Managing warehouse efficiency and worker discomfort through enhanced storage assignment decisions

José Antonio Larco, René de Koster, Kees Jan Roodbergen & Jan Dul

Taylor & Francis
Taylor & Francis Group

Check for updates

Managing warehouse efficiency and worker discomfort through enhanced storage assignment decisions

José Antonio Larco[a], René de Koster[b], Kees Jan Roodbergen[c]* and Jan Dul[b]

[a] *Faculty of Industrial Engineering, Universidad de Ingeniería y Tecnología, Barranco, Lima, Peru;* [b] *Rotterdam School of Management, Erasmus University, Rotterdam, The Netherlands;* [c] *Faculty of Economics and Business, University of Groningen, Groningen, The Netherlands*

Humans are at the heart of crucial processes in warehouses. Besides the common economic goal of minimising cycle times, we therefore add in this paper the human well-being goal of minimising workers' discomfort in the context of order picking. We propose a methodology for identifying the most suitable storage location solutions with respect to both goals. The first step in our methodology is to build data-driven empirical models for estimating cycle times and workers' discomfort. The second step of the methodology entails the use of these empirically grounded models to formulate a bi-objective assignment problem for assigning products to storage locations. The developed methodology is subsequently tested on two actual warehouses. The results of these practical tests show that clear trade-offs exist and that optimising only for discomfort can be costly in terms of cycle time. Based on the results, we provide practical guidelines for taking storage assignment decisions that simultaneously address discomfort and travel distance considerations.

Keywords: order picking; warehousing; cycle time; discomfort

1. Introduction

Material handling operations have received considerable attention in the literature with a particular focus on order-picking operations. To maximise the efficiency of order picking, several approaches have been proposed in the literature (De Koster, Le-Duc, and Roodbergen 2007) that generally aim at maximising efficiency by minimising travel distances. However, as walking distances decrease by various means, the relative importance of other activities will increase. Specifically, most papers on order picking do not consider the time spent on retrieving and searching for products, even though these activities may account for 35% of total picking time (Tompkins et al. 2010). Typically, it is assumed explicitly or implicitly that each pick requires the same amount of time, regardless of the height of the pick location, the quantity picked or the volume and mass of the product.

Only few papers have recognised the influence of rack height on order-picking times. These papers mostly use a Golden Zone strategy, implying that frequently retrieved products should be located at a height between the waist and the shoulders of 'average' pickers (Saccomano 1996; Jones and Battieste 2004; Petersen, Siu, and Heiser 2005). The economic justification for this is that locations within the Golden Zone are expected to take less time to identify and retrieve products from than locations outside this zone. However, the effectiveness of storage location selection on picking efficiency has not been tested empirically for different contexts nor is it known how to quantify such effect.

Proper positioning of products also has a social justification when considering the well-being of order pickers since this may reduce the incidence of working in uncomfortable postures (Jones and Battieste 2004). Discomfort felt by employees is a pervasive problem in warehouses and has been found to be a predictor for future long-term muscular pain (Hamberg-van Reenen et al. 2008) as well as occupational disorders such as the so-called low back disorders (LBDs). The importance of LBDs is highlighted by reports of an American insurer that LBD-related claims account for 33% of total worker claim costs (Webster and Snook 1994). Reducing discomfort is then directly related to the well-being of workers and may also yield long-term economic benefits through higher productivity (Kuijt-Evers et al. 2007), reduced worker health-related costs, absenteeism and drop-out rates. The social justification of storage location selection has not been empirically tested. It remains to know what the effect of locating products is with respect to discomfort measures in various contexts.

In this paper, we first propose for warehouse order picking a unified methodology to quantify and balance two potentially conflicting criteria: (1) the short-term economic criterion of minimising total order-picking time, and (2) the human well-being criterion of minimising average discomfort ratings. Note that, as indicated above, long-term health savings due to the

second criterion may positively impact the first criterion, however, this is difficult to quantify and therefore disregarded in this paper. Our proposed methodology enables an explicit consideration of both goals and the interplay between them in the warehouse design practice.

Secondly, this paper presents results from applying the proposed methodology to two order-picking operations from practice. The first situation considered is at the fast-moving products area in the parts distribution center of Yamaha Motors. The second application is at the Sorbo Distribution Center, an importer and distributor of non-food products for supermarkets. The results from these two studies serve to point out the importance of the presented methodology. Furthermore, based on the results, we give guidelines to obtain solutions that balance both economic and well-being goals with the use of simple decision rules that require only a modest effort in collecting data.

The novelty of our methodology is visible in a number of dimensions. Firstly, we provide an interface between insights of operations management and insights from human sciences in a warehousing context. In a sense, this paper can be considered a response to the challenge posited by Boudreau et al. (2003), Gino and Pisano (2008), and Grosse et al. (2015) to include human aspects in conventional operations decisions. In warehousing operations, only few other papers have aimed to incorporate human aspects into relevant decisions, see Lodree (2008), De Koster, Stam, and Balk (2011), Doerr and Gue (2013), and Battini et al. (2015).

Furthermore, our approach is novel since we present a methodology that can be used to capture, quantify and analyse the actual effects of human factors in practice, such that they are directly usable in common operational modelling constructs. Our research goes beyond modelling papers, such as Petersen, Siu, and Heiser (2005), by providing an empirically grounded, quantitative basis instead of presuming certain human well-being effects to be present. And we extend beyond our preliminary studies in Larco et al. (2008), by describing a structured methodology, improved estimation methods and new trade-off results.

It must be noted that the proposed combination of goals may be non-trivial as has been often claimed in practice and in the scientific literature. Peacock (2002) suggests that a tension between human-centred criteria and operational performance criteria exists, in particular if the operational performance is only short-termed. On the other hand, Dul et al. (2004) show that well-being goals expressed in ergonomic standards may yield economic benefits, which suggests a certain degree of alignment between well-being and economic goals. This paper sheds light on this debate in the warehousing context.

The paper is organised as follows. In Section 2, we present an overview of the proposed methodology for storage location decisions, combining an empirical study and a multi-criteria optimisation study. Next, in Section 3, we present the results of applying our methodology on two distinct warehouses. Based on the empirical results of Section 3, Section 4 presents a heuristic to obtain good location solutions that balance the economic and well-being objectives. Furthermore, Section 4 explores the trade-off between economics and well-being further and provides insights on when such trade-off may be more significant. Finally, in Section 5, we give conclusions and discuss limitations of this study that offer further research opportunities.

2. Methodology

We restrict our study to picker-to-part order picking systems where workers walk and retrieve products from shelves (cf. Tompkins et al. 2010). These systems are often organised into zones, where orders are partially picked in each zone and transferred between zones via conveyors; these are typically referred to as pick-and-pass configurations. These systems are used by numerous warehouses (De Koster, Le-Duc, and Roodbergen 2007) and are typically characterised by many picks per time unit with a relatively limited amount of walking. Furthermore, each picker only picks one or a few products per order from multi-levelled shelves, since the remainder of the order is picked by other pickers in other zones.

For the purpose of making storage location decisions that consider both an economic and a well-being objective, we present a two-phase methodology. First, in the *effect quantification phase*, the effects of location and product factors on cycle time and discomfort are determined using data from the warehouse management system (WMS) as well as data that needs to be actively collected. Second, in the *multi-criteria optimisation phase*, we use results from the effect quantification phase to construct two two-entry matrices, one for estimated cycle time and another for estimated discomfort ratings for every possible product-to-location assignment. This results in a bi-objective assignment model where each product is to be assigned to a single storage location. With this assignment model, the efficient frontier of cycle time and discomfort goals can be identified, along with the set of non-dominated solutions that generate such an efficient frontier. It is then up to the warehouse manager to choose from the non-dominated solutions. In the following, we describe our methodology in detail.

2.1 *Quantifying effects on cycle time*

We define the cycle time as the time lapse from the receipt of an orderline at the picker's terminal until dropping the products in a bin at the depot. The picking cycle can be broken down into the following activities: receipt of a new orderline, walking to the picking location, searching the specific location and retrieving the units of a product from such location, walking back

Table 1. Picking time breakdown analysis.

