handling equipment is used mainly for handling smaller, so-called ‘totable’ goods. The buffers for received goods (totable and non-totable) and the stock keeping of non-totables is done on pallets in conventional semi-high bay pallet racking.

Growth in the stock keeping of holdback stock and e-commerce stock, and the need for a fast and efficient picking process, were the key-drivers in the choice of a mini load system with an integrated pick-by-light system and an extra goods-to-man (Pick To Tote) picking station. Initially, only fashion goods were held in the mini load, most of which have all the physical features necessary to be marked “totable”, and a high turnover, making them suitable for stock keeping in a mini load.

The pick-by-light zone (PBL) can hold 1600 SKUs. Previously, the flow racks in the PBL zone were permanently occupied by basic goods. With the integration of Aduard, this was changed into a more dynamic system that involves the daily refreshing of flow racks with replenishments (pull) articles, making it possible to increase productivity. The PTT-station was designed to handle customer orders. From the start it became a pick place for store orders with slim order lines. The mini load can hold 40,000 totes.

Five parallel sorters were designed for the sortation-to-stores of purchase orders. Each sorter has an infeed line where boxes from pallets are destacked onto an infeed conveyor and sent to a pick station. Each sorter contains a loop for the store totes. This concept makes it possible to plan the workload on a product group level per sorter and to work in parallel. The five sorters make easy scaling up or down possible. The empty cartons are retained and conveyed out on a special system into shredders and containers. This makes it possible to ship goods out according to the service levels mentioned above.
All store-ready, totable push and pull goods are buffered in four carrousels, with a capacity of 3400 totes. A day’s production is held and released to the loading area in sync with the transportation schedule. The carrousel is also used for buffering any pre-produced totes with special promotion articles until the required selling dates. After their release from the carrousel (on store level), the totes are sorted again per floor level in the store, then stacked onto dollies (max 12 per dolly). Each tote is labelled before entering the carrousel. These labels contain information for the stores regarding the content of the tote (i.e. the department, promotion code, floor level).

The trolley system and the conveyor system both have exits to a future e-commerce fulfilment area. The WCS for the whole system is delivered by Schaefer.

For the sortation of non-totable goods, the SAP-EWM functionalities are used, with the addition in 2014 of the manual Screen-to-Sort sortation system from Equinox, which was re-used from Aduard. The goods are shipped to the stores in roll containers, according to the same rules as the totables and the hanging garments.

The handling of cross dock flows is a manual process, with administrative support by the SAP-EWM functionality. Orders (unopened cartons) are consolidated into V&D rolling cages or totes on dollies.

Administrative functionalities throughout the warehouse are processed by RF-scanning.

So far, the new warehouse has proven to be very efficient, allowing faster response to stores and higher availability of stock, with a reduction of transportation costs.
4. The emergence of material handling research in The Netherlands

Kees-Jan Roodbergen,
University of Groningen

In the late 1970’s and early 1980’s, academic research into the optimization of warehouse processes started to appear regularly in scientific literature. The Material Handling Research Center at Georgia Tech, established in 1982, was one of the world’s frontrunners. One of the classic papers of this time is the routing paper by Ratliff and Rosenthal (1983) which, according to the WebOfScience database, is currently the second most cited article in order-picking research, with five citations fewer than De Koster et al. (2007).

Research output on warehousing in the Netherlands began to emerge about a decade later. The oldest English-language scientific journal article from a Netherlands-based author on warehouse processes that I was able to find concerned the performance approximation of pick-to-belt order-picking systems (De Koster, 1994). The fourth Material Handling Research Colloquium, held in 1996 in Den Bosch, also clearly marked this period.

Many of the following publications looked at one of the most traditional warehouse systems: manual order picking from racks where workers walk or drive along the racks to retrieve products. Order picking is a warehouse process, the importance of which cannot easily be overestimated. It is at the heart of the order fulfillment process.

De Koster et al. (2007) defined order picking simply as “the process of retrieving products from storage (or buffer areas) in response to a specific customer request.” However the true complexity of organizing the order-picking process becomes apparent when trying to list all possible control strategies for all types of systems. Batching, zoning, routing, storage location assignment, sorting; many control strategies are involved in tuning an order-picking system. And each control strategy influences the effectiveness of all other control strategies.

This notion is further strengthened by the findings of Faber et al. (2002): “We found that the number of orderlines to be processed per day and the number of stock-keeping units are the two main observable aspects of warehouse complexity; that the more complex the warehouse is, the more tailor-made the planning and control structure should be.” It therefore may not come as a
The surprise that warehouse control strategy optimization nowadays attracts the attention of numerous researchers world-wide.

In the 1990’s, several of the earliest studies in the Netherlands were performed by students working on their Master theses (e.g., De Koster and Van der Poort, 1998; De Koster et al., 1999a). Their work, as described in De Koster et al. (1999b), together with the simultaneously published work of Vaughan and Petersen (1999), while both not highly cited, marked the start of a stream of publications.

These studies introduced a wider variety of layouts in the scientific literature on manual order picking. Instead of the commonly studied single block layout, this research examined multiple-block layouts. Later work demonstrated the added value of multiple-block layouts compared to single-block layouts, derived performance estimators, and designed routing methods (Roodbergen, 2001). In current research, multiple-block layouts continue to appear on a regular basis in order picking studies. Only for unit-load handling was a significantly better layout alternative found several years later by Gue and Meller (2009).

The period of the 1990’s and early 2000’s was marked mostly by publications that studied one single aspect of warehouse control, such as routing of order pickers (Roodbergen and De Koster, 2001) or vehicle dispatching (De Koster et al., 2004). With the advance in Operations Research techniques and computing power, more comprehensive models for material handling control began to appear in the literature, combining properties of multiple control strategies into a single modeling construct.

For example, De Koster and Le-Duc (2005) presented a model that aims to determine the best storage locations for products in a manual order picking system, based on their demand frequency. This problem of “storage location assignment” in the literature was thus far mostly addressed by predefining several possibilities for storage location assignments, after which the various possibilities were compared, for example, by simulation.

The paper of De Koster and Le-Duc (2005) explicitly investigated (near)optimal storage assignments based on properties of the routing problem. After all, the best storage locations are those where the routes involve the minimum distance. This trend towards an ever-increasing scope is continuing to expand even today. For example, a recent contribution (Roodbergen et al., 2015) presents a unified approach for the simultaneous determination of an area layout, the routing method, as well as the storage location assignment.
In 2007, an overview article was published (De Koster et al., 2007) reviewing all state-of-the-art order-picking research. Part of our argumentation for the importance of order-picking research in this article was underlined by the sentence: “In manufacturing, there is a move to smaller lot sizes, point-of-use delivery, order and product customization, and cycle time reductions. In distribution logistics, in order to serve customers, companies tend to accept late orders while providing rapid and timely delivery within tight time windows (thus the time available for order picking becomes shorter).”

While this statement is still very true, it also shows how difficult it can be to predict the near future; we did not even mention e-commerce. This is not to suggest that warehouse processes have become less important due to e-commerce. On the contrary, warehousing has become a cornerstone of fulfillment strategies for e-commerce companies and, along with that, research into material handling is on an ever more challenging journey.

The article (De Koster et al., 2007) also notes that scientific literature is dominated by studies into automated systems, with manual systems being discussed in only about 30% of publications. Though the absolute number in publications on manual systems has since increased, I have no reason to believe that the percentage has changed much. Nevertheless, in practice the need for control strategies for manual systems is still large, since a majority of e-commerce companies is employing manual systems, mostly due to their scalability and low investment costs. The increased consumer demand for fast delivery and quality requires more real-time decision making as, for example, discussed in Gong and De Koster (2008).

In e-commerce we see a trend, however, towards mechanization and one that is detectable in practice, for example, with the shuttle system that is currently being implemented at Wehkamp in the Netherlands and several Amazon warehouses that are running KIVA systems.

Research conducted over the past few years has shown many interesting studies into optimizing automated systems. Pick-and-pass systems have been analyzed (Yu and De Koster, 2008), and new systems such as 3D compact storage systems have been receiving more attention (Yu and De Koster, 2009).

Looking back over the past 20 years of material handling research both in the Netherlands and the world, it is evident that it is a field within which practice and academic research are easily blended. Numerous algorithms developed at
universities have found their way into WMS software, as well as some of the methods mentioned above, while ideas and new systems from practice have quickly led to scientific articles devoted to investigating their potential.

The Material Handling Forum, initiated by De Koster, has played an important role in the interactions between practice and academia and will hopefully continue to do so in the future. In the 1990’s, nobody could have predicted where we are now, both in practice as well as research. Knowledge within the field has increased tremendously.

Yet the challenges that need to be addressed have also grown. When we look towards the future we see just how much more work there is to be done. There is a famous saying, “Today is the first day of the rest of your life”, and the same holds true for the field of material handling research.

References


5. Developments in material handling research

Ivo Adan,
Manufacturing Networks Group, Department of Mechanical Engineering, Eindhoven University of Technology

It is a truly a pleasure to write a few words on my collaboration over the years with René on material handling research.

Let me start with my first encounter with René’s research, some 30 years ago. René had completed his PhD thesis on capacity analysis and the design of production systems [4], under the supervision of Jacob Wijngaards. In his PhD thesis he proposed an innovative modeling framework for finite capacity discrete production systems, namely stochastic fluid flow models, which attempt to model the discrete flow of products as a continuous fluid flowing through the workstations, connected by finite buffers (or fluid reservoirs).

The model René formulated is natural in high-volume production environments and in the (chemical) process industry, but it also seems to work surprisingly well in discrete production such as the assembly of truck cabins. To assess the throughput performance of the production system, René developed an elegant and efficient approximation based on buffer aggregation: the fluid buffers are aggregated one-by-one until a simple single-buffer system remains that can be solved exactly.

René’s pioneering work spurred many other researchers, including myself, to study and analyze stochastic fluid models with finite buffers and to apply these models in production. Since then, a huge body of literature has appeared on this topic, leading to efficient decomposition-based approximations [1, 2], but also to new mathematical tools, such as matrix-analytic methods for fluid models with finite buffers [5] which have been successfully applied in many production environments, for example, to support decision making in setting conveyor capacities and machine speeds in bottling (fluid!) lines, and steering daily maintenance operations in semiconductor assembly plants [7], see Fig. 1.
This example shows the extent of René’s intuition to find and analyze problems that are both scientifically and practically relevant, with the latter often inspired by his close collaborations with companies. Below I will mention two more examples that illustrate this.

Since the advent of the Internet, the field of e-commerce has been growing rapidly. In particular, the number of online retailers has drastically increased over the last decade. Think, for example, of Amazon.com, bol.com, AH.nl and Wehkamp.nl. Fast response to customer orders is critical in this market, and one of the possibilities to reduce response times is to employ dynamic order picking: the order picker travels through the warehouse picking all outstanding order lines according to a constantly updated pick list, including order lines arriving online during and downstream this pick route.

To estimate the response-time performance of dynamic order picking, René realized that a dynamic order picking system could be modeled as a polling system: a multi-queue single-server model, where queues correspond to storage locations, customers to order lines at these locations, and the server to the order picker traveling along the queues, see Fig 2.
Polling systems are among the best and most intensively studied models in the mathematical theory of queues, with many applications in computer, communication and manufacturing networks. But its application to warehousing systems is completely new. By picking the appropriate results from the vast polling literature, René’s observation resulted in an award-winning journal paper [3], written together with Yeming Gong.

Fig. 2. Polling system (first) and multi-segment zone picking system (second)

More recently René, together with Jelmer van der Gaast and others [6], studied zone-picking systems, among the most popular internal transport and order picking systems in practice due to their scalability, flexibility and high-throughput ability.
Zone-picking systems are picker-to-parts systems, where the warehousing or storage area is divided into zones, where in each zone one or more order pickers are responsible for picking the order lines in that part of the warehouse. Zones are often grouped in segments and connected by circular conveyors, employing a dynamic block-and-recirculate protocol to prevent congestion as much as possible.

These systems can be modeled as finite-capacity queuing networks, which are, however, highly complex and intractable. René and his collaborators developed a queuing framework to accurately and efficiently estimate the throughput performance of zone picking systems, based on classical product-form results for so-called jump-over networks and Norton’s aggregation theorem. This framework provides a powerful tool to develop design principles for multi-segment zone picking systems.

We are now in the midst of the fourth industrial revolution – referred to in the Netherlands as Smart Industry, and driven by far-reaching digitization, flexibilization and automation in industry. In warehousing operations, this is, among others, visible in the developments of novel autonomous vehicle and conveyor technologies, and robotic systems. Examples are the KIVA system and the novel shuttle technology ADAPTO, which has recently been developed by Vanderlande Industries, see Fig. 3.

Fig. 3. KIVA system (top) and the shuttle system ADAPTO (bottom)
The combination of the use of novel warehousing technologies and digitization imposes enormous challenges for smart design and the control of automated warehousing systems to achieve the highest performance in terms of throughput, utilization and availability. In addressing these challenges I am confident that René will once again be able to play a leading role.

References


6. Compact Storage Systems: a new research area

Yugang Yu
Professor of Logistics and Supply Chain Management
Executive Dean, University of Science and Technology of China, Hefei, PRC

Storage systems play an importation role in warehouses and distribution centres in manufacturing and logistics, but also in automated car parking (particularly found in congested East-Asian cities), container stacking at terminals, libraries in educational institutes, dispensing machines for pharmaceuticals in hospitals, and so on.

These systems can be found everywhere and are indispensable; they fill a crucial role in bringing products and related services to the customer. They are particularly important both in China, the largest manufacturer and exporter in the world, and the Netherlands, the main gateway to Western Europe.

My research with René de Koster examines storage facilities with a focus on management problems in compact storage systems. Compact storage systems solve the problem of space shortage around large cities like Beijing, Shanghai, and Hefei in China, as well as in the western part of the Netherlands. They can be used for parking vehicles, or storing and retrieving loads. They save space by removing the travel aisles present in conventional systems and utilizing the available space at much higher utilization rates (of up to 98%) compared with aisle-based systems (which utilize up to about 85% of the storage locations). Aisle-space is saved in compact systems by storing loads both multi-deep and multi-high (3D). In the case of parking systems, space can be reduced by up to 1/6 of the space of a traditional car park.

In our research, we have studied systems with autonomously operating internal transport (for example, shuttles carrying parked cars), making every load rapidly accessible. In the example system sketched in Fig.1 (a “puzzle-based” system, or PBS), at all levels each load is stored on a shuttle which transports it in x and y directions. A requested load can be moved to the output location by simultaneously circulating this load and other loads on the path to the output location, which is possible as long as there is at least one open location (Fig.1a). If multiple open locations are available a virtual aisle can be created (Fig.1.b).

Lifts or conveyors move loads and shuttles between different levels. Compact autonomous storage systems come in different configurations, like PBSs, general
shuttle-lift based systems, conveyor-lift based systems, conveyor-crane based systems, and others. The full automation allows for 24/7 operations, often making them more efficient than aisle-based systems, with shorter response times, while the operational cost can go down.

There are also implications for the environment resulting from the need for much smaller buildings, although this might be partly offset by higher energy use. Particular PBSs, the newest type of compact storage system, are challenging to study and optimize, as the movement of loads depends on the number and location of empty locations, and the locations of loads at a level is effected by the movements of other loads.

![Figure 1. Work mechanism for moving a load to the gate on each floor](image)

Overall, my research with René de Koster aims to develop optimized designs and operations that will improve the performance of different types of compact autonomous storage systems. We have defined two prime objectives, and have received several research grants in our bid to meet them, including grants from NOW-Veni, KNAW in the Netherlands and two from NSF in China.

Our first objective, to propose optimized storage polices for locating loads to minimize system response time in compact autonomous storage systems, has led to the following publications:


The studies listed above have taught us how to segment the storage area of the system and where to store various loads. Different real-life applications have been studied, including public car parking systems, storage and cross-dock systems for fresh produce or frozen goods, and container terminal systems.

Our second objective, which proposes scheduling algorithms for optimally moving loads with minimum operational costs or makespan, has resulted in the following publications:


As a result of our research in relation to this second objective, we now understand when to move which loads and how to move them individually or together. We have also challenged class-based storage in our research by considering a finite number of items that can only share storage space to a limited extent, within their storage class.

The extant literature assumes space sharing is always and fully possible (implicitly assuming the number of stored items is infinite). We developed a travel-time model and used it for optimizing the number and the boundaries of storage classes (Guo et al, 2015; Yu et al., 2015). Contrary to the results revealed in this literature, our findings illustrate that a small number of classes is often optimal, and that the full turnover-based storage is often suboptimal.

The following articles represent this stream of research:


My research with René de Koster forms part of a long-term research partnership in the area of warehouse management and material handling, which is a prime research theme at the Rotterdam School of Management, and one of the three key themes defined in the School of Management at USTC. Our joint research has proven to be very productive and successful, and I look forward to an equally rewarding and fruitful collaboration in the future.
Past and Future
Perspectives on Material Handling

Part 2
Selected academic contributions and publications
Warehouse assessment in a single tour

Abstract

This paper presents an assessment method for warehouses based on a single facility tour and some Q&A. The method helps managers and students that visit a facility to get more information from tour visits through a simple and rapid assessment form. Since its inception, it has been applied to a number of cases, successfully identifying weak and strong points of the operations.
1. Introduction

Over the last decades, many companies have offshored manufacturing activities to Asia Pacific and Eastern Europe. Since the consuming markets have not moved, this has put an increasing burden on the distribution operations of such companies. Companies have centralized warehouse operations in few, but often large facilities responsible for distributing products over a large region. Managing efficiency and effectiveness (service) is a great challenge for managers of such facilities. As a result, they feel a great need to benchmark warehouse operations, not only their own, but also their competitors’. However, assessing the performance of a distribution facility is a tricky business. Even after having visited a large number of them, it is still difficult to tell after a visit, whether this was a best-in-class operation, just above-average, or even relatively poor performing. Nevertheless, even short tour visits can reveal a lot of information to the trained eye.

This paper proposes a method to help managers getting more information from tour visits, through a simple and rapid assessment form. The form should be filled out immediately after the visit. The evaluation has been inspired by the ideas of Gene Goodson in Harvard Business Review on rapid plant assessment (Goodson, 2002). Since its development, the method has been successfully applied in several visits, with different groups of managers (with and without warehouse experience), and students.

The major functions of a warehouse are to store products in order to make an assortment for customers, to assemble customer orders, sometimes to add value to the orders by customization activities, organize transport to the customers, and ship orders timely, in the way desired by the customer. Warehouse performance therefore, has multiple dimensions. Often, performance is measured in terms of ratios of output and input factors. Output factors include production (shipped orders, lines and units), quality (for example, order completeness, error-free and on-time delivery), flexibility (possibility to cope with changes in customer demand), agility (process adaptation to changed environment), and innovativeness (use of new supply-chain concepts yielding competitive advantage). Inputs are the resources used to achieve the outputs. These include the number of full-time equivalents (work hours used per year), investment in systems, buildings and IT infrastructure, process organization (i.e. the management), or the assortment carried.
Some researchers have tried to develop benchmark tools for warehouses (McGinnis et al., 2002; Hackman et al., 2001; De Koster and Balk, 2008). One such tool is DEA (data envelopment analysis), which expresses the warehouse efficiency as a ratio of weighed output and weighed input factors, normalized on a 0 to 1 scale. Although DEA is a powerful tool, it is usually difficult to obtain the necessary data at the required accuracy level. Also, for every factor that is included in the efficiency analysis, more cases are needed in order to have statistically meaningful results. Furthermore, the warehouses should be comparable, which in practice may be difficult to realize. It is also difficult to compare warehouses in different countries, even when they operate in the same industry branch (think of cultural differences, or just of the number of working hours per full-time employee). Finally, it is difficult to include factors in DEA that are not measured on interval scales, or more subjective assessments (like teamwork, motivation, safety, cleanliness).

As an alternative, or addition, to more quantitative analyses, this tool is based on a single facility tour and can be carried out in a few hours, including some Q&A. It is not necessary to have deep insight in the operations. The main objectives of the tool are to discern the warehouse’s strengths and weaknesses after some elementary training on how to use the tool. The tool can also be used to evaluate operations of logistics service providers, operating public or dedicated warehouses. This is not to say that the tool can be a substitute for due diligence and care when analyzing company performance. In particular, financial performance is not part of the tool. However, all too often managers ignore visual signals that can be easily acquired in favor of seemingly objective data, like quantities processed, inventory turns or company profits (which are rarely directly attributable to a warehouse).
2. The Assessment Method

The tool is based on a factor-rating method (see, for example, Heizer and Render 2004) and consists of 11 areas that have to be assessed, each on a six-category scale (see Exhibit 3). Seven areas (1 to 5, 8 and 10) are more or less generally applicable to industrial facilities and have been adapted to warehouse environments from Goodson (2002). Areas 6 and 7 (storage and order picking systems) form the heart of any warehouse (Tompkins et al., 2003) and must therefore be included in an assessment. Areas 9 (level and use of IT) and 11 (managing efficiency and flexibility) are equally important in an assessment. To aid filling out the assessment form, a number of yes/no questions have been formulated (Exhibit 2), which serve the purpose of conveying the opinions on the area and aiding area scoring. A score is measured on a 6 category ordinal scale and ranges from poor (1 point) to excellent (9 points) with an additional category ‘best in class’ (11 points). Best-in-class means that there is no better. We first discuss the areas in more detail and then discuss results as well as further validation of the method.

Area 1: Customer Satisfaction

Customer satisfaction is difficult to rate in a facility visit. However, all people in the facility – and particularly workers – should clearly know who the customers are, both internal and external. Management can take care of this by explicitly showing external quality performance indicators to the workers. Signboards with picking or shipping errors, customer complaints and returns over time, quality guidelines for workers, and so on, indicate sensitivity to wishes of customers and quality assurance. Try asking an order picker, packer or dispatcher: “What is the impact for customers when you make an error?” When this person answers that it will result in a complaint (or return, or a customer credit note), it should lead to a higher score than when the employee has no idea at all, or when she or he deems there are no clear consequences.

Even when products are picked by article (batched over multiple customer orders), the person should have an idea of the customers’ wishes, whether there are deadlines for the (batch) order to be shipped (many large warehouses work with fixed departure schedules in order to reach their customers timely) and what the consequences are for not finishing the work in a timely manner.

Questions 1, 4, 14 and 21 are related to this area.
**Area 2: Cleanliness, Environment, Ergonomics, Safety, Hygiene (HACCP)**

This is an area that is relatively easy to assess. If a facility is clean, it usually indicates that management organizes the processes well. In clean facilities, items do not get lost, inventory accuracy is higher (as well as order fulfillment accuracy), and there is an overall sensitivity to orderliness. Order picking warehouses (where case and item picking occur) typically generate much waste (pallets have to be unwrapped, boxes have to be opened) and workers have to be able to get rid of it in an easy way. In well-run warehouses, one can find waste baskets in front of the racks, where waste can be separated immediately at the source by type (which is compulsory in the EU). In a well-run facility, the air is clean, noise levels are low, and it is well-lit. In short, it is comfortable to work in. All location codes are easily readable (also from a distance) and barcoded, such that there is no confusion as to which code refers to which location (particularly for the lower beams in a pallet rack, or in a shelf area where location sizes are often tiny). Worker positions should have been designed with attention for ergonomics. As much of the work is repetitive, or strenuous, ill-designed work positions lead to high absence rates and labor turnover.

In many warehouses, pickers do not have fixed work positions, because they drive trucks or walk with pick carts. Even in such cases ergonomics pays off. The use of tiny screens and buttons on mobile terminals leads to low productivity and even to errors (reduction of which often was the main reason for the use of such terminals). In the European warehouse of a large Japanese manufacturer of consumer electronics, pickers use mobile terminals to receive pick instructions and confirm the picks. When they were asked about the contents of their work, it appeared that for a single order (of a few units) about 20 entries had to be made to confirm this. If 20 cases of the same product had to be picked from a pallet, labeled, scanned and put on a conveyor belt, it might take minutes to confirm this via the RF-terminal / scanner in the information system. Workers obviously find workarounds (do first and confirm when convenient), which may compromise the system integrity.

Safety is of utmost importance in many warehouses, especially where heavy pallet lifting or order picking trucks or cranes are used. Order-pick and forklift trucks may weigh up to several tons and can drive at considerable speeds. Warehouses should have safe travel paths for pedestrians and safety collision protection devices. Workers on foot should not work in narrow aisles together with heavy order-pick trucks. Unsafe working conditions can be discerned from
the amount of damage at the racks, at the trucks or signboards indicating the number of accidents, or if people smoke in a battery charging area. Unsafe working conditions should lead to a low score on this criterion.

Hygiene (based on hazard analysis and critical control points) is of particular importance for warehouses which process (pet) foods, drugs, or raw materials for such products. If deep-frozen products wait for a considerable time in an insufficiently conditioned receiving or shipping area, the condition of the product may deteriorate. Questions 2a, 2b, 3, 17 and 21 relate to this area.

Area 3:
Use of Space, Condition of Building and Technical Installations

Although (particularly in distribution warehouses) labor is the most important ingredient of operational cost (in particular the order pickers, see Tompkins et al., 2003), facility cost (including technical installations) is a close second. Whether buildings and technical installations are owned, rented or leased is irrelevant. Therefore, space should not be wasted. Excessively large warehouses do not only lead to high cost, but often also to inefficient processes, due to long travel times for storage, order picking or cross-dock. In case of storage of large numbers of loads of slow-moving products, high-bay stacking is preferred. There is, of course a difference between countries in the costs of land and labor. If labor and land are relatively cheap (USA), buildings are usually lower. If land is expensive (Japan), buildings are higher.

On the other hand, insufficient space may prevent a process from being executed effectively and efficiently. If products have to be dropped at temporary locations because of lack of space in the proper area, if products have to be dug up because they are stored at wrong locations, or if much waiting and delays occur because maneuvering spaces are used by other workers, this area receives a low score. It may be necessary that multiple persons work in the same area (for example order pickers and replenishers in a pallet storage area); nevertheless blocking and congestion should be avoided. This can be enforced by having one-directional traffic or distribution of fast-moving articles over multiple storage zones.

Many facilities have undergone natural expansion: gradually, more and more buildings and systems have been added. In many cases this leads to suboptimal logistic processes. Warehouses spread over multiple locations lead to necessary transport movements between the parts. How is this process
organized? Can inventory get lost while in transport? If not handled properly it should lead to a low score for this area.

The technical state of buildings, doors, floors, dock levelers, dock shelters, sprinkler installation, heating, cooling installations is fairly easy to assess during a visit. The quality of floors (i.e. flatness, and absence of pits and ramps) is particularly important if forklifts, reach trucks and high-bay trucks are used for discrete transport.

The basic facility layout is important for achieving top performance. U-shaped layouts, where dock doors are mainly located along one façade, usually lead to better performance (greater expansion possibilities, more flexible use of dock doors and receiving/shipping personnel, less crossing flows, shorter average travel distances) than layouts with dock doors on opposite sides of the buildings (I-shaped layout).

Questions 5a, 5b, 6a, 6b, 15, and 21 support the assessment of this area.

**Area 4:**
**Condition and Technical State of Material Handling Equipment**

Although it may seem wise at first sight to use a special truck for every different type of work, multiple brands of material handling equipment lead to less flexibility, higher risk of unavailability and higher maintenance cost. Material handling equipment that breaks down frequently or batteries that do not charge sufficiently may lead to an inefficient operation and missed deadlines. Even old trucks can work properly, if well maintained. You might try to ask a driver whether (s)he experiences any problems with the trucks. While asking this in a warehouse of a Serbian food retailer, it appeared that the batteries of one of the narrow-aisle pallet trucks charged insufficiently. This made the truck unavailable for a substantial part of the day, leading to orders that could not be filled completely on time.

Proper working material handling equipment shows from maintenance recorded on the equipment, the looks of the equipment and few failure records or performance obstructions in the operation. Question 16 supports this area.
Area 5: 
Teamwork, Management and Motivation

As Bartholdi and Eisenstein (1996) and Bartholdi et al. (2000) showed, bucket brigades, a teamwork order-picking concept, can lead to substantial performance (particularly throughput) improvements in picker-to-parts order picking systems. Although the bucket-brigade concept is only applicable under special circumstances, people working as a team will perform better than as individuals. This is particularly true in order picking, receiving and shipping. If people are multi-skilled and rotate in different areas of the warehouse, this might be an indicator of team spirit. If people are proud of their work and the company, this is a positive indicator. One might try to discern this factor by asking questions to the employees and management.

Questions 1, 12, 21 support this area.

