Int J Adv Manuf Technol (2013) 65:473–484
DOI 10.1007/s00170-012-4186-0

ORIGINAL ARTICLE

The berth allocation problem with stochastic vessel handling times

Jeffery Karafa · Mihalis M. Golias · Stephanie Ivey · Georgios K. D. Saharidis · Nikolaos Leonardos

Received: 3 January 2011 / Accepted: 19 April 2012 / Published online: 13 May 2012

Abstract In this paper, the berth allocation problem with stochastic vessel handling times is formulated as a bi-objective problem. To solve the resulting problem, an evolutionary algorithm-based heuristic and a simulation-based Pareto front pruning algorithm is proposed. Computational examples show that the proposed approach provides solutions superior to the ones where the expected value of the vessel handling times is used.

J. Karafa · M. M. Golias · S. Ivey
Department of Civil Engineering, University of Memphis,
3815 Central Avenue,
Memphis, TN 38152, USA

J. Karafa
e-mail: jefferykarafa@yahoo.com

M. M. Golias
e-mail: mihalisdgolias@gmail.com

S. Ivey
e-mail: ssalyers@memphis.edu

G. K. D. Saharidis
Mechanical Engineering Department, University of Thessaly,
Leoforos Athinon, Pedion Areos,
38834 Volos, Greece

G. K. D. Saharidis (✉)
Kathikas Institute of Research and Technology,
4 Hatziarfanou st,
Paphos 6815, Cyprus
e-mail: saharidis@gmail.com

N. Leonardos
Department of Computer Science, Rutgers University,
100 Brett Road,
Piscataway, NJ 08854, USA
e-mail: nikolaos.leonardos@gmail.com

Keywords Berth allocation · Stochastic · Optimization · Simulation

1 Introduction

The Berth allocation problem (BAP) can be described as the problem of allocating berth space for vessels at container terminals and is a critical function of marine container terminal operations. Vessels arrive over time and the terminal operator needs to assign them to berths to be served (loading and unloading containers) as soon as possible. Ocean carriers, and therefore vessels, compete over the available berths and different factors (discussed in detail later) affect the berth and time assignment of each vessel. Among models found in the literature [15], there are four most frequently observed cases: (a) discrete vs. continuous berthing space, (b) static vs. dynamic vessel arrivals, (c) static vs. dynamic vessel handling times, and (d) variable vessel arrivals. In the discrete problem, the quay is viewed as a finite set of berths. In the continuous problem, vessels can berth anywhere along the quay. The majority of the published research considers the discrete case [9, 10, 11]. In the static arrival problem, at the time of scheduling, all vessels are already at the port whereas in the dynamic arrival problem, only a portion of the vessels to be scheduled are present, with arrival times for vessels not present known in advance. The majority of the published research in berth scheduling considers the latter case. In the static handling time problem, vessel handling times are considered as input, whereas in the dynamic handling problem, vessel handling is a variable; usually a function of the quay cranes that will operate on the vessel and the distance of the vessels' berthing position from a location in the yard. Finally, in the last case, the vessel arrival times are considered as variables and are optimized. Technical restrictions such as berthing draft and intervessel and end-berth clearance distance are further assumptions that have been adopted in some of the studies dealing with the BAP, bringing the problem formulation closer to real-world conditions. Introducing technical restrictions to existing berth allocation models is rather straightforward and is therefore not attempted here.

In this paper, we deal with the discrete space and dynamic vessel arrival berth allocation problem (DDBAP), which can be formulated as the machine scheduling problem [5, 16, 18]. The DDBAP continues to receive increased attention from the research community as it is a problem that marine container terminal operators deal with on a daily basis [13]. To our knowledge, the majority of berth scheduling models have not accounted for the stochastic nature of the vessel handling times; a stochasticity that stems from the fact that quay cranes (QCs) and internal transport vehicles (ITVs) serving the vessels do not have a deterministic productivity (e.g., random down time of QCs, unpredicted congestion in the yard, etc.). To our knowledge, the only exceptions have been five separate studies by Moorthy and Teo [12], Zhou et al. [20], Zhou and Kang [19], Golias et al. [7], and Du et al. [4] accounting for less than 10 % of the available berth allocation literature. Unlike the model presented herein, Zhou et al. [20], Zhou and Kang [19], and Du et al. [4] did not present a formal methodology to handle different handling time distributions. Golias et al. [7] only focused on online conceptual formulations, while Moorthy and Teo [12] proposed an approach (as stated by the authors) relevant only when a substantial number of vessels arrive periodically. We note that Zhou et al. [20] and Zhou and Kang [19] formulations constrain the vessel waiting times and may lead to: (a) infeasibility (i.e., strict waiting time limits) or (b) low quality solutions (i.e., high waiting time limits); issues that were not addressed or discussed by the authors.