	Main effects								Interaction effects		
	Location factors				Product factors						
Picking sub-activities	MA	CA	CN	$L^{(k)}$	LB	M	V	Q	$M * L^{(k)}$	$V * L^{(k)}$	$Q * L^{(k)}$
New orderline receipt	-	-	-	-	-	-	-	-	-	-	-
Walk to picking location	x	x	x	x	-	?	?	?	?	-	-
Retrieve & search item	-	-	-	x	x	x	x	x	?	?	?
Return to depot	x	x	x	x	?	?	?	?	-	-	-
Drop picked item(s)	-	-	-	-	-	?	?	?	-	-	-
Confirm orderline	-	-	-	-	-	?	?	x	-	-	-

with the units to the depot, dropping off the picked products at the depot, and finally confirming the pick at the depot. A number of drivers can influence each of the order picking activities. We classify these drivers either as location factors or as product factors. The *location factors* consider the position of a product in the shelves of a warehouse with a standard layout of parallel shelf racks and one cross aisle. The *product factors* include the mass of the product, the volume of the product and the quantity picked in a single orderline. All location and product factors can usually be obtained from the WMS (WMS). We use the following notation to denote the location and product factors that may have an impact on order picking cycle times:

Location factors

MA : Section number in the main aisle.
CA : Section number in the cross aisle (if applicable).
CN : Cross aisle number (if applicable).
K : Picking levels set $K = \{1, 2, 3\}$ where 1 is the lowest level and 3 the highest.
$L^{(k)}$: 1 if picked at level $k \in K$; 0 otherwise.
LB : type of bin, $LB = 1$ if picked from a large bin; $LB = 0$ for a small bin.

Product factors

Q : Quantity picked.
M : Unit mass of the product.
V : Unit volume of the product.

It is natural to assume that while two-dimensional (2D) location factors (i.e. MA, CA and CN) affect walking times, the different picking levels only influence retrieving and searching times. The product factors mainly influence retrieving times and possibly walking times due to greater difficulty in carrying products. It may be possible, though, that product factors also influence the time it takes to drop off products, the time to return to the depot or even the time to confirm an order. In particular, the quantity picked may influence the time to confirm an order as products are counted in this activity. Finally, in the case of retrieving times there are possible interaction effects between the pick level and the product factors. For example, it may well be that it is additionally difficult to manipulate heavy and/or voluminous products at a top storage level than at an intermediate level, however this is to be verified empirically. We summarise our hypotheses in Table 1 marking likely effects with an 'x' and potential effects with a '?'.

One approach to estimate cycle times concerns the use of predetermined motion time systems (PDMTS). The use of PDMTS, like MOST (Zandin 1990), requires to observe actual work and then disaggregate the cycle of a job into motions, specifying characteristics of each motion. Products factors are not accounted for. Hence, to include product factors, pairwise comparisons would be needed. For example, if one is interested in the effect of retrieving a 'heavy' product vs. a 'light' product from a Golden Zone position, then the cycle time of both would need to be determined separately and then subtracted. Hence, separate measurements would be required for every combination of factors: height, location, product mass and volume. Since such information cannot be derived from WMS data, this method demands an extensive effort of firms, requiring them to record order-picking operations for all the combinations of factors needed in a study. Furthermore, PDMTS makes fundamental assumptions (Genaidy, Mital, and Obeidat 1989) that may not hold in a warehousing context, and should be tested empirically for each new study. Notably, the expected time taken to execute a motion is assumed independent of other motions. As activities like searching and walking may interact, this assumption of motion independence may likely be violated.

Alternatively, information from the WMS can be used as input for a linear regression to estimate cycle times. Using this method, it is possible to obtain a large set of observations with little effort as the day-to-day operational data is automatically stored in the system and the impact of several cycle time drivers can be quantified simultaneously. The method allows us to use observations taken under normal operating conditions without any interference of a video camera or a researcher, hence possible distortions on the data-set are minimised. However, using data directly from a WMS also has its limitations, such as the inability to control variables. By taking a large number of observations from the WMS, this potential limitation can be mitigated since all situations are then likely to occur. Another limitation is the fact that cycle times may not be measured directly in the WMS. Often only the time lapse between the starting times of two consecutive picks is known, but not the actual time for the activity. Hence, the raw data may also include breaks, waiting times related to system breakdowns, and idle times due to slack capacity in the system. The impact of such outliers must then be minimised by means of additional statistical methods (Wisnowsky 1999).

Although both approaches are feasible, we opted for regression modelling because of the widespread availability of WMS data and the fact that this puts a lower burden on warehouses to collect data. Furthermore, in Curseu et al. (2009) the use of linear regression has already been proven effective to estimate drivers of cycle time in retail material handling operations. Using Table 1, we formulate the picking cycle time, CT, in terms of the hypothesised effects. We detail how our model is built as follows. As the 2D location factors directly influence the walking distances, it is logical to assume that these factors (MA, CA and CN) linearly influence cycle times, where b_1, b_2 and b_3 are the corresponding linear coefficients to be estimated. The different levels are modelled as dummy variables, $L^{(k)}$, with one level, level k^*, as the reference level to avoid perfect multicollinearity, with $\alpha^{(k)}$ as coefficients. The quantity to be picked is also assumed to influence the retrieving and dropping-off times linearly. This is actually an approximation of a more complex relationship as certain products may be grabbed in batches. Note that the incremental pick quantity over 1 is used in the model and not the quantity picked itself. We can thus interpret coefficient b_4 as the additional time required to pick one additional unit. The main effects are completed by not making *a priori* assumptions on the effects of mass and volume, using general functions $f(M)$ and $g(V)$ as these are unknown and several functional forms must be tested. Further, to account for possible interaction effects of quantity, mass and volume effects with heights $k \in K$, we introduce linear coefficients $\beta^{(k)}$, $\gamma^{(k)}$ and $\lambda^{(k)}$, respectively. The formulation for cycle time is given by:

$$CT = b_0 + b_1 MA + b_2 CA + b_3 CN + \sum_{k \in K, k \neq k^*} \alpha^{(k)} L^{(k)}$$
$$+ b_4(Q - 1) + b_5 f(M) + b_6 g(V) + b_7 LB + IN + \varepsilon, \tag{1}$$

where IN contains the interaction effects given by the following relationship:

$$IN = (Q - 1) \sum_{k \in K, k \neq k^*} \beta^{(k)} L^{(k)} + f(M) \sum_{k \in K, k \neq k^*} \gamma^{(k)} L^{(k)} + g(V) \sum_{k \in K, k \neq k^*} \lambda^{(k)} L^{(k)} \tag{2}$$

The procedure we propose to counter the effects of possible outliers is as follows. First, conduct a limited observational study to obtain a (course) estimate of cycle times, and to identify the approximate sizes of included waiting times. The first step is to set cut-off times to eliminate 'obvious' outliers from the preliminary study. These cut-off times are based on the maximum observed cycle times in the preliminary study, typically with a safety margin. The second step is statistical in nature and can then be used for the remaining observations.

To statistically address the outliers, a number of techniques are available. Classical identification techniques for outliers that use common distance measures such as Mahalanobis or Cook's fail in our study, because the computed distance measures are based on the covariance matrix of the observations which may be already biased towards the outliers (Wisnowsky 1999). As a result, these techniques suffer from either masking errors where outliers are falsely classified as inliers or swamping errors where inliers are classified falsely as outliers.

More advanced techniques that deal with multiple outliers exist. Although each technique has its advantages and drawbacks and no specific multiple-outlier analysis technique has been deemed superior under all circumstances, robust regression remains a widely used and flexible technique for dealing with multiple outliers (Rousseeuw and Leroy 1987). We thus select this technique for our methodology. In particular we use, M-robust regression (Huber 1964) which is suitable for the type of outliers present in WMS data: outliers in the Y-space, meaning response values that are deemed to be different than the response values of interest (Hampel et al. 1986). M-robust regression addresses the case of multiple outliers by assigning less weight to probable outliers in an iterative procedure.

The algorithm starts by running a simple OLS regression and then iteratively reduces the weights of the observations with greater residuals. The procedure is repeated until the change in the regression coefficients is negligible. There are several functions available for the objective function and the weights w_t. In particular, we use the method of Huber (1964). In this

case, the weights are given by $w_t = 1$ if $|e_t| \leq c$, and $w_t = c/|e_t|$ otherwise. Here $c = 1.345\sigma$ is a constant, with σ the standard deviation of the errors, and e_t is the residual of the t-th observation obtained from the previously applied weighted regression.

2.2 *Quantifying effects on discomfort*

To evaluate overall physical discomfort we use Borg's CR-10 scale (Borg 1982; Dul, Douwes, and Smitt 1994) which is commonly used in the ergonomics field. The scale combines desirable ratio and categorical properties by assigning labels for values from 0 to 10. In this way, 0 stands for no discomfort at all, 2 for weak discomfort, 3 for moderate discomfort, 5 for strong discomfort and 10 for the maximum discomfort, which requires the person to immediately stop the work activity. Values can be obtained from direct feedback of the workers on the job collected by an evaluator. There are two main advantages of directly inquiring after a pick for the perceived discomfort of the pick. First, the picker can concentrate fully on his task without having to write down the ratings himself which would interfere with a normal work-flow. Second, the picker is urged to state his rating immediately, thus avoiding any ex-post rationalisations of his rating and enhancing the recall of the picking experience. At the same time, the evaluator records relevant location factors (shelf height) and product factors (mass, volume and quantity). Depending on the situation these can be obtained from the WMS or recorded manually.