Areas 6 and 7:
Storage and Order Picking Methods

Storage and order picking form the heart of most warehouse operations. Warehouse efficiency depends to a large extent on the methods used for storing products and picking the orders. The question is whether the appropriate methods are used. This is probably difficult to assess, particularly for inexperienced visitors. Also, great varieties of storage and picking technologies are available on the market. The choice of these also depends on the volume to be picked, the variety in the assortment and quantity to be stored and the labor cost rate. Higher labor costs and larger throughput volumes justify more automated storage and picking systems, and a higher level of order picking aids, like scanners, mobile terminals, or voice-recognition equipment. In low volume warehouses, i.e. with few orders, the preferred way is picking by order. Although multiple workers can work on the same order, the order is kept intact: it does not have to be split and sorted, but can, after possible order assembly, immediately be packed for shipping. In very high throughput volume warehouses, picking by order is impossible. Instead, orders are picked by article (in batch) after which the items are sorted and grouped by order.
Area 6:  
Storage Systems and Strategies and Inventory Management  

In order to assess the methods used the visitor might pay attention to the following elements.  
• Are products stored at their appropriate locations? This includes storage based on physical properties (conditioning, dimensions, weight, and theft-proneness) and turn-over speed: fast-moving items should be located on easily accessible locations at short distances from the dispatch position (Q7a).  
• Are locations used dynamically? In many warehouses fixed locations are used, from which products are picked. Even when products are initially assigned to these locations on turnover frequency (to reduce travel time), such an assignment will be far from optimal if not regularly maintained (like reassignment every month). Few companies do this. Companies that use dynamic locations, taking into account dynamic turnover frequency, score better than companies with fixed locations and little reassignment (Q7b).  
• Is the number of different storage systems (with different racks, material handling systems and storage logic) justified? Warehouses often store large numbers of products. The idea is to create the highest throughput efficiency possible, with the fewest systems used. These are often contradictory requirements, but a balance between the two should be struck. In case many different storage systems are used consideration should be given to merging two of them, without decreasing order picking efficiency or where few storage systems are used part of the assortment could be taken from a system and stored separately to increase efficiency and homogeneity of handling (Q7a, Q7b, Q8).  
• Is the inventory of certain products split into bulk storage and forward pick storage? If items are picked in a condensed forward storage area, the order picking lead times are reduced considerably and storage activities can be decoupled from order picking. Such systems can be designed for box picking (bulk stored on pallets, lower pallet locations used for picking the boxes), or item picking (bulk stored on boxes on pallets, shelves used for item picking). Particularly if bulk quantities tend to be large and order pick quantities are small, splitting inventory pays off and outweighs the replenishment efforts (Q9a).  
• Is family grouping applied in storage with the objective of making processes efficient? Many forms exist, such as grouping items that are frequently ordered together. Grouping methods that do not immediately lead to higher efficiency (such as products of the same supplier together, or products of the same owner together) score lower (Q7a, Q7b).
• Is inventory managed appropriately? Are inventory levels appropriate? It may be difficult to answer these questions, but clear visible signals should not be ignored. For example, in a company with short product life cycles, there should be an explicit program to get rid of “old” products. Look for a corner in the warehouse where seemingly non-movers are stored. These can be recognized by little pick activity, great product inhomogeneity, and sometimes small quantities stored per product. Inventory levels (ask for inventory turnover rate) depend on product properties, where suppliers are located and on the degree of supply chain cooperation. If suppliers are located further and products are cheaper, higher inventory levels are justified. Expensive products with short life cycles should have low inventory levels (Q9b, Q19).

**Area 7: Order Picking Systems and Strategies**

Before making an assessment, the order picking methods used (often more than one!) should be classified. A typical classification and explanation of methods can be found in Exhibit 1. Have the weak points of the order picking systems used been addressed adequately and sufficiently? Every order picking system has strengths and weaknesses. The strengths are usually immediately visible in a visit (apparently, the system works); weaknesses are more difficult to discern. Batch picking, followed by sorting on an automated sorter, requires that all items (including the last items, which usually are missing) are picked in time for the sorter to start. Is this handled adequately? Order throughput times in picker-to-parts systems can sometimes be very long. Is this controlled sufficiently? For example in Océ’s parts warehouse (Océ is a manufacturer of professional copiers and printers), which supplies parts overnight directly to technicians in Western Europe, orders are picked in batches (of orders for technicians in the same country) of about 60-120 order lines per order picker. The throughput time can be very long and is difficult to predict. Also, pickers can decide themselves on the number of lines they want to work on. This makes it difficult to guarantee that the fixed departure times of the trucks can be realized, requiring extra control effort (regular progress checking and emergency help) to guarantee this. The European warehouse of Yamaha Motor Parts uses a zoned pick-by-order system. A conveyor passes the order bins between the zones. As there are many zones (about 60), and orders can sometimes be large, orders queue before every zone, making order throughput times close to unpredictable at busy moments. Yet, Yamaha has a fixed truck departure schedule for all customer destinations. The problem was solved by
batching multiple small orders into the same order bin, thereby strongly reducing queuing. In C-Market’s warehouse (a supermarket chain) pickers on order pick trucks travel long distances in a large pallet warehouse to pick orders for a single supermarket. In competitors’ warehouses, pickers on long-fork trucks pick two or three stores simultaneously in roll containers in one warehouse zone only, which leads to a large increase in productivity.

The following questions (see also questions 10, 11a, 11b, and 20) might guide the evaluation of the order picking process:

• Are throughput times sufficiently controlled?
• Does avoidable double handling occur?
• Are obvious improvements possible in the picking process? You might think of some improvements and ask the pickers for their evaluation.
• How is the progress of the order picking process monitored and controlled?
• Are the used picking aids (order lists, labels, RF terminals, scanners, picking carts) well designed and of help to increase quality and efficiency?
• Have measures been taken to make the picking process sufficiently ergonomic?

**Area 8: Supply Chain Coordination**

The degree of supply chain coordination is visible at the shop-floor in several areas. At the yard, inbound trucks may be waiting to be allocated to a dock door, due to inability to properly coordinate arrival times. In the receiving area, trailers and containers must be unloaded and goods must be processed. Is this a rapid, well-organized process, or very time-consuming because the product carriers are wrong and products have to be restacked, information cannot be found or is incomplete, boxes of the same product are spread over multiple pallets or over the entire container? In case much paperwork is necessary to check incoming shipments, this is also not a sign for well-tuned processes. You might also ask what happens in case of wrong, under or over receipts. Does this happen often? Does it delay the process? Attention also has to be paid to the frequency of supply and the drop size. Drop size might be identified at a visit, frequency not without asking. If you see small drop sizes, ask the receivers the frequency of supply of these suppliers. At some warehouses, powerful customers try to reduce their inventories by JIT policies: frequently ordering small quantities. Although this leads to inventory reduction at the customer’s facility, it leads to high handling and transportation cost for the supplier, which might retaliate against the customer.
In an extreme case, we asked a US wholesaler where the customer returns were handled. In response to that question we were taken to a warehouse at the other side of the street, where an endless heap of mostly damaged boxes were waiting to be processed. These were the returns of mainly one customer, who returned “suddenly” a few dozen truckloads of excess stock. This was representative for the company’s entire receiving process.

Even if products are loosely stacked in sea-containers, it is still possible to have an efficient receiving process if adequate agreements have been made with suppliers. In the warehouse of Zeeman, a textile hard-discounting retailer mainly receiving products in sea containers from East-Asian suppliers, the boxes are grouped by product in the container, and box-sizes are standardized. This allows rapid manual unloading of the containers using extendible conveyors, after which the boxes are automatically counted, labeled and palletized. Conversely Schuitema, a franchise retail organization, has to restack all of Unilever’s pallets (a main supplier), because they do not fit into the storage slots.

The level of supply chain coordination is also visible in the shipping area. An abundance of paperwork needed to ship products is an indicator, as well as the carriers on which products are shipped. If products are shipped on product carriers that return (for example pool pallets, or closed-loop bins), this often indicates an efficient distribution and collection process, coordinated with the recipients. It saves one-way packaging materials which, particularly in Europe, are expensive, not only because of material cost, but also because fees have to be paid to green-dot systems in different countries to organize proper recycling of these materials. If products are shipped in sea containers on slipsheets (loads on flat carton ‘pallets’ that can be pushed into the container by ‘push-pull’ trucks) this saves space in the container and it suggests advanced coordination with the receiving customer (who also needs such a truck).

Question 19 refers to this area.

Area 9: Level and Use of IT

Nowadays, warehouses do not run without a sufficient level of information systems. Best-in class warehouses use systems for electronic information exchange with suppliers, customers, carriers, customs authorities, and brokers in the supply chain. They use a warehouse management system for managing the warehouse processes and they use appropriate tools and aids to support
important warehouse processes. Warehouse management systems come in a great variety, varying from simple spreadsheet applications, to standard modules of ERP software packages, specialized WMS packages or tailor-made applications. In general, the more complex the operation (mainly measured in number of order lines, assortment size, different processes and uncertainty in demand and supply, see Faber et al., 2002), the more justified or even necessary specific or tailor-made software becomes. A warehouse management system is necessary to find the best location where an incoming load can be stored, the best location from which an order line can be picked, the right person to pick an order line (in the right sequence, minimizing travel time), the regular update of article-to-location assignments (based on turnover frequency) to internally move products to make sure that articles are cycle counted regularly without disturbing the main work flows, and so on. Tools that can be used to speed up processes and reduce errors include pick-to-light and put-to-light systems and use of the right communication means with drivers and pickers to guarantee real-time monitoring of work progress. Bakker, a mail-order company which specializes in flower bulbs, uses a put-to-light system for distributing bulbs that have been pre-picked over the right customer order bins. A graphical screen helps the picker, as it shows visually which bins have to be addressed. These aids increase productivity significantly.

Question 20 reflects this area.

**Area 10:**
**Commitment to Quality**

Commitment to quality can be derived from a number of factors in a facility. First, from the design itself, at which points is it easily possible to make errors? If an operator can determine where to store an incoming load and later provide confirm, this is an obvious source for errors. Storage errors are very serious, as they potentially impact multiple customer orders. The same is true for picking: can an operator easily pick the wrong item or the wrong quantity? Best-in class operations do not ensure quality by building in additional checks of the picked orders. Instead, they take measures that prevent people from making obvious errors (‘poka-yoke’, or fool proofness principle). In the previously mentioned warehouse of Yamaha, pickers at a miniload workstation have to pick a unit from a compartmented bin containing multiple products. In order to prevent errors, the computer screen is divided in the same way as the bin, with the proper part illuminated. On top of this, a battery of spotlights illuminates exactly the right compartment of the bin.
Second, is continuous process improvement actively propagated in the facility? Are workers stimulated to improve their processes and can proof be found for this? Indicators for this can be an idea-box, implementation of six-sigma improvement projects or the number of master black-belts, or the number of process improvements recently realized. You might try and ask about this. In a recent tour of the European distribution center of a US manufacturer, we were told that people could be promoted to management level only if they at least owned a six-sigma green-belt.

This area is addressed with questions 4, 11a, 11b, 12, 13, 14, 17 and 20.

**Area 11:**
*Managing Efficiency and Flexibility, as a Function of Volume, Assortment and Variety*

It is very difficult – if not impossible – to manage a large number of orders, together with a large assortment and a variety of customer wishes efficiently, in a manner that is flexible enough to accommodate late changes. Process automation and mechanization, with multiple solutions for different storage areas can help for efficiency, but usually bring down flexibility. Logistics service providers with public warehouses and short-term contracts usually opt primarily for flexibility and sacrifice efficiency to some extent. Flexibility is expressed as the ease to which different customer order patterns (large versus small orders), different customer wishes (product and order customization) can be accommodated, the processes expanded or shrunk, assortment changes handled. During a visit attention can be paid to what extent any of these principles have been sacrificed. If processes seem very efficient you might ask whether the above-mentioned flexibility features can be accommodated. In case an operation seems very flexible, it is interesting to estimate whether customers are really willing to pay for the inefficiency. If a right balance seems to have been struck a company scores higher than when there are obvious flaws. This is addressed with question 18.
3. Results and Validation

The assessment has been carried out with several groups of managers and students. Within a group the areas are divided over different group members. Immediately after the visit, each group filled out the warehouse rating sheet as a team effort. Exhibit 4 shows the outcomes of some assessments carried out in 2004 and 2005 with different groups of international people (in total 96 persons from 22 countries participated, about 30-40 people per visit, with and without warehousing experience). For every facility, the maximum score is 121. The results show a clear distinction between high and low-ranking facilities. Low ranked facilities nearly always score ‘NO’ on question 21; high-ranked facilities ‘YES’. The outcomes of area ratings are quite varied as well, although “Customer satisfaction” (area 1) obviously scores fairly high in general.

In order to validate the method, basically three different methods were used. First, we independently benchmarked the warehouses using data envelopment analysis (DEA), based on a database of 71 warehouses. Second we compared the standard deviations of area and total scores among groups. If these standard deviations are moderate, we can at least say that the scoring is reliable. Third, we asked the managers method for feedback on the scores per area (the method was mailed to them prior to the visit).

In order to benchmark the warehouses with DEA, we asked the warehouse or logistics manager to fill out a questionnaire, addressing performance in the areas of shipment quality, production (volume and variety) and flexibility (for a full description of the method see De Koster and Balk, 2008). The resulting efficiency scores (the maximum efficiency to be obtained is 100%) can be found in Exhibit 4. Although the factor rating and benchmarking methods look at different indicators, the correlation between the two scores is quite high: 64% for the companies listed in Exhibit 4, indicating that the assessment method is a good forecaster of performance (albeit the number of included warehouses is still small).

Exhibit 4 also displays the standard deviation of total and area scores. The maximum standard deviation of the total score is within 16% of the average. For individual area scores, the average standard deviation varies between 1.4 and 1.7 (less than 25% of the average area score). Usually there are 1 or 2 areas of some disagreement between groups, with standard deviations up to 2.6. No areas consistently showed a higher standard deviation in the scoring. The score reliability improves when the assessment is done with more experienced
people: having seen more facilities obviously helps in calibrating one’s judgment. However, it should be emphasized that all facilities were also visited by such inexperienced people, leading to the above-mentioned moderate standard deviations of scores.

After every visit, the warehouse manager was confronted with the area scores. In all cases, they agreed with the relative ranking of their scores. Obviously, warehouse management is often aware of weak points, but it is not always easy to improve. For example a weak layout cannot easily be changed by the management; such a conclusion should serve as input for the company’s facility development staff.
4. Conclusion

The method presented in this paper may help managers and students to rapidly assess warehouse facilities. The method serves as an addition to more quantitative methods, like financial analysis. We have validated the method with DEA benchmarking. Although the number of warehouses benchmarked with both methods is still small, first results indicate that indeed the method shows some value in an assessment. Total and area scores are reasonably homogeneous among the different groups (although every warehouse so far shows one or two areas with standard deviations higher than 2, which may be as much as 40% of the average area score). It is helpful, in this respect, that the assessors have applied the method more than once.

In conclusion, if a warehouse appears to score well, based on the visual information and Q&A, it usually is. If it scores poorly, there definitely is room for improvement, particularly in the low-ranked areas.

Exhibit 1: Order-picking Methods

The next figure shows different order picking methods that can be found in warehouses (for a description of some of these methods, see Tompkins et al., 2003). In many warehouses multiple methods are used. The large majority employs humans for order picking. Among those, the picker-to-parts system, where the picker walks or rides along the items, is most common. Parts-to-picker systems include automated storage and retrieval systems (AS / RS), mostly using aisle-bound cranes that retrieve one or more unit loads (bins: miniload system, or pallets) and bring it to a pick position. At this position the picker takes the number of pieces required by the customer order, after which remaining load is stored again. Other systems use vertical lift modules (VLM), or carousels that also offer unit loads to the picker, who is responsible for taking the right quantity. Put systems are positioned between the picker-to-parts and parts-to-picker systems, because they often combine the two principles. First, inventory has to be retrieved, which can be done in a parts-to-picker or picker-to-parts manner. Second, the carrier (usually a bin) with these parts is offered to a picker who distributes the items over customer orders. Put systems are particularly popular in case a large number of customer order lines have to be picked in a short time window (for example at the Amazon German warehouse) and can result in about 500 packages on average per picker hour (for small packages) in well-managed systems.
Picker-to-part systems are the most common. The basic variants include picking by article (sometimes called batch picking) or pick by order. In the case of article picking, multiple customer orders (the “batch”) are picked simultaneously by a picker. Many in-between variants exist: picking multiple orders followed by immediate sorting (on the pick cart) by the picker (“sort-while-pick”), or “pick-and-sort” in which case the sorting takes place after the pick process has finished. Another basic variant is zoning, which means that a logical storage area (this might be a pallet storage area, but also the entire warehouse) is split in multiple parts, each with different pickers. The pickers can work sequentially, traveling along the locations in their zone and pass the product carrier with pick instruction to pickers in the next zone, or they can work in parallel, and work on the same orders. If this is the case, the order parts have to be assembled before they can be packed and shipped. Parallel and batch picking speed up the picking process, at the cost of additional sorting and order assembly work. The term “wave picking” is used if orders for a common destination (for example, departure at a fixed time with a certain carrier) are released simultaneously for picking in multiple warehouse areas. Usually it is combined with batch picking.
**Exhibit 2:**
**Questionnaire**

The total number of yeses on this questionnaire is an indicator of the warehouse’s overall performance. The more yeses, the better the performance. A question should be answered a yes only, if the warehouse obviously adheres to the principle implied by the question. In case of doubt, answer no.

<table>
<thead>
<tr>
<th>Warehouse</th>
<th>Date visit</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group:</strong></td>
<td>Yes</td>
</tr>
<tr>
<td>1</td>
<td>Are visitor welcomed and given information about warehouse operation, customers and products?</td>
</tr>
<tr>
<td>2a</td>
<td>Is the facility clean, safe, orderly and well lit?</td>
</tr>
<tr>
<td>2b</td>
<td>Is the air quality good and noise level low?</td>
</tr>
<tr>
<td>2c</td>
<td>Is the environment attractive to work in?</td>
</tr>
<tr>
<td>3</td>
<td>Are the work processes ergonomically well-thought over?</td>
</tr>
<tr>
<td>4</td>
<td>Do the employees appear committed to quality?</td>
</tr>
<tr>
<td>5a</td>
<td>Is the warehouse laid out in a U-shape, rather than an I-shape?</td>
</tr>
<tr>
<td>5b</td>
<td>Does the layout prevent major crossing flows?</td>
</tr>
<tr>
<td>6a</td>
<td>Is material moved over the shortest / best possible distances?</td>
</tr>
<tr>
<td>6b</td>
<td>Is double handling prevented and are appropriate product carriers used?</td>
</tr>
<tr>
<td>7a</td>
<td>Are products stored on their right locations?</td>
</tr>
<tr>
<td>7b</td>
<td>Do storage strategies lead to operational efficiency?</td>
</tr>
<tr>
<td>7c</td>
<td>Are locations used dynamically?</td>
</tr>
<tr>
<td>8</td>
<td>Is the number of different storage systems (with different racks, material handling systems and storage logic) justified?</td>
</tr>
<tr>
<td>9a</td>
<td>Is appropriate (non-)splitting of inventory in bulk and forward pick stock applied?</td>
</tr>
<tr>
<td>9b</td>
<td>Is there an effective process management for introducing new products, getting rid of non-movers, and internal relocations?</td>
</tr>
</tbody>
</table>
Continued

<table>
<thead>
<tr>
<th>No.</th>
<th>Question</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Is the organization of the picking process well-designed without obvious improvement possibilities?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11a</td>
<td>Are storage and receiving processes monitored and controlled on-line?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11b</td>
<td>Is the response to mistakes and errors immediate?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Are work teams trained, empowered and involved in problem solving and ongoing improvements?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Are up-to-date operational goals and performance measures for those goals prominently posted?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Are ratings for customer satisfaction and shipping errors displayed?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Are the buildings, floors and technical installations in good quality and well-maintained?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Are the material handling systems used, the racks and the product carriers in good operating condition and well-maintained?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Are inventories accurate?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Has a right balance been struck between order customization, process flexibility and efficiency?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Are receiving and shipping processes, and inventory levels tuned with suppliers and customers?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Is the level of IT, picking and storage technologies adequate for the operation?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Is this a warehouse you would like to work in?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total yes / no: **67**
### Exhibit 3: Warehouse Rating Sheet

<table>
<thead>
<tr>
<th>Group</th>
<th>Area</th>
<th>Related questions</th>
<th>Poor (1)</th>
<th>Below average (3)</th>
<th>Average (5)</th>
<th>Above average (7)</th>
<th>Excellent (9)</th>
<th>Best in class (11)</th>
<th>Total (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Customer satisfaction</td>
<td>1, 14, 21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Cleanliness, environment, ergonomics, safety, hygiene</td>
<td>2a, 2b, 3, 17, 21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Use of space, condition of building and technical installations</td>
<td>5a, 5b, 6a, 6b, 15, 21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Condition and maintenance of material handling equipment</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Teamwork, management &amp; motivation</td>
<td>1, 12, 21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Storage systems &amp; strategies, inv. man.</td>
<td>7a, 7b, 8, 9a, 9b, 19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group:</td>
<td>Area</td>
<td>Related</td>
<td>Poor</td>
<td>Below</td>
<td>Average</td>
<td>Above</td>
<td>Excellent</td>
<td>Best in</td>
<td>Total</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>---------</td>
<td>------</td>
<td>-------</td>
<td>---------</td>
<td>-------</td>
<td>-----------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>questions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Order picking</td>
<td>10, 11a, 11b, 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Supply chain coordination</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Level and use of IT</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Commitment to quality</td>
<td>4, 11a, 11b, 12, 13, 14, 17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Managing efficiency and flexibility</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Total score**

**Date visit**

**Warehouse**

**RENE DE KOSTER PAST AND FUTURE: PERSPECTIVES ON MATERIAL HANDLING – PART 2 – SELECTED ACADEMIC CONTRIBUTIONS AND PUBLICATIONS**
### Exhibit 4:
Some Examples of the Tool’s Results. N = Number of Groups

<table>
<thead>
<tr>
<th>Warehouse</th>
<th>Description</th>
<th>Ave. total rating (N)</th>
<th>Std. dev. (max) std dev. per area</th>
<th>Average (max) std dev. per area</th>
<th>DEA efficiency score</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Multinational interior-decoration retailer</td>
<td>65.9 (8)</td>
<td>10.8</td>
<td>1.6 (2.5)</td>
<td>58.8 %</td>
</tr>
<tr>
<td>B</td>
<td>Automotive manufacturer, spare parts</td>
<td>82.5 (8)</td>
<td>8.9</td>
<td>1.7 (2.6)</td>
<td>95.5 %</td>
</tr>
<tr>
<td>C</td>
<td>National wholesaler supermarket products</td>
<td>76.3 (6)</td>
<td>3.5</td>
<td>1.4 (2.0)</td>
<td>100 %</td>
</tr>
<tr>
<td>D</td>
<td>National food retailer</td>
<td>59.2 (9)</td>
<td>7.2</td>
<td>1.5 (2.1)</td>
<td>-</td>
</tr>
<tr>
<td>E</td>
<td>Multinational hard-discounting non-food retail chain</td>
<td>64.0 (6)</td>
<td>10.0</td>
<td>1.6 (2.5)</td>
<td>66.2 %</td>
</tr>
<tr>
<td>F</td>
<td>Multinational fashion products manufacturer / wholesaler / retailer</td>
<td>73.0 (6)</td>
<td>3.1</td>
<td>1.4 (2.2)</td>
<td>44.2 %</td>
</tr>
</tbody>
</table>

---

1 In calculation, we interpreted the scores as measured on an interval scale.
5. Literature


Heizer, J. and B. Render, Operations management, Pearson-Prentice-Hall, Upper Saddle River, 2004


Design and control of order picking

Abstract

Order picking has long been identified as the most labour-intensive and costly activity for almost every warehouse; the cost of order picking is estimated to be as much as 55% of the total warehouse operating expense. Any under-performance in order picking can lead to unsatisfactory service and high operational cost for its warehouse, and consequently for the whole supply chain. In order to operate efficiently, the order-picking process needs to be robustly designed and optimally controlled. This paper gives a literature overview on typical decision problems in design and control of manual order-picking processes. We focus on optimal (internal) layout design, storage assignment methods, routing methods, order batching and zoning. The research in this area has grown rapidly recently. Still, combinations of the above areas have hardly been explored. Order-picking system developments in practice lead to promising new research directions.

Keywords: Order picking; Warehouse management; Logistics
1. Introduction

As more companies look to cut costs and improve productivity within their warehouses and distribution centres, picking has come under increased scrutiny. Order picking – the process of retrieving products from storage (or buffer areas) in response to a specific customer request – is the most labour-intensive operation in warehouses with manual systems, and a very capital-intensive operation in warehouses with automated systems (Goetschalckx and Ashayeri 1989, Drury 1988, Tompkins et al. 2003). For these reasons, warehousing professionals consider order picking as the highest-priority area for productivity improvements.

Several recent trends both in manufacturing and distribution have made the order-picking design and management become more important and complex. In manufacturing, there is a move to smaller lot-sizes, point-of-use delivery, order and product customisation, and cycle time reductions. In distribution logistics, in order to serve customers, companies tend to accept late orders while providing rapid and timely delivery within tight time windows (thus the time available for order picking becomes shorter). Many smaller warehouses are being replaced by fewer large warehouses to realise economies of scale. In these large warehouses, the daily pick volume is large and the available time window is short. In order to be more responsive to customers, many companies have adopted a postponement strategy (Van Hoek 2001) leading to various value-adding activities (like kitting, labelling, product or order assembly, customised packaging or palletisation) that take place in the distribution centre and which have to be scheduled and integrated in the order-picking process. Warehouses are also involved in recovering products, materials, and product carriers from customers in order to redistribute them to other customers, recyclers, and original-equipment manufacturers (De Koster et al., 2002).

The organisation of order-picking operations immediately impacts the distribution centre’s and thereby the supply chain’s performance. Between the time an order is released to the warehouse and the time it takes to reach its destination, there is ample opportunity for errors in both accuracy and completeness, not to mention time lost. There is also room for improvement. Industry has come up with innovative solutions, making it possible to attain productivity up to 1,000 picks per person hour. Science is also progressing rapidly. Over the last decades, many papers have appeared studying order picking processes. New problems have been studied and new models have been
developed. Still, there is a gap between practice and academic research, since not all new picking methods have been studied and the optimal combinations of layout, storage assignment, order clustering, order release method, picker routing and order accumulation have been addressed to a minor extent only. This paper presents a systematic overview of these recent developments in academic literature. We structure typical decision problems in design and control of order-picking processes by focusing on optimal (internal) layout design, storage assignment methods, routing methods, order batching, and zoning. Several areas appear to have received only little attention from researchers. Innovations from practice also lead to new research challenges.

The remainder of the paper is organised as follows. In the next section, we briefly highlight warehouse missions and functions and give an overview of order-picking systems. In Sections 3 to 8, we review recent literature on design and control of order-picking processes, focussing on layout design, storage assignment, batching, picker routing, and order accumulation. We conclude and discuss potential research directions in Section 8.
2. Warehouses and Order Picking

According to ELA / AT Kearney (2004), warehousing contributed to about 20% of the surveyed companies’ logistics costs in 2003 (other activities distinguished are value added services, administration, inventory costs, transportation and transport packaging). Warehouses apparently form an important part of a firm’s logistics system. They are commonly used for storing or buffering products (raw materials, goods-in-process, finished products) at and between points of origin and points of consumption. The term ‘warehouse’ is used if the main function is buffering and storage. If additionally distribution is a main function, the term ‘distribution centre’ is commonly used, whereas ‘transhipment’, ‘cross-dock’, or ‘platform’ centre are often used if storage hardly plays a role. As we focus on order picking from inventory, we use the term ‘warehouse’ throughout the paper. Lambert et al. (1998) state that more than 750,000 warehouse facilities exist worldwide, including state-of-art, professionally managed warehouses, as well as company stockrooms and self-store facilities. Warehouses often involve large investments and operating costs (e.g. cost of land, facility equipment, labour ...). So, why do warehouses exist? According to Lambert et al. (1998) they contribute to a multitude of the company’s missions, like

- Achieving transportation economies (e.g. combine shipment, full-container load).
- Achieving production economies (e.g. make-to-stock production policy).
- Taking advantage of quality purchase discounts and forward buys.
- Supporting the firm’s customer service policies.
- Meeting changing market conditions and uncertainties (e.g. seasonality, demand fluctuations, competition).
- Overcoming the time and space differences that exist between producers and customers.
- Accomplishing least total cost logistics commensurate with a desired level of customer service.
- Supporting the just-in-time programs of suppliers and customers.
- Providing customers with a mix of products instead of a single product on each order (i.e. consolidation).
- Providing temporary storage of material to be disposed or recycled (i.e. reverse logistics).
- Providing a buffer location for trans-shipments (i.e. direct delivery, cross-docking).
In some special situations (e.g. lean manufacturing, ‘virtual’ inventory, cross-docking), storage functions in a supply chain can be reduced. But, in almost all supply chains, raw materials, parts, and product inventories still need to be stored or buffered, implying that warehouses are needed and play a critical role in the companies’ logistics success.

2.1 Warehouse flows

Figure 1 shows the typical functional areas and flows within warehouses. The main warehouse activities include: receiving, transfer and put away, order picking / selection, accumulation / sortation, cross-docking, and shipping.