In this paper, we propose a formulation for scheduling vessels to the available berths where vessel handling times are assumed as stochastic parameters with known probability distributions. These distributions can be obtained from historical data (i.e., berth assignment, number of QCs and ITVs, breakdown rates of QCs, utilization of yard, vessel handling volumes, etc.) using data mining algorithms, but, in this paper, are assumed to be known for all the vessels at all the berths. Based on these distributions, a given berth schedule risk measure is proposed and minimized. The proposed risk measure considers the variability of the vessel service start and finish times. Incorporating this type of risk requires the calculation of the probability distribution and percent point functions for the service start and finish time of the predecessor of each vessel. This can be a complex task that depends on the distributions of the random variables involved. In this paper, we present a discussion on the implications that different probability distributions may have and how they can be addressed. To account for berth productivity, we also introduce a second objective function that minimizes the total service time for all the vessels. The second objective is the most commonly considered in the related literature [11] as it is a basic measure of berth productivity. For further discussion on other objectives for the berth allocation problem, we refer to Meisel [11], Saharidis et al. [14], and Theofanis et al. [15].

The two objectives introduced (i.e., minimization of the risk and total service time) are conflicting and improvement in one objective will cause the degradation of the quality of the other [1]; thus, the terminal operator needs to select a schedule that balances between the two objectives. Berth schedules with a high berth throughput (i.e., small total service time for all the vessels) have a greater degree of risk (i.e., risk of matching the total service time when the stochastic vessel handling times are realized). On the other hand, berth schedules with a lesser degree of risk (decreased berth throughputs) provide more confidence to the terminal operator that the resulting assignment will be stable in terms of the handling times for each vessel [3, 17] and thus deviations from the initial schedule will be minimized in case rescheduling is needed [13]. The proposed model formulation provides the terminal operator with the berth schedule that balances between the two objectives. We choose to introduce the risk measure in contrast to formulating a stochastic optimization problem as the inherent combinatorial complexity of such a model would make it impossible to construct a meaningful heuristic that would efficiently search through the extremely large set of vessel handling time scenarios.

Existing exact solution algorithms for bi-objective scheduling problems rely on iterative-type procedures. These procedures employ exact algorithms and solve single-objective problems, equivalents of the bi-objective formulation. These algorithms cannot be efficiently applied to our problem, as a single objective formulation of the bi-objective problem formulation proposed herein is intractable. To tackle this issue, an evolutionary algorithm (EAs)-based heuristic and a simulation-based Pareto front pruning algorithm are proposed to solve the resulting problem.

The remainder of this paper is organized as follows: the next section presents the model formulation, the third section presents the solution algorithm, the fourth section presents a number of computational examples and examines the validity of the proposed approach, and the final section concludes the paper and proposes future research areas.

2 Model formulation

In this section, we present the mathematical formulation of the problem. The model is partially based on the formulation introduced by Golias et al. [8]. Before we present the motivation behind the model, we define the following:

Sets

I	set of berths
J	set of vessels

Decision variables

$x_{ij}, i \in I, j \in J$	=1 if vessel j is served at berth i and zero otherwise
$y_{ab}, a, b \in J$	=1 if vessel b is served as the immediate successor of vessel a at the same berth and zero otherwise
$f_j, j \in J$	=1 if vessel j is the first vessel to be served at its assigned berth and zero otherwise
$l_j, j \in J$	=1 if vessel j is the last vessel to be served at its assigned berth and zero otherwise

Auxiliary variables

$st_j, j \in J$	service start time for vessel j (stochastic variable)
$ft_j, j \in J$	service finish time for vessel j (stochastic variable)

Parameters

$c_{ij}, i \in I, j \in J$	handling time of vessel j at berth i (stochastic variable with a known distribution)
$A_j, j \in J$	arrival time of vessel j
Ξ	confidence level
M	large positive number