To estimate the effects of location and product factors on discomfort, we propose to use an ordinary least squares model. Besides the variables to quantify the effects of location and product factors, we use additional dummy variables to control for individual differences in evaluating ratings caused by different personal traits including mood and sensitivity to discomfort. These dummy variables are denoted by $E^{(r)}$ where r is the identifier of the employee such that $r \in R$, and r^* is the employee taken as reference.

To predict the discomfort rating for a product picked at a certain location, we use a similar formulation as was used for predicting the cycle time, we estimate a linear relationship with main effects where b_1, b_2, b_3, b_4, b_5 and b_6 are the linear coefficients to be estimated using the following equation:

$$D = b_0 + \sum_{k \in K, k \neq k^*} \alpha^{(k)} L^{(k)} + b_1 HM + b_2 MV + b_3 HV + b_4 MQ + b_5 HQ + \sum_{r \in R, r \neq R^*} b_6 E^{(r)}$$
$$+ IND + \varepsilon. \tag{3}$$

We choose low mass, low volume and small pick quantities as reference values and focus on additional discomfort generated by products with high mass (HM), medium and high volume (MV, HV) and medium and high pick quantities (MQ, HQ). The term IND allows for possible interaction effects with picking at different levels and assumes that the nature of such interactions is linear with coefficients $\beta^{(k)}$, $\gamma^{(k)}$, $\lambda^{(k)}$, $\eta^{(k)}$ and $\zeta^{(k)}$ to be estimated.

The term IND is given by:

$$IND = HM \sum_{k \in K, k \neq k^*} \beta^{(k)} L^{(k)} + MV \sum_{k \in K, k \neq k^*} \gamma^{(k)} L^{(k)} + HV \sum_{k \in K, k \neq k^*} \lambda^{(k)} L^{(k)}$$
$$+ MQ \sum_{k \in K, k \neq k^*} \eta^{(k)} L^{(k)} + HQ \sum_{k \in K, k \neq k^*} \zeta^{(k)} L^{(k)}. \tag{4}$$

2.3 *Analysing storage-location trade-offs*

In contrast to most storage location problems in the literature, we seek not only the economic objective of minimising the cycle time, but also the social objective of minimising the perceived average discomfort. This implies we have two potentially contradicting objectives. We therefore aim for identifying a set of non-dominated solutions from which a decision-maker may choose.

To formulate the model we first define the following:

Sets

I : the set of all products to be stored,
J : the set of all possible storage locations.

Variables

x_{ij} : equals 1 if product i is stored at location j; and 0 otherwise.

Model parameters

D_{ij} : is a function that assigns for each possible combination of products and storage locations $(i, j) \in I \times J$, an expected discomfort measure such that $D_{ij} \in [0; 10]$,

CT_{ij} : is a function that assigns for each possible combination of products and locations $(i, j) \in I \times J$ an expected cycle time,

p_i : probability that whenever there is a pick, the product picked is $i \in I$.

We can now formulate the problem as follows:

$$z_1 = \min \sum_{i \in I} \sum_{j \in J} p_i CT_{ij} x_{ij} \tag{5}$$

$$z_2 = \min \sum_{i \in I} \sum_{j \in J} p_i D_{ij} x_{ij} \tag{6}$$

s.t.

$$\sum_{i \in I} x_{ij} \leq 1 \qquad \forall j \in J \tag{7}$$

$$\sum_{j \in J} x_{ij} = 1 \qquad \forall i \in I \tag{8}$$

$$x_{ij} \in \{0, 1\} \qquad \forall i \in I, \forall j \in J \tag{9}$$

The economic objective given by Equation (5) minimises the expected cycle time by multiplying the respective cycle times of a given product in a given location by the probability that such product is picked. This implies single-command cycles, i.e. the picker returns to the depot after each pick as is the case for unit-load operations (Ang, Lim, and Sim 2012). Analogously, the social objective in Equation (6) minimises the expected average discomfort rating by multiplying the respective discomfort rating of a given product-to-location assignment by the probability that such a product is picked. To obtain the discomfort rating D_{ij} for each product-to-location assignment we use Equation (3) with the coefficients found via the ordinary least-squares regression method. Constraints 7 require that at most one type of product can be stored in a single storage position. Constraints 8 require that every product be assigned to only one location. Implicitly we assume that there must be at least as many locations as products to be located, i.e. $|I| = \sum_{j \in J} \sum_{i \in I} x_{ij} \leq |J|$, otherwise the model would be infeasible.

The bi-objective assignment problem is known to be $\mathcal{NP}$-complete (Ehrgott 2000). However, when considering only one objective, the problem reduces to a classical assignment problem which can be solved in polynomial time. We use the Jonker and Volgenant (1987) algorithm to solve different assignment problem instances. Our main interest is to characterise the trade-off relationships between the economic and social objectives. Hence, we are interested only in non-dominated solutions that can be obtained as a convex combination of objective functions. To find such solutions we use the procedure used in Przybylski, Gandibleux, and Ehrgott (2008) modified to only find the vertex set of the convex hull of the decision space and thus, increase its computation speed. The procedure is given in Appendix 1.

3. Case results

To test our methodology, we applied it in two distinct warehouses that in some aspects are similar, but in other aspects distinct. The warehouses were selected for having the order-picking activities organised in zones with limited walking distances such that the retrieving times are an important component of the cycle time. In both warehouses products are stored in totes at multiple levels. On the other hand, we also ensured differences in layout and product factors to enable identification of common insights that may have the potential to apply to a larger class of order picking systems.

The first warehouse is the main distribution centre of Yamaha Motor Europe for motorised vehicles' spare parts. The warehouse has a large assortment of over 150,000 different products. We conducted a study in the main area for fast-moving products that is sub-divided into 32 zones. Within this area, each pick route visits exactly one location. Each zone uses a pick-to-light system and has a computer terminal next to the depot, where the picker scans the product, confirms the pick and views the next orderline to pick.

The second warehouse is that of Sorbo, an importer and distributor in the Netherlands of non-food products for supermarkets. Sorbo's warehouse is organised in 24 zones, where one picker per zone is responsible for picking products. Sorbo also uses a pick-to-light system, however, picked products need not be scanned at the depot. Confirmation of the picks
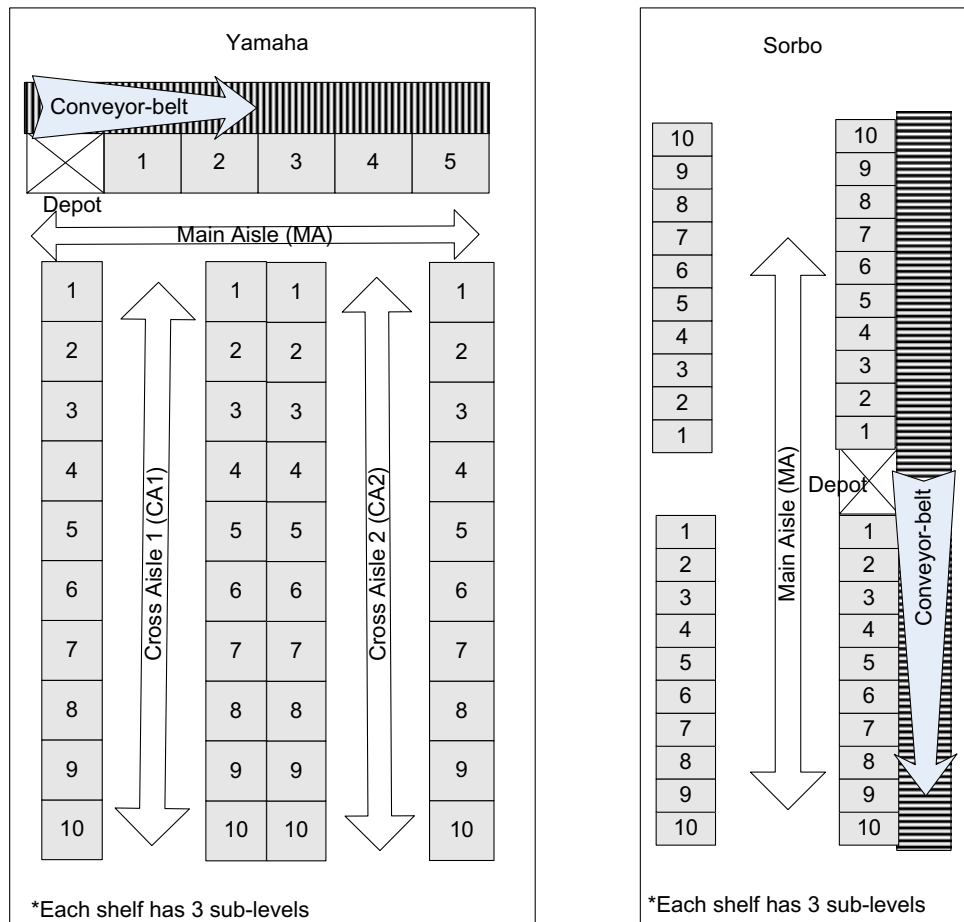
Figure 1. Layout of a typical picking zone at Yamaha and Sorbo, respectively.

is achieved by indicating the number of picked units at the picking locations. Most of the routes, about 85%, visit only a single location.