Figure 1. Typical warehouse functions and flows (Tompkins et al. 2003)

The receiving activity includes the unloading of products from the transport carrier, updating the inventory record, inspection to find if there is any quantity or quality inconsistency. Transfer and put away involves the transfer of incoming products to storage locations. It may also include repackaging (e.g. full pallets to cases, or standardised bins), and physical movements (from the receiving docks to different functional areas, between these areas, from these areas to the shipping docks). The order picking / selection is the major activity in most warehouses. It involves the process of obtaining a right amount of the right products for a set of customer orders. The accumulation / sortation of picked orders into individual (customer) orders is a necessary activity if the orders have been picked in batches. In such a case the picked units have to be grouped by customer order, upon completion of the pick process. After picking, orders often have to be
packed and stacked on the right unit load (e.g. a pallet). Cross-docking is performed when the received products are transferred directly to the shipping docks (short stays or services may be required but little or no order picking is needed).

### 2.2 Order picking

Order picking involves the process of clustering and scheduling the customer orders, assigning stock on locations to order lines, releasing orders to the floor, picking the articles from storage locations and the disposal of the picked articles. Customer orders consist of order lines, each line for a unique product or stock keeping unit (SKU), in a certain quantity. In Figure 1, order lines are split, based on quantity and product carrier of the SKU, in pallet picks, case picks and broken case (unit) picks. Many different order-picking system types can be found in warehouses. Often multiple order-picking systems are employed within one warehouse, for example in each of the three zones of Figure 1. Figure 2 distinguishes order-picking systems according to whether humans or automated machines are used. The majority of warehouses employ humans for order picking. Among these, the picker-to-parts systems, where the order picker walks or drives along the aisles to pick items, are most common (De Koster 2004).

We can distinguish two types of picker-to-parts systems: **low-level** picking and **high-level** picking. In low-level order-picking systems, the order picker picks requested items from storage racks or bins (bin-shelving storage), while travelling along the storage aisles. Other order-picking systems employ high storage racks; order pickers travel to the pick locations on board of a lifting order-pick truck or crane. The crane automatically stops in front of the appropriate pick location and waits for the order picker to perform the pick. This type of system is called a high-level or a **man-aboard** order-picking system.

**Parts-to-picker** systems include automated storage and retrieval systems (AS / RS), using mostly aisle-bound cranes that retrieve one or more unit loads (pallets or bins; in the latter case the system is often called a mini-load) and bring them to a pick position (i.e. a depot). At this position the order picker takes the required number of pieces, after which the remaining load is stored again. This type of system is also called a **unit-load** or **end-of-aisle** order-picking system. The automated crane (also: **storage and retrieval (S / R) machine**) can work under different operating modes: **single**, **dual** and **multiple command cycles**. The single-command cycle means that either a load is moved from the depot to a rack location or from a rack location to the depot. In the dual-command mode, first a load is moved from the depot to the rack location and next another load is...
retrieved from the rack. In multiple command cycles, the S/R machines have more than one shuttle and can pick up and drop off several loads in one cycle. For example, in a four-command cycle (described in Sarker and Babu 1995), the S/R machine leaves the depot with two storage loads, stores them and returns with two retrieved loads. Other systems use modular vertical lift modules (VLM), or carousels that also offer unit loads to the order picker, who is responsible for taking the right quantity.

Put systems, or order distribution systems (see Figure 2) consist of a retrieval and distribution process. First, items have to be retrieved, which can be done in a parts-to-picker or picker-to-parts manner. Second, the carrier (usually a bin) with these pre-picked units is offered to an order picker who distributes them over customer orders (‘puts’ them in customer cartons). Put systems are particularly popular in case a large number of customer order lines have to be picked in a short time window (for example at the Amazon Germany warehouse, or flower auctions) and can result in about 500 picks on average per order picker hour (for small items) in well-managed systems (De Koster 2004). Newly developed systems indicate that up to 1000 put handlings per picker hour are feasible.

Figure 2 also shows several organisational variants of picker-to-parts systems. The basic variants include picking by article (batch picking) or pick by order (discrete picking). In the case of picking by article, multiple customer orders (the batch) are picked simultaneously by an order picker. Many in-between variants exist, such as picking multiple orders followed by immediate sorting (on the pick cart) by the order picker (sort-while-pick), or the sorting takes place after the pick process has finished (pick-and-sort). Another basic variant is zoning, which means that a logical storage area (this might be a pallet storage area, but also the entire warehouse) is split in multiple parts, each with different order pickers. Depending on the picking strategy, zoning may be further classified into two types: progressive zoning and synchronised zoning, depending on whether orders picked in a zone are passed to other zones for completion or picked in parallel. The term wave picking is used if orders for a common destination (for example, departure at a fixed time with a certain carrier) are released simultaneously for picking in multiple warehouse areas. Usually (but not necessarily) it is combined with batch picking. The batch size is determined based on the required time to pick the whole batch completely, often between 30 minutes to 2 hours (see Petersen 2000). Order pickers pick continuously the requested items in their zones, and a next picking wave can only start when the previous one is completed.
Automated and robotised picking is only used in special cases (e.g. valuable, small and delicate items).

Figure 2. Classification of order-picking systems (based on De Koster 2004)

In this paper we concentrate on low-level, picker-to-parts order-picking systems employing humans (and with multiple picks per route). These systems form the very large majority of picking systems in warehouses worldwide (based on the authors’ experience: over 80% of all order-picking systems in Western Europe). Surprisingly, academic order-picking literature focuses more on high-level picking and AS/RS systems. Although not the main topic of this paper, we will briefly mention some of the latter type of literature as well.

The design of real order-picking systems is often complicated, due to a wide spectrum of external and internal factors which impact design choices. According to Goetschalckx and Ashayeri (1989) external factors that influence the order-picking choices include marketing channels, customer demand pattern, supplier replenishment pattern and inventory levels, the overall demand for a product, and the state of the economy. Internal factors include system characteristics, organisation, and operational policies of order-picking systems. System characteristics consist of mechanisation level, information availability and warehouse dimensionality (see Figure 3). Decision problems related to these factors are often concerned at the design stage. The organisation and operational policies include mainly five factors: routing, storage, batching, zoning and order
release mode. Figure 3 also shows the level of complexity of order-picking systems, measured by the distance of the representation of this problem in the axis system to the origin. In other words, the farther a system is located from the origin, the harder the system is to design and control.

![Figure 3. Complexity of order-picking systems (based on Goetschalckx and Ashayeri 1989)](image)

2.3 **Order picking objectives**

The most common objective of order-picking systems is to maximise the service level subject to resource constraints such as labour, machines, and capital (Goetschalckx and Ashayeri 1989). The service level is composed of a variety of factors such as average and variation of order delivery time, order integrity, and accuracy. A crucial link between order picking and service level is that the faster an order can be retrieved, the sooner it is available for shipping to the customer. If an order misses its shipping due time, it may have to wait until the next shipping period. Also, short order retrieval times imply high flexibility in handling late changes in orders. Minimising the order retrieval time (or picking time) is, therefore, a need for any order-picking system.

Figure 4 shows the order-picking time components in a typical picker-to-parts warehouse. Although various case studies have shown that also activities
other than travel may substantially contribute to order-picking time (Dekker et al. 2004, De Koster et al. 1999a), travel is often the dominant component. According to Bartholdi and Hackman (2005) ‘travel time is waste. It costs labour hours but does not add value’. It is, therefore, a first candidate for improvement.

**Figure 4. Typical distribution of an order picker’s time (Tompkins et al. 2003)**

![Diagram of order picker's time distribution](image)

For manual-pick order-picking systems, the travel time is an increasing function of the travel distance (see, for example, Jarvis and McDowell 1991, Hall 1993, Petersen 1999, Roodbergen and De Koster 2001a,b, Petersen and Aase 2004). Consequently, the travel distance is often considered as a primary objective in warehouse design and optimisation. Two types of travel distance are widely used in the order-picking literature: the average travel distance of a picking tour (or *average tour length*) and the total travel distance. For a given pick load (a set of orders), however, minimising the average tour length is equivalent to minimising the total travel distance.

Clearly, minimising the average travel distance (or, equivalently, total travel distance) is only one of many possibilities. Another important objective would be minimising the total cost (that may include both investment and operational costs). Other objectives which are often taken into consideration in warehouse design and optimisation are to:

- minimise the throughput time of an order
- minimise the overall throughput time (e.g. to complete a batch of orders)
- maximise the use of space
- maximise the use of equipment
- maximise the use of labour
- maximise the accessibility to all items
Companies make decisions on design and control of order picking systems at tactical or operational level, with a different time horizon (Rouwenhorst et al. 2000). Common decisions at these levels are:

- layout design and dimensioning of the storage system (tactical level)
- assigning products to storage locations (storage assignment) (tactical and operational level)
- assigning orders to pick batches and grouping aisles into work zones (batching and zoning) (tactical and operational level)
- order picker routing (routing) (operational level)
- sorting picked units per order and grouping all picks of the orders (order accumulation / sorting) (operational level)

In realising the above objectives, decisions made at the various levels are strongly interdependent. For example, a certain layout or storage assignment may perform well for certain routing strategies, but poorly for others. However, including all decisions (with obvious different decision horizons) in one model is intractable. Researchers, therefore, limit themselves to one or few decision areas simultaneously. In practice, decisions are also made sequentially, or variations are simply not considered. In the following sections we therefore subsequently treat these decision areas and mention area interactions observed by authors when appropriate. We first give an introduction to the problem and then briefly mention the concerned literature. Issues in design and planning of warehousing systems have been discussed in Ashayeri and Gelders (1985), Cormier and Gunn (1992), Cormier (1997), Van den Berg (1999), Van den Berg and Zijm (1999) and Rouwenhorst et al. (2000). Issues in design and control of order-picking processes in particularly are mentioned in Goetschalckx and Ashayeri (1989), Choe and Sharp (1991), Roodbergen (2001) and Wäscher (2004). An extensive bibliography on order-picking systems is gathered in Goetschalckx and Wei (1994) and Roodbergen (2001). As many papers on the order-picking problem have appeared recently, most of the above-mentioned overview publications are not up-to-date. Wäscher (2004) chooses a similar approach to ours and discusses storage assignment, order batching and picker routing problems, but treats only a small fraction of the available literature.
3. Layout Design

In the context of order picking, the layout design concerns two sub-problems: the layout of the facility containing the order-picking system and the layout within the order-picking system. The first problem is usually called the *facility layout problem*; it concerns the decision of where to locate various departments (receiving, picking, storage, sorting, and shipping, etc.). It is often carried out by taking into account the activity relationship between the departments. The common objective is minimising the handling cost, which in many cases is represented by a linear function of the travel distance. We refer to Tompkins et al. (2003) for a description of several efficient layout design procedures and to Meller and Gau (1996) for a general literature overview on this subject. Furthermore, Heragu et al. (2005) give a model and heuristic for sizing of areas and assignment of products to areas. In this paper, we focus on the second sub-problem, which can also be called the *internal layout design* or *aisle configuration problem*. It concerns the determination of the number of blocks, and the number, length and width of aisles in each block of a picking area (see Figure 5). The common goal is to find a ‘best’ warehouse layout with respect to a certain objective function among the layouts which fit a given set of constraints and requirements. Again, the most common objective function is the travel distance.

Literature on layout design for low-level manual order-picking systems is not abundant. An early publication, albeit focussing on unit loads, is by Bassan et al. (1980). They compare two different parallel-aisle layouts for handling (including travel) and layout costs. Rosenblatt and Roll (1984), using both analytical and simulation methods, study the effect of storage policy (i.e. how to assign products to storage locations) on the internal layout of warehouse. Rosenblatt and Roll (1988) examine the effect of stochastic demands and different service levels on the warehouse layout and storage capacity. Recently, Roodbergen (2001) proposed a non-linear objective function (i.e. average travel time in terms of number of picks per route and pick aisles) for determining the aisle configuration for random storage warehouses (including single and multiple blocks) that minimises the average tour length. Also considering minimisation of the average tour length as the major objective, Caron et al. (2000) consider 2-block warehouses (i.e., one middle cross aisle) under the COI-based storage assignment (see Section 4 for a discussion of storage assignment methods), while Le-Duc and De Koster (2005b) focus on the class-based storage assignment. For both random and volume-based storage assignment methods, Petersen (2002) shows, by using simulation, the effect of the aisle length and number of
aisles on the total travel time. Much of the existing knowledge on warehouse layout is captured in the Erasmus-Logistica website (http://www.fbk.eur.nl/OZ/LOGISTICA) that can be used to interactively optimise warehouse layouts for various storage and routing strategies.

Compared to manual-pick order-picking systems, the layout design problem for unit-load (mainly AS / RS) systems has received much attention. For a literature review on throughput time models for AS / RS, we refer to Sarker and Babu (1995), Johnson and Brandeau (1996), Van den Berg (1999), and Le-Duc (2005). We briefly mention here the literature focussing on designing the picking face. For random storage assignment, Bozer and White (1984) show that a square-in-time rack (i.e. a rack where the ratio of height to length equals the ratio of the S / R machine vertical to horizontal velocity) is optimal for single and dual-command cycles. Larson et al. (1997) use a heuristic approach to layout a unit-load warehouse and to assign product classes to locations, with the objective of increasing floor space utilisation and decreasing travel distance. Eldemir et al. (2004) give estimates for storage requirements. Few papers deal with laying out three-dimensional unit-load systems. Park and Webster (1989) are an exception. They deal with the problem of finding rack locations for product turnover classes to minimise the travel time. De Koster and Le-Duc (2005) extend Bozer and White’s (1984) method to determine the optimal dimensions of a three-dimensional rack of given capacity that minimise the unit-load retrieval time.


4. Storage Assignment

Products need to be put into storage locations before they can be picked to fulfill customer orders. A storage assignment method is a set of rules which can be used to assign products to storage locations. Before such an assignment can be made, however, a decision must be made which pick activities will take place in which storage system.

4.1 Forward – reserve allocation

In order to speed up the pick process, it is in many cases efficient to separate the bulk stock (reserve area) from the pick stock (forward area). The size of the forward area is restricted: the smaller the area, the lower the average travel times of the order pickers will be. It is important to decide how much of each SKU is placed in the forward area and where in the area it has to be located. Figure 2 shows three areas in which a single SKU can be stored and picked, depending on the storage and pick quantity. Dividing a SKU’s inventory over multiple areas implies regular internal replenishments from the reserve to the forward area. One of the trade-offs to be made is then to balance additional replenishment efforts over extra pick effort savings. It may even be advantageous to store some of the SKUs only in the reserve area, for example if demand quantities are high or if demand frequencies are low. Furthermore, replenishments are often restricted to times at which there is no order picking activity, which gives additional constraints. The decisions concerning the problems described here are commonly called the forward-reserve problem. Literature includes Frazelle et al. (1994) Hackman and Platzman (1990), and Van den Berg et al. (1998). In their book *Warehouse science*, Bartholdi and Hackman (2005) devote a full chapter to this problem.

A concept closely related to the forward-reserve problem is dynamic storage. It aims at making the pick area very small in order to reduce travel time, and bringing the SKUs to the storage locations dynamically, just in time for the pick (by an automated crane, carousel, or VLM). The number of locations available in the forward area is usually smaller than the total number of SKUs. As these systems are capable of achieving very high picker productivity, they are becoming more and more popular (according to the authors’ knowledge at least 15 implementations the last few years in Western Europe). The interesting decision problems are in the interaction of the grouping of orders in a batch (more orders means fewer replenishments, but simultaneously more SKUs are
needed implying larger travel distances), the assignment of SKUs to locations, the timing of the replenishments, and scheduling of the automated crane. This area is still virgin ground for academics.

### 4.2 Storage assignment policies

There are numerous ways to assign products to storage locations within the forward and reserve storage areas. We describe five frequently used types of storage assignment: random storage, closest open location storage, dedicated storage, full turnover storage and class based storage. For random storage every incoming pallet (or an amount of similar products) is assigned a location in the warehouse that is selected randomly from all eligible empty locations with equal probability (see e.g. Petersen, 1997). The random assignment method results in a high space utilisation (or low space requirement) at the expense of increased travel distance (Choe and Sharp, 1991). The random storage policy will only work in a computer-controlled environment. If the order pickers can choose the location for storage themselves we would probably get a system known as closest open location storage. The first empty location that is encountered by the employee will be used to store the products. This typically leads to a warehouse where racks are full around the depot and gradually more empty towards the back (if there is excess capacity). Hausman et al. (1976) argue that closest open location storage and random storage have a similar performance if products are moved by full pallets only.

Another possibility is to store each product at a fixed location, which is called dedicated storage. A disadvantage of dedicated storage is that a location is reserved even for products that are out of stock. Moreover, for every product sufficient space has to be reserved such that the maximum inventory level can be stored. Thus the space utilisation is lowest among all storage policies. An advantage is that order pickers become familiar with product locations. In retail warehouses often the product-to-location assignment matches the layout of the stores (De Koster and Neuteboom, 2001). This can save work in the stores because the products are logically grouped. Finally, dedicated storage can be helpful if products have different weights. Heavy products have to be on the bottom of the pallet and light products on top. By storing products in order of weight and routing the order pickers accordingly, a good stacking sequence is obtained without additional effort. Dedicated storage can be applied in pick areas, with a bulk area for replenishment that may have, for example, random storage. In this way, the advantages of dedicated storage still hold, but the disadvantages are only minor because dedicated storage is applied only to a small area.
A fourth storage policy is full-turnover storage. This policy distributes products over the storage area according to their turnover. The products with the highest sales rates are located at the easiest accessible locations, usually near the depot. Slow moving products are located somewhere towards the back of the warehouse. An early storage policy of this type is the cube-per-order index (COI) rule, see Heskett (1963, 1964). The COI of an item is defined as the ratio of the item’s total required space to the number of trips required to satisfy its demand per period. The algorithm consists of locating the items with the lowest COI closest to the depot. See also Kallina and Lynn (1976), Malmborg and Bhaskaran (1987, 1989, 1990) and Malmborg (1995, 1996). A practical implementation of full-turnover policies would be easiest if combined with dedicated storage. The main disadvantage is that demand rates vary constantly and the product assortment changes frequently. Each change would require a new ordering of products in the warehouse resulting in a large amount of reshuffling of stock. A solution might be to carry out the restocking once per period. The loss of flexibility and consequently the loss of efficiency might be substantial when using full-turnover storage. The adoption of COI-based storage assignment, or other assignments based on demand frequency generally require a more ‘information intensive’ approach than random storage, since order and storage data must be processed in order to rank and assign products (Caron et al. 1998). In some cases this information may not be available, for example, because the product assortment changes too fast to build reliable statistics (see De Koster et al., 1999a).

4.3 Class-based storage

The concept of class-based storage combines some of the methods mentioned so far. In inventory control, a classical way for dividing items into classes based on popularity is Pareto’s method. The idea is to group products into classes in such a way that the fastest moving class contains only about 15% of the products stored but contributes to about 85% of the turnover. Each class is then assigned to a dedicated area of the warehouse. Storage within an area is random. Classes are determined by some measure of demand frequency of the products, such as COI or pick volume. Fast moving items are generally called A-items. The next fastest moving category of products is called B-items, and so on. Often the number of classes is restricted to three, although in some cases more classes can give additional gains with respect to travel times.

Based on simulation experimental results, Petersen et al. (2004) show that with regards to the travel distance in a manual order-picking system, full-turnover storage outperforms class-based storage. The gap between the

---

2 It is based on an observation of the Italian sociologist and economist Vilfredo Pareto: “85% of the wealth of the world is held by 15% of the people.”
two depends on the class partition strategy (i.e. number of classes, percentage of the total volume per class) and the routing method used. However, they suggest using the class-based method with 2 to 4 classes in practice as it is easier to implement than the volume-based method; it does not require a complete list of the items ranked by volume and it requires less time to administer than the other dedicated methods do. While for AS / RS, Yang (1988) and Van den Berg and Gademann (2000) found that (in their studies) 6-class is the best among other options. The advantage of this way of storing is that fast-moving products can be stored close to the depot and simultaneously the flexibility and low storage space requirements of random storage are applicable. Graves et al. (1977) observe that in order to enable an incoming load to be stored in the correct class region, empty slots must be available, thus increasing space requirements with the number of classes. Accordingly, class-based storage requires more rack space than randomised storage.

Most research on class-based storage has been performed in the context of AS / RS. Hausman et al. (1976) consider the problem of finding class regions for an AS / RS using the class-based storage assignment method and the single-command operating mode. The authors prove that L-shaped class regions where the boundaries of zones accommodating the corresponding classes are square-in-time are optimal with respect to minimising the mean single-command travel time. They also analytically determine optimal storage class-sizes for two product classes. Rosenblatt and Eynan (1989) extend these results and establish optimal class boundaries for any given number of classes in a square-in-time rack. Eynan and Rosenblatt (1994) extend this method further to any rectangular rack. For S / R machines with dual-command cycles and class-based storage racks, Graves et al. (1977) show by simulation that the L-shaped regions with square-in-time boundaries are not necessarily optimal. However, an L-shaped class allocation will in general be no more than 3% above the optimum. For multi-command cycles with class-based storage, Guenov and Raeside (1992) compare three zone shapes. Ashayeri et al. (2003) develop a travel time calculation method used for designing storage-class shapes in both square-in-time and non-square-in-time racks with single and multiple depots.

Various possibilities exist for positioning the A-, B- and C-areas in low-level picker-to-part systems. Jarvis and McDowell (1991) suggest that each aisle should contain only one class, resulting in the within-aisle storage as depicted in Figure 6. Petersen (1999, 2002), Petersen and Schmenner (1999), Petersen and Aase (2004) and Petersen et al. (2004) compare multiple configurations among which across-aisles storage (also depicted in Figure 6). Roodbergen (2005) compares various
storage assignment policies for warehouse layouts with multiple cross aisles. Based on a closed form travel-time estimate for the return routing policy, Le-Duc and De Koster (2005c) optimise the storage-class positioning. They claim that the across-aisle storage method is close to optimal. Le-Duc (2005) extends these results for other routing policies. The optimal storage strategy depends on the routing policies (and on warehouse size and number of SKUs per pick route). In the warehousing literature, there is no firm rule to define a class partition (number of classes, percentage of items per class, and percentage of the total pick volume per class) for low-level picker-to-part systems.

Figure 6. Illustration of two common ways to implement class-based storage

4.4 Family grouping

All storage assignment policies discussed so far have not entailed possible relations between products. For example, customers may tend to order a certain product together with another product. In this case, it may be interesting to
locate these two products close to each other. An example of this is called family-grouping, where similar products are located in the same region of the storage area. Clearly, grouping of products can be combined with some of the previously mentioned storage policies. For example, it is possible to use class-based storage and simultaneously group related items. However, the decision in which class to locate the products has to depend on a combination of the properties of all products in the group. Roll and Rosenblatt (1983) compare the space requirements for the random and grouped storage for a port warehouse and show that the grouped storage assignment increases space requirements compared to random storage assignment. Rosenblatt and Roll (1984) set up a model for warehousing costs, taking the effect of space requirements into account.

To apply family grouping, the statistical correlation between items (e.g. frequency at which they appear together in an order, see Frazelle and Sharp 1989 and Bryner and Johansson 1996) should be known or at least be predictable. In the literature, two types of family grouping are mentioned. The first method is called the complementary-based method, which contains two major phases. In the first phase, it clusters the items into groups based on a measure of strength of joint demand (‘complementary’). In the second phase, it locates the items within one cluster as close to each other as possible (Wäscher 2004). Rosenwein (1994) shows that the clustering problem can be formulated as a p-median problem. For finding the position of clusters, Liu (1999) suggests that the item type with the largest demand should be assigned to the location closest to the depot (volume-based strategy), while Lee (1992) proposes to take into account also the space requirement (COI-based strategy). The second type of family-grouping method is called the contact-based method. This method is similar to the complementary method, except it uses contact frequencies to cluster items into groups. For a given (optimal) routing solution, a contact frequency between item type i and item type j is defined as the number of times that an order picker picks either item type i directly after item type j, or item type j directly after item type i. However, the routing decision is dependent on the location of the item types, which demonstrates the strong interrelationship between item location and routing. Due to the fact that finding a joint optimal solution for both problems is not a realistic approach, at least not for problem instances of the size encountered in practice, contact-based solution methods alternate between the two problem types (Wäscher 2004). The contact-based method is considered, for example, in Van Oudheusden et al. (1988) and Van Oudheusden and Zhu (1992).
5. Zoning

As an alternative to single order picking, the order picking area can be divided into zones. Each order picker is assigned to pick the part of the order that is in his assigned zone. Compared to other planning issues, the zoning problem has received little attention despite its important impact on the performance of order-picking systems. Possible advantages of zoning include the fact that each order picker only needs to traverse a smaller area, reduced traffic congestion, and furthermore the possibility that order pickers become familiar with the item locations in the zone. The main disadvantage of zoning is that orders are split and must be consolidated again before shipment to the customer. Two approaches can be used to cope with this. The first approach is that of progressive assembly of an order. Using this approach one order picker starts on the order. When he finishes his part, the tote and pick list (or any other means that are used) are handed to the next picker, who continues the assembly of the order. Hence an order (or batch of orders) is only finished after having visited all relevant zones. This system is also called pick-and-pass. The second approach for zoning is parallel (or synchronised) picking, where a number of order pickers start on the same order, each order picker in his own zone. The partial orders are merged after picking. In practice, zoning is partially based on product properties, like size, weight, required temperature and safety requirements.

Little literature on zoning is available. A generic discussion on zoning is given in Speaker (1975). De Koster (1994) models a zoned pick-and-pass system as a Jackson queuing network which allows rapid estimation of order throughput times and average work-in-process. Results are compared with simulations. The estimates can be used to determine the number of zones and the system size. Mellema and Smith (1988) examine the effects of the aisle configuration, stocking policy and batching and zoning rules by using simulation. They suggest that a combination of batching and zoning can significantly increase the productivity (pieces per man-hour). Also, using simulation, Petersen (2002) shows that the zone shape (number of aisles per zone, the aisle lengths), the number of items on the pick-list and the storage policy have a significant effect on the average travel distance within the zone. Choe et al. (1993) study the effects of three strategies in an aisle-based order-picking system: single-order-pick, sort-while-pick, and pick-and-sort. They propose analytical tools for a planner to quickly evaluate various alternatives without using simulation. In Malmborg (1995) the problem of assigning products to locations is studied with zoning constraints. Brynzer and Johansson (1995) describe a case study with zoning and batching. An important issue, particularly in progressive zoning is that the workload must be equally
distributed (balanced) over the order pickers. Analogous to manufacturing flow lines, imbalance can cause serious deterioration of order throughput and order throughput time. Jane (2000) proposes several heuristic algorithms to balance the workloads among the order picker and to adjust the zone size for order volume fluctuation in a progressive zoning system. Jane and Laih (2005) consider the problem of heuristically assigning products to zones in a synchronised system. The method is based on co-appearance of items in the same order (i.e. items appear in the same order are stored in the same zone). Jewkes et al. (2004) tackle the product assignment problem (as well as zone sizing and picker home base location) for a progressive system. Their method is based on dynamic programming. Using a mixed-integer linear program, Le-Duc and De Koster (2005a) determine the optimal number of zones in a synchronised zoning system such that the total order-picking and assembly time is minimised.

An alternative for progressive zoning with fixed zone sizes, would be a more dynamic way of zone sizing and assigning order pickers to zones. The bucket-brigades concept is an example of this. It coordinates workers who are progressively assembling products along a flow line. The idea is roughly as follows. There is one rack from which the products are to be retrieved. One order picker starts an order at the far left of the rack. He picks a number of products and at some point gives the partially fulfilled order to the next order picker, who continues picking the products along the line. The order is handed from picker to picker until it reaches the far right of the line, where it is put on a conveyor for further transport. The special feature of the bucket brigades is the way in which it is determined when an order is handed from one order picker to the next. Suppose, at some point in time all order pickers are working on separate orders, if the order picker closest to the end of the line deposits his finished order, he walks back along the line towards the starting point. If he meets another order picker, he then takes over the order from the other person and continues picking this order. The order picker from which the order was taken moves back along the line until he meets another order picker, and so on. One order picker starts all orders. The order pickers have to be in sequence of their respective speed of working for the system to function adequately. The main advantage of bucket brigades is that they are self-balancing with respect to workload. See Bartholdi (1993), Bartholdi et al. (1999, 2005) and Bartholdi and Eisenstein (1996, 2005a,b). Bartholdi et al. (2001) report the implementation of bucket brigades in a distribution centre and show that bucket brigades increased the throughput rate and reduced management efforts.
6. Batching

When orders are fairly large, each order can be picked individually (i.e. one order per picking tour). This way of picking is often referred as the single order picking policy (or discrete or pick-by-order). However, when orders are small, there is a potential for reducing travel times by picking a set of orders in a single picking tour. Order batching is the method of grouping a set of orders into a number of sub-sets, each of which can then be retrieved by a single picking tour. According to Choe and Sharp (1991), there are basically two criteria for batching: the proximity of pick locations and time windows.