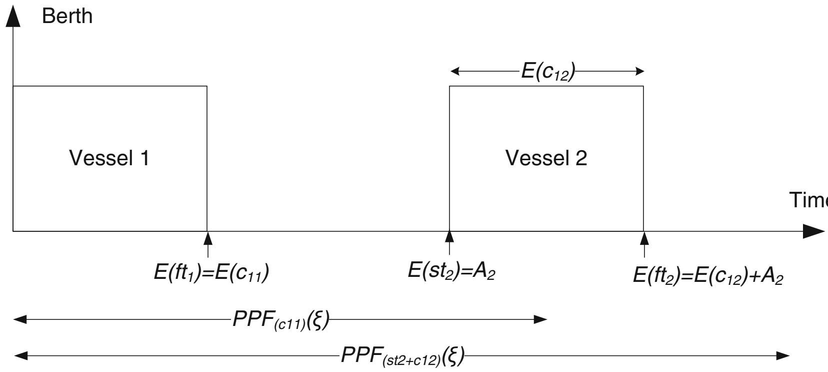
Definition 1 Let $M_{(Z)}$, $\Sigma_{(Z)}$, and $\mathrm{PPF}_{(Z)}$ be the mean, standard deviation, and the percent point function (respectively) of a stochastic variable $Z = \sum_k \pi_k, k \in \Pi$, where Π is a set of stochastic variables with known and closed form probability distributions.

Definition 2 Let $E(g)$ denote the expectation of the value of a stochastic variable g.

Given a vessel to berth assignment, the probability that a vessel's service start and/or finish time will be delayed depends on two factors: (a) that the proceeding vessels at the same berth will be delayed (i.e., the combined distribution of their handling times will exceed the mean and thus the vessel might wait more) and/or (b) the vessel's handling time will exceed its mean handling time (and thus the total service finish time will exceed the expected service finish time). To illustrate this, we present a simple example with two vessels and one berth shown in Fig. 1. In this example, vessel 1 is served first with an expected service finish time of: $\mathrm{E}(c_{11})$ and vessel 2 s with an expected start time of: A_2 (in this example we assume that $A_2 > E(ft_1)$) and an expected service finish time of: $A_2 + E(\mathrm{c}_{12})$. Vessels' 1 service finish time is a linear combination of its handling time. Therefore, there is an ξ percent probability that vessels' 1 service finish time will be less than or equal to: $\mathrm{PPF}_{(c_{11})}(\xi)$. Thus, vessels' 1 service finish time can be delayed (with an ξ percent probability) by (at most): $\mathrm{PPF}_{(c_{11})}(\xi) - E(\mathrm{ft}_1)$. As vessel 1 is served first (and assuming that the vessel arrival times are deterministic), the probability that its service start time will be delayed is zero. Service start time of vessel 2 is a linear combination of its arrival time and vessels' 1 service finish time while its service finish time a linear combination of its service start and handling times. For vessel 2, its service start time can be delayed up to: $\mathrm{PPF}_{(c_{11})}(\xi)$and its service finish time up to: $\mathrm{PPF}_{(st_2+c_{12})}(\xi)$. Thus, vessels' 2 service start and finish times may exceed the expected by: $\max\left(0, \mathrm{PPF}_{(c_{11})}(\xi) - A_2\right)$ and $\mathrm{PPF}_{(st_2+c_{12})}(\xi) - E(\mathrm{ft}_2)$, respectively. In this paper, we define the summation of these possible delays for all the vessels as the risk of a berth schedule.