In both warehouses there are three equally spaced picking levels at heights ranging from 0.25 to 1.90 m in the case of Yamaha and 0.20 to 1.40 m in the case of Sorbo. Figure 1 presents the layout of a typical zone for each of the warehouses. While Yamaha has 145 locations available per zone, Sorbo has 120 locations per zone and a simpler layout. Yamaha has typically heavier and less voluminous products than Sorbo as Yamaha's products are mostly made of metal and Sorbo's products of plastic.

3.1 *Empirical results of cycle time estimation*

For Yamaha, we obtained 15,190 observations over a period of three days with two shifts per day from 20 different order pickers. In the case of Sorbo, we obtained 21,866 observations over a period of two days from 24 different pickers working in one shift. Following our methodology, we first established cut-off values to remove obvious outliers. For Yamaha, it was determined that main aisle picks do not exceed 52 s and that cross aisle picks do not exceed 55 s. For Sorbo, the cut-off time was established at 26 s. These cut-offs result in 13% (9%) of the observations to be deleted for Yamaha (Sorbo). This removes mostly the picks with waiting times caused by disruptions, which is evident from the fact that 80% of the removed observations exceeded twice the cut-off time. Once the more 'obvious' outliers were removed from the sample, the number of observations remaining for Yamaha and Sorbo are 13, 216 and 19, 898, respectively. At Yamaha (Sorbo) the average cycle time equals 25.139 s (14.485 s) with a standard deviation of 8.722 (5.312). For Yamaha the average quantity picked (Q) equals 1.694 units, the average unit volume of the products (V) is 1.128 dm^3, and the average unit mass of the product (M) is 0.19 kg, while for Sorbo these number are, respectively, 1.525 units, 0.708 dm^3 and 0.460 kg.

Table 2. Empirical cycle time, raw (i.e. Raw b) (in seconds) and standardised coefficients (i.e. Std. b) shown for Yamaha and Sorbo.

| | Yamaha | | | | Sorbo | | | |
| | OLS | | MRR | | OLS | | MRR | |
Variables	Raw b	Std. b	Raw b	Std. b	Raw b	Std. b	Raw b	Std. b
(Intercept)	18.105***	.	17.378***	.	8.310***	.	7.556***	.
MA	0.762***	0.164	0.621***	0.134	0.590	0.397	0.629	0.423
CA	1.257***	0.396	1.381***	0.435	.	.	.	.
CN	1.401***	0.179	1.551***	0.198	.	.	.	.
$L^{(1)}$	1.084***	0.029	1.203***	0.058	−0.131	−0.014	0.055	0.006
$L^{(3)}$	0.272	0.017	0.303*	0.019	0.068	0.003	0.172*	0.008
$Q - 1$	0.748***	0.156	0.963***	0.191	1.083***	0.549	1.222***	0.680
LV	−0.391*	−0.018	−0.537***	−0.025	−0.094	−0.010	−0.368**	−0.038
HV	0.182	0.005	0.522	0.013	1.031***	0.139	0.880***	0.119
LB	0.860***	0.040	0.894***	0.041	.	.	.	.
$(Q - 1) * L^{(1)}$	0.623**	0.026	0.524***	0.022	0.007	0.001	−0.103*	−0.019
$(Q - 1) * L^{(3)}$	0.045	0.006	0.036	0.005	0.121	0.121	0.033	0.001
$LV * L^{(1)}$	0.169	0.004	0.265	0.007	−0.501*	−0.018	−0.217	−0.008
$LV * L^{(3)}$	1.368***	0.074	1.335***	0.072	0.347	0.014	0.315	0.013
$HV * L^{(1)}$	1.299	0.019	0.824	0.012	0.820***	0.041	0.581***	0.029
$HV * L^{(3)}$	−0.016	−0.002	−0.105	−0.016	1.305***	0.037	0.993***	0.028
R^2	0.202	.	.	.	0.306	.	.	.

Notes: OLS: Ordinary Least Squares, M-R.R.: M-Robust regression method.
Significance levels: $p \leq 0.05$(*), $p \leq 0.01$(**), $p \leq 0.001$(***). Time given in seconds.

Using Equation (1), we tested the significance of the location and product factors and interaction effects. We used piecewise linear regression for $f(M)$ and $g(V)$ to explore possible non-linearities of mass and volume, varying categories to find the best fit to the data. The most parsimonious solution for both data-sets was to include only volume as a potential factor, thus excluding mass. The model then leaves LV and HV to represent high and low volumes, where $LV = 1$ if $V \leq 0.05$ dm^3 and $LV = 0$ otherwise, and $HV = 1$ if $V > 5$ dm^3 and $HV = 0$ otherwise. The results showed that volume is a better proxy than mass to handle complexities such as easiness to grab a product or easiness to retrieve a product from its location. This is confirmed by the fact that we did not find significant results when different mass categories were included, not even at a 0.1 significance level. For the range of product volumes studied, it appeared that in general, the larger the product picked, the more time it took to retrieve it. The final empirical cycle time estimation model for Equation (1) is given in Table 2 for the two warehouses, both using ordinary least squares (OLS) and M-robust regression (MRR).

The raw coefficients in Table 2 indicate the additional time in seconds if the corresponding variable is increased by a unit with respect to the reference value. In this way, for example, picking from a low level ($L^{(1)}$), implies an additional 1.203 s compared to the reference, middle level ($L^{(2)}$).

Both data-sets show that the OLS and MRR methods agree in the main effects, and to a certain extent in the magnitude. Hence the results are moderately robust, suggesting that the impact of the outliers is limited. As input for the next phase we continue with the results of MRR, knowing that we are likely to have remaining outliers in the data-set. Furthermore, MRR yields a higher fit than OLS (R^2 is 30.6% vs. 20.2% and 40.4% vs. 22.4% for Yamaha and Sorbo, respectively). In evaluating the fit of the model, it is important to note that a significant part of the variations in the dependent variable remains unexplained by the model. A variety of reasons can exist for this. Notably, the data remain to contain several unobserved effects, including employees taking micro pauses, employees correcting small mistakes in the confirmation of picks, small variations in the method used for retrieval actions, and brief delays in information processing at the terminals. We verified if fatigue may influence the results by incorporating dummy variables representing the shift intervals and found no significant effects. To account for heteroskedasticity, we used White's heteroskedasticity-consistent errors to determine p-values (Hayes and Cai 2007).

The results in Table 2 show for important commonality across both warehouses in the significance of location and product factors that account for cycle time. To rank the effects in descending order of importance, we make use of the standardised coefficients shown in Table 2 as these measure the 'the proportion of the greatest likely variation in the dependent variable

that can be accounted for by the greatest likely variation in the independent variable' (Luskin 1991, 1035). In this way, we observe consistency across both warehouses when ranking the factors in terms of importance. We mention them here in descending order: (1) 2D location factors (i.e. MA, CA, CN), (2) quantity to be picked, (3) height level of location, (4) product volume and (5) interaction factors.

Noteworthy is the effect of picking levels on cycle time for both data-sets. This result confirms that picking outside the Golden Zone does require additional retrieving time. However, the effects are more evident for Yamaha than for Sorbo because Yamaha has a larger range of heights. Moreover, at Sorbo the slanted design of the racks makes the lower level closer to the picker and thus compensates partly for the extra retrieving time of picking from such level. This may explain why no significant effect was found for picking at the lower level for Sorbo as opposed to Yamaha.