Proximity batching assigns each order to a batch based on proximity of its storage location to those of other orders. The major issue in proximity batching is how to measure the proximities among orders, which implicitly assumes a pick sequencing rule to visit a set of locations. Gademann et al. (2001) consider the proximity order-batching problem in a manual-pick wave-picking warehouse. The objective is to minimise the maximum lead-time of any batch (this is known as a common objective in wave picking). They show that the order-batching in this case is an NP-hard problem. They propose a branch-and-bound algorithm to solve this problem exactly for small instances and a 2-opt heuristic procedure for large instances. Furthermore, they claim that the 2-opt heuristic provides very tight upper bounds and would suffice in practice. Also for a manual picking system, Gademann and Van de Velde (2005) consider the order-batching problem with a more general objective: minimising the total travel time. They show that the problem is still NP-hard in the strong sense when the number of orders per batch is greater than 2. A branch-and-price algorithm is designed to solve instances of modest size to optimality. For larger instances, it is suggested to use an iterated descent approximation algorithm. Chen and Wu (2005) measure the proximity of orders by taking into account the level of “association” between orders (orders having more similar items have a high association and may form a batch). They develop a clustering model based on 0-1 integer programming to maximise the total association of batches. A data-mining approach is given in Chen et al. (2005) and genetic algorithms are employed in Hsu et al. (2005).

As order batching is an NP-hard problem, many studies focus on developing heuristic methods for solving it. For manual picking systems, we can distinguish two types of order-batching heuristics: seed and savings algorithms. Seed algorithms construct batches in two phases: seed selection and order congruency. Seed selection rules define a seed order for each batch. Some examples of a
seed selection rule are: (a) a random order; (b) an order with large number of positions; (c) an order with longest pick tour; (d) an order with most distantly-located (i.e. furthest from the depot); (f) an order with the largest difference between the aisle number of the right-most and the left-most aisle to be visited (see De Koster et al. 1999b for more seed selection rules). Order congruency rules determine which unassigned order should be added next into the current batch.

Usually, an order is selected, to be included in a batch, based on a measure of the ‘distance’ from the order to the seed order of the batch. Examples are: (a) the number of additional aisles which have to be visited if the order is added; (b) the difference between the gravity centre of the order and the gravity centre of the seed order; (c) the sum of the travel distances between every location of an item in the order and the closest location of item in the seed order. Seed algorithms are considered in Elsayed (1981), Elsayed and Stern (1983), Hwang et al. (1988), Hwang and Lee (1988), and Pan and Liu (1995) for single aisle man-on-board AS / RS systems, and in Gibson and Sharp (1992), Rosenwein (1994), Ruben and Jacobs (1999) and De Koster et al. (1999b) for multiple aisle systems. Saving algorithms are based on the algorithm of Clarke and Wright (1964) for the vehicle routing problem: a saving on travel distance is obtained by combining a set of small tours into a smaller set of larger tours. Elsayed and Unal (1989) propose four batching heuristics of which the SL algorithm (combine Small with Large orders), which classifies orders as ‘large’ or ‘small’ ones before assigning them to different batches, generates smallest travel distances.

De Koster et al. (1999b) perform a comparative study for the seed and time savings heuristics mentioned above for multiple-aisle picker-to-parts systems. The performance of the algorithms is evaluated using two different routing heuristics. The batching heuristics are compared for travel time, number of batches formed and also for the applicability in practice. They conclude that: (a) even simple order batching methods lead to significant improvement compared to the first-come first-serve batching rule; (b) the seed algorithms are best in conjunction with the S-shape routing method and a large capacity of the pick device, while the time savings algorithms perform best in conjunction with the largest gap routing method and a small pick-device capacity.

Under *time window batching*, the orders arriving during the same time interval (fixed or variable length), called a time window, are grouped as a batch. These orders are then processed simultaneously in the following stages. If order splitting is not allowed (thus each order picker picks a group of complete orders in one picking tour), it is possible to sort items by order during the picking process. This picking strategy is often referred as the sort-while-pick
picking strategy. If order splitting is possible, a further effort is needed to sort the items after picking (the pick-and-sort strategy). Tang and Chew (1997), Chew and Tang (1999) and Le-Duc and De Koster (2003a, 2003b) consider variable time window order batching (i.e. number of items per batch is ‘fixed’) with stochastic order arrivals for manual picking systems. They model the problem as a batch service. For each possible picking batch size, they first estimate the first and second moments of the service time. Then using these moments, they can find the time in system of a random order. The optimal picking batch size is then determined in a straightforward manner. Results from the simulation experiments show this approach provides a high accuracy level. Furthermore, it is simple and can be easily applied in practice.

All publications mentioned above do not take into account the order due time and the penalty of violating the due time. Elsayed et al. (1993) and Elsayed and Lee (1996) consider the order-batching problem in a man-aboard system with minimising of the penalties and tardiness as respective objectives. They propose a heuristic which first establishes batches and then determines the release times for the batches. Won and Olafsson (2005) focus on customer response times by jointly considering the batching and picking operation.
7. Routing Methods

The objective of routing policies is to sequence the items on the pick list to ensure a good route through the warehouse. The problem of routing order pickers in a warehouse is actually a special case of the Travelling Salesman Problem, see also Lawler et al. (1985). The travelling salesman problem owes its name to the problem described by the following situation. A salesman, starting in his home city, has to visit a number of cities exactly once and return home. He knows the distance between each pair of cities and wants to determine the order in which he has to visit the cities such that the total travelled distance is as small as possible. Clearly, the situation of the travelling salesman has many similarities with that of an order picker in a warehouse. The order picker starts at the depot (home city), where he receives a pick list, has to visit all pick locations (cities) and finally has to return to the depot. An example layout of a warehouse with pick and a corresponding graph representation is given in Figure 7.

Some differences exist between the classical Travelling Salesman Problem and the situation of order picking in warehouses. First of all, if we look at the graph in Figure 7, a number of nodes do not have to be visited (indicated with
white circles). These nodes are the cross points between aisles and cross aisles. The order picker is allowed to visit them, but does not have to. The black circles represent the pick locations and the depot; these nodes must be visited. It is permissible to visit the pick locations and depot more than once. The problem of order picking classifies as a Steiner Travelling Salesman Problem because of the two facts that some of the nodes do not have to be visited and that the other nodes can be visited more than once. The difficulty with the (Steiner) Travelling Salesman Problem is that it is in general not solvable in polynomial time. However, for type of warehouse shown in Figure 7, it was shown by Ratliff and Rosenthal (1983) that there does exist an algorithm that can solve the problem in running time linear in the number of aisles and the number of pick locations.

In Cornuéjols et al. (1985) it is shown that the algorithm of Ratliff and Rosenthal (1983) can be extended to solve the Steiner Traveling Salesman Problem in all, so-called, series-parallel graphs. In De Koster and Van der Poort (1998) and Roodbergen and De Koster (2001) the algorithm by Ratliff and Rosenthal (1983) is extended to different warehouse situations that cannot be represented as series-parallel graphs. The algorithm from De Koster and Van der Poort (1998) can determine shortest order picking routes in a warehouse of one block with decentralised depositing. Decentralised depositing means that order picker can deposit picked items at the head of every aisle, for example on a conveyor. Instructions for the next route are given via a computer terminal. Roodbergen and De Koster (2001b) developed an algorithm for a warehouse with three cross aisles, one in the front, one in the back, and one in the middle.

### 7.1 Routing heuristics

In practice, the problem of routing order pickers in a warehouse is mainly solved by using heuristics. This is due to some disadvantages of optimal routing in practice. Firstly, it must be noted that an optimal algorithm is not available for every layout. Secondly, optimal routes may seem illogical to the order pickers who then, as a result, deviate from the specified routes (Gademann and Van de Velde 2005). Thirdly, a standard optimal algorithm cannot take aisle congestion into account, while with heuristic methods it may be possible to avoid (or at least to reduce) the aisle congestion (i.e. the S-shape method has a single traffic direction if the pick density is sufficiently high). Hall (1993), Petersen (1997) and Roodbergen (2001) distinguish several heuristic methods for routing order pickers in single-block warehouses. Example routes are shown in Figure 8.
One of the simplest heuristics for routing order pickers is the S-shape (or traversal) heuristic. Routing order pickers by using the S-shape method means that any aisle containing at least one pick is traversed entirely (except potentially the last visited aisle). Aisles without picks are not entered. From the last visited aisle, the order picker returns to the depot. Another simple heuristic for routing order pickers is the return method, where an order picker enters and leaves each aisle from the same end. Only aisles with picks are visited. The midpoint method essentially divides the warehouse into two areas (see Figure 8). Picks in the front half are accessed from the front cross aisle and picks in the back half are accessed from the back cross aisle. The order picker traverses to the back half by either the last or the first aisle to be visited. According to Hall (1993), this method performs better than the S-shape method when the number of picks per aisle is small (i.e. one pick per aisle on average).

The largest gap strategy is similar to the midpoint strategy except that an order picker enters an aisle as far as the largest gap within an aisle, instead of the midpoint. The gap represents the separation between any two adjacent picks, between the first pick and the front aisle, or between the last pick and the back aisle. If the largest gap is between two adjacent picks, the order picker performs a return route from both ends of the aisle. Otherwise, a return route from either the front or back aisle is used. The largest gap within an aisle is therefore the portion of the aisle that the order picker does not traverse. The back aisle can only be accessed through either the first or last aisle. The largest gap method always outperforms the midpoint method (see Hall 1993). However, from an implementation point of view, the midpoint method is simpler. For the combined (or composite) heuristic, aisles with picks are either entirely traversed or entered and left at the same end. However, for each visited aisle, the choice is made by using dynamic programming (see Roodbergen and De Koster 2001a).

Petersen (1997) carried out a number of numerical experiments to compare six routing methods: the S-shape, return, largest gap, mid-point, composite and optimal in a situation with random storage. He concludes that a best heuristic solution is on average 5% over the optimal solution. A route improvement method using Lin and Kernighan’s (1973) k-opt methodology is presented by Makris and Giakoumakis (2003).
All above-mentioned methods were originally developed for single-block warehouses, however, they can be used for multiple-block warehouses with some modifications (see Roodbergen and De Koster, 2001a). Methods specifically designed for multiple-block warehouses can be found in Vaughan and Petersen (1999) and Roodbergen and De Koster (2001a). The latter paper compared six
routing methods (optimal, largest gap, S-shape, aisle-by-aisle, combined and combined+), in 80 warehouse instances, with the number of aisles varying between 7 and 15, the number of cross aisle between 2 and 11 and the pick-list size between 10 and 30. They report that the combined+ heuristic gives the best results in 74 of the 80 instances they analysed.

### 7.2 Other routing issues

All articles discussed so far assume that the aisles of the warehouse are narrow enough to allow the order picker to retrieve products from both sides of the aisle without changing position. In Goetschalckx and Ratliff (1988b) a polynomial-time optimal algorithm is developed that solves the problem of routing order pickers in wide aisles. Another problem with routing may arise if products are stored at multiple locations in a warehouse. In this case a choice has to be made from which location the products have to be retrieved. A model for the problem of simultaneous assignment of products to locations and routing of order pickers is given in Daniels et al. (1998). Furthermore, heuristics are given to solve the problem. A further routing problem is that of allowing the order picker to do multiple picks per stop. That is, the order picker travels through a warehouse with a vehicle. He stops the vehicle and walks back and forth to a number of pick locations to retrieve products. Then he continues to the next stop location, and so on. The trade-off is between the time to start and stop the vehicle and the distance walked by the order picker. This problem was analysed and solved optimally in Goetschalckx and Ratliff (1988a).

### 7.3 Analytical estimation of routing time

Part of the research on routing consists of travel time estimation. Using techniques from statistics and operations research an attempt is made to give an estimate of how much time (or distance) it takes to collect an order. Many results are known for systems where the vehicle is confined to a single aisle, see for example Bozer and White (1984). For travel time estimates of single and dual command cycles in multiple aisle systems see e.g. Bassan et al. (1980), Francis (1967), Larson et al. (1997) and Pandit en Palekar (1993). A recent interesting addition is Bozer and Cho (2005) where analytical expressions are presented for an AS/RS under stochastic demand.

Few researchers have looked for travel time estimates for picking in systems with multiple aisles and multiple picks per route. Kunder and Gudehus (1975) give travel time estimations for three routing heuristics in a warehouse.
consisting of one block. This work is extended in Hall (1993) with more advanced routing heuristics for one block warehouses. Furthermore, a lower bound on travel time for the optimal algorithm from Ratliff and Rosenthal (1983) is given. Formulations for average travel time in a situation with decentralised depositing are given in De Koster et al. (1998). From Hall’s analyses it appears that largest gap outperforms S-shape when the pick density is less than about 4 picks per aisles. De Koster and Van der Poort (1998) and De Koster et al. (1998) use simulation to compare the optimal and S-shape methods for several single-block random storage warehouses. They find that the S-shape provides routes which are, on average, between 7% and 33% longer than the optimum solutions. Even though no paper has included this notion, it would actually be fairly simple to prove on the basis of statistical properties that for random storage situations an upper bound for the average length of S-shape routes equals two times the average length of optimal routes.

Both Hall (1993) and Kunder and Gudehus (1975) assume that pick locations are distributed randomly over the order picking area according to a uniform distribution. In Jarvis and McDowell (1991) travel time estimates are determined and used to determine which products (fast moving, slow moving) should be located in which aisles. Le-Duc and De Koster (2004) develop similar travel time estimates for the return heuristic. A travel time analysis for a general product-to-location assignment is given in Chew and Tang (1999) and Tang and Chew (1997). That is, demand rates for products can vary throughout the warehouse. They use the travel time estimates to evaluate batching strategies. Expected travel distances for two routing methods in a warehouse consisting of two blocks are given in Caron et al. (1998). They locate the depot between the two blocks and items are assumed to be distributed according to the cube-per-order index. Hwang et al. (2004) present analytical expressions for three routing methods (return, S-shape, midpoint) under various COI-based storage rules.
8. Order Accumulation and Sorting

When batching and/or zoning is applied, usually some additional effort is needed to split the batch and to consolidate the items per customer order or per destinations to which orders will be shipped. These processes are often called accumulation/sorting (A/S).

Figure 9. A typical accumulation/sorting (A/S) system

Figure 9 shows an example of a typical A/S system (mentioned in Meller, 1997 and Johnson, 1998). Items of a group of orders (a pick-wave) that are to be loaded onto a certain number of trucks are picked from the picking area. In general, items from the same order are assigned to multiple order pickers (to maintain high order picker efficiency) and the order pickers follow pre-specified routes to pick the items assigned to them. After picking, order pickers place their items on the transportation conveyor and the items are transported to the sorter. Owing to the assignment of orders to more than one order picker, the items of each order arrive at the sorter in a random sequence. Items are released onto the circulation conveyor of the sorter and enter the assigned shipping lane if all items of the preceding order assigned to that lane have already entered. If not, the items re-circulate around the circulation conveyor. Orders are released from shipping lanes as needed by the trucks and the lane capacity is made available for the next sort-group. The throughput of an A/S system depends not only on the equipment capacity (i.e., sorter capacity and conveyor speed) but also on operating policies like assignment of orders to shipping lanes (see Figure 9).
The order-to-lane assignment problem is critical for most A/S systems as usually the number of shipping lanes is less than the number of orders, which may cause a blocking of orders at the entrance of the lanes.

The number of publications on A/S systems is limited. By simulation, Bozer and Sharp (1985) examine advantages of using a recirculation loop to avoid lane blocking in an A/S system when a shipping lane is full, assuming that each lane is assigned to one order. Considering A/S systems where multiple orders can be are assigned to one lane, Bozer et al. (1988) and Johnson (1998) recommend that assigning orders to shipping lanes just before the orders arrive at the circulation desk of the sorter is a better than any static fixed-assignment rule. Johnson and Lofgren (1994) describe an A/S system used at Hewlett-Packard. Meller (1997) proposes an integer formulation for the order-to-lane assignment problem in an A/S system. He claims that the problem can be solved efficiently for small instances (in terms of the number of lanes) by solving a number of minimum-cardinality sub-problems. Russell and Meller (2003) present a model to aid in the decision whether or not to automate the sorting process. Le-Duc and De Koster (2005a) present an integer-programming model to minimise the total picking and order accumulation time. Although the general problem is NP-hard, it solves to optimality in reasonable time for a real-life problem.
9. Conclusions

We can draw the following conclusions from the literature. First, in spite of their dominance in practice, pickers-to-parts order-picking systems have received less research attention compared to parts-to-picker order-picking systems. Less than 30 percent of the about 140 papers we considered concerns pickers-to-part order-picking systems. The reasons for this may have something to do with the complexity and diversity of picker-to-parts order-picking systems. Furthermore, parts-to-picker systems are often fully or partly automated, thus catch the attention of researchers.

Second, although the number of publications in the areas of layout, batching, zoning, storage strategies (like forward-reserve allocation, family grouping, and dynamic storage), and accumulation and sorting is still limited, their number is growing. Particularly, the areas of storage assignment and routing appear to have matured the last decade. Few authors address combinations of the decision problems. Yet, this is necessary as there is obvious interdependency in their impact on the order picking objectives. New developments in practice yielding unprecedented picker productivities like dynamic storage and put (order distribution) systems have not yet led to attention from academics.

Third, existing studies in picker-to-parts order-picking systems mainly focus on random storage assignments. Analytical models for optimising dedicated and class-based storage assignment manual-pick order-picking systems are still lacking. Furthermore, storage assignment has an impact on the performance of the routing method. However, this effect seems to be largely neglected in the literature. Instead, many authors focus on random storage assignment to discuss about the performance of routing methods.

Fourth, almost all research in order picking treats demand as given (or known in advance). Certainly, this is not true, especially in fast picking environments (e.g. small orders arrive on line and need to be shipped within a tight time window). These order-picking situations are becoming more and more daily practice, particularly for mail order companies which sell products online. Optimisation problems arising from these order-picking systems, therefore, should be considered as stochastic optimisation problems, not deterministic ones.
Finally, most of the research focuses on a specific order picking situation or decision problem. However, it is not straightforward to apply methods developed for a specific situation to another situation. ‘General’ design procedures and ‘global’ optimisation models for order picking are still lacking.
10. References


Clarke, G. and Wright, W., Scheduling of vehicles from a central depot to a number of delivery points, Operations Research, 1964, 12, 568-581.


De Koster, R., How to assess a warehouse operation in a single tour, Report, RSM Erasmus University, the Netherlands, 2004.


Roodbergen, K.J. and De Koster, R., Routing order-pickers in a warehouse with a middle aisle, European Journal of Operational Research, 2001b, 133, 32-43.


Rosenwein, M.B., An application of cluster analysis to the problem of locating items within a warehouse. IIE Transactions, 1994, 26(1), 101-103.


Yang, M.H., Analysis of optimization of class-based dedicated storage systems, Report, Material Handling Research Center, Georgia Institute of Technology, Atlanta, Georgia, 1988.

Online sources

Warehouse automation: developments in practice and in science

De Koster, R. (2013), Magazijnautomatisering, in: R. Jansen and A. van Goor (Eds.), 40 jaar logistiek, 40 jaar VLM, pp. 41-44.
The first automated high-bay warehouses were introduced some 50 years ago. Since then, developments have continued at a rapid pace. Initially, automation was mainly focused on pallet warehouses with bulk storage facilities. A major reason was to increase the storage density, which could be achieved by making the warehouses higher. Later, mini-load warehouses and order picking warehouses were also automated. In this chapter we will discuss the different types of automated systems as well as a number of scientific results that are now known about such systems. We will first discuss storage systems for unit loads (bins and pallets). This will be followed by order picking systems from which individual packages can be picked. Finally, we will provide our future expectations of warehouse automation.
1. Systems or Unit Loads

1.1 Automated unit load warehouses with aisle-captive cranes.

These warehouses have already been around since the sixties. Many variants have been developed since then. Figure 1 shows an example of such a warehouse.

Figure 1. Automated high-bay warehouse for pallets with aisle-captive cranes.
Source: Daifuku America

In such a warehouse (AS/RS: Automated Storage and Retrieval System), an aisle-captive crane picks up a load needing storage, usually from a conveyor at the pick-up and delivery point, and automatically stores it in the racks at heights of up to 30 m. Driving and lifting take place simultaneously. Retrieval is exactly the opposite. It is also possible to carry out so-called dual command cycles, wherein a storage command is combined with a retrieval command, which saves one movement per dual command cycle. If the stored unit loads consist of bins, the system is also referred to as mini-load. Many studies have been conducted into such systems. One of the first scientific articles was written by Bozer and White and dates from 1984. Among other things, they calculated the cycle time of the crane for single command cycles, for the situation in which any location within the rack is equally likely for the crane to travel to.

The average cycle time for a single command cycle (back and forth) is

\[ E[T] = \left(1 + \left( \frac{t_y}{3} \right)^3 \right) t_x, \]

whereby \( t_x \) 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. It is assumed that the travel time to the farthest location is longer than the lifting time to the highest location and the fact that the crane drives and lifts simultaneously has
been taken into account. So if the travelling distance of an aisle is 100 sec and the lifting height is 100 sec, the average travel time for a single command cycle (without picking up and dropping off the load) is 133 sec. This formula can be applied to obtain all kinds of partial results, for example, to calculate the optimal ratio between the length and height of an aisle. This proves to be square in time, meaning the travel time to the farthest location and the lifting time to the highest location are identical. Since a crane travels approximately four times faster than it lifts, the aisle should be four times longer than high to achieve a minimal cycle time. The formula can also be used to determine the optimal number of aisles and cranes, including optimal dimensions, at a given storage capacity and throughput.

Over time, the formula has been adjusted to match different storage strategies (such as ABC storage), dual command cycles, different location of the depot: the pick-up and delivery point (the above formula assumes one such point, at the head of the rack), multiple load handling devices, combination with order picking stations, etc. In the case of ABC storage, the items are divided into classes (e.g., 3: A, B, C), based on turnover rate. The locations are also divided into groups based on distance to the pick-up and delivery point. This ensures that the items from the class with the highest turnover rate are located closest to that point.

Figure 2. Overview of AS / RSs (AS / R systems). Source: Roodbergen and Vis (2009)

Figure 2 provides an overview of AS / R systems, based on the type of crane, the type of loads and the types of racks that may be used. The overview also includes carousels and mobile racks, which are not usually considered part of the classical AS / R systems.
An important impetus for cycle time calculations with ABC storage dates from 1976, by Haussman, Schwarz and Graves. They calculated the optimal class boundaries for known ABC demand curves, for example, 20/70 demand curves, whereby 20% of the products are responsible for 70% of the demand. In the calculation they took account of product replenishment according to a continuous review \(<s, Q>\) ordering policy, with the order quantity \(Q\) being equal to the optimal order quantity. However, they did not take account of the fact that the fewer storage classes there are, the more items are stored in the classes, and the less space is required per item, since the space within the classes can be shared by the items. If there are more classes and fewer products are stored per class, it is more complicated to share the space between products and more space is required per product. This means that an optimum number of storage classes can be distinguished. In practice, the optimal number of classes is small (about 3 to 5) 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 the class. However, the required number of locations on top of the average stock level quickly amounts to an additional 30% (Yu et al., 2015).

1.2 Compact storage

Cranes equipped with satellites

The 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. Often satellites are used if even deeper storage is required. Cranes (or trucks) equipped with satellites have been on the market for several decades and are widely used in refrigerated and frozen storage, where savings in the space to be cooled immediately reduce costs. Other forms of multi-deep storage are racks containing mounted conveyors (flow or driven conveyors), or multi-deep racks with multiple, independent satellites. The latter systems are also referred to as AVS/R (#automated vehicle-based storage and retrieval) systems, and are discussed below. For crane systems with multi-deep storage, the first question is how deep the storage locations should be (number of unit loads behind each other). racks with locations that are too deep have a low location occupancy rate. This effect is also called “honeycomb” (see Tompkins et al., 2010). Insufficiently deep locations require too many locations per product, as a result of which the aisles will be too long, thus increasing the travel time of the crane. In general, all the loads of one product and batch are stored in one location.
A lot of research has been conducted into such aisle-captive cranes with multi-deep storage. It is known, for instance, how to calculate the optimal location depth, how to calculate the crane travel time, what the optimal storage strategies are and in which order the commands should be performed in order to minimise the total time.

Storage on shuttles

Recently, new multi-deep systems have been developed, whereby each load is placed on a shuttle that can move in both directions. Such systems are used in parking garages in East Asia, for example, on locations where parking is expensive. Figure 3 shows an example of the use of such shuttles. The advantage of shuttle-based storage is that multiple shuttles can be moved at the same time, thus achieving a high throughput. In addition, it saves a lot of space, since no transport aisles are necessary. It is known how to calculate the cycle time, what the effect is of the class-bound storage and what the optimal length, width and height ratio of these systems is (Zaerpour, 2013). If the lift is located in a corner of the system, a cube-shaped warehouse (measured in time) almost (but not quite) minimises the cycle time.

Figure 3. An automated, compact parking garage. Source: avgparking.com
1.3 **AVS / R systems**

AVS / R ("autonomous vehicle-based storage and retrieval") systems do not use cranes, but shuttles, which can drive in the x and the y-direction on any level in the aisle, and lifts that can move shuttles (or unit loads) between the levels. Such systems are increasingly popular because nowadays, the investment is similar to that of AS / R systems, while they offer a much higher retrieval capacity and are significantly more flexible in capacity. By using additional shuttles the capacity can easily be increased, and by removing shuttles, capacity can be decreased. Figure 4 shows an example of such a system.

![AVS / R system diagram](image)

An increasing number of studies is being conducted into these systems. The performance (throughput, average throughput time of an order) can be estimated reasonably well as a function of the number of vehicles (see Roy, 2011). This can be used to calculate the optimum length to depth ratio of the racks, or an adequate distribution of the shuttles across the system. Meanwhile, shuttle-based systems for compact (multi-deep) storage have also been developed (see Tappia et al., 2015).
2. Order Picking Systems

In addition to systems for unit loads, an increasing number of systems for the (semi) automatic picking of goods from racks is being developed. A-frames, fully automatic machines capable of placing products in a passing bin, have been in use in the pharmaceutical industry for some time now. The machines can achieve a very high picking capacity for products that are suitable for automatic dispensing. This section consecutively discusses automated replenishment systems, Kiva systems and systems for automatic (roll)container loading.

2.1 Automated replenishment systems

In the Netherlands, there are several dozen warehouses equipped with automatic replenishment of the pick locations. Products are stored in bins in a mini-load warehouse, which complements a flow rack warehouse. Figure 5 shows a picking station with flow racks, whereby the locations are automatically replenished by a mini-load system.

Only fast-moving products are stored permanently in the picking stations. The other products are delivered just-in-time by the mini-load system, in which empty bins and items that are not needed for the next batch are retrieved by the mini-load crane. Some studies have already been conducted for such systems, especially regarding good system designs (layouts, required number of picking
stations), depending on the necessary storage capacity and the order profile. The design is especially complicated when picking stations are connected in series and order bins must visit several stations. As the number of picking stations increases, the number of stored products per station decreases as well as the size of the station, thus decreasing the walking distance. This results in fewer, but also faster, picks per bin per station. On the other hand, a bin should visit several stations and probably wait more often or longer in total. This trade-off results in an optimum number of stations. The allocation of products to stations also affects the performance. Research into this has led to models that can quickly and accurately calculate and even optimise different configurations. See Van der Gaast et al. (2013).

2.2 Robot system

A new development are systems with mobile robots that take products to the picker. There are already many different systems available on the market. A recently introduced robotic system is the AutoStore of Swisslog, in which mobile robots move above the rack and take product bins from the rack slots to the picking stations.