To estimate the risk function, we need to be able to estimate the values of the PPF for a given ξ for each vessel's service start and finish times. At each berth, given a vessel-to-berth assignment, we can estimate these functions assuming that the handling times of all vessels follow a normal or a Poisson distribution. The service start time of each vessel will be a linear combination of the handling times of all the preceding vessels. The service finish time will be a linear combination of its service start and handling times. Using these linear combinations, we can estimate the service start and finish time probability distribution functions (PDFs) and use a simple iterative numerical procedure to estimate the PPF values with a specific probability ξ. We will use a simple example (Fig. 2) to illustrate how we estimate the risk of berth schedule with $n+1$ vessels. Without loss of generality in the example shown in Fig. 2, we assume that vessels are served at a single berth in increasing order of their identification number. For a random vessel (in this example vessel $n+1$), we can expect three different cases. In the first case, the vessel arrival time is smaller than its expected start time and the service start and finish time risk values are set equal to: $\mathrm{PPF}_{(\mathrm{f}t_n)}(\xi) - E(\mathrm{f}t_n)$ and $\mathrm{PPF}_{(\mathrm{ft}_{n+1})}(\xi) - E(\mathrm{ft}_{n+1})$, respectively (although by definition of the PPF the service start and finish times are less than or equal to:$\mathrm{PPF}_{(\mathrm{ft}_n)}(\xi)$and $\mathrm{PPF}_{(\mathrm{ft}_{n+1})}(\xi)$, respectively). In the second case, the $n+1$ vessels' arrival time is greater than:$\mathrm{PPF}_{(\mathrm{ft}_n)}(\xi)$and we assume that there is zero risk for its service start time (although there is a probability of $1-\xi\%$ that the arrival time might be less than: $\mathrm{PPF}_{(\mathrm{ft}_n)}$