Interestingly, quantity and volume show significant interaction effects with the level. This means that the cycle time increases beyond the main effects of quantity and volume categories if picks occur outside the 'Golden Zone'. Hence, product factors are not only important for estimating cycle times but also for location decisions. The interactions are, however, present in different ways for both warehouses. Additional time is needed for at Yamaha for small products at upper levels and at Sorbo for large products at the lower and upper levels. The potentially counter-intuitive result at Yamaha can be explained by the fact that small products at the upper level are not always visible and need to be located by touch in the rack.

3.2 *Empirical results of discomfort estimation*

For the discomfort study, for a number of picks the perceived discomfort of the picker on the Borg CR-10 scale was recorded manually, supplemented with product and layout factors. The data were collected during two days in each warehouse, observing each employee for a full day. Data of five employees were collected for a total of 235 observations at Yamaha. For Sorbo, data of seven employees were collected for a total of 749 observations. The sample sizes and number of pickers involved are typical for discomfort studies (Kadefors and Forsman 2000). Moreover, as the objective of the discomfort study is to derive empirical relationships that are internally valid for the operation at hand, the sampling need only to involve the actual workers in the picking zones. If the demographics (age, sex) and body height of the workforce changes, however, a new discomfort study should be considered.

Dutch law requires prior approval by an ethics committee in case persons are subjected to treatment or are required to follow a certain behavioural strategy (WMO 1998). For our research, no tasks or activities of the persons involved have been altered by the researchers in any way. Persons have been solely observed and interviewed in their normal working environment while performing their normal daily activities. Hence, our research did not alter the state of mind or behaviour of persons and required no screening by one of the Dutch METC ethics committees. All persons first received an explanation of the data-collection methods and were then asked whether they were willing to participate; the options not to participate or to withdraw during the process were explicitly offered. No personally identifying information was recorded.

We classified the product as of moderately high volume (MV) if the volume is between 1 dm^3 and 5 dm^3, and of high volume (HV) if the volume is greater than 5 dm^3. Similarly, we classify picks to be of moderately high quantity (MQ) if it has more than three units, and of high quantity (HQ) if it has more than seven units in a single orderline. Thirdly, heavy products (HM), having a mass over 3 kg, are identified by the pickers themselves and communicated to the evaluator who checks the actual mass of the pick. Of course, for other warehousing contexts other classifications may be appropriate, which can easily be incorporated in the presented methodology.

The recorded mean (standard deviation) of CR-10 discomfort ratings in the study are 4.10 (2.02) and 2.98 (2.41) for Yamaha and Sorbo, respectively. Table 3 provides an overview of the number of observations for the studied factors across picking levels in the discomfort study. The dummy variables we introduced for controlling for individual differences between pickers did not impact the results significantly. For this reason and for the sake of conciseness, we report in Table 4 the results omitting the dummy variables of individuals even though these variables were included. The results appear to show consistency among pickers as evidenced by the high significance of the effects. Moreover, there is also consistency within individuals as they tended to rate the same type of actions similarly. This suggests that pickers can indeed distinguish between differences in discomfort.

The analysis of discomfort ratings using OLS is given in Table 4. The raw coefficients can be interpreted similar to the cycle time study. For example, picking at the lowest level ($L^{(1)}$) at Yamaha gives 1.274 points more on the Borg CR-10 scale compared to picking from the reference picking level ($L^{(2)}$). This interpretation of the marginal contribution of raw coefficients to discomfort ratings, is possible because the Borg CR-10 scale has been designed such that it allows for ratio calculations (Borg 1982).

The results show that picking height, quantity picked and mass are factors that significantly contribute to discomfort across both warehouses. At Sorbo also the volume factor and an interaction effect for picking heavy products from a low level were found to be significant. The limited number of observations on medium to large size products (49 observations

Table 3. Observations count across product factor categories and picking levels.

			Yamaha				Sorbo			
		Range	$L^{(1)}$	$L^{(2)}$	$L^{(3)}$	Total	$L^{(1)}$	$L^{(2)}$	$L^{(3)}$	Total
Quantity										
Low quantity	LQ	$x \leq 3$ units	46	79	69	194	266	157	225	648
Medium quantity	MQ	$3 < x \leq 7$ units	7	7	4	18	45	23	26	94
High quantity	HQ	$x > 7$ units	12	5	6	23	3	1	3	7
Total			65	91	79	235	314	181	254	749
Mass										
Reference mass	RM	$x \leq 3$ kg	63	88	74	225	269	157	226	652
High mass	HM	$x > 3$ kg	4	7	9	20	45	24	28	97
Total			65	91	79	235	314	181	254	749
Volume										
Low volume	LV	$x \leq 1$ dm^3	49	76	61	186	141	89	102	332
Medium volume	MV	1 dm$^3 < x \leq 5$ dm^3	12	11	12	35	161	84	149	394
High volume	HV	$x > 5$ dm^3	4	4	6	14	12	8	3	23
Total			65	91	79	235	314	181	254	749

Table 4. Empirical discomfort model for Yamaha and Sorbo.

	Yamaha			Sorbo		
Variables	Raw b	Std. error	Std. b	Raw b	Std. error	Std. b
(Intercept)	1.595***	0.167	.	1.880***	0.150	.
$L^{(1)}$	1.274***	0.338	0.225	0.722***	0.169	0.176
$L^{(3)}$	1.176***	0.336	0.211	0.842***	0.176	0.197
HM	2.965*	1.427	0.345	1.171**	0.390	0.210
MV	0.335	0.348	0.037	1.046***	0.365	0.258
HV	−0.175	0.802	−0.029	3.286***	0.390	0.286
MQ	2.008***	0.533	0.222	1.161***	0.189	0.209
HQ	1.773**	0.550	0.219	2.143***	0.333	0.086
$HM * L^{(1)}$	2.059	2.332	0.061	1.075*	0.476	0.123
$HM * L^{(3)}$	−0.974	1.758	−0.064	0.395	0.443	0.037
R^2	0.279	−.−	−.−	0.311	−.−	−.−

Significance levels: $p \leq 0.05$(*), $p \leq 0.01$(**), $p \leq 0.001$(***).

for MV and HV) may explain that volume was not significant for Yamaha. Also few high mass (HM) observations were present in the sample of Yamaha, making it difficult to assess interactions between high mass and picking levels. Nonetheless, it is important to recognise that the low prevalence in the sample implies that these events occur relatively infrequently and hence can therefore only have a limited effect on the total rating.

3.3 *Empirical results for the storage location model*

To illustrate our method, we determine storage locations for one picking zone of each warehouse. To this end, we applied the bi-objective optimisation procedure described in Section 2.3, using Equations (1) and (3) for respectively CT_{ij} and D_{ij}. The coefficients for Equations (1) and (3) are taken from Tables 2 and 4, respectively.

Figures 2 and 3 show the convex hull of the non-dominated solutions at the two zones in the bi-objective assignment problem. Each figure shows two dominated solutions indicated by '$W(Z_1, Z_2)$' and '$W(Z_2, Z_1)$' that correspond to the worst case of solving a single-objective assignment problem by considering either the cycle time or the discomfort criterion. Additionally, we include two lexicographic solutions marked with '$L(Z_1, Z_2)$' and '$L(Z_2, Z_1)$' in which we first optimise for one criterion and then, fixing the first criterion, optimise the second.
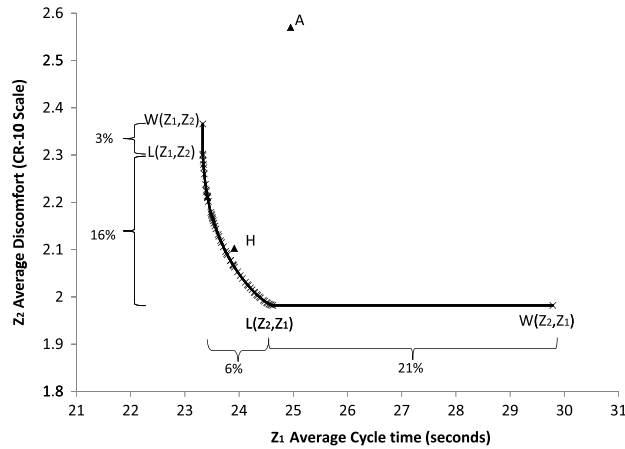
Figure 2. Efficient frontier for cycle times and workers' discomfort at Yamaha.
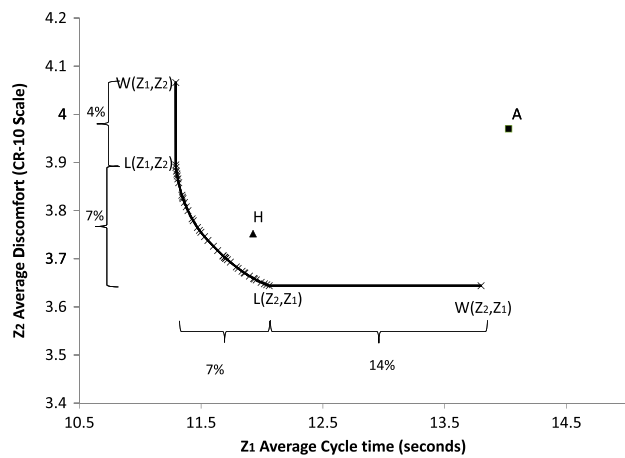

Figure 3. Efficient frontier for cycle times and workers' discomfort at Sorbo.