Most remarkable of all, however, are the Kiva systems, whereby complete racks ("pods") are moved by robots. If a product is requested by a customer, a robot moves to the rack in which the product is located, picks it up and takes it to a picking station. There the robot awaits its turn with the rack. The picker takes the requested products and adds them to the bin with the customer order waiting at a different rack, also on top of a robot. The robot then returns the rack with the product to storage, at a location that takes account of the expected moment the item will be needed again. The storage location is therefore fully dynamic. In principle, the layout can be fully adapted both dynamically and automatically to the product and order characteristics. The performance of Kiva systems has hardly been studied scientifically. The systems are ideally suited for Internet retailers who require the picking of relatively small orders (meaning not too many items that should be consolidated per order) from a wide product range. This explains why Amazon recently acquired Kiva Systems and wants to equip its new warehouses with this system. Meanwhile, other providers have also entered the market with mobile racks in combination with robots.
2.3 **Automatic loading of roll containers**

Retail warehouses have been experimenting with the automatic picking and loading of store orders in roll containers for years. In the Netherlands, Albert Heijn in Pijnacker was one of the first retail warehouses with fully automated order picking: the triple-O system. Products received on pallets are stored in an automated bulk warehouse. Pallets are then automatically retrieved, the units are offloaded from the pallets and the packages are stored per product on a conveyor belt in a storage aisle, from where they are automatically ejected on a belt which takes them to a roll container loading station. At Albert Heijn, the roll containers are still loaded manually, however, several manufacturers offer solutions to perform this process automatically. The first fully automated retail warehouse is possibly the Edeka warehouse in Hamm (Germany), where the roll containers are also loaded automatically. This requires accurate sequencing of the products in the loading of the roll container, which takes account of family grouping in the store, stackability of products and the loading degree of the roll containers. Many studies have also been conducted in these areas, the results of which can be found in the stacking algorithms. Meanwhile, several retailers are working on largely automating their warehouses.
3. The Future

Is the full automation of both storage and picking the future? The main advantages of automation are mainly the saving of space (an automated warehouse can be built on a smaller area), savings on labour costs (a 24/7 operation can be achieved relatively easily and inexpensively), 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. Automation of storage and order picking, however, still has limitations: it requires considerable scale and a long-term vision (investments can only be earned back in the medium term). Furthermore, a part of the process still needs to be carried out manually. This part of the process is usually not the most interesting work. Especially not if a person is considered an extension of the automation solution. In short, manual warehouses will continue to exist for the time being, despite the new developments, even in economies with high labour costs. Logistics remains a people business and well-managed warehouses perform significantly better in the areas of productivity, process innovation, quality and safety (see De Koster et al., 2011, and other literature).
4. References


Warehouse math

De Koster, M.B.M. (2010), Warehouse Math, in:
L. Kroon, T. Li, R. Zuidwijk (eds.), Liber amicorum.
In memoriam Jo van Nunen, pp. 179-186, Dinalog Breda.
Abstract

In this chapter I show that elementary math models reveal powerful insights that can help understand better design and management decisions in facility logistics operations. Subsequently I develop models for designing a truck-operated unit-load warehouse (sizing it, calculating the load orientation, and determining the warehouse length-width ratio). Next the optimal stacking depth is calculated and for automated unit-load warehouses the optimum rack size, number of cranes and optimum storage zone boundaries are determined. Finally the impact of the number of inducts on sorting system throughput is determined.
1. Introduction

Warehouses and distribution centres play an important role in the supply chains of companies as they decouple supply from demand. Moreover, warehouses have a role in forming an assortment and as the prime point in the supply chain to group material flows to create transport efficiencies. Companies have moved their push-pull boundaries as far downstream as possible, implying many warehouses are involved in value-adding activities. According to the report of HIDC and TNO (2009), warehousing activities (including value addition) directly contributes € 15.7 billion annually to the Dutch GDP. Indirect services (like material handling systems, consultancy, leasing services) are not included in that number.

In an attempt to minimize operational cost many companies operating warehouses have invested in largely automated warehouse operations. Storage of unit loads was already automated in the sixties with automated storage and retrieval systems (AS / RS), usually employing high-bay racks with rail-mounted cranes. Automation of the order picking process of individual selling units is known since about 15 years in sectors with limited variety in product shapes, such as pharmaceuticals. In the last decade it became possible to automatically store, replenish and order pick dry groceries from a fairly large and complicated assortment. The warehouses of Edeka and Lidl in Germany are the first in the grocery retail sector to receive, store, order pick, sequence products, load roll containers and prepare loads for shipments in largely robotized warehouses. Scarcity of labour, 24-hour operations, an urge to work in relatively smaller (and hence cheaper) facilities, sufficiently large shipping volumes per facility, and cheaper technologies have driven these developments.

In order to design, plan, and operate such facilities, models are needed. Software algorithms taking care of order grouping and release, storage slot allocation, pick location allocation, timing of replenishments between reserve and forward storage locations, balancing work loads over pick stations, planning shipments, sequencing pick instructions, and routing order pickers are used to realize the desired performance. These algorithms are the fruits of academic research. Although models for certain configurations can easily become large and intractable, it is possible to derive simple models that yield much insight in system performance. As examples I will treat design models focusing on unit load warehouses. Unit loads are pallets, totes, roll cages, or any other product carrier handled as a single unit. Unit-load storage systems can be found in many warehouses. They are used to store inbound materials, which are later moved to
forward areas or pick stations where the orders for single cases or units are picked. Figure 1 sketches product flows in a typical warehouse, where the “reserve storage” area represents the unit-load warehouse.

Figure 1. Typical warehouse functions and flows (Tompkins et al. 2003)
In the next sections the following warehouse-design questions are covered subsequently, using elementary models.

Rack-based unit-load warehouse design:
• How many unit loads should the warehouse cater for?
• Should unit loads be stored deep or wide?
• What is the optimum size (length / width ratio) of a unit-load warehouse?

Block-stack based unit load warehouse design:
• What is the optimum storage depth in a multi-deep block stack?

AS / RS-based unit-load warehouse design:
• What is the optimum size (length / height ratio) of a unit-load storage rack in an AS / RS?
• How many storage / retrieval cranes are needed to achieve a certain throughput?
• What are the optimum zone boundary shapes for turnover-class based storage?

Sorting system design:
• How many inducts should the sorter have?
2. Rack-based Unit Load Warehouse Design

The first question asks how many loads should be stored in the warehouse. Only then can the warehouse and the racks be sized. We treat these two topics in the following subsections.

2.1 Determining the number of unit loads to be stored

The number of loads the storage system should be able to cater for depends on the following factors:

- The number of products, \( N \), stored (average and maximum simultaneously). Every product, even when stored in a small quantity needs a separate storage slot, capable of storing the unit load. In determining \( N \) not only current products should be included, but also the assortment that should be catered for in the medium-term future.

- The number of slots needed per product.

- The size and weight of the products: how many product units fit on the unit load?

- The amount of slack the storage space should have. Every warehouse needs slack space to warrant a smooth operation. According to a rule of thumb explained by Tompkins et al. (2003, p.403), the slack space \( \beta \) should be around 20%. The slack space depends on the technology used. In more automated warehouses less slack space is needed.

- The used storage strategy. In case a dedicated storage strategy is followed every product, or even stronger, every unit load of a product, has a fixed storage location, which usually is based on turnover speed of the load. In case of a shared storage policy unit loads can share storage slots, albeit not at the same point in time. In practice combinations are often used, for example a fixed slot for the first load and shared slots for the remaining loads of all products.

The number of loads to be stored per product depends on its replenishment policy. Assuming a continuous review \( \langle s, Q \rangle \) ordering policy, with \( s \) the reorder level and \( Q \) the economic order quantity, the reorder level \( s \) will have to cover the mean demand over lead time.

In formula, \( s = \mu_d \cdot L + Z_\alpha \sigma_{dl} \) where \( \mu_d \) is the product’s mean demand per day, \( L \) is the (stochastic) supply lead time in days, \( Z_\alpha \) is the safety factor for a service level (fill rate) \( 1-\alpha \) of the product. The term \( Z_\alpha \sigma_{dl} \) is called the safety stock, \( ss \). In case lead time is constant and demand is iid, then the standard
deviation of demand over lead time, \( \sigma_{dl} \), equals \( \sigma_d \sqrt{L} \), with \( \sigma_d \) the standard deviation of daily demand. The safety factor can be determined if the product’s demand distribution is known. The economic order quantity \( Q \) equals \( \sqrt{2DK} \), where \( D \) is the annual demand and \( K \) is the ratio of ordering over annual unit inventory holding cost. In an \( <s, Q> \) system, it is straightforward to see the average stock of a product equals \( Q/2 + ss \) and the maximum stock equals \( Q + ss \).

The total number of storage slots \( C \) needed will therefore fall between

\[
\sum_{i=1}^{N} \left[ \frac{Q_i}{2} + ss_i \right] (1 + \beta) \quad \text{and} \quad \sum_{i=1}^{N} \left[ Q_i + ss_i \right] (1 + \beta),
\]

where all quantities are expressed in unit loads and \( \beta \) is the system’s slack factor. The notation \( \lceil x \rceil \) indicates \( x \) should be rounded upwards (to full unit loads). Although this analysis partly answers the first question, some intricate further questions remain. The gap between the lower and upper bound may be substantial: think of a warehouse for 1,000 products where all products are replenished in lots of 2 pallets, and the average space needed is 1 pallet per product. Should the warehouse accommodate for 1,200 or 2,400 pallets? As in most practical environments some degree of storage slot sharing is possible, the number of slots needed is usually closer to the lower than to the upper bound. Recently, in a paper by Yu at el. (2015), it has been shown that, with three storage classes, the needed number of locations is about 30% above the average.

### 2.2 Determining the unit load storage orientation

The footprint of most unit loads is not square, but rectangular, meaning they have a long and short side. Euro pallets and block pallets have a size of 120 \( \times \) 80 and 120 \( \times \) 100 cm, respectively. Euro-sized totes have a footprint of 60 \( \times \) 40 cm. As most loads can be retrieved from any direction, an important design decision is whether loads have to be stored wide or deep. In conventional unit load warehouses the building, with fixed installations, may take up to 70% of the total operational costs (rent, interests, write-offs, maintenance, energy; De Koster 1996). The most important criterion for load orientation therefore is space consumption. Figure 2 sketches the space occupied by two loads stored on opposite sides of a travel aisle, stored wide or deep. In the calculation we assume the loads are Euro pallets, and the truck we use is a pallet stacker (a small type of reach truck, capable of lifting about 4 m high) which needs a turning circle of 2.1 m.
In case the loads are stored wide, the aisle width + two pallet depths + safety space equals $b = 2.1 + 2 \times (0.8 + 0.2) = 4.1 \text{ m}$. Therefore, the space occupied by two pallets plus aisle, $O_w$, becomes $b \times l = 4.1 \times (1.2 + 0.2) = 5.7 \text{ m}^2$. In case loads are stored deep, the net aisle width becomes 2.5 m, since the pallet has now become 0.4 m longer, leading to a space of $O_d = 5.3 \times (0.8 + 0.2) = 5.3 \text{ m}^2$, implying a space saving of 7%. This explains why loads in unit load facilities should always be stored deep; if not, this calls for a serious inquiry.

### 2.3 Determining the length and width of the warehouse

Figure 3 shows two different layouts of the warehouse, the left one has a single input / output (I / O) point whereas the right one has split input and output points. We aim to determine the length $L$ and the width $B$ of the warehouse, under the assumption the storage capacity $C$ is given (from Equation (1)). We assume the storage strategy employed in the warehouse is random (i.e. each location is equally likely to contain the requested unit load), storage and retrieval requests arrive on line and the warehouse truck used to store and retrieve unit loads carries out single cycles only. In case of a storage job, such a single cycle consists of fetching the load at position $X$, storing it at some storage location $Y$ and returning empty to location $X$. The cycle is similar in case of a retrieval. We furthermore assume for now the number of storage levels $n$ is given, hence we divide $C$ by $n$ to obtain the capacity per level. In optimizing $L$ and $B$, we have to realize warehouse trucks do not drive and lift simultaneously for safety reasons, i.e. they always – should – drive with the forks in the lowest position. Upon arrival at the storage location, the truck makes a 90° curve and starts lifting until it reaches the desired slot and stores or retrieves the load. After this it lowers its
forks and travels to the output point. Therefore, the only element of the storage/retrieval cycle influenced by $B$ and $L$ is the travel time. In both layouts, the mean travel time to carry out a single cycle equals (approximately) $B + \frac{L}{2}$. This is to be minimized under the condition $L \cdot B \geq C$.

Substituting the constraint into the objective, and applying first order conditions yields $L=2B$. In other words the warehouse should be twice as wide as deep. If the number of unit loads to be stored $P$ is given, it is now straightforward to calculate the real warehouse dimensions, since $P = 2n \frac{L}{b}$.

Note that $\frac{B}{b}$ represents the number of aisles in the warehouse and $\frac{L}{l}$ represents the number of storage sections per aisle side.

Combining this with $L=2B$ yields: $B = \frac{1}{2} \sqrt{\frac{Pbl}{n}}$, $L = \sqrt{\frac{Pbl}{n}}$. (2)

In determining the real $L$ and $B$ values from (2) we have to round the values found such that $\frac{B}{b}$ and $\frac{L}{l}$ are integers. In practice this is not a great restriction as the cycle travel time is greatly insensitive to the exact warehouse shape ratio, as shown in Figure 4. This may explain why so many warehouses in practice have suboptimal shape factors. I suspect though there may be other explanations.
The remaining question is how to determine the number of storage levels $n$. The number of levels is primarily determined by the building costs and the number of trucks and people needed. In general driving a truck is much speedier than lowering or lifting, implying cycle times go up for higher buildings. Therefore, for a given throughput we might need more trucks and people in a high building. However, high buildings are much cheaper than low buildings of the same cube. In view of the dominant part of building-related costs within the operational costs higher buildings are usually to be preferred. However, in order to find the exact trade-off curve and determine the optimal value of $n$, the precise real costs should be calculated as function of $n$. An impression on how to do this can be found in De Koster (1996).
3. Block-stack Based Unit Load Warehouse Design

Many warehouses store bulk unit loads in block stacks, drive-in or drive-through racks. Such block stacks are particularly attractive if few products need to be stored in large quantities per product. Think of beer and soft-drink manufacturers or retail warehouses storing such products. Block stacks are usually arranged around transport aisles and have a fixed depth (the lane length). A major question in such stacks is to determine the proper storage lane length. If lanes are too deep, the average space usage will be very poor (this effect is called ‘honeycombing’, see Tompkins et al. 2003). However, if lanes are short, we may need multiple lanes per product which increases warehouse size. The main objective is again to minimize the space usage to store all unit loads needed. Figure 5 sketches a block stack with 3-deep lanes accessible from two sides (like using drive-through racks). Our prime objective is to determine the number of unit loads, $x$, to be stacked behind each other in a lane. We assume product $i$ is received in the warehouse in batches of $q_i$ unit loads.

Figure 5. Top view of a 3-deep block stack lane, with access on two sides, through travel aisles of width $a$. The load has size $d \times l$ (including safety space)

For a given lane of depth $x$ (in number of unit loads) the average total space occupied by all products plus half of the two aisles equals $(dx + a) \cdot l \sum_{i=1}^{N} \left[ \frac{q_i}{2x} + \frac{1}{2} \right]$, which approximately equals $(dx + a) \cdot l \sum_{i=1}^{N} \left[ \frac{q_i}{2x} \right]$.

Applying first order conditions yields $x = \sqrt{\frac{a}{dN} \sum_{i=1}^{N} q_i}$.

If products can be stacked on top of each other (say $n_i$ levels for product $i$), we have to divide $q_i$ by $n_i$ in this formula. If access is possible from one side only (like in the case of drive-in racks) we have to divide $a$ by 2.
4. AS / RS-based Unit-load Warehouse Design

In an AS / RS (like a miniload, or high-bay pallet warehouse) each storage aisle has its own aisle-captive crane. We first calculate the cycle time (time to store or retrieve a load) of the crane assuming a random storage strategy and use this to determine the optimum size (length / height ratio) of a unit-load storage rack in an AS / RS. From this we can determine the number of cranes (and aisles) and the we can find the optimum zone boundary shapes.

4.1 Calculating the optimum rack shape

In order to calculate the optimum rack shape we have to realize a crane drives and lifts simultaneously. The crane speeds are $v_x$ and $v_y$, in horizontal and vertical direction, respectively. The driving time to the farthest location is now $t_x = \frac{L}{v_x}$; the lifting time to the highest location $t_y = \frac{H}{v_y}$ (see Figure 6).

We aim to find $E[W]$, where $W$ is the driving time to a random location $(X,Y)$, and back to the depot.

Figure 6. Side view of the rack, with the depot in the lower left corner. Right part: normalized rack

In order to simplify the calculations, following Tompkins et al. (2003), we define $T = \max\{t_x, t_y\}$, $b = \min\{\frac{t_x}{T}, \frac{t_y}{T}\}$.

We assume the rack is longer (in time) than it is high (i.e. $t_x \geq t_y$), and we normalize the rack by dividing the time dimensions by $T$. In the normalized rack (see Figure 6), the horizontal storage location $X$ and vertical storage location $Y$ are now stochastic variables uniformly distributed on $[0,1]$ and $[0,b]$, respectively.

Let $Z = \max\{X,Y\}$, then $E[W] = 2E[Z] \cdot T$. 
Since the crane drives and lifts simultaneously, 
\[ F_z(z) = P[Z \leq z] = P[\max\{X, Y\} \leq z] = P[X \leq z] \cdot P[Y \leq z], \]
with \( P[X \leq z] = \begin{cases} z, & \text{if } 0 \leq z \leq 1 \\ 1, & \text{if } z > 1 \end{cases} \) and \( P[Y \leq z] = \begin{cases} \frac{z}{b}, & \text{if } 0 \leq z \leq b \\ 1, & \text{if } z > b \end{cases} \).

In other words, \( F_z(z) = \begin{cases} \frac{z^2}{b}, & \text{if } 0 \leq z \leq b \\ z, & \text{if } b < z \leq 1 \\ 1, & \text{if } z > 1 \end{cases} \).

By taking the derivate of \( F_z \), we find the probability density function \( f_z \), and from this we can obtain the expected value of \( Z \) by calculating \( E[Z] = \int_{b}^{1} z f_z(z) dz \).

From \( E[Z] \) we find \( E[W] = (1 + \frac{b^2}{3})T \). \( \quad (3) \)

As an example, assume the rack is 50 sec long and 50 sec tall, then \( b = 1 \) and \( T = 50 \). The driving time to a random location and back to the depot is therefore 66.7 sec (and not 50 sec!). We can now use formula (3) to optimize the rack shape as follows:

\[
\text{min}! E[W] = (1 + \frac{(t_x / t_y)^2}{3}) T, \quad \text{s.t.} \quad t_x \cdot t_y = C.
\]

We can eliminate one of the variables by substituting the constraint into the objective. We then apply first order conditions yielding \( t_x = t_y (= \sqrt{C}) \). In other words, \( b = 1 \); the optimum rack is square in time. However, since cranes usually drive much faster than they lift (in a ratio of about 4 to 1), optimum racks should be rectangular with this same ratio (of 4 to 1). It is fairly easy to see from the outside of a high-bay warehouse whether it (approximately) has this optimal shape. The building should be about 4 times longer than the height.

### 4.2 Calculating the number of cranes

Cranes are very expensive. Therefore we use formula (3) to calculate the minimum number of cranes needed, for a given storage capacity (using formula (i)) and throughput, expressed in number of unit loads to be stored and retrieved per time unit. As an example assume we need 4000sec\(^2\) of storage space and a throughput of 200 pallets per hour. The problem can be formulated as: \( \text{min}! N \), such that \( 2N \cdot t_x \cdot t_y = 4000 \), and \( \frac{N \cdot 3600}{(1 + \frac{(t_x / t_y)^2}{3}) T} \geq 200 \).
\( \delta \) is the – constant – load pick-up or drop-off time, about 18 sec in pallet warehouses. This problem can easily be solved iteratively by starting with 1 crane, calculating the resulting (square) rack size and then calculating the throughput using formula (3). If the throughput constraint is met we are done, otherwise the number of cranes (and aisles) is increased by 1, etcetera. For the example given, the optimum number of cranes and aisles is 4, leading (for \( \delta = 18 \)) to a throughput of 218.8 pallets / hour.

4.3 The optimum storage-zone shape

Many warehouses do not apply random, but class-based storage. The idea is to divide the products in turnover-based classes, with the fast-moving products (A-products) stored close to the depot and slower-moving products farther away. Within a storage class, products are stored randomly. The question is what the optimum storage zone looks like. It is fairly easy to see the optimum shape of a storage zone is a square \( L \)-shape in the time dimension, as sketched in Figure 6 (right part). Since the crane drives and lifts simultaneously, all points on the boundary of such an \( L \)-shape have exactly the same travel time. Therefore they must belong to the same travel time class. In conclusion, the optimum zone shapes are rectangular in meters, in a ratio of about 4 to 1. Although this might seem fairly obvious I had a recent visit with students to a high-tech fully automated carpet storage warehouse. The carpet rolls were stored vertically, retrieved by an overhead crane moving in \( x \)- and \( y \)-direction simultaneously. The zone boundaries were not square in time, but circular around the depot.
5. The Impact of the Number of Sorter Inducts on Throughput

Many warehouses pick orders in batch and sort them per customer order using sorting machines. Figure 1 gives an impression of the position of this process in the total product flow. Sorting machines have strongly improved over the last decades; they have become faster and cheaper, more ergonomic, less noisy, and more generally useable. Figure 7 shows pictures of a loop sorter, where products can circulate if the destination chute is full.

The sorter has line capacity $C$ (products per time unit) at any intersection. However, the number of products actively sorted by the sorter can be larger than $C$, depending on the number of inducts the sorter has. Assume the sorter has $N$ inducts, equally spaced around the sorter, with the sorting chutes equally divided along the sorter. The maximum inbound flow at a sorter induct $X$, is now limited by the products on the sorter fed by other inducts, but not yet sorted. For two inducts with flows $X$ and $Y$ we find: $X + \frac{1}{2} Y = C$ and $Y + \frac{1}{2} X = C$, implying $X = Y$. Therefore $X = \frac{2}{3} C$.

The total throughput capacity of the sorter is therefore $2X = \frac{4C}{3}$, or 33% more than the line capacity, achieved with one induct. In general, with $N$ inducts we find $\frac{1}{N} \sum_{i=1}^{N} i = C$, or $X = \frac{2}{N+1} C$. Hence the throughput is $NX = \frac{2N}{N+1} C$. 

![Loop sorter sorting to two sides with one (left) and two inducts (right). Source: Vanderlande](image)
Note that for large $N$, we can nearly double the sorter’s line capacity! This also works in practice. During a company tour at a discount retailer, it was mentioned the sorter was a bottleneck for further store expansion. The sorter had three inducts clustered together, effectively acting as one induct. After making a remark on this, I noticed the next year the sorter had not been upgraded, but a new induct had been bought, placed diagonally opposite the first induct group. This had substantially increased sorting capacity.
6. Conclusions

The above examples show elementary math models can be used to provide insight into warehouse design questions. Many extensions of the above-mentioned models exist, albeit often with more complicated analyses.
7. References


Self-storage warehousing

De Koster, R. (2013),
Boosting revenues in the self-storage warehouse industry,
RSM Insight 16, 8-10.
“...self-storage operators should seriously consider taking steps to maximise expected revenue at a stable cost-level.”

Self-storage is a booming industry, one that has seen remarkable growth in the United States and Europe, and across the rest of the world. While an obvious success story, new research shows that by taking a unique facility design approach, warehouse operators can do much to maximise their revenue potential.

There is good reason to be optimistic about the global self-storage warehouse business. In the United States, this type of warehousing generated some US$22 billion annually, with one in ten households renting self-storage space. Between 2007 and 2008, European countries reported an increase of between 19 and 117 per cent in self-storage warehouses. And in the rest of the world, this business is also growing rapidly.

The self-storage business focuses on providing private individuals, households and small businesses with public, temporary storage at centrally located facilities. Importantly, with the self-storage customers handle storage operations themselves and without the intervention of warehouse personnel. A typical self-storage warehouse contains storage spaces of varying sizes and qualities. A customer rents a storage unit of an appropriate size for one or more months. Such a facility is of course ideal for university students, for instance, who need a temporary place to store their belongings during their summer break (when they are required to vacate their dormitories until the start of the next scholastic year).
1. Designing for Maximum Revenue

With decades of management experiences in cost-control, this industry has been able to manage costs well. For example, large self-storage facilities in the USA only employ an average of 3.5 employees per facility, where traditionally in warehouses labour is the largest cost-component.

However, there are good reasons for improvement. For instance, the existing storage sizes or the number of storage units available per size may not fulfil the needs of the market: certain storage types may be scarce, while others plentiful. This results in either lost customers and revenue, or inefficient utilisation of capacity of one type, and the potential loss for another. In short, these self-storage operators should seriously consider taking steps to maximise expected revenue at a stable cost-level.

Now, the design of self-storage warehouses differs from other facility designs in its focus on revenue maximisation. An obvious question is therefore whether it is possible to design such a facility so that it offers a better fit between storage design (types and numbers) and market demand, and at the same time maximises the revenue based on limited and somewhat “perishable” capacity. If it is indeed possible, the next question is: how do you design self-storage facilities so that they fit market segments and accommodate volatile demand and thus maximise revenue?

In seeking answers, there are two key issues to consider: customers who cannot be accommodated with a storage-unit size of their choice can be either rejected or upgraded to a larger space; and demand forecasts (based on historical data) should trigger a space allocation and division into the necessary unit sizes that reflect this demand.
2. Change is just a Snap

In general, warehouse facility design is a tactical decision: once a facility has been designed and built, it is difficult to adapt it to a changed environment. This is of course true of “traditional” warehouses. With self-storage, warehouse designs appear to be more flexible, thanks to the wide application of modular steel-base products, like modular corridors, standardised internal wall panels, standardised swing doors, and roller doors.

In particular, the internal panel has a special patented “snap together” interlocking seam rather than fixed jointing, which makes repartitioning of warehouse space easier. Most self-storage warehouses use a limited number of storage sizes (in the United States, usually eight types) and most sizes are an integer multiple of a standard size. It is usually possible to remove (or add) non-supporting walls, for example, to create one 9 m² room from a 3 m² and a 6 m² one. (Admittedly, there are some constraints: it is impossible to merge split rooms while they are still occupied; and adding a wall still requires that both newly created rooms have an access door). This space flexibility enables self-storage operators to adapt the layout of their facility to a changing demand at relatively short notice (like six months).
3. With no Reservations

We started our study by formulating three different model sales policies where:

1. The warehouse company simply rejects customers (typically in a high-demand situation) if the requested space is not available. In this case, the company does not try to convince customers to upgrade to a larger, more expensive space.
2. The company does try to convince customers to upgrade, and extra spaces of the popular types (based on historical data) are reserved.
3. The company tries to convince customers to upgrade, but no extra spaces are reserved.

Thanks to 54 self-storage warehouses in Europe, Asia and the USA, we had access to historical business data (including demand, contract type, price, and storage-space configurations) to help us investigate our three sales policies to determine the best at maximising total revenues based on optimised space allocation and division. We did this by inputting this data into mathematical models to which we subsequently applied dynamic programming (a useful mathematical technique for creating a sequence of interrelated decisions) to determine the optimal combination of decisions.

Let me use an example to illustrate how we conducted our experiments. If we received a company’s data over, say, the last four years, we used the data over the first three to forecast demand in the fourth year, and to optimise space allocation and division accordingly. We then tested all three sales policies using the scenarios we created and subsequently compared the “simulated” results to the actual fourth-year ones.

Experiment results showed that a new space-allocation design could indeed improve expected revenue of self-storage warehouses. Furthermore, small changes in load and upgrade-acceptance probabilities will not affect the overall design. Therefore, we recommend that self-storage warehouses review and adapt storage-space distribution as frequently as possible in order to benefit from changing demand rates. Considering how easy it is to redesign the storage space of a self-storage warehouse, there is no reason not to do it.

In addition, we discovered that the “upgrade without prior reservation” policy yields larger revenues than the one with prior reservation, although the
differences in revenue are slight. However, a prior-reservation policy may slightly outperform the no reservation policy for low demand and occupancy times that increase with unit size. In both cases, the difference in revenue decreases with increased load. Hence, selecting the prior-reservation policy may be justified due to its simplicity and near-optimal revenue.

However, because there are so many factors (some unknown) to consider, the best course of action may not always be that obvious. For example, a warehouse we investigated is in Chicago, close to the central business district and a university. Surprisingly (or perhaps not), this warehouse did not allocate and divide its storage space according to the needs of its two main customers with contrasting storage demands: small for students, and large for small businesses. By optimising space allocation based on customer demand and applying the rejection sales policy (policy number 1) our experiment achieved a revenue increase of 14.7 per cent.
4. Wider Applications

Notably, this study is one of the first to apply capacity and revenue management to facility design and management, taking into consideration market segmentation and uncertainty of data. In fact, our design approach can be applied to other fields as well, particularly to hotel design and management (in deciding which room types to build), parking-lot businesses (a parking lot layout has features similar to the layout of a self-storage warehouse), or to restaurant revenue management (to determine the optimal table mix). Equally interesting is applying it to construction equipment leasing. This is a huge market, since most civil construction engineering companies tend to rent such equipment as bulldozers, shovels, and cranes. Here our algorithms can help these leasing companies to determine the types and quantities of equipment they need to stock.

This article is based on the paper <c>Increasing the revenue of self-storage warehouses by facility design</c>, written by Yeming (Yale) Gong, René B. M. de Koster, J.B.G. (Hans) Frenk, and Adriana F. Gabor, and published in Production and Operations Management Vol. 22, No 3, May-June 2013, pp. 555-570. DOI: 10.1111/j.1937-5956.2012.01380.x

René de Koster is Professor of Logistics and Operations Management, Department of Technology and Operations Management, Rotterdam School of Management, Erasmus University. E-mail: rkoster@rsm.nl

“...this study is one of the first to apply capacity and revenue management to facility design and management, taking into consideration market segmentation and uncertainty of data.”
Accidents will happen

De Koster, R., Stam, D., Balk, B. (2011),
Accidents will happen: do hazard-reducing systems help?,
RSM Insight 5, 8-11.
In the summer of 2009, soon after the winners of the annual Safest Warehouse of the Year Awards were being lauded at an industry conference, journalist Marcel te Lindert wondered out loud in his regular column for the Dutch ‘Logistiek’ magazine why it was that there were more questions raised about safety issues than there were answers.

Getting straight to the heart of the matter, he asked how those involved in warehouse management knew which safety measures worked versus those that did not. Surely, he said, it is about time for a serious scientific study into the effectiveness of safety systems.

He was right, no research of substance had been conducted into warehousing accidents and the impact of the systems that are meant to prevent them. So, with the support of Marcel te Lindert’s publisher and the cooperation of the Dutch organisation of manufacturers and importers of material handling equipment (BMWT), we took up the challenge.

Statistically, the warehousing sector does not have a very good reputation when it comes to occupational safety. If you look at data for workplace accidents in the Netherlands, then the construction sector typically features as the most hazardous. Warehousing, although the Dutch Centraal Bureau voor de Statistiek (CBS) does not identify it as a standalone sector for statistical purposes, comes a close second.

**Table: Occupational accidents and deaths in the Netherlands**

*Between the years 2000-2007, the number of occupational deaths ranged from 87 to 147 annually.*

*In 2007, the number of occupational accidents leading to injury and absence from work totalled 219,000.*

*The medical costs of those occupational accidents requiring hospital treatment in 2007 amounted to €94 million.*

*Employee absence caused as a direct result of these occupational accidents cost €220 million.*

*In 2008, 1,700 serious workplace injuries were caused by accidents involving forklift trucks.*
Part of the reason for the high number of occupational accidents in warehouses is that work can be conducted under a lot of pressure. There are delivery times to be met, regardless of the volume of orders. It doesn’t matter if there are 1,000 orders to be fulfilled in one day or 10,000 the next, they still need to be picked, packed and processed in time for the scheduled collections. That’s one side of the story.

The other side of the story is that in many cases the systems used to achieve the time-driven goals involve forklift trucks and heavy moveable machinery weighing up to as much as eight tonnes. Working in the same space are workers on foot and as we know only too well, accidents happen, and sometimes with deadly consequences. The problem here is that it is difficult to create an environment where man and machinery do not mix.

After a review of what little scientific literature exists in this area, we identified two existing constructs, that of Safety-Specific Transformational Leadership (SSTL) and Worker Safety Consciousness (WSC). SSTL defines the ways and means by which managers are able to transfer safety issues to the workforce and motivate their safety consciousness. In both constructs, how managers lead in promoting safety can or should have a strong impact. This, at least, was our main hypothesis.

The first stage of our research was to measure the number of accidents, which we did from three and a half year’s worth of data, and place them into five already defined accident categories, the three most serious of which have to be reported to the Labour Inspection department of the Ministry of Social Affairs:

1. Near occupational accidents;
2. Occupational accidents resulting in injury but not leading to absence;
3. Occupational accidents resulting in injury and minimal absence from work of one day;
4. Occupational accidents resulting in hospital admission after a visit to the Emergency Department of a hospital;
5. Fatal occupational accidents.

Secondly, we looked at the number and type of safety systems used in warehouses. As a sector, warehouses implement numerous safety-enhancing systems that include a diverse range of safety procedures and safety equipment (for example, anti-collision devices, globe mirrors, safety signs and personal protective equipment). So diverse are these that a handbook published by the
BMWT advises of a bewildering 300 different safety-enhancing measures that warehouse managers can utilise.

Armed with the information on safety measures and statistical data, a survey was developed with the purpose of (1) defining what we called Hazard Reducing Systems (HRS) – the systems available to managers in order to enhance warehouse safety – and for the first time making them measurable, (2) defining safety performance, (3) demonstrating that safety performance is driven by managerial leadership, accident registration and safe storage procedures; factors that are mediated by safety consciousness, and (4) deriving managerial insights by making explicit which measures help to reduce accidents in warehouses.

From the 300 or so different safety-enhancing measures outlined in the BMWT handbook, we condensed them into 70 key HRS and placed them in four groups. The starting point of our survey was to ask participants to what degree the HRS had been implemented.

By analysing the statistical data and the results of the survey, for which we received input from both managers and employees, we were able to look for insights from the following variables: 1) the number and type of accidents, 2) the safety leadership abilities of managers as expressed by employees, 3) the perceived safety consciousness of the workforce, and 4) the hazard-reducing systems currently in place.
How Leadership Influences Safety

In Fig 1, we outline our hypotheses that SSTL positively influences safety performance, an effect that is mediated by safety consciousness. What helps drives this leadership more than anything is the introduction of HRS, something that managers must take responsibility for. One way of looking at it then is that from the perspective of the workforce the manager is the most crucial link in the safety chain. It is the manager who helps to develop and instil the environment of safety consciousness in the workplace that combined with SSTL and HRS impacts on overall safety performance.

We grouped HRS into four factors thus: safe traffic systems, hygiene, safety training, and safe storage systems. Safe traffic systems relate to the separation of people and machinery flows. High hygiene standards go hand in hand with high safety standards and this is recognised in the second factor. The third factor identifies the level and frequency of safety training. Standards and procedures for the correct storage of stock – empty pallets, equipment, machinery and tools, for example – fall into our final category.

Our findings show that of these four factors, safe storage systems have the greatest impact upon the effectiveness of safety leadership. In turn, the safety leadership of managers, partially mediated by the safety consciousness of the workforce, has the greatest impact on the number of accidents in warehousing facilities.

Serious and sustained attention to safety brought about by strong safety leadership makes workers more conscious of risks and so reduces accidents. Safety is therefore not a one-time issue, but is something that requires constant managerial attention. Another important driver is the careful registration of near and minor accidents that health and safety legislation does not require employers to record. Our findings indicate that recording these lesser events further fosters a culture of awareness of potential dangers and so provides managers with opportunities for even more hazard reduction thus exemplifying proactive safety leadership. This is an action we see taken by warehousing facilities with the best safety records.

In conclusion, we have four variables that impact upon the number of accidents: safety leadership, accident registration, workplace safety consciousness, and safety storage procedures. Linking to these is the manager, the most important factor of all. The effectiveness of managers in safety
leadership is strongly influenced by the safety procedures and systems that are in place. With a proactive manager and strong safety leadership even the least modernised and most accident-prone of warehouses can become a safe place to work.

To be an effective safety leader, appropriate safety-related incentives need to be offered to the workforce. If the right incentives are not offered then the workers will be less willing to comply with safety procedures and overall safety consciousness in the warehouse is not improved. So how can managers encourage employees to become more aware of safety issues through incentives? One way is to incorporate safety into appraisals. We don’t mean once a year personnel appraisals, but instead team performance evaluations specifically aligned to safety conducted at regular intervals, even weekly.

To do so will increase employee awareness of safety issues, compliance with which should be rewarded, as should increases in safety standards. Rewards should not be monetary. Instead rewards should be shown through an appreciation of employee efforts. Aligned with that should be worker empowerment. This way management drives to improve safety standards are more than just top down efforts, and instead continuous improvements are instigated and developed by those at the sharp end of safety matters.

Looking at the shortlist of nominations for the 2010 Safest Warehouse of the Year Award, one of the companies, Boston Scientific, has its warehousing staff divided into teams. These teams have been empowered to develop their own Key Performance Indicators (KPI) and this includes the whole area of safety.

The workforce at the spare parts facilities of Nissan in the Netherlands, another high performing warehouse, is also divided into teams. All are empowered to a very large degree and each is responsible for creating and meeting their own KPIs, one of which relates to improvements and innovations in processes leading to lower costs, higher quality and increased safety. Every month managers present the innovations put forward by their teams and the best ones are implemented right across the board.

Smart managers are realising the consequences of laxity in safety leadership. Not only does it have a negative impact on the workforce and lead to higher direct and indirect costs, but it can also damage company reputation: reputation as an employer and as a company with which to do business. Those
same managers also acknowledge that a proactive attitude to safety leadership and hazard reduction can lead to reduced costs for the organisation, increase employee satisfaction, and improve productivity and quality of work. The benefits to businesses are therefore obvious.

Figure 1. Hypotheses on variables driving Safety performance

Figure 2. Results of path model
Container terminal operations: overview

Gharehgozli, A., Roy, D. & De Koster, M.B.M. (2015),
Sea container terminals: New Technologies and OR models Maritime
Economics and Logistics, doi:10.1057/mel.2015.3

Amir Hossein Gharehgozli
(Rotterdam School of Management, Erasmus University, Netherlands)
AGharehgozli@rsm.nl

Dejit Roy
(Indian Institute of Management, Ahmedabad, India)
Debjit@iimahd.ernet.in

René de Koster
(Rotterdam School of Management, Erasmus University, Netherlands)
RKoster@rsm.nl
Abstract

Due to a rapid growth in world trade and a large increase in the flow of containerized goods, sea container terminals play a vital role in globe-spanning supply chains. Container terminals should be able to handle large ships, with large call sizes within the shortest time possible, and at competitive rates. In response, terminal operators, shipping lines, and port authorities are investing in new technologies to improve container handling and operational efficiency. Container terminals face challenging research problems which have received much attention from the academic community. The focus of this paper is on highlighting recent developments in container terminals, which can be categorized into two areas: (1) innovative container terminal technologies, and (2) new OR directions and models for existing research areas. By choosing this focus, we complement existing reviews on container terminal operations.

KEYWORDS: Container terminal; literature review; optimization; heuristic; simulation
1. Introduction

Since the introduction of the container in April 1956, when Malcolm McLean moved fifty-eight 35 foot containers from Newark to Houston by a refitted oil tanker, container flows have increased continuously. According to the US Customs and Border Protection (2011) report, in 2011, about 108 million cargo containers were transported annually through seaports around the world, constituting the most critical component of global trade. Between 1990 and 2015, the total number of full containers shipped internationally is expected to grow from 28.7 million to 177.6 million (United Nations: ESCAP, 2007). A simple calculation shows that there are enough containers on the planet to build more than two 8-foot-high walls around the equator (Taggart, 1999).

Containerization has become the main driver for intermodal freight transport, which involves the transportation of freight in containers of standard dimensions (20 ft equivalent unit (1 TEU), 40 ft. (2 TEU), 45 ft. (high-cube)), using multiple modes of transportation such as ships, trucks, trains, or barges without any handling of the freight itself when changing modes (Crainic and Kim, 2007). Stuffing freight in containers reduces cargo handling, and thereby improves security, reduces damages and losses, and allows freight to be transported faster (Agerschou et al., 1983). In the chain of intercontinental transport, container terminals are of special importance since all containers pass through at least one of them during their drayage. Container terminals are the nodes where different modalities meet to transport containers.

Container terminals have received increasing attention from the academic community due to the opportunities and challenges they offer in research. A number of reviews have been published in the last decade, focusing on the use of operations research models for handling containers (Vis and De Koster, 2003; Steenken et al., 2004; Günther and Kim, 2005; Murty et al., 2005; Stahlbock and Voß, 2008a; Gorman et al., 2014). Parallel to this paper, three focused reviews, focusing on seaside, transport, and stackside storage operations have appeared (Carlo et al., 2013, 2014a, 2014b). However, this paper provides an integrated view of container terminal operations, including gate operations and connections with the hinterland, based on recent literature published since 2008. We specifically focus on new technological developments and OR models. The scope of this study is limited to internal operations at a container terminal, as well as hinterland operations.
2. Container Terminal Operations

Container handling equipment includes quay cranes (QCs), yard cranes (YCs), automated guided vehicles (AGVs), and straddle carriers (SCs). These systems are shown in Figures 1a-d, and are used to transship containers from ships to barges, trucks and trains, and vice versa. Other, newer, equipment is introduced in the next sections. Containers can be transshipped directly from one mode of transportation to another. Alternatively, containers can be stored for a certain period in a stack, before they are transferred to another mode. Material handling equipment used at a terminal is very expensive, regardless of whether it is automated or manned. The investment in a single modern container terminal can be as high as €1 billion or more and the payback period ranges between 15-30 years (Wiegmans et al., 2002; De Koster et al., 2009).

Figure 1. A top view of a container terminal and cargo handling equipment
(Source: Europe Container Terminals (ECT), 2012)

Sea container terminals are divided into several areas such as seaside, landside, stacking, and internal transport areas that cater to seaside and landside operations (see Figure 2). At an automated container terminal, QCs load and unload containers from ships berthed along the quay at the seaside.
QCs pick up or drop off containers on AGVs which transport containers from the seaside to the stacking area where YCs take over. Finally, SCs transport containers either between the YCs and trucks or between the YCs and trains at the landside. In more traditional container terminals, SCs are also used to stack containers.

Containers are stacked in container stacks. Each stack consists of multiple rows, tiers, and bays. Containers arrive or leave the terminal at the seaside or landside and spend a period of time in these stacks. Input / output (I / O) points are located at each stack end and a single YC is used to stack and retrieve containers in that stack. A container’s storage position within the container stack is mainly determined by the loading sequence onto the ships. This sequence depends on the container’s ship departure time, its port of destination, and weight.

Obviously, containers have to be retrieved from the stack in the sequence of the departure of their corresponding ships. Furthermore, containers have to be loaded onto the ship in a reverse order of the sequence of destination. Containers with a later destination have to be loaded first. Finally, containers have to be loaded according to their weight. In order to ensure ship stability, heavier containers should be loaded before lighter ones. Many other practical constraints need to be considered while loading a ship (i.e., dynamic stability, container sizes, containers with hazardous materials, reefer containers, etc.). However, some flexibility may be permitted in the yard while retrieving containers from the stack to load a ship, because multiple QCs load a ship in parallel, and vehicles can be rescheduled while bringing the containers. In addition, stacks constantly change due to pre-marshalling.
A large terminal handles millions of containers annually (Drewry, 2011). Container terminals in the Port of Rotterdam handled more than 11.6 million TEU in 2013 while those in Shanghai handled more than 33.8 million TEU in the same year (Port of Rotterdam Authority, 2012; Shanghai International Port Group, 2014). Because many containers have to be stacked temporarily, more land is needed for the related supply chain activities. Lack of space has driven container terminal operators to build higher container stacks. In addition, ships have grown larger over the past decades with larger call sizes at ports. The largest Post-Panamax ships can carry about 18,000 TEU, compared to the first generation ships, which had a capacity of about 400 TEU. Shipyards are planning new ships of even larger sizes. Large ships can only berth in ports with adequate draft, at terminals with sufficiently wide gantry cranes, with adequate terminal material handling systems, and with adequate hinterland connections. This limits the number of ports of call and increases the drop size per terminal visited. Thus, larger ships spend more time in port than smaller ships. For instance, an 8,000 TEU ship spends 24% of its roundtrip time in port compared to 17% for a 4,000 TEU Panamax ship (Midoro et al., 2005). An idle 2,000 TEU ship costs $20,000-$25,000 per day (Agarwal and Ergun, 2008). Container terminal managers are constantly looking for new technologies and methodologies to efficiently handle all the containers arriving and leaving their terminals.

In recent years, port authorities and companies in several countries have started to integrate supply chain and transportation activities in particular by extending the sea terminal gate into the hinterland (Veenstra et al., 2012; Iannone, 2012). Previously, integrated hinterland terminals were introduced as dry ports. Different firms in multimodal hinterland networks, such as terminal operators, freight forwarders, information service providers, infrastructure managers, shippers, and consignees play a role. All these firms aim to contribute to a better performance of the overall supply chain. Terminal operators, for instance, are more and more involved in linking sea terminals with inland terminals. This enables them to better connect with shippers and receivers in the network. This change comes with serious and yet unexplored challenges, but it also provides one with opportunities to develop a sustainable and competitive advantage. The seamless flow of goods from seaports to locations far into the hinterland can prevent negative external effects from transport, such as congestion in seaports, or on motorways due to too much trucking.
In the subsequent sections of this paper, we discuss recent papers studying these topics using operations research tools. In addition, we try to identify new and important topics which are still pristine and offer opportunities to operations researchers. We start with seaside operations in the next section, and then discuss internal transport and stack operations in two subsequent sections, respectively. Previous studies generally categorize papers related to container terminal operations into three groups: seaside, stacking area, and landside (initially suggested by Steenken et al. (2004a) and Stahlbock and Voß (2008a)). We expand this framework to a seaside – stacking area – landside – hinterland framework. Due to lack of space, pollution, and long waiting times, integrated hinterland terminals have become an essential part of sea container terminals. Therefore, a survey on container terminal operations should include the recent developments and literature on hinterland operations. We do this in a Hinterland operations section. Each section, devoted to a specific container terminal process (seaside, transport, stacking area, and hinterland), is composed of two subsections. We first discuss the new technologies and then we describe the new developments in OR models. Finally we draw conclusions.
4. Seaside Operations

Seaside operations planning consists of ship berthing operations (berth planning and quay crane scheduling), and loading and unloading of containers onto ships. Further, the stowage planning, where the sequence of loading and unloading containers in a ship is optimized, plays a critical role in seaside operations planning. In this section, we discuss technological advancements in QCs and also review some of the recent work on sea operations.

New technologies

Recently, a new generation of fully automated (remote controlled) QCs has been developed. As shown in Figures 3a-c, these QCs are equipped with two trolleys, each capable of handling two or even three TEUs at the same time. In some designs, QCs are equipped with shuttles on the boom to reduce the horizontal handling time, or with trolleys that can rotate 90 degrees, as respectively shown in Figures 3d-e. In the section on stacking area operations, we discuss other designs in which QCs spread over an indented berth, or in which the QCs float on the water to build artificial temporary space.

Since the new designs can be used more flexibly with higher capacity compared to traditional QCs, the existing models may have to be adapted to these new developments. For example, Xing et al. (2011) analyse the problem of dispatching AGVs in container terminals equipped with tandem lift QCs, requiring two AGVs to be ready simultaneously to unload containers. The problem is formulated by a mixed-integer linear programming model and a decomposition method is used to solve the problem.

New terminal emulation systems that use terminal operating systems for input control allow dry testing of equipment control rules, remote quay crane control, as well as stack storage methods. For instance, Boer and Saanen (2012) developed an emulation tool – called CONTROLS (which stands for CONtainer TeRminal Optimised Logistics Simulation). All core processes in a terminal are supported by a TOS (Terminal Operations System), including quay side planning, vessel planning, yard planning, equipment control, and gate management. Emulation allows the user to experiment with the real TOS without the risk of negatively affecting real operations.
Figure 3. New generation of QCs
5. OR Models

Quay crane and berth operations planning

When a ship arrives, several tactical and operational decisions are made such as allocating berthing space, berthing time, and assigning QCs to load and unload containers at minimum terminal cost and delay. The first problem is commonly known as the berth allocation problem (BAP). The optimal allocation of berths to incoming ships is very complex because of spatial constraints such as the draft requirement for ship berthing, ship size, space availability, and the distance between the berthing location to the stacks where the ship’s containers are stacked. The complexity of the problem is further increased due to temporal constraints (static vs. dynamic arrival of ships). The second problem is related to assigning QCs to the ship. Modelling challenges such as addressing the interference between QCs (which all move on a common rail) and improving crane productivity, makes this problem interesting from both a research and a practical viewpoint. The third problem is scheduling QCs to unload or load containers from / to the ship by adhering to task precedence constraints. Until recently, the research community has mostly addressed these problems in isolation. However, due to interactions among the decisions, new algorithms and heuristic approaches have been developed to solve these problems within an integrated framework.

Figure 4 illustrates a berth plan with four ships. In this figure, the x and y axes denote the ship berthing time and the ship berthing space respectively. In Figure 4b, the QCs are assigned to each ship. Note that QC 2 is reassigned from ship 1 to ship 2, after its process is complete.

Figure 4. Illustration of (a) the berth allocation problem and (b) the QC assignment problem (adapted from Bierwirth and Meisel, 2010)
We now discuss these problems in more detail and review the recent OR modelling contributions. For a comprehensive survey on berth allocation and QC scheduling problems including papers prior to 2010, see Bierwirth and Meisel (2010).

**Berth Allocation Problem (BAP):**

To minimize the sum of ship waiting and handling times (port stay times), optimization models have been developed by fixing the choice of spatial, temporal, and handling time attributes. The spatial attribute denotes whether the quay area is partitioned into discrete or continuous berths (Buhrkal et al., 2011). The temporal attribute indicates the restriction imposed on the ship berthing time or departure time. Likewise, the handling time attribute indicates if this time is fixed or dependent on berthing position, QC assignment, or QC schedules. Hansen et al. (2008) solve the dynamic BAP problem by taking into account the service costs of ships depending on the berth they are assigned to, in addition to the handling times. The continuous dynamic BAP with both fixed and berth-position dependent handling times has received considerable attention from researchers (Wang and Lim, 2007). Many other versions of the BAP have also been considered. Hendriks et al. (2010) study a robust BAP in which cyclically calling ships have arrival time windows, instead of specific arrival times. They minimize the maximum amount of QC capacity required in different scenarios. In a later study, Hendriks et al. (2012) work on a similar problem in which cyclically calling ships have to be processed in different terminals of the same port. They minimize the amount of inter-terminal movement, and balance the QC workload at different terminals and time periods. Xu et al. (2012) study the BAP considering the water depth and tidal constraints. They model the problem in a static mode (all ships are available) and a dynamic one (ships arrive over time). They develop efficient heuristics to solve the problems. Nowadays, environmental issues are also considered in BAP models. Du et al. (2011) develops an elaborate berth allocation strategy that also minimizes the fuel consumption by the vessel. The nonlinear intractability introduced by the consideration of fuel consumption is overcome by reformulating the original MINLP model as a mixed integer second order cone programming (MISOCP) model.

**QC Assignment Problem (QCAP):**

After allocating a berth space to a ship, QCs are assigned to the ship such that crane productivity is maximized by reducing the number of QC setups and QC travel times. The two problems, QCAP and BAP, are closely interrelated, since once the QCs are allocated, the ship handling times are affected. In practice, the QCAP
is solved using rules of thumb and has received little attention from researchers (Bierwirth and Meisel, 2010).

Giallombardo et al. (2010) propose two formulations for combining the BAP and QCAP: a mixed-integer quadratic program and a linearization which reduces to a mixed-integer linear program. To solve the problem, they develop a heuristic which combines tabu search methods and mathematical programming techniques. Han et al. (2010) consider a similar problem, but with stochastic ship arrival time and handling time. They formulate the problem as a mixed-integer programming model and solve it by a simulation-based Genetic Algorithm. Chang et al. (2010) study the problem in a rolling horizon fashion. They solve the problem by a parallel genetic algorithm in combination with a heuristic algorithm.

**QC Scheduling Problem (Q CSP):**

In terminal operations, QCs are typically the most constrained resources. Hence, optimal schedules can maximize throughput, and minimize ship handling time (ship makespan). Several constraints need to be satisfied during the schedule generation process, such as preventing crane crossovers (by structural constraints imposed on cranes and the crane trajectory), maintaining a minimum distance between cranes (neighbourhood constraint), time separation of containers that need to be stacked in the same location (job separation constraint), and ensuring that unloading transactions within a ship bay precede loading transactions (precedence constraint defined by the stowage plan). Multiple optimization formulations have been developed with variations in task attributes (single or multiple bays), crane attributes (initial and final positions of the cranes, operational time windows), and interference attributes. Recently, container reshuffling and stacking area attributes (congestion constraints) have also been included in the models (Meisel and Wichmann, 2010; Choo et al., 2010). Legato et al. (2012) consider most of these constraints in a rich mixed-integer programming model. They solve the problem by a modified branch-and-bound algorithm which is based on the one developed by Bierwirth and Meisel (2009). Initial studies in this area generate QC schedules (unidirectional schedules) that consider non-crossing of cranes; i.e., all QCs move in the same direction throughout the service. For instance, Lim et al. (2007) generate unidirectional SC schedules for complete bays. They model the Q CSP using constructs from an m-parallel crane scheduling problem and develop a backtracking algorithm based on dynamic programming that generates optimal QC schedules for average-size problems. Another stream of research allows the cranes to share the workload of bays, and develops optimal
QC schedules for container groups. Lu et al. (2012) consider such a problem and solve it by developing an efficient heuristic which has a polynomial computational complexity. Queuing network models are also used to study the QCSP (Canonaco et al., 2008). The solutions of such models are usually evaluated based on simulation. Meisel and Bierwirth (2011) develop a unified approach for evaluating the performance of different model classes and solution procedures.

**Stowage planning**

To achieve economies of scale and better ship utilization, ships sail from one port to another (up to 20 ports) through a fixed route. At each port, thousands of containers may be loaded, unloaded, or repositioned. While such container movement plans reduce the transportation cost per container, it poses a difficult operational problem known as the container stowage problem (CSP). A stowage plan includes the placement of a container at a ship slot described by a combination of the row number, bay number, and tier number. The objectives of a good stowage plan are to minimize handling time, ensure stability, obey ship stress limits, and maximize QC utilization. Several constraints have to be taken into account, such as container size, weight, height, port of unloading, and container type (reefer, danger class). The complexity of developing high quality stowage plans has further increased with the launching of mega-ships with a carrying capacity of 18,000 TEU or more (for instance, see Maersk’s Triple E series plan, Maersk Line, 2011).

Wilson and Roach (2000) classify the methodologies developed for addressing the CSP into five categories: 1) simulation based upon probability, 2) heuristic driven, 3) mathematical modelling, 4) rule-based expert systems, and 5) decision support systems. They also indicate that the existing solution methods either relax some of the important constraints or do not generate high quality solutions in a short time (see also Avriel et al., 2000). Further, existing exact models do not scale beyond very small feeder ships.

To deal with the complexity of the CSP, successful studies decompose the problem hierarchically into a multi-port master planning phase and a slot planning phase (Delgado et al., 2012). In the first phase (Master planning), the hatch-overstowage and crane utilization measures are optimized by determining the number of 20ft and 40ft containers that need to be stowed in a location. The integer programming model, which is shown to be NP-hard, is solved using a relaxed MIP formulation. The second phase (slot planning) refines the master
plan by assigning the containers associated with each location to a specific slot in that location. A constraint-based local search (CBLS) approach is used to solve the optimization problem.

Other potential research topics for quay and berth planning include: 1) joint continuous / discrete quay berthing and QC planning with uncertainties in arrival and inhomogeneous quay cranes (think of cranes of different sizes or specializations for e.g. barges); 2) different objective functions, not just minimization of total berthing time, but also minimization of violation of agreed (contracted) time windows; 3) berth planning with cooperating QCs; and 4) berth planning and QC scheduling with mixed vessel types of different priorities, including barges. All these topics have until now hardly been studied.
The internal seaside transport process connects the seaside and the stacking. Vehicles are used both in the unloading process by transporting containers from seaside to the stacking areas and in the loading process by transporting containers from the stacking area to the seaside area.

Internal transport vehicles have varying degrees of automation and functionalities. We first review different types of vehicles. We then examine vehicle guide path types. The guide path has a significant impact on vehicle travel times and overall throughput performance. Further, we present innovations in information and communication technologies, such as vehicle tracking and tracing, that can help to improve coordination among vehicles. We then classify the different design decisions that affect vehicle transport performance, and discuss how OR tools can be deployed to analyse and to improve internal transport performance.

New technologies

Types of vehicles

Internal transport vehicles can be broadly classified into two categories: human-controlled and automated systems. Further, depending on the vehicle and crane transfer interface, the vehicles are classified as coupled (C) or decoupled (DC). Trailer-trucks and SCs are manual transport vehicles used in several container terminals in Asia (such as JNPT, India, and Northport, Malaysia). Automated lifting vehicles (ALVs) and AGVs are used in automated container terminals such as the Patrick container terminal in Australia, the ECT container terminal in Rotterdam, and the container terminal in Nagoya. Lift AGVs (L-AGVs) are a recent innovation in the AGV family, which will be deployed at the new APM terminal at Maasvlakte II, Rotterdam (Gottwald Port Technology GmbH, 2012). We briefly describe these internal transport vehicles below.

Automated Lifting Vehicles (DC):

These automated straddle carriers, also known as automated lifting vehicles (ALVs), decouple the container handling process between the seaside and the stacking area (Figure 5a). Due to their self-lifting capability, they are used in the unloading process, and pick up the containers from one of the several buffer lanes located beneath the QC5s and transport them to the YC buffer locations. The new automated lifting vehicles can lift up to two containers at a time.
Lift-Automated Guided Vehicles (C/DC):

Lift automated guided vehicles (L-AGVs) decouple the transport of containers to the stacking area processes (Figure 5b). The L-AGV is an AGV with electrically operated lifting platforms. These enable the vehicle to raise its load and deposit it independently and automatically on handover racks in the stacking crane interface zone and to pick up containers from those racks. Gottwald Port Technology GmbH (2012) claims that fleet size can be considerably reduced as a result of the increased working frequency; the overall number of vehicles required to service each QC can be reduced by up to 50% compared with conventional AGVs.

L-AGVs and ALVs operate in a very similar way; they both transport containers and use decoupled interfaces at the stack and quay side. However, ALVs can also stack containers, while L-AGVs cannot.

Figure 5. New internal transport vehicles: (a) Automated Lifting Vehicle, (b) L-AGV (Source: Gottwald Port Technology GmbH (2012))
Vehicle guide-path types

Automated vehicles travel along guide-paths in the yard area. Two types of guide-path networks are typically found: closed-loop and cross-lane. The closed-loop guide-path is composed of several large circular guide-paths for vehicles to follow (see Figures 6a-b). While a uni-directional closed loop travel path allows a simplified control of vehicles, it may increase vehicle travel time due to long travel distances. To reduce vehicle travel time, most automated terminals now use guide-path networks with multiple cross-lanes (see Figures 6c-d). A cross-lane path is composed of parallel travel paths with several big, small or mixed (both big and small) crossings. In cross-lane guide-paths, a vehicle adopts the shortest travel path (using shortcut paths) from the quay buffer lane to the stack buffer lane and vice versa. Hence, cross-lane guide-path networks can significantly reduce AGV travel distances, but the complexity of controlling traffic (and hence, chance of blocking) at the intersections of paths increases.