The first important observation from Figures 2 and 3 is that lexicographic solutions can provide improvements. Lexicographic solutions yield up to 3% and 4% improvements in discomfort for the studied picking zones of Yamaha and Sorbo, respectively, compared to optimising for cycle time only. Although the improvements are modest, these are without economic cost. When we select lexicographic solutions with discomfort as the first criterion, we can obtain a 21% improvement in terms of cycle time for the picking zone at Yamaha, while for the picking zone at Sorbo a maximum improvement of 14% cycle time can be obtained. Multiplied by the 32 and 24 picking zones of Yamaha and Sorbo and assuming similar saving percentages, this result may already imply a difference of 7 and 4 FTE (full-time employees) of productive capacity.

It thus appears that optimising for discomfort only can have stronger negative effect on cycle time than the reverse. This can be intuitively explained by the fact that the cycle-time criterion also considers that the Golden Zone levels not only reduce discomfort but also are faster to use (see Table 2). Optimising for discomfort on the other hand, does not consider the travel distances at all as relevant factor. As this reasoning is valid in the absence of strong interaction effects, this observation is bound to be valid for other warehouses with similar ranges of mass and volume in their products.

A decision-maker may also decide to select an intermediate non-dominated solution. At Yamaha 16% improvement in discomfort costs only 6% of cycle time. This means that better values for discomfort can be obtained with a fairly low impact on cycle time. At Sorbo a potential improvement of 7% in cycle time has to be traded off against an improvement of 7% in terms of discomfort ratings. It must be noted that these numbers reflect the shape of the efficient frontier. In many cases, a warehouse may actually be operating at a point above and to the right of the curve. The solutions marked with an 'A' in Figures 2 and 3 show the current configurations of the selected picking zones, which are clearly dominated solutions to the

Table 5. Empirical studies summary: a comparison of relative importance of location factors.

		Yamaha				Sorbo			
		Cycle time		Discomfort		Cycle time		Discomfort	
Factor category	Factor	Std. Coef.	Imp.	Std. Coef.	Imp.	Std. Coef.	Imp.	Std. Coef.	Imp.
2D location	MA	0.134^{***}	P	$-.-$	n.a.	0.423^{***}	P	$-.-$	n.a.
2D location	CA	0.435^{***}	P	$-.-$	n.a.	$-.-$	n.a.	$-.-$	n.a.
2D location	CN	0.198^{***}	P	$-.-$	n.a.	$-.-$	n.a.	$-.-$	n.a.
Level	$L^{(1)}$	0.058^{***}	S	0.225^{***}	P	0.006	n.s.	0.176^{***}	P
Level	$L^{(3)}$	0.019^{***}	S	0.211^{**}	P	0.008^{*}	S	0.197^{***}	P
Interaction	$LV * L^{(1)}$	0.007	n.s.	$-.-$	n.s.	-0.008	n.s.	$-.-$	n.s.
Interaction	$LV * L^{(3)}$	0.072^{***}	S	$-.-$	n.s.	0.013	n.s.	$-.-$	n.s.
Interaction	$HV * L^{(1)}$	0.012	n.s.	$-.-$	n.s.	0.029^{***}	S	$-.-$	n.s.
Interaction	$HV * L^{(3)}$	-0.016	n.s.	$-.-$	n.s.	0.028^{***}	S	$-.-$	n.s.
Interaction	$(Q-1) * L^{(1)}$	0.022^{***}	S	$-.-$	n.s.	-0.019^{*}	S	$-.-$	n.s.
Interaction	$(Q-1) * L^{(3)}$	0.005	n.s.	$-.-$	n.s.	0.004	n.s.	$-.-$	n.s.
Interaction	$HM * L^{(1)}$	$-.-$	n.s.	0.061	n.s.	$-.-$	n.s.	0.123^{*}	P
Interaction	$HM * L^{(3)}$	$-.-$	n.s.	-0.064	n.s.	$-.-$	n.s.	0.037	n.s.

Notes: Importance: P: Primary (i.e. St. Coef. > 0.1), S: Secondary (i.e. St. Coef. ≤ 0.1), n.a.: not applicable. n.s.: not significant. Significance levels: $p \leq 0.05(*)$, $p \leq 0.01(**)$, $p \leq 0.001(***)$. Time given in seconds.

storage location problem. Starting from such a situation, initially large savings are possible for both goals simultaneously. Only once the efficient frontier has been reached, will trade-offs between cycle time and discomfort arise. Our model can aid in reaching the efficient frontier in the first place. Then after identifying the frontier, the model can be used to give insights about the trade-offs, which can be moderate, as we found for Yamaha, or more significant, as we found for Sorbo. Without the model, a warehouse may be able to measure the current status of discomfort and cycle times, but it would not be possible to predict the effects of a reconfiguration on the two criteria and their interplay.

4. Implications for practice

A possible drawback of the proposed methodology is that it is data intensive. A heuristic that is less data intensive and that yields solutions close to the efficient frontier may then be desirable to apply in practice. We aim at exploiting commonalities between the two warehouses from our empirical study to develop an effective product-to-location assignment heuristic. To illustrate the potential, Table 5 shows the standardised coefficients of the factors affecting cycle time and discomfort for both warehouses. Note that product factors that do not interact with location factors may affect the estimation accuracy, but have no influence on determining where a product should be stored. Therefore, these factors can be omitted for the purpose of the heuristic.

In Table 5, we establish a cut-off value of 0.1 for the standardised coefficient as a way to distinguish whether a factor may be considered primary (P) or secondary (S) for determining cycle time and discomfort. As there are more options for location assignments in terms of horizontal travel (MA, CA, CN) than in the picking level (K), we opt to position products first favourably in term of horizontal travel (i.e. favourable for cycle time) and second favourably in picking levels (i.e. favourable for cycle time and discomfort). We exclude the location interaction effects as none of them have been identified as of prime or secondary importance for *both* warehouses.

As a result we propose the following simple heuristic that combines the two criteria and the popularity of a product.

(1) Rank every location according to its horizontal distance from the depot. Allow for ties in the case of locations in the same section (i.e. a whole column of locations).
(2) Assign locations in the Golden Zone a rank of 1 and locations outside the Golden Zone a rank of 2.
(3) For every rank at steps 1 and 2, divide the rank by the maximum rank obtained for each step. Next, sum both ratios (for steps 1 and 2) to obtain a location score.
(4) Sort the location scores in ascending order. Then, sort products in descending order popularity and then assign the most popular products to the locations with the lowest scores.
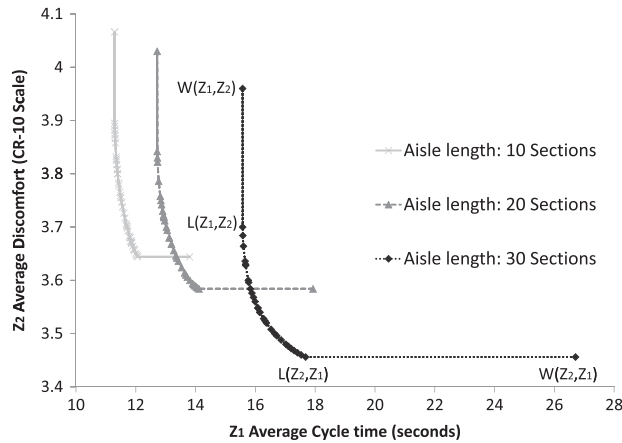
Figure 4. Exploring the trade-off curve for varying aisle lengths.