Vehicle coordination and tracking

Better coordination and real-time control of AGVs has multiple benefits for internal transport operations. A smaller fleet size can be used, and (empty) travel times can be reduced. Further, it allows one to deal with inherent operational variability in the system. Real time resource status information obtained by automatic context capturing devices such as sensor networks may be used by the terminal operators to re-plan the schedule. Today, several techniques exist for vehicle tracking and tracing, including the use of transponders, laser scanners, or GPS in combination with RFID. Ngai et al. (2011) develop an intelligent context-aware prototype for resource tracking at container terminals. Ting
(2012) discusses the feasibility of applying RFID for vehicle tracking purposes. Hu et al. (2011) discuss RFID related tracking solutions for orderly balancing and seamlessly connecting different operational processes at the terminal gate.

**OR models**

**Internal transport management**
Optimization formulations have been developed to determine optimal fleet size and to decide on vehicle routing and operation schedules. Jeon et al. (2011) adopt a Q-learning technique to determine the shortest-time routes for internal transport using AGVs. Note that their approach also considers the expected waiting times that result from vehicle interference and the shortest-path travel times, to determine the optimal routes. Vis and Roodbergen (2009) consider the problem of scheduling SCs to process container storage and retrieval requests in the yard area. The two components of the problem are assigning transport requests to the vehicles and scheduling these requests for each vehicle. By using a combination of a graph-theoretic and dynamic programming approach, they solve the problem to optimality. Nguyen and Kim (2009) develop a mixed integer model for a terminal which uses ALVs to handle containers at the seaside. The objective is to minimize the total travel time of the ALVs and the total delays of QCs. They transform constraints regarding the buffer space under the QCs to time window constraints and propose a heuristic algorithm to solve the model.

Analytical models based on queueing theory have been also put to practice to study internal transport management. Kang et al. (2008) develop a cyclic queue model of container unloading operations that provides a steady-state throughput measure and can estimate the optimal fleet (cranes and trucks) size. The model assumes exponentially distributed service times in order to obtain closed-form analytical results. They also develop a Markovian decision problem (MDP) model that can dynamically allocate a transport fleet based on general service time distributions. Finally, through simulations, researchers have evaluated design choices and operational policies. Petering (2010) develops a simulation model to study the real-time dual-load yard truck control in a transshipment terminal.

**Gate operations planning**
Terminal gates are the decoupling points of internal and external transport. Gate management is important, since the massive number of containers arriving and leaving terminals at the landside creates congestion. Many trucks and trains show up at the terminal gates for inland container transport. Trains have
fixed schedules which are set externally. Violating these time windows is costly and container terminal operators prioritize trains so they can be easily loaded and unloaded to leave the yard. On the other hand, trucks are a more flexible, more efficient mode in door-to-door service for containers, but they bring higher cost (Wang and Yun, 2011). Modelling truck scheduling problems and their interaction with container terminal operations offers interesting challenges for researchers.

Long queues of trucks at terminals lead to delays and cause emissions, congestion, and high costs. In the past several years, a growing number of studies have addressed truck congestion. One of the solutions is to carefully manage truck arrival times, using an appointment system in which a terminal operator announces the time periods that trucks can enter the terminal. Huynh and Walton (2008) develop a model to determine the maximum number of trucks a terminal can accept per time window. Huynh and Walton (2011) extend this model by additionally scheduling the trucks. Namboothiri and Erera (2008) study how a terminal appointment system affects the management of a fleet of trucks providing container pickup and delivery services to a terminal. The objective is to minimize transportation costs. Chen and Yang (2010) propose a ship-dependent time window optimization method, which involves partitioning truck entries into groups serving a specific ship and assigning different time windows to the groups. They use an integer programming model to optimize the position and the length of each time window and develop a genetic algorithm heuristic to solve the problem. Chen et al. (2013a) use several metaheuristic methods to solve the problem. Unlike the other studies, Lang and Veenstra (2010) consider congestion at the seaside. They develop a quantitative arrival scheduling simulation (centrally controlled by the terminal) to determine the optimal approach speed for the arriving vessels. Their cost function includes both fuel and delay costs.

Besides mixed-integer programming models, some studies use conventional stationary queueing models to analyse the gate system at container terminals (Guan and Liu, 2009; Kim, 2009). However, stationary queueing models should not be used to analyze a queueing system that is non-stationary in nature. The gate system at a container terminal is typical non-stationary, because the truck arrival rate varies from hour to hour and the gate service rate may change over time (Guan and Liu, 2009). Therefore, Chen et al. (2011b) propose a two-phase approach to find a desirable pattern of time varying tolls that leads to an optimal truck arrival pattern, by combining a fluid based queueing and a toll pricing model.
In general, terminal appointment systems have mixed performance. For example, Giuliano and O’Brien (2007) report on unsuccessful application of a terminal appointment system at the Ports of Los Angeles and Long Beach. As a result, in addition to a gate appointment system, the OR community should find new solutions to the gate congestion problem at terminals. Over the past decades, truck scheduling and storage allocation in port operations have been studied extensively as two separate sub-problems. However, from the operational point of view, they are highly interdependent. Researchers might find better solutions when studying the two problems together. Van Asperen et al. (2011) have conducted a simulation experiment to evaluate the impact of truck announcement time on online stacking rules. The longer the announcement period, the better the performance of the stacking rules is. Similarly, Borgman et al. (2010) use simulation to compare the effect of different stacking rules on the number of reshuffles. They use given container departure times to minimize the number of reshuffles. Zhao and Goodchild (2010) assess how truck arrival information can be used to reduce the number of reshuffles when containers are retrieved to be loaded on the truck. The results demonstrate that significant container re-handle reductions can be achieved by using the truck arrival sequence obtained from the terminal appointment system, even if the sequence information does not cover all the trucks.

In addition to trucks, trains are also handled at terminals. However, the literature on train transportation focuses mainly on scheduling and routing trains outside the terminal, which is outside the scope of this review (see, for example, Wang and Yun, 2011; Woxenius and Bergqvist, 2011; Almotairi et al., 2011; Leachman and Jula, 2012; Newman and Yano, 2000; Yano and Newman, 2001; Cordeau et al., 1998). On the other hand, a handful of papers discuss train loading and unloading operational problems. Most of these papers focus on handling trains at transshipment yards which are designed to move containers from trains to trucks and vice versa, which is again outside the scope of this review (see, for example, Jaehn, 2012).

Potential research topics for internal transport operations include: 1) studying the impact of new vehicle types such as Lift-AGVs on the terminal throughput capacity for loading and unloading operations; and 2) studying the effect of the number of buffer positions at the quay side on the vehicle throughput capacity.
7. Stacking area operations

The stacking area is one of the most important areas at a container terminal, since almost every container spends a period of time there. In the past, container terminals used traditional container handling equipment such as straddle carriers and reach stackers to stack and retrieve containers in the stacking area. However, these types of equipment cannot serve the huge number of containers that nowadays arrive and leave terminals. Today, most large new sea terminals use yard cranes to handle containers in the stacking area. Therefore, we focus on YC operations in the following sections.

New technologies

Yard cranes

New container terminals use two or three YCs per stack lane to retrieve and stack containers, as shown in Figures 7a-c (Li et al., 2009; Vis and Carlo, 2010; Li et al., 2012). Depending on the design of the stacks and YCs, the YCs can or cannot pass each other. There is a fixed safety distance between YCs during stacking operations. Twin lifting QC s have been used in automated container terminals for several years. However, YCs with twin lifting capabilities have only been introduced recently. Zhu et al. (2010) study a new type of YC designed by Shanghai Zhenhua Heavy Industries that can lift two 40 feet containers at a time (see Figure 7d).

Figure 7. Twin, Double, triple, and twin lifting YCs on a container stack
New container terminal layouts and stacking systems

In new automated container terminals, containers are generally stacked in container stacks which can be perpendicular or parallel to the quay. Obviously, the size of the stacks, their number, and the type of container handling equipment can be different. Many papers have recently studied the effect of these layout variables on the performance of the terminal. Kim et al. (2008) develop an integer programming model to determine the layout type, the outline of the yard, and the numbers of vertical and horizontal aisles. Wiese et al. (2011) develop a decision support model for the design of yard layouts of SC-based terminals. However, the dominant methodology in such papers is discrete event simulation because it is difficult to capture all elements and find an optimal solution. Petering and Murty (2009) develop a simulation model for a transshipment yard. They find out that in order to keep QCs busy and minimize the makespan of ships, the block length should be limited between 56 and 72 TEU. Furthermore, the movements of the YC should be restricted to one block. Petering (2011) extended the simulation study to include decision support for yard capacity, fleet composition, truck substitutability, and scalability issues. Kemme (2012) develops a simulation study to evaluate the effects of four Rail-mounted gantry (RMG) crane systems and 385 yard block layouts, differing in block length, width, and height, on the yard and terminal performance. Lee and Kim (2013) compare two terminal layouts that differ in the orientation of the stack blocks with respect to the quay: a perpendicular layout and a parallel layout. They consider different cost factors including the construction cost of the ground space, the fixed overhead cost of yard cranes, and the operating costs of yard cranes and transporters. The effect of various design parameters on the throughput capacity and storage space capacity of the designs is evaluated. Both optimization and analytical models are proposed to analyze optimal parallel stack layouts. For example, Lee and Kim (2010) propose optimization models to determine the stack block size at a container terminal based on different objective function such as minimizing the weighted expected Yard Crane cycle time for various operations subject to the minimum block storage capacity provided, and minimizing the weighted expected truck waiting time for various operations subject to the minimum block storage capacity provided. On the other hand, Roy et al. (2014) use an integrated analytical model for the seaside operations to obtain the optimal parallel stack layout.

In all the papers discussed so far, containers are loaded and unloaded by QCs from only one side of the ship. At a terminal with an indented berth, containers can be loaded and unloaded from the ship at both sides (see Figure
8a). Some of these terminals have special QCs spreading over the indented berth (see Figure 8b). Furthermore, terminal operators can use floating QCs to load or unload containers from both sides of the ship (see Figure 8c). Chen et al. (2011a) develop an integer programming model to schedule QCs loading and unloading containers at such an indented berth. They propose a Tabu search to solve the problem. Vis and Van Anholt (2010) develop a simulation model for a similar setting. They argue that an indented berth results in more flexibility. If all equipment is scheduled properly, an indented berth can lead to shorter makespan of the ship. However, the QCs can be used less flexibly compared to conventional quays, as they cannot easily move to other quay positions. In addition, using many cranes per ship may lead to a low productivity per crane, due to blocking from two sides of the ship, as well as by losing the advantages offered these days by the larger hold hatches. It may therefore be difficult to financially justify such operations. Recently, Imai et al. (2013) studied terminals with different indented berth designs servicing both feeders and mega-ships where mega-ships have priority to feeders. They conclude that a linear berth performs better in terms of reducing the handling time of feeders. On the other hand, an indented berth design, where ships can enter from one side and exit from the other, performs better in terms of reducing the handling time of mega-ships.

Figure 8. Different indented berth designs
Another layout suggested to increase performance is to add a chassis exchange terminal to a terminal or a group of terminals (Dekker et al., 2012b). In such a terminal, containers are stored on a chassis and can rapidly be moved by terminal trucks. This reduces congestion at the main terminals, as external trucks are handled elsewhere. During the night, import containers are collected from the terminal and loaded on chassis. During the day, these chassis are exchanged with chassis loaded with export containers. Since trucks can quickly charge or discharge chassis, the capacity of the terminal increases substantially. However, Guan and Liu (2009) argue that due to land requirements, more feasibility studies are necessary to justify its application.

Reviewing the literature on warehouse layout design reveals that in order to obtain more flexibility and a higher performance, designs in which stacks and transfer lines are diagonal to the quay, or in which stacks are divided into smaller stacks with different sizes and I/O points in the middle should be studied (Öztürkoğlu et al., 2012; Gue and Meller, 2009; Gue et al., 2012). In warehouses, such new layouts have achieved a reduction in vehicle travel time of up to 20% (Öztürkoğlu et al., 2012). Recently, container terminal managers have started to adopt new stacking systems stemming from warehouse literature, such as rack-based compact storage, or overhead grid rail systems. Ez-Indus of South Korea has built a prototype of an ultra-high container warehouse (UCW) system (see Figure 9a). The UCW is a high-rise rack-based automatic system that can theoretically save 90% of the space by stacking containers up to 50 tiers high. Containers are delivered to the UCW where they are placed on shuttles. These take containers into the UCW elevator which takes them to a slot in the rack. Shuttles can also be used to transfer containers between the seaside and landside, as shown in Figure 9b. These shuttles can move containers to the UCW or to the traditional container stacks. In the overhead grid system shown in Figure 9c, containers can be handled using the cranes hanging from the overhead grid. The new systems impact the entire terminal operations and may result in a higher terminal efficiency. However, the cost of implementing such systems is very high, as terminals deal with large and heavy containers compared to small and light totes in a warehouse. The question is therefore whether such systems can become profitable in the long term.
Yard crane operations planning

At a container terminal, yard cranes (YCs) move retrieval containers from the block to the input and output (I/O) points, and move storage containers from the I/O points to the block. Due to the huge number of containers handled at terminals, YCs often deal with a queue of containers waiting to be stacked or retrieved. It is therefore important to minimize the makespan and total delay time of all requests to be carried out by a YC, by optimally sequencing them. In the following, we review some of the recent OR models developed for scheduling YCs.

Yard crane scheduling:

In general, YCs can be classified into two types: rail-mounted gantry (RMG) cranes and rubber-tired gantry (RTG) cranes. RTG cranes are manned and can move freely from one stack to another. RMG cranes can be automated or manned, and their movements are limited to one or a few adjacent stacks in a row. Automated RMG cranes are sometimes called automated stacking cranes (ASCs), according to Stahlbock and Voß (2008b). A survey by Wiese et al. (2010) using the data of 114 container terminals, worldwide, shows that 63.2% of all terminals use YCs for stacking. In Asia, this is 75.5%. 
Most of the papers dealing with YC scheduling do not specify any special type of YC (i.e., manual or automated RMG or RTG crane), and as such the models and solution methods developed are applicable to all sorts of YC. However, the assumptions considered often show that the models are more suitable for a special type of crane and usually need to be modified for another type of crane. Table 1 summarizes recent integer programming models (other than the ones mentioned in Stahlbock and Voß, 2008a) on YC scheduling including the objective functions and constraints considered. We try to indicate the type of crane considered based on the constraints mentioned in the papers.

Recently, Sharif and Huynh (2012) employed agent-based models to formulate the YC scheduling problem. Agent-based models can dynamically adapt to real-time truck arrivals, making them better suited for real-life operations. Vidal and Huynh (2010) also use an agent-based approach to schedule YCs with a specific focus on assessing the impact of different crane service strategies on drayage operations. In their work, they model the cranes as utility maximizing agents and develop a set of utility functions to determine the order in which individual containers are handled. Finally, Petering et al. (2009) develop a simulation model for real-time YC control in transshipment terminals.

**Scheduling two or three YCs dedicated to a stack:**

In the previous section, we focused on papers in which each stack has a single YC or in which stacks share YCs. We now review papers in which stacks with double or triple YCs are considered.

Li et al. (2009) introduce a discrete time model to schedule twin YCs carrying out storage and retrieval requests in a single stack. The YCs cannot pass each other and must be separated by a safety distance. The requests have different due times and the objective is to minimize a weighted combination of earliness and lateness of all requests, compared to their due times. They introduce a rolling horizon algorithm in which a horizon of a specific length is defined, and all requests falling within this horizon are considered and optimized by CPLEX. The horizon is updated whenever all its requests have been scheduled. In a recent paper, Li et al. (2012) extend the model to a continuous time model. The results show a significant improvement compared to a previous discrete model. Gharehgozli et al. (2014a) also schedule twin YCs for a large number of requests using adaptive-large neighbourhood search. Park et al. (2010) consider container rehandling (see the stacking area operations section) in their mixed-integer programming model used for scheduling twin YCs in a rolling horizon mode. Containers that have to be rehandled are considered as independent requests.
and are assigned to any idle YC. This approach results in balancing the workload of YCs and reducing the waiting times of trucks and AGVs. Vis and Carlo (2010) consider a double YC problem in which the YCs can pass each other but cannot work on the same bay simultaneously. In their problem, requests do not have any due time and can be scheduled in any sequence. They formulate the problem as a continuous time model and minimize the makespan of the YCs. They solve it by a simulated annealing algorithm and use the single-row method proposed by Vis (2006) to compute a lower bound. Cao et al. (2008) propose an integer model for a similar problem. They develop two heuristics and a simulated annealing algorithm to solve the problem. Stahlbock and Voß (2010) perform a simulation study to investigate to what extent double YCs can help to improve container terminal efficiency. They evaluate different online algorithms for sequencing and scheduling requests. The experiments are based upon real world scenarios (from the Container Terminal Altenwerder, CTA, Hamburg, Germany).

Recently, Container Terminal Altenwerder (CTA) in the Port of Hamburg installed three cranes per stack to handle stacking operations (Dorndorf and Schneider, 2010). Two cranes are smaller so that the larger crane can pass. Dorndorf and Schneider (2010) model the scheduling problem of these cranes as an integer programming model. The objective is to maximize the productivity of the crane system under peak load while preventing delays in the transport of import and export containers. They solve the problem in a rolling horizon scheme using a beam search method. The results show that the method performs better than nearest neighbor and first-come-first-served request selection heuristics by more than 20%.

**Minimizing container reshuffling**

New technologies and methods for managing the stacking area of container terminals reduce container throughput time. However, a discussion about improving the efficiency of container stacking would be incomplete without considering container reshuffling. A reshuffle is an unwanted movement of a container stacked on top of the one which has to be retrieved (Kim et al., 2000; De Castillo and Daganzo, 1993; Caserta and Voß, 2009b). Reshuffling is time consuming and it increases a ship’s berthing time. Few systems, such as the UCW discussed above, allow direct access to all containers. However, not much technological innovation can be seen in this regard. On the other hand, many new methods have been designed to reduce or avoid reshuffling. Papers dealing with container reshuffling study three main subjects: (1) pre-marshalling, (2) relocating methods while retrieving containers, and (3) stacking methods.
The common objective in all these papers is to reduce the number of reshuffles.

Pre-marshalling:

Some researchers focus on how to reduce the number of reshuffles by pre-marshalling containers in a way that fits the ship’s stowage plans. Pre-marshalling is the repositioning of containers of the stack so that no, or few, reshuffles are needed when containers are loaded onto the ships. Lee and Hsu (2007) propose an integer programming model for a container pre-marshalling problem preventing reshuffles. They develop a multi-commodity network flow model for obtaining a plan on how to pre-marshal containers stacked in a single bay and solve it by replacing some of the constraints and relaxing others. They also propose a simple heuristic for large-scale problems. Lee and Chao (2009) develop a neighbourhood-based heuristic model to pre-marshall containers of a single bay of a container stack in order to find a desirable final layout. Caserta and Voß (2009b) propose a dynamic programming model to pre-marshall containers of a single bay. In order to quickly find the solution, they propose a corridor method. In this local search method, a pre-marshalled container of a specific pile can only be stacked in a corridor which consists of the next few predecessor or successor piles of the bay with a specific limit on the number of empty locations. Expósito-Izquierdo et al. (2012) develop an instance generator which creates instances with varying degrees of difficulty. The difficulty of the instances is determined based on the occupancy rate and the percentage of containers with high priority that are located below those with low priority. Bortfeldt and Forster (2012) propose a tree-based heuristic to solve the pre-marshalling problem. In the tree search procedure, the nodes of the tree represent layouts. The root node corresponds to the initial layout and each leaf node corresponds to a final layout. Finally, Huang and Lin (2012) work on two different types of container pre-marshalling problems, and develop two heuristics to solve them. They obtain better solutions than the solutions present in the literature (i.e., Lee and Hsu, 2007) in a shorter time.

Relocating:

The problem of minimizing the number of reshuffles of a container stack while containers are retrieved is called the block (stack) relocation problem (BRP), which is proven to be NP hard by Caserta et al. (2012). Given a retrieval sequence of containers in the BRP, the decision is where to locate reshuffled containers to obtain the minimum number of reshuffles when all containers in the retrieval sequence are retrieved. Caserta et al. (2011) formulate the problem as a dynamic programming (DP) model and use a corridor method similar to the one proposed by Caserta and Voß (2009b) to solve it. Since the quality of the
solution depends on the length and height of the corridor, in two later papers, Caserta and Voß (2009a) and Caserta and Voß (2009c) incorporate heuristic algorithms to tune these variables. Caserta et al. (2009) propose a greedy heuristic algorithm to solve the BRP. In their heuristic, a reshuffled container will be stacked either in a pile which is empty or in a pile of containers that have lower retrieval priority. If such a pile is not available, the container will be stacked in a pile of containers that have a retrieval time nearest to the container that has to be stacked. Forster and Bortfeldt (2012) propose a tree search algorithm similar to the one developed by Bortfeldt and Forster (2012) for the pre-marshalling problem. However, they use a finer move classification scheme, different rules for branching and bounding, and require an additional greedy heuristic. Finally, Lee and Lee (2010) consider a multi-bay generalization of the BRP and describe a solution approach that combines heuristics with integer programs.

Stacking:

Although pre-marshalling and relocating help to minimize the number of reshuffles while containers are retrieved, a good stacking policy significantly decreases the handling effort in later stages. Some papers focus on how to avoid reshuffling by proposing methods to properly locate incoming containers in a container stack. Dekker et al. (2007) investigate different stacking policies, using simulation based on real data. They allocate containers to the stack based on the container’s expected duration of stay. Kim and Park (2003) also propose a heuristic algorithm based on the container’s duration of stay to locate export containers. Kim et al. (2000) propose a stochastic DP model for determining storage positions of export containers in a single bay of a stack. To avoid solving a time consuming DP model for each incoming container, they build decision trees, using the optimal solutions of the DP model. The trees decide where to store an incoming container. The validity of the recursive function of the DP model is proven by Zhang et al. (2010). Their method is adapted to handle large realistic stacks by Gharehgozli et al. (2014c). Sauri and Martin (2011) propose three new strategies to stack import containers. They also develop a model to compute the expected number of reshuffles based on the container arrival times. They compare their strategies based on different criteria including the size of terminals and container traffic. Yu and Qi (2013) consider a similar problem. They propose two models of which one is used to allocate import containers to the stack after they are unloaded from a ship, and the other one is used to pre-marshalling containers. Through simulation, they validate the models and find out that segregating the space and pre-marshalling enhance the efficiency of the terminal in terms of reducing truck waiting times.
<table>
<thead>
<tr>
<th>Paper</th>
<th>Request type</th>
<th>Objective function (to minimize)</th>
<th># stacks</th>
<th># YCs</th>
<th>Specific assumptions</th>
<th>Solution method</th>
<th>Crane Type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>adjacent stacks in a row</td>
<td></td>
<td>retrieval sequence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Narasimhan and Palekar (2002)</td>
<td>Retrieval</td>
<td>Makespan</td>
<td>Single</td>
<td>Single</td>
<td>Containers with different types and a retrieval sequence</td>
<td>B&amp;B, heuristic with a worst-case performance ratio of 1.5</td>
<td>RMG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>adjacent stacks in a row</td>
<td></td>
<td>retrieval sequence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen and Langevin (2011)</td>
<td>Retrieval</td>
<td>Makespan</td>
<td>Adjacent stacks in a row</td>
<td>Multiple</td>
<td>Containers with different types and a retrieval sequence, non-passing cranes</td>
<td>Genetic algorithm and tabu search</td>
<td>RMG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guo and Huang (2011)</td>
<td>A set of jobs including retrieval and storage</td>
<td>Minimize the average vehicle job waiting time at yard side</td>
<td>Adjacent stacks in a row</td>
<td>Multiple</td>
<td>Crane interference (2005) constraints, comparing the results with Ng</td>
<td>A combination of simulation and optimization</td>
<td>RMG</td>
</tr>
<tr>
<td>Froyland et al. (2008)</td>
<td>Retrieval and storage</td>
<td>Smooth utilization of RMGs</td>
<td>Adjacent stacks in a row</td>
<td>Multiple</td>
<td>Other objectives including minimizing long dwell time and maximum number of SCs</td>
<td>Decomposition heuristic</td>
<td>RMG</td>
</tr>
<tr>
<td>Paper</td>
<td>Request type</td>
<td>Objective function (to minimize)</td>
<td># stacks</td>
<td># YCs</td>
<td>Specific assumptions</td>
<td>Solution method</td>
<td>Crane Type</td>
</tr>
<tr>
<td>---------------------------</td>
<td>---------------------------</td>
<td>---------------------------------------------------------------------</td>
<td>----------</td>
<td>------</td>
<td>-------------------------------------</td>
<td>--------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Gharehgozli et al. (2014b)</td>
<td>Retrieval and storage</td>
<td>Then makespan</td>
<td>Single</td>
<td>Multiple</td>
<td>Crane interference</td>
<td>B&amp;B</td>
<td>RTG</td>
</tr>
<tr>
<td>Zhang et al. (2002)</td>
<td>Retrieval and storage</td>
<td>The total delayed workload</td>
<td>Multiple</td>
<td>Multiple</td>
<td>Crane interference constraints</td>
<td>Lagrangian relaxation</td>
<td>RTG</td>
</tr>
<tr>
<td>Cheung et al. (2002)</td>
<td>Retrieval and storage</td>
<td>The total delay time of requests</td>
<td>Multiple</td>
<td>Multiple</td>
<td>Crane interference constraints</td>
<td>Lagrangian decomposition and successive piecewise-linear approximation</td>
<td>RTG</td>
</tr>
<tr>
<td>He et al.</td>
<td>Retrieval and storage</td>
<td>The total delay time of requests, the number of times that YCs move among stacks</td>
<td>Multiple</td>
<td>Multiple</td>
<td>Crane interference constraints</td>
<td>Genetic algorithm employing heuristic rules</td>
<td>RTG</td>
</tr>
<tr>
<td>Chang et al. (2011)</td>
<td>Retrieval and storage</td>
<td>The total delay time of requests</td>
<td>Multiple</td>
<td>Multiple</td>
<td>Crane interference constraints</td>
<td>Heuristic algorithm</td>
<td>RTG</td>
</tr>
</tbody>
</table>
Finally, Casey and Kozan (2012) develop a dynamic mixed integer programming model to determine where to stack containers while others are retrieved in each period. The objective is to minimize the total amount of time that containers are stacked in the block. Several constraints are developed to calculate the amount of required time for reshuffling, loading and unloading containers from the equipment systems handling the containers.

Other potential research topics for stack side operations include: 1) understanding the impact of multiple stack cranes on interference delays, and 2) studying how the stack crane scheduling can be done to minimize delays in both seaside and landside operations.
8. Hinterland Operations

Large deep-sea terminals face many challenges such as congestion, delays, and pollution. This has driven container terminals to transform their supply chains to increase their competitiveness and robustness (Vervest and Li, 2009; Heinrich and Betts, 2003). By closely collaborating with hinterland terminals, deep-sea terminals can balance flows and workload more efficiently over time. As a result, not all the value adding activities, such as container inspection and container delivery to end customers, need to be done at deep-sea terminals, but it can be relegated to hinterland terminals.

New technologies

Close cooperation of deep-sea and hinterland terminals is a recent development, caused by the substantial increase of the number of containers handled (Heaver et al., 2001; Notteboom and Winkelmans, 2001; Notteboom, 2002; Robinson, 2002; Van Klink and Van den Berg, 1998; Roso et al., 2009). At first, the goal was to increase terminal capacity by adding dry ports to the network of main deep-sea terminals. During the last decade, the amount of authority delegated to these hinterland terminals has increased. Veenstra et al. (2012) describe an example project in Rotterdam which integrates supply and transportation by extending the sea terminal gate into the hinterland. An extended gate is an inland intermodal terminal directly connected to the seaport terminal(s) with high capacity transport mean(s), where customers can leave or pick up their containers as if directly at a seaport (including customs and security inspections), and where the seaport terminal operator can choose to control the flow of containers between the terminals. Iannone (2012) empirically studies a similar example in the Campania region in Southern Italy.

OR models

Network configurations give rise to several strategic and operational challenges such as information sharing, modal split, inter-terminal transportation, repositioning of empty containers, asset light solutions, and barge operations which were not a matter of concern previously. We discuss these topics in more detail below. Due to their novelty, the operations research-related literature on this topic is still limited, providing opportunities for future research.
**Information sharing**

One of the crucial conditions for the development of efficient networks is the availability of reliable information on containers (arrival, departure times, content, and modes of final transport). Terminal operators usually possess estimated arrival and departure times. More exact information is needed though, to better stack containers; minimize internal travel time to the proper pick up points; and avoid tardiness or earliness of loading and unloading different means of transport. Some examples were discussed in the previous sections (see, for example, Van Asperen et al., 2011; Borgman et al., 2010). Douma et al. (2009) propose a decentralized multi-agent system to align barge operators with terminal operations. They compare their approach with a central approach, where a trusted party coordinates the activities of all barges and terminals. The results indicate that, in spite of the limited information available, their approach performs quite well compared to the central approach. In their later study, Douma et al. (2011a) examine the effect of different degrees of cooperativeness on the efficiency of the barge handling process.