Figures 2 and 3 show the solution that this proposed heuristic yields, indicated by the label 'H'. Although the heuristic solutions are dominated, these are close to the set of non-dominated solutions for all practical purposes. Moreover, in the case of Yamaha, the heuristic improves current average cycle time by 4% and current average discomfort by 22%. Similarly, for Sorbo the heuristic improves the current average cycle time by 18% and the current average discomfort by 6%. Hence, for the cases studied we find that the proposed method results in solutions that provide important benefits in terms of cycle time and discomfort. Although this heuristic may be used as a rule of thumb in practice, we caution that this method is designed based on the cases studied and therefore is particularly tailored to picking from shelves in environments where most of the picks involve only one orderline and where products do not exceed a mass of 3 kg. In the case of higher mass, the importance of interaction effects may need to be re-assessed.

Further insights for practice can be obtained by extrapolating from the data of the cases studied. In particular, we explore how the shape of the convex hull of non-dominated solutions may vary with two key variables (1) the length of a picking aisle and (2) the popularity distribution of products. We explore the effects of these two variables for Sorbo, as the effects may be easier to interpret due to the simpler layout.

To analyse the effect of aisle length, we consider the original aisle length consisting of 10 sections, and aisle lengths of 20 and 30. Longer aisles provide more storage space, for which we replicated the original assortment to provide a sufficient number of products for storage, keeping the same popularity distribution of products. Figure 4 shows the effect of increasing the aisle length at Sorbo. The main effect shown in Figure 4 is that the trade-off between both objectives changes, becoming relatively more costly in terms of cycle time to improve discomfort at the efficient frontier. This can be seen from the curves for longer aisles being 'less steep'.

The popularity distribution of products can be represented by various functions; see Yu, De Koster, and Guo (2015) for a comparison of their impact on storage zones. We use the parametrisation of Bender (1981), which is given in Equation (10). Here, x represents the proportion of all products ordered in descending order of popularity, $F(x)$ is the corresponding cumulative popularity for that same proportion of products and s is the shape parameter of the curve. For example, taking the 20% most popular products, lower values of B imply a higher popularity of these products compared to the other 80% of the products. When s approaches infinity, all products are equally likely to be ordered.

$$F(x) = \frac{(1+s)x}{s+x}, 0 \leq x \leq 1, s \geq 0 \text{ and } s + x \neq 0 \tag{10}$$

We test different values of s in the original Sorbo layout, using $s = 0.02857$, $1/15$, $1/3$ and 10^5 such that the ratios of products to cumulative popularity are: 20%/90%, 20%/80%, 20%/50%, 50%/50%, respectively.

Figure 5 shows that the effects of changes in the popularity distribution of products on the efficient frontier. The main effect of lowering s is that it stretches the efficient frontier in both dimensions. This is explained by the fact that switching assigned locations between products with a greater difference in popularity, implies greater differences in cycle times and discomfort ratings, thus stretching efficient frontier values. It is also important to observe that when all products have the same probability of being picked (50%/50%) only three points on the efficient frontier are identifiable. This implies, that in terms of product factors, the main driver of storage location decisions is the popularity of a product, reflecting the structure
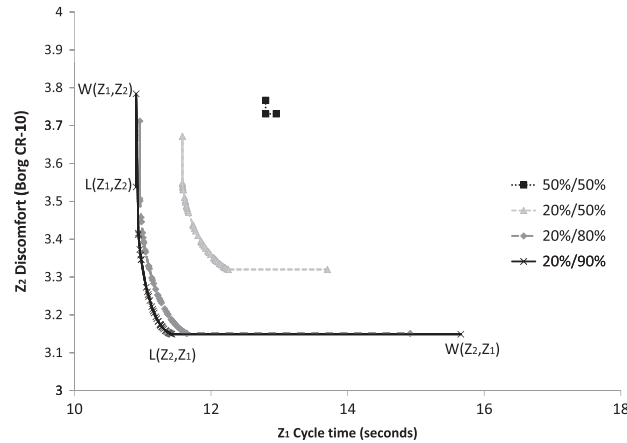
Figure 5. Exploring the trade-off curve for varying skewness of the ABC curve.

of the empirical results summarised in Table 5 where interaction effects between location and product factors were mostly of secondary importance for storage location decisions. This means that managers must assess the trade-offs more carefully the longer aisles are and the less skewed the popularity distribution of products is.

5. Conclusions

This study presents a method for storage allocation decisions that can be used in any warehouse where the context is of order picking from shelves and most picks involve single orderlines. This method goes further than current storage allocation decision models in two main respects. First, we explicitly model the effect of location factors on cycle time using actual data. Second, we introduce the criterion of improving the workers' well-being by minimising their discomfort. Our method highlights the value of data stored in WMSs. Furthermore, it shows that direct inquiry to pickers about their level of discomfort is an effective way of determining their preferences.

From the empirical studies, we find that horizontal distance from the depot and picking heights are main drivers for cycle times and discomfort. Product factors such as quantity and volume were also significant factors contributing to cycle time and discomfort. From the trade-off analysis, we conclude that optimising only for discomfort may be a costly option in terms of increased cycle time and is thus not advisable. Optimising only for cycle time seems relatively less costly in terms of discomfort. We also found that the two analysed warehouses currently operate outside the efficient frontier. This means that the decision of well-being vs. economic benefit may be a false dichotomy even in the short-term in the cases studied and in other cases where it is possible that firms operate outside the efficient frontier. Based on the similarity of the empirical results, it was possible to propose a heuristic that does not require extensive use of data to obtain good solutions that balance both criteria.

We also found an important insight when designing picking zones: extending the length of aisles with more picking positions stretches the efficient frontier in the direction of cycle time implying relatively more costly trade-offs of improvements of discomforts in terms of cycle time. In addition, the stronger the differences in demand popularity between products, the more there is the need to do a trade-off analysis.

The main limitation of this study is that it focuses on short-term effects of location and product factors on cycle time and discomfort. Thus, it implicitly assumes that discomfort and cycle time are not influenced by the past history of picks but only the current picks. Re-assessing this assumption and investigating the long-term links of location decisions with health outcomes like Low Back Pain reports, absenteeism rates as well as long term fatigue, are worthy of future research.

References

Ang, M., Y. F. Lim, and M. Sim. 2012. "Robust Storage Assignment in Unit-load Warehouses." *Management Science* 58 (11): 2114–2130.

Borg, G. 1982. "A Category Scale with Ratio Properties for Intermodal and Interindividual Comparisons." In *Psychophysical Judgment and the Process of Perception*, edited by H. G. Geissler and P. Petzold, 25–34. Berlin: VEB Deutscher Verlag der Wissenschaften.

Battini, D., M. Calzavara, A. Persona, and F. Sgarbossa. 2015. "Order Picking System Design: The Storage Assignment and Travel Distance Estimation (SA&TDE) Joint Method." *International Journal of Production Research* 53 (4): 1077–1093.

Bender, P. 1981. "Mathematical Modeling of the 20/80 Rule: Theory And Practice." *Journal of Business Logistics* 2 (2): 139–157.

Boudreau, J. W., W. J. Hopp, J. O. McClain, and L. J. Thomas. 2003. "On the Interface between Operations and Human Resources Management." *Manufacturing & Service Operations Management* 5 (3): 179–202.

Curseu, A., T. van Woensel, J. Fransoo, K. van Donselaar, and R. Broekmeulen. 2009. "Modelling Handling Operations in Grocery Retail Stores: an Empirical Analysis." *Journal of the Operational Research Society* 60 (2): 200–214.

De Koster, R. B. M., D. A. Stam, and B. M. Balk. 2011. "Accidents Happen: The Influence of Safety-specific Transformational Leadership, Safety Consciousness, and Hazard Reducing Systems on Warehouse Accidents." *Journal of Operations Management* 29 (7–8): 753–765.

De Koster, R., T. Le-Duc, and K. J. Roodbergen. 2007. "Design and Control of Warehouse Order Picking: A Literature Review." *European Journal of Operational Research* 182 (2): 481–501.

Doerr, K. H., and K. R. Gue. 2013. "A Performance Metric and Goal-setting Procedure for Deadline-oriented Processes." *Production and Operations Management* 22 (3): 726–738.

Dul, J., H. De Vries, S. Verschoof, W. Eveleens, and A. Feilzer. 2004. "Combining Economic and Social Goals in the Design of Production Systems by Using Ergonomics Standards." *Computers & Industrial Engineering* 47 (2–3): 207–222.

Dul, J., M. Douwes, and P. Smitt. 1994. "Ergonomic Guidelines for the Prevention of Discomfort of Static Postures based on Endurance Data." *Ergonomics* 37 (5): 807–815.

Ehrgott, M. 2000. *Multicriteria Optimization*. Berlin: Springer.