**Modal split and service network design**

One of the most important challenges in a container handling and transport network is the modal split. Besides ships, terminal operators deploy trucks, trains, and barges to transport containers. Barges and trains have less negative environmental and societal impact than road transport. However, compared to trucks, barges and trains usually have longer transit times, and do not connect directly to any final destination (Groothedde et al., 2005). Port authorities are strongly urging container terminals to adopt more environmentally friendly modes of transport. For example, the Port of Rotterdam needs to move from the current truck / barge / rail split of 45 / 40 / 15 percent to 35 / 45 / 20 percent by 2035 (Veenstra et al., 2012; Port of Rotterdam Authority, 2012). Synchromodal transport connections (i.e. connections where multiple transport modes can be used on different legs of the transport path and where switching between modes is possible, can be instrumental in achieving this.

Over time, seaport terminals are being integrated with inland terminals, by means of frequent services of high capacity transport modes such as river vessels (barges) and trains. The multi-modal transport operators typically face three interrelated decisions: (1) determine which inland terminals can act as extended gates of the seaport terminal, (2) determine capacity of the transport modes and frequency of service, and (3) set prices for the transport services on the network (Ypsilantis and Zuidwijk (2013)). Van Riessen et al. (2013) propose an integer-programming model for the design of such networks. The model uses a
combination of a path-based formulation and a minimum flow network formulation that penalizes overdue deliveries and combines both self-operated and subcontracted services. Sharypova et al. (2012) address the minimum cost service network design problem by developing a continuous-time mixed-integer linear programming model. Using this model, they are able to accurately determine transportation events and the number of containers to be transshipped by vehicles. Ypsilantis and Zuidwijk (2013) propose a bi-level programming model to jointly obtain the design and price of extended gate network services for profit maximization.

Inter-terminal transportation

Inter-terminal transportation is the movement of containers between close-by terminals (by sea, rail, or otherwise), often in the same port area. When developing new terminals and container ports, the movement of containers between terminals has to be taken into consideration. Previous work in the area of design and evaluation of inter-terminal transportation commonly deals with simulation (see, for example, Ottjes et al., 2007). Recent work uses optimization methods to study the problem. Tierney et al. (2013) develop a model combining vehicle flows and multi-commodity container flows for inter-terminal movements. They solve the model to optimality using real data from the Port of Hamburg, Germany, and the Maasvlakte 1 & 2 area of the Port of Rotterdam, The Netherlands. Lee et al. (2012) develop a mixed-integer programming model for assigning ships to terminals within a terminal hub and allocating corresponding containers. The objective function minimizes the total inter-terminal and intra-terminal handling costs. They develop a two level heuristic algorithm to solve the problem. Minimizing intra-terminal handling costs is also considered in the BAP problem studied by Hendriks et al. (2012).

Barge transportation

During the past decades, truck transport has been the dominant mode of inland transportation compared to train and barge. To reduce pressure on road infrastructure as well as to reduce greenhouse gas emissions, port authorities aim for a modal shift from road to barge or train. For a country with easy access to waterways, barge transport is a competitive alternative to road and rail transport due to its ability to offer cheap and reliable transport services. One of the crucial conditions for successful barge freight transportation is the alignment of terminal and barge operators. The barge handling problem (BHP) consists of routing and scheduling barges to visit different terminals in a port. A centralized decision making method, where a trusted party coordinates the activities of all barges and terminals, is unacceptable by both terminal and
barge operators. Generally, they are not willing to share information, wanting to be autonomous. Therefore, online decentralized decision making methods are much more suitable in this case. To achieve this goal, Douma et al. (2009) model the problem using agent based planning systems and compare their approach with a central approach. The results indicate that, in spite of the limited information available, their approach performs quite well compared to the central approach. The authors extend this idea in their later studies (Douma et al., 2011a, b). The insights from these studies are currently embedded in a project entitled barge terminal multi-agent network (BATMAN) which will be implemented in the Port of Rotterdam (Mes, 2012). Despite the intriguing findings and clear contributions, the present studies do not yet capture all aspects of such operations in practice, leaving room for further study.
9. Conclusions

During the last decade, container terminals have witnessed rapid developments that have led to the design of more automated, responsive, cost and energy-efficient, and secure terminals. Operations research models encompassing new constraints and objective functions, needed by such advancements, are required to efficiently manage container terminals. The operations research community needs to revisit and update the previous studies on container terminal operations. This paper discusses the new developments in container terminal technologies and OR models, and reviews the related literature. Although the study is limited to container handling operations performed inside a terminal, this paper shows that there is a substantial body of research on the related topics and there is enough room for further research.

Acknowledgements: We would like to thank the reviewers for their suggestions, which greatly improved the quality of this paper and SmartPort Research Centre at Erasmus University (www.eur.nl/smartport) for funding this research.
10. References


Port of Rotterdam Authority, ed. (2014) Abridged annual report 2013, Port of Rotterdam Authority, Rotterdam.


Optimal design of container terminal layout

Optimal design of container terminal layout, in:
B. Kuipers and R. Zuidwijk (Eds.),
Smart Port Perspectives. Essays in honour of Hans Smits (pp. 129-140).
Erasmus Smartport, Rotterdam.
Through this paper we would like to express our gratitude to the Port of Rotterdam Authority, and its CEO Hans Smits in particular, for making SmartPort possible. This paper is one of the first tangible SmartPort research results. We are confident more will follow. We wish Hans lots of success and good health in the coming future!

Debjit Roy  
Indian Institute of Management Ahmedabad Gujarat, India

Rene de Koster  
Rotterdam School of Management, Erasmus University Rotterdam, The Netherlands
Abstract

Due to rapid growth in foreign trade using sea vessels, there is a growing focus in improving the infrastructure and operational efficiencies at the container terminals. Particularly, the operational responsiveness of loading and unloading of containers, affects the vessel idle times and profitability of the shipping liners. In this research, we determine optimal stack layout design, which minimizes the container unload times using Automated Guided Vehicles (AGVs). To analyse alternate stack layout designs, we develop integrated queuing network models that capture the stochastic interactions among the container terminal processes (quayside, vehicle transport, and stackside), and provides realistic estimates of expected container unload throughput times.
1. Background and Motivation

Due to growth in international trade and better accessibility to the major seaports via deep-sea vessels, containerization has become the preferred mode for maritime shipping and inland transportation. Between 1990 and 2008, container traffic has grown from 28.7 million TEU to 152.0 million TEU, an increase of about 430% ([3]). Currently, several new deep-sea as well as inland container terminals are being designed across continents. Several of the larger ones will be automated.

The design of the container terminal includes strategic design choices such as the terminal layout at the stackside, choice of equipment for handling containers at the seaside and landside, and type of vehicles for container transport between seaside and the landside. However, the process to arrive at an optimal design is extremely complex due to several reasons. They are: 1) physical constraints such as variations in ground conditions and topology of the terminal area, 2) large number of design parameters and corresponding solution search space, and 3) stochastic interactions among the three processes (quayside, vehicle transport, stackside). In this research, we analyse container terminal operations at the seaside using AGVs. Figure 1a shows an aerial view of a container terminal that includes vessels berthing at the quayside and the stackside whereas Figure 1b describes AGVs transporting containers in the yard area.

Figure 1. (a) Aerial view of a container terminal (Courtesy: marineinsight.com) and (b) AGVs transporting containers in the yard area (Courtesy: porttechnology.org)
Due to significant investments involved in the development of a container terminal, an optimal design of the terminal is crucial. Traditionally, the main research focus has been on building simulation and optimization models to address strategic and tactical issues such as the container stowage problem at ships and in the stack, as well as on operational issues such as vehicle dispatching rules and quay crane scheduling ([2], [5]). Practitioners also develop detailed simulation models to design new terminals or improve the efficiency of existing terminal operations. While simulation provides detailed performance measures, it limits the extent of the design search procedure due to associated model development time and costs. In this research, we develop analytical models, which enable the terminal operator to analyse alternate configurations rapidly.

Analytical models have also been built to analyse specific system design aspects, for instance, Canonaco et al. [1] developed a queuing network model to analyse the container discharge and loading at any given berthing point. Hoshino et al. [4] proposed an optimal design methodology for an Automated Guided Vehicles (AGV) transportation system by using a closed queuing network model. However, in literature, integrated analytical models for analysing the performance of loading and unloading operations by considering some of the stochastic inputs are scarce ([8], [7]). For instance, Vis et al. [9] assume deterministic AGV travel times while estimating the number of AGVs in a semi-automated container terminal.
New automated terminals typically adopt Automated Guided Vehicles (AGVs) for vehicle transport. AGVs do not have self-lifting capabilities and they need to be synchronized with the quay cranes at the quayside and with the stack cranes at the stackside to pick up or drop off the containers. In this research, we analyse alternate terminal layout configurations by varying the stackside configuration (number of stacks, bays, and height), and vehicle transport configuration (number of AGVs and travel path dimensions and topology) using analytical models. Each configuration may also impact the vehicle guide path and hence the travel times. For instance, by increasing the number of stack blocks, the length of the vehicle guide path also increases (refer Figure 2). Therefore, the stacking time per stack may decrease whereas the vehicle transport time may increase. Therefore, the configuration of an optimal stack layout is not clear.

Figure 2. Alternate terminal layout configurations (a) small number of stacks and large number of bays (b) large number of stacks and small number of bays
Our work closely aligns with the analytical model developed by [4]. However our research differs from their work in several aspects:

1. We develop a semi-open queuing network model of the terminal system, which considers the synchronization of the AGVs and the containers waiting at the vessel to be unloaded. In reality, on some occasions, an AGV would be waiting for a container to be unloaded while during other times, a container would be waiting in the vessel for unloading operations. In a closed queuing network (such as in [4]), synchronization effects are not considered.

2. We consider realistic vehicle travel paths with multiple shortcuts that decrease the average travel times and improve vehicle capacity. Previous models do not consider the effects of multiple short cuts.

3. We develop protocols for handling containers at the quayside and the stackside that allow us to model the vehicle synchronization effects at the quay and the stack area.

4. We adopt our model to analyse alternate terminal layouts by varying the number of stacks, bays, and vehicle path dimensions, and arrive at a layout that minimizes throughput times and costs.

In this research, we develop an integrated analytical model for the unloading of containers at the *seaside* by considering the queuing dynamics at the quayside operations, vehicle transport operations, and stackside operations. Each quay crane is modelled as a single server station with general service times. The travel times associated with vehicles are modelled using Infinite server stations with general service times. Similarly, each stack crane is modelled as a single server station with general service times. Containers that wait to be unloaded may wait for an available vehicle, at the quayside. However, due to capacity limitations of the quay crane, a vehicle may also wait for a container arrival. This interaction between vehicles and containers is precisely modelled using a synchronization station and the queuing dynamics in the vehicle transport is modelled using a semi-open queuing network (SOQN) with $V$ vehicles. The performance measures from the analytical model are validated using detailed simulations. Using the analytical tool, which can be evaluated rapidly, we analyse alternate terminal layout configurations and arrive at an optimal configuration. We believe that the stochastic model of the container handling operations can be used for rapid design conceptualization for container port terminals and improve container handling efficiencies.
The rest of this paper is organized as follows. The terminal layout adopted for this study is described in Section 2. The queuing network model for terminal operations with AGVs along with the solution approach is provided in Section 3. The results obtained from numerical experimentation and model insights are included in Section 4. The conclusions of this study are drawn in Section 5.
2. Description of Terminal Layout

Figure 3 depicts the top view of a part of a container terminal, which includes the quayside, transport and the stackside area (stack blocks with cranes, transport area with vehicles, QCs). The design of this layout is motivated from practice (see [2]). We focus on the space allowing berthing of one jumbo vessel with a drop size of several thousands of containers. A large container terminal may contain several of such identical berthing positions. The number of stacks is denoted by $N_s$ and each stack crane is referred as $SC_i$ where $i \in \{1, \ldots, N_s\}$. Similarly, the number of QCs is denoted by $N_q$ and each crane is referred as $QC_j$ where $j \in \{1, \ldots, N_q\}$. There is one shortcut path after each QC (referred as $SP_j$ where $j \in \{1, \ldots, N_q\}$) that connects the quayside and the stackside areas. Both stacks and QCs have a set of buffer lanes, which are used by the vehicles to park during loading or unloading containers. The number of buffer locations at each QC and SC are denoted by $N_{qb}$ and $N_{sb}$ respectively. The other notations present in Figure 3 indicate path dimensions, which are used later to estimate the vehicle travel times.

![Figure 3. Layout of the container terminal used in this research](image)

The container unload operation using an AGV is explained now. Due to hard coupling between the AGVs and the QCs, the containers that are waiting to be unloaded need to first wait for an AGV availability (waiting time denoted by $W_v$). When an AGV is available and the container needs unloading, it travels to the
quayside (travel time denoted by $T_{v1}$). Then the AGV may wait for the QC to be available after which the QC repositions the container from the vessel to the AGV (the waiting time and repositioning time denoted by $W_q$ and $T_q$ respectively). Then the AGV, loaded with a container, travels to the stackside, may wait for the SC availability. Once a SC is available, the crane travels to the stack buffer lane and picks the container from the AGV. The container is then stored in the stack area. The AGV travel time to the stackside, waiting time for the SC, and the crane travel times are denoted by $T_{v2}$, $W_s$ and $T_s$ respectively. Using these travel and wait time components, the throughput time for the unload operations with the AGVs is expressed using Equation 1.

\[
CT_u = W_v + T_{v1} + W_q + T_q + T_{v2} + W_s + T_s
\]

To determine the optimal layout of the terminal, the number of storage locations, number of vehicles ($V$), and the number of quay cranes ($N_q$) are fixed; we vary the number of stacks ($N_s$), number of rows per stack ($N_r$), bays per stack ($N_b$), and tiers per stack ($N_t$). By varying the four parameters, $N_s$, $N_r$, $N_b$, and $N_t$, the length of the vehicle guide path is also altered (Figure 2), which affects the unload throughput time, $CT_u$. The optimization formulation to determine the optimal combination of the four design variables is presented in Equation 2. The objective function is to minimize $E[CT_u]$, subject to the network throughput ($X(V)$) stability constraint with $V$ vehicles, fixed locations constraint ($C$), vehicle utilization constraint ($U(V)$), and upper and lower bound constraints for the decision variables. To determine the optimal terminal layout configuration for unloading operations with AGVs, we analyse alternate configurations for different combinations of design parameter settings using the integrated queuing network model (described in the following section).

\[
\begin{align*}
\text{minimize} & \quad E[CT_u] (N_q, N_s, N_r, N_b, N_t, V) \\
\text{subject to} & \quad X(V) \geq \lambda_u \\
 & \quad N_t N_s N_r N_b = C \\
 & \quad U(V) \geq U_{\text{min}} \\
 & \quad N_{\text{min}} \leq N_t \leq N_{\text{max}} \\
 & \quad N_{\text{min}} \leq N_r \leq N_{\text{max}} \\
 & \quad N_{\text{min}} \leq N_b \leq N_{\text{max}} \\
 & \quad N_{\text{min}} \leq N_s \leq N_{\text{max}} \\
 & \quad N_t, N_s, N_r, N_b, N_t \in \mathbb{Z}^+ 
\end{align*}
\]
3. Queuing Network Model for Terminal Operations with AGVs

In this section, we develop the model of the unloading operations at a container terminal using AGVs. In an AGV-based system, both the QC and the SC drops-off (picks-up) the container on (from) the top of the vehicle. Therefore there is a hard coupling between the vehicle and the QC / SC. We first discuss the protocols that we develop to model the AGV-based terminal operations.

1. **Synchronization protocol at the quayside:** For the unloading operation, the QCs begin their operation only when an empty AGV has arrived at the buffer lane to transport the container. Similarly, for the loading operation, the QCs begin their operation only when an AGV loaded with a container has arrived at the quay buffer lane from the stackside.

2. **Synchronization protocol at the stackside:** For the unloading operation, the SCs begin their operation only when an AGV loaded with a container has arrived at the stack buffer lane to store the container. Similarly, for the loading operation, the SCs begin their operation only when an empty AGV has arrived at the stack buffer lane to transport the container to the quayside.

We now list the modelling assumptions for the three processes.

**Quayside process:** We assume that there is one trolley / QC. Further, there is infinite buffer space for parking vehicles at the QC location. The dwell point of QCs is the point of service completion. Containers arrive in single units with exponential inter-arrival times. Further, containers are randomly assigned to a QC.

**Vehicle transport process:** Each AGV can transport only one container at a time. The dwell point of the vehicles is the point of service completion. The vehicle dispatching policy is FCFS and the blocking among vehicles at path intersections is not considered. Further, vehicle acceleration and deceleration effects are ignored.

**Stackside process:** We assume that the stack layout is perpendicular to the quay and there is one crane per stack. The dwell point of cranes is the point of service completion. Similar to the quayside, we also assume infinite buffer space for parking vehicles at the SC location. Containers are randomly assigned to a SC.
3.1 Model Description

The inputs to the queuing network model are the first and second moment of the container inter arrival times \( \left( \lambda_a^{-1}, c_a^2 \right) \), and the service time information at the resources. Each QC is modelled as a single server FCFS station with general service times. Like-wise each SC is modelled as a single server FCFS station with general service times. The components of the AGV travel times are modelled as IS stations \( VT_1 \) and \( VT_2 \). The AGVs circulate in the network processing container movements.

We now describe the routing of the AGVs and containers in the queuing network model with respect to the unloading operations. Figure 4 describes the queuing network model of the container unloading process with AGVs. The containers that need to be unloaded, wait for an available vehicle at buffer \( B_1 \) of the synchronization station \( J \). Idle vehicles wait at buffer \( B_2 \). The physical location of the vehicles waiting in buffer \( B_2 \) would correspond to the stackside buffer lanes. Once a vehicle and a container is available to be unloaded, then the vehicle queues at the IS station \( VT_1 \). The expected service time at \( VT_1, \mu_{t1}^{-1} \), denotes the expected travel time from its dwell point (point of previous service completion) to the QC buffer lane. After completion of service, the vehicle queues at the QC station \( QC_i, i = 1,\ldots,N_q \) to pick up the container. The expected service time at \( QC_i, \mu_{qi}^{-1} \), denotes the expected movement time of the QC to reach the container in the vessel, container pickup time, movement time to reach the AGV, and container dropoff time. Then, the vehicle queues at the IS station: \( VT_2 \). The expected service time at \( VT_2, \mu_{t2}^{-1} \) denotes the expected travel time from the QC buffer lane to the SC buffer lane. After completion of service at \( VT_2 \), the vehicle queues at the SC station \( SC_i, i = 1,\ldots,N_s \) to dropoff the container. The expected service time at \( SC_i, \mu_{si}^{-1} \), denotes the expected travel time of the SC from its dwell point to the stack buffer lane and the container pickup time. Once the container is picked up from the AGV, the AGV is now idle and available to transport the containers that are waiting to be unloaded at the quayside.
Note that due to random assignment of containers to a QC and random storage of a container at a stack block, the routing probabilities from station $VT_1$ to $QC_i$ ($i=1,\ldots,N_q$) and from $VT_2$ to $SC_i$ ($i=1,\ldots,N_s$) are $\frac{1}{N_q}$ and $\frac{1}{N_s}$, respectively.

The queuing network in Figure 4 is a semi-open network model because the model possesses the characteristics of both open as well as closed queuing networks. The model is open with respect to the transactions and closed with respect to the vehicles in the network. Due to non-product form nature of the integrated network, an approximate procedure is developed to evaluate the network. First, a sub-network of the original network is replaced by a load-dependent server. The service rates correspond to the throughput of a closed queuing network (sub-network). The reduced model is evaluated using a continuous time Markov chain (CTMC). This approximate procedure provides substantial computational advantage in evaluating the integrated queuing network and estimating performance measures. By accounting for the stochastic interactions among quay cranes, vehicles, and stacking cranes, realistic estimates of system performance measures such as throughput capacity, resource utilization, the container waiting times for resources, and the expected cycle times are obtained. The expressions for the service times at various nodes and detailed description of the solution methodology are included in our working paper (6).
4. Numerical Experiments and Insights

We considered a container terminal scenario with a quay crane capacity of 30 cycles / hr, 40 AGVs, each stack has 6 rows, 40 bays, and 5 tiers. The total number of container storage locations is fixed at 48000, which corresponds to the capacity of the stacking lanes to serve a deep-sea vessel at the ECT terminal at Rotterdam. The travel velocity of the AGV and the SC are assumed to be 6 m/s and 3m/s respectively. The area of the AGV path is 540m × 90m. There are 5 buffer lanes per stack block.

We validate the analytical model for the container terminal with AGVs using detailed simulations. The average percentage absolute errors in the expected queue lengths and the expected throughput times are less than 7%. To determine the optimal terminal layout configuration we varied the design parameters in the following manner: number of stack blocks is varied between 20 and 120 with an increment size of 20, number of rows / stack is varied between 4 and 10 with an increment of 2, number of tiers / stack is either 3 or 5.

The expected throughput times are determined for all possible layout combinations. Table 1 includes five poor layout choices whereas Table 2 includes five good layout choices. The results suggest that a small number of stack blocks in combination with a small number of bays / block are a better design choice than either a large number of stack blocks in combination with a small number of bays / block or a small number of stack blocks in combination with a large number of bays / block.

Table 1. Poor terminal layout design choices

<table>
<thead>
<tr>
<th>$N_2$</th>
<th>$N_r$</th>
<th>$N_b$</th>
<th>$N_t$</th>
<th>$U_v$</th>
<th>$E[T_{u}]$ (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>6</td>
<td>89</td>
<td>3</td>
<td>26.8%</td>
<td>1338.7</td>
</tr>
<tr>
<td>110</td>
<td>8</td>
<td>11</td>
<td>5</td>
<td>95.9%</td>
<td>1399.9</td>
</tr>
<tr>
<td>110</td>
<td>8</td>
<td>19</td>
<td>3</td>
<td>95.9%</td>
<td>1413.1</td>
</tr>
<tr>
<td>20</td>
<td>10</td>
<td>80</td>
<td>3</td>
<td>24.6%</td>
<td>1772.4</td>
</tr>
<tr>
<td>20</td>
<td>6</td>
<td>80</td>
<td>5</td>
<td>20.0%</td>
<td>2000.4</td>
</tr>
</tbody>
</table>

Table 2. Good terminal layout design choices

<table>
<thead>
<tr>
<th>$N_2$</th>
<th>$N_r$</th>
<th>$N_b$</th>
<th>$N_t$</th>
<th>$U_v$</th>
<th>$E[T_{u}]$ (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>10</td>
<td>32</td>
<td>5</td>
<td>34.0%</td>
<td>787.9</td>
</tr>
<tr>
<td>40</td>
<td>8</td>
<td>30</td>
<td>5</td>
<td>38.6%</td>
<td>803.4</td>
</tr>
<tr>
<td>30</td>
<td>8</td>
<td>40</td>
<td>5</td>
<td>30.4%</td>
<td>813.3</td>
</tr>
<tr>
<td>40</td>
<td>10</td>
<td>24</td>
<td>5</td>
<td>43.5%</td>
<td>817.5</td>
</tr>
<tr>
<td>40</td>
<td>6</td>
<td>40</td>
<td>5</td>
<td>33.7%</td>
<td>817.8</td>
</tr>
</tbody>
</table>
5. Conclusions

In this research, we develop an integrated analytical model for the unloading operations in the container terminal using Automated Guided Vehicles. Numerical experiments suggest that stack configuration with a small number of stacks and a small number of bays (30 stacks, 30 bays) yields better throughput performance than a small number of stacks and a large number of bays (20 stacks, 80 bays). We believe that the stochastic models of the container handling operations can be used for rapid analysis of multiple design configurations for container port terminals and improve container-handling efficiencies.

Acknowledgements

The authors thank the SmartPort Centre at Erasmus University for partly funding this research.
6. References


Publication list
René de Koster
Article / Letter to the editor


Yu, Y., Koster, M.B.M. de & Guo, X. (2015). Class-based storage with a finite number of item: using more classes is not always better. Production and Operations Management, Accepted. doi: http://dx.doi.org/10.1111/poms.12334


Yu, Yugang & Koster, René de (2009). Optimal zone boundaries for two-class-based compact three-dimensional automated storage and retrieval systems. IIE Transactions, 41 (3), 194-208. doi: http://dx.doi.org/10.1080/07408170802375778

Gong, Y. & Koster, René de (2008). A polling based dynamic order picking system for online retailers. IIE Transactions, 40 (11), 1070-1082. doi: http://dx.doi.org/10.1080/07408170802167670


Lindert, M. te & Koster, René de (2003). Eerlijk delen is noodzakelijke voorwaarde voor factory gate pricing. LogistiekKrant, 16 (14), 10-11.


Part of book - chapter


PhD candidates

supervised by René de Koster
Meer, J.R. van der (2000, September 28). Operational Control of Internal Transport. Erasmus University Rotterdam (180 pag.) (Rotterdam: Erasmus Research Institute of Management (PhD Serie 001)). Prom./coprom.: René de Koster & prof.dr.ir. R. Dekker.
Senior Project Manager, UNC Plus Delta, Amstelveen

Professor, University of Groningen

Professor, University of Groningen

De Brito, Marisa (2004, February 12), Managing Reverse Logistics or Reversing Logistics Management? Erasmus Universiteit Rotterdam (166 pag.) (Rotterdam: Erasmus Research Institute of Management (PhD Serie 035)). Prom./coprom.: Rommert Dekker & René de Koster.
Senior Lecturer/Researcher “Events & Placemaking”, NHTV Breda University of Applied Sciences

Associate Professor, Electric Power University, Vietnam

Manufacturing and Supply Chain Manager, Akzo Nobel, Hochiminh city, Vietnam
Senior Scientist, TNO Transpor & Mobility, Delft

Supply Chain Analyst, Biomet Global Supply Chain Center

Associate Professor, EMLYON Business School, University of Lyon, France

Profesor a Tiempo Parcial, UTEC - Universidad de Ingeniería y Tecnología, Lima, Peru

Assistant Professor, Texas A&M University, Galveston, USA

Zaerpour, N. (2013, February 22). Efficient management of compact storage systems. Erasmus University Rotterdam (222 pag.) (Rotterdam: Erasmus Research Institute of Management (PhD Serie 276)). Prom./coprom.: prof.dr.ir. M.B.M. de Koster & Dr. Y. Yu.
Assistant Professor, VU University Amsterdam
Data Scientist, Siili Solutions, Helsinki, Finland

Faber, N. (2015, September 17). Structuring Warehouse Management: Exploring the fit between warehouse characteristics and warehouse planning and control structure, and its effect on warehouse performance. Erasmus University Rotterdam (175 pag.) (Rotterdam: Erasmus Research Institute of Management (PhD Serie 336)). Prom./coprom.: prof.dr.ir. M.B.M. de Koster & prof.dr.ir. A. Smidts. Associate professor logistics and supply chain management, Dutch Defense College, Breda

Li, Dong (2015, September 18). Supply Chain Contracting for After-Sales Service and Product Support. Erasmus University Rotterdam (163 pag.) (Rotterdam: Erasmus Research Institute of Management (PhD Serie 347)). Prom./coprom.: prof.dr.ir. M.B.M. de Koster & dr. N. Mishra. Research Fellow, Singapore University of Technology and Design (SUTD), Singapore

J. de Vries, Behavioral Operations in Logistics, promotor: M.B.M. de Koster, copromotor: Dr. D.A Stam, promotion: February 2016 (expected), Assistant professor, VU University Amsterdam

J.P. van der Gaast, Stochastic models for order picking systems, promotores: M.B.M. de Koster, Prof.dr.ir. I.J.B.F. Adan, promotion, February 2016 (expected)
Erasm us Research Institute of Managem ent -
On the occasion of René de Koster’s 20 years full professorship
Rotterdam School of Management, Erasm us University


Erasm us Research Institute of Managem ent (ERIM) is the Research School (Onder-zoekschool) in the field of managem ent of the Erasm us University Rotterdam. The founding participants of ERIM are the Rotterdam School of Management (RSM), and the Erasm us School of Econo mics (ESE). ERIM was founded in 1999 and is officially accredited by the Royal Netherlands Academy of Arts and Sciences (KNAW). The research undertaken by ERIM is focused on the managem ent of the firm in its environment, its intra- and interfirm relations, and its business processes in their interdependent connections.

The objective of ERIM is to carry out first rate research in managem ent, and to offer an advanced doctoral programme in Research in Managem ent. Within ERIM, over three hundred senior researchers and PhD candidates are active in the different research programmes. From a variety of academic backgrounds and expertises, the ERIM community is united in striving for excellence and working at the forefront of creating new business knowledge.

Inaugural Addresses Research in Managem ent contain written texts of inaugural addresses by members of ERIM. The addresses are available in two ways, as printed hard-copy booklet and as digital fulltext file through the ERIM Electronic Series Portal.