Genaidy, A. M., A. Mital, and M. Obeidat. 1989. "The validity of predetermined motion time systems in setting production standards for industrial tasks." *International Journal of Industrial Ergonomics* 3 (3): 249–263.

Gino, F., and G. Pisano. 2008. "Behavior in operations management: Assessing recent findings and revisiting old assumptions." *Manufacturing & Service Operations Management* 10 (4): 676–691.

Grosse, E. H., C. H. Glock, M. Y. Jaber, and P. Neumann. 2015. "Incorporating human factors in order picking planning models: framework and research opportunities." *International Journal of Production Research* 53 (3): 695–717.

Hamberg-van Reenen, H., H. A. J. van der Beek, B. M. Blatter, M. P. van der Grintenand, and P. M. Bongers. 2008. "Does musculoskeletal discomfort at work predict future musculoskeletal pain?" *Ergonomics* 51 (5): 637–648.

Hampel, F. R., E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel. 1986. *Robust Statistics*. The Approach Based on Influence Functions. New York: John Wiley and Sons.

Hayes, A., and L. Cai. 2007. "Using Heteroskedasticity-consistent Standard Error Estimators in OLS Regression: An Introduction and Software Implementation." *Behavior Research Methods* 39 (4): 709–722.

Huber, P. J. 1964. "Robust Regression of a Local Parameter." *Annals of Mathematical Statistics* 35 (1): 73–101.

Jones, E. C., and T. Battieste. 2004. "Golden Retrieval." *Industrial Engineer* 36 (6): 37–41.

Jonker, R., and A. Volgenant. 1987. "A Shortest Augmenting Path Algorithm for Dense and Spare Linear Assignment Problems." *Computing* 38 (2): 325–340.

Kadefors, R., and M. Forsman. 2000. "Ergonomic Evaluation of Complex Work: A Participative Approach Employing Video-computer Interaction, Exemplified in a Study of Order Picking." *International Journal of Industrial Ergonomics* 25: 435–445.

Kuijt-Evers, L. F. M., T. Bosch, M. A. Huysmans, M. P. De Looze, and P. Vink. 2007. "Association between Objective and Subjective Measurements of Comfort and Discomfort in Hand Tools." *Applied Ergonomics* 38 (5): 643–654.

Larco, J. A., K. J. Roodbergen, R. De Koster, and J. Dul. 2008. "Optimizing Order Picking Considering Workers' Comfort and Picking Speed." In *Progress in Material Handling Research: 2008*, edited by K. Ellis, R. Meller, M. K. Ogle, B. A. Peters, G. D. Taylor and J. Usher, 467–478. Charlotte, NC: Material Handling Institute.

Lodree, E. H. 2008. "Scheduling approaches to order picking." In *Progress in Material Handling Research: 2008*, edited by K. Ellis, R. Meller, M. K. Ogle, B. A. Peters, G. D. Taylor and J. Usher, 357–369. Charlotte, NC: Material Handling Institute.

Luskin, R. C. 1991. "Abusus Non Tollit Usum: Standardized Coefficients, Correlations, and R2s." *American Journal of Political Science* 35 (4): 1032–1046.

Peacock, B. 2002. "Measuring in Manufacturing Ergonomics." In *Handbook of Human Factors Testing and Evaluation*, edited by S. G. Charlton and T. G. O'Brien, 157–180. Mahwah, New Jersey: Lawrence Erlbaum.

Petersen, C. G., C. Siu, and D. R. Heiser. 2005. "Improving Order Picking Performance Utilizing Slotting and Golden Zone Storage." *International Journal of Operations and Production Management* 25 (10): 997–1012.

Przybylski, A., X. Gandibleux, and M. Ehrgott. 2008. "Two Phase Algorithms for the Bi-objective Assignment Problem." *European Journal of Operational Research* 185 (2): 509–533.

Rousseeuw, P. J., and A. M. Leroy. 1987. *Robust Regression and Outlier Detection*. New York: John Wiley and Sons.

Saccomano, A. 1996. "How to Pick Picking Technologies." *Trade World* 13: 157–180.

Tompkins, J. A., J. A. White, Y. A. Bozer, and J. M. A. Tanchoco. 2010. *Facilities Planning*. New York: Wiley.

Webster, B. S., and S. H. Snook. 1994. "The Cost of Workers' Compensation Low Back Pain Claims." *Spine* 19 (10): 1111–1116.

Wisnowsky, J. W. 1999. *Multiple Outliers in Linear Regression: Advances in Detection, Methods, Robust Estimation, And Variable Selection*. PhD diss.: Arizona State University.

WMO. "Wet medisch-wetenschappelijk onderzoek met mensen." 1998. http://wetten.overheid.nl/BWBR0009408

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

Zandin, K. 1990. *MOST Work Measurement Systems*. New York: Marcel Dekker.

Appendix 1.

Below we describe our adapted version of the procedure from Przybylski, Gandibleux, and Ehrgott (2008) to find the vertex set of the convex hull of the decision space.

(1) Obtain the lexicographic solution $x^{(1)} = \arg \operatorname{lex} \min_{x \in X}(z_1(x), z_2(x))$:

    (a) Solve the problem for one objective obtaining a solution such that $x' = \min_{x \in X} z_1(x)$.

    (b) Use solution x', to construct a new auxiliary problem and find the lexicographic solution: $x^1 = \arg \min_{x \in X} \lambda_1 z_1(x) + \lambda_2 z_2(x)$ where $\lambda_1 = z_2(x') + 1$ and $\lambda_2 = 1$.

(2) Obtain the lexicographic optimum $x^{(2)} = \arg \operatorname{lex} \min_{x \in X}(z_2(x), z_1(x))$ using the same procedure as in *Step 1*, but exchanging the order of the objectives.

(3) Add $x^{(1)}$ and $x^{(2)}$ to the set of non-dominated solutions.

(4) Initialise with $x_{L1} = x^{(1)}$ and $x_{L2} = x^{(2)}$.

(5) Solve the auxiliary convex combination problem, obtaining solution $x^{(t)}$ such that $x^{(t)} = \arg \min_{x \in X} \lambda_1 z_1(x) + \lambda_2 z_2(x)$ with $\lambda_1 = z_2(x_{L1}) - z_2(x_{L2})$ and $\lambda_2 = z_1(x_{L2}) - z_1(x_{L1})$.

(6) Recursive dichotomous search procedure

    If $\lambda_1 z_{x^{(t)}} + \lambda_2 z_{x^{(t)}} \leq \lambda_1 z_{x_{L1}} + \lambda_2 z_{x_{L1}}$ then

    (a) Add $x^{(t)}$ to the set of non-dominated solutions

    (b) Update $x_{L1} = x_{L1}$ and $x_{L2} = x^{(t)}$. Solve auxiliary convex combination problem, obtaining solution $x^{(t)}$ such that $x^{(t)} = \arg \min_{x \in X} \lambda_1 z_1(x) + \lambda_2 z_2(x)$ with $\lambda_1 = z_2(x_{L1}) - z_2(x_{L2})$ and $\lambda_2 = z_1(x_{L2}) - z_1(x_{L1})$ Execute *Step 6*.

    (c) Update $x_{L1} = x^{(t)}$ and $x_{L2} = x_{L2}$. Solve auxiliary convex combination problem, obtaining solution $x^{(t)}$ such that $x^{(t)} = \arg \min_{x \in X} \lambda_1 z_1(x) + \lambda_2 z_2(x)$ with $\lambda_1 = z_2(x_{L1}) - z_2(x_{L2})$ and $\lambda_2 = z_1(x_{L2}) - z_1(x_{L1})$ Execute *Step 6*.

    End

For finding the lexicographic minima, the algorithm makes use of the fact that all solutions to the assignment problem are integer. Thus, by multiplying the cost matrix of one criterion with a previously defined upper bound of the other objective plus one unit: $\lambda_1 = z_2(x') + 1$, whereas the other cost criterion is multiplied by $\lambda_2 = 1$, then a clear hierarchy of solving is guaranteed with cost function $z_1(x)$ being optimised first and $z_2(x)$ second.

To guarantee that all supported solutions are found, a dichotomous search is performed by varying the weights of λ_1 and λ_2 to solve convex combinations. The weights, λ_1 and λ_2 are chosen such that they define the normal vector of a hyperplane that has level curves parallel to two already identified non-dominated points in the z_1, z_2 plane. As the level curves are parallel to two already identified non-dominated points, the hyperplane is then guaranteed to identify an intermediate supported solution if such a solution exists. When the identified solution lies along the line defined by the previously identified two dominated points, the algorithm stops searching for further non-dominated solutions in that segment as no more non-dominated supported solutions located in a vertex exist.