Fig. 1 Illustration of one-berth two-vessel risk calculation



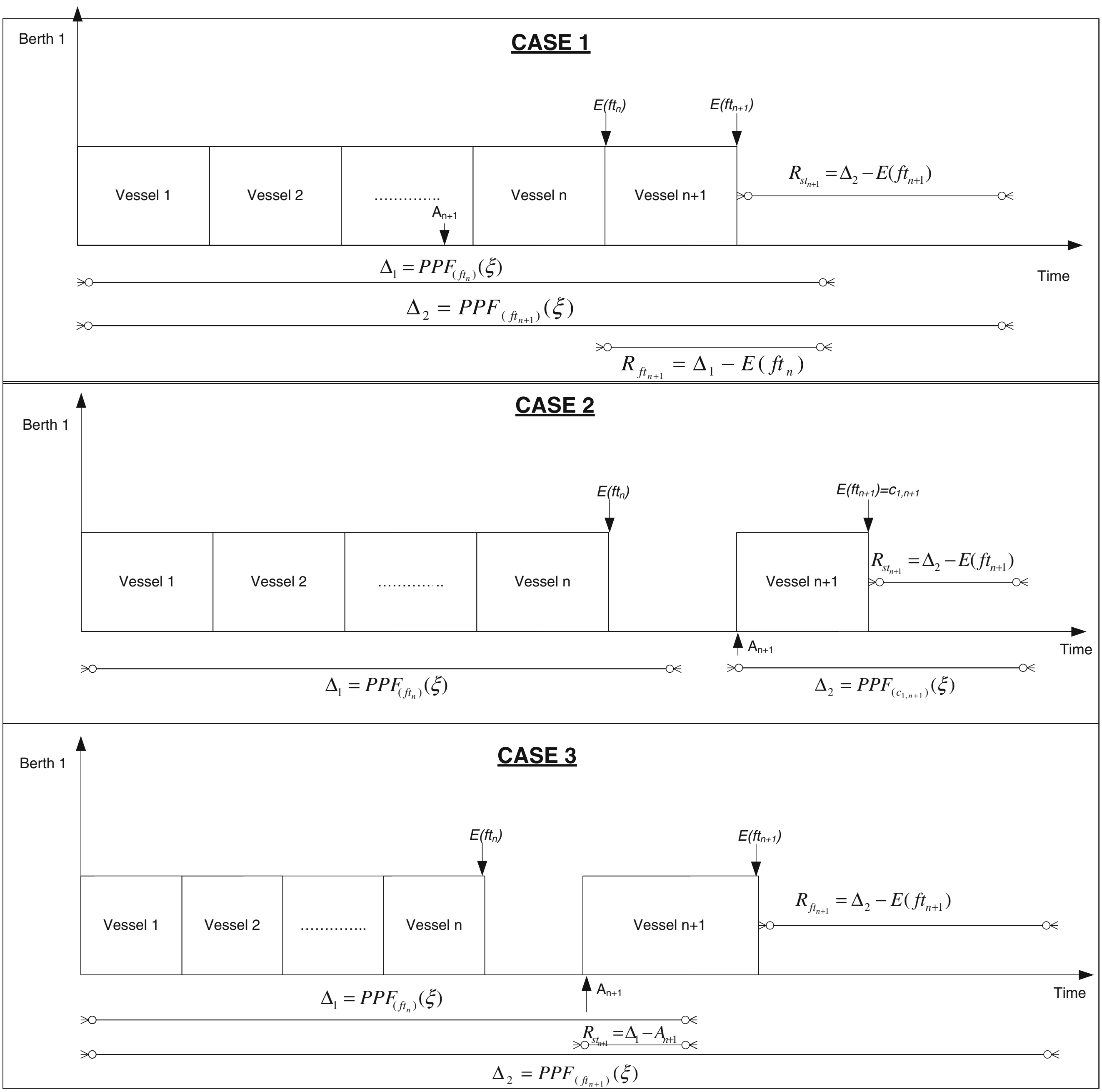


Fig. 2 Schematic illustration of the risk function calculation for the three possible vessel-to-berth assignments between the nth and nth+1 vessel

Berth	1						2					
Vessel	2	4	5	0	0	0	1	3	6	0	0	0

Fig. 3 Illustration of chromosome representation

(ξ)). In the second case, the same service finish time risk as with the first case is considered. In the third case, where $n+1$, vessels' arrival time is greater than the expected finish time of its immediate predecessor but less than: $\mathrm{PPF}_{(\mathrm{ft}_n)}(\xi)$ we set the service start and finish time risk values of vessel $n+1$ equal to:$\mathrm{PPF}_{(\mathrm{ft}_n)}(\xi) - A_{n+1}$ and $\mathrm{PPF}_{(\mathrm{ft}_{n+1})}(\xi) - E(\mathrm{ft}_{n+1})$, respectively.

The bi-objective model minimizing the vessel total service time and risk (from now on referred to as RSBM) is formulated as follows:

$$\min\left[f_1(x) = \sum_{j\in J} E(\mathrm{st}_j) + \sum_{i\in I}\sum_{j\in J} E(c_{ij})x_{ij}\right] \tag{1}$$

$$\min\left[f_2(x) = \sum_{j\in J} \left(R_{\mathrm{st}_j} + R_{\mathrm{ft}_j}\right)\right] \tag{2}$$

Subject to:

Decision variable constraints

$$\sum_{i\in I} x_{ij} = 1, \forall j \in J \tag{3}$$

$$f_b + \sum_{a\neq b\in J} y_{ab} = 1, \forall b \in J \tag{4}$$

$$l_a + \sum_{b\neq a} y_{ab} = 1, \forall a \in J \tag{5}$$

$$f_a + f_b \leq 3 - x_{ia} - x_{ib}, \forall i \in I, a, b \in J, a \neq b \tag{6}$$

$$l_a + l_b \leq 3 - x_{ia} - x_{ib}, \forall i \in I, a, b \in J, a \neq b \tag{7}$$

$$y_{ab} - 1 \leq x_{ia} - x_{ib} \leq 1 - y_{ab}, \forall i \in I, a, b \in J, a \neq b \tag{8}$$

Vessel-expected service start and finish time estimation

$$E(\mathrm{st}_j) \geq A_j, \forall j \in J \tag{9}$$

$$E(\mathrm{st}_j) \geq E(\mathrm{st}_a) + \sum_{i\in I} E(c_{ia})x_{ia} - M\left(1 - y_{aj}\right), \forall a, j \in J, a \neq j \tag{10}$$

$$E(\mathrm{ft}_j) = E(\mathrm{st}_j) + \sum_{i\in I} E(c_{ij})x_{ij}, \forall j \in J \tag{11}$$

Vessel service start and finish time upper bounds

$$\mathrm{st}_j = \max\left(E(st_j), \mathrm{PPF}_{\left(\mathrm{st}_a + c_{ia}x_{ia}y_{aj}\right)}(\xi)y_{aj}\right), \forall a, j \in J, a \neq j \tag{12}$$

$$\mathrm{ft}_j = \mathrm{PPF}_{\left(\mathrm{st}_j + c_{ij}x_{ij}\right)}(\xi), \forall j \in J \tag{13}$$

Risk estimation

$$R_{\mathrm{st}_j} \geq \mathrm{st}_j - E(\mathrm{st}_j), \forall j \in J \tag{14}$$

$$R_{\mathrm{ft}_j} \geq \mathrm{ft}_j - E(\mathrm{ft}_j), \forall j \in J \tag{15}$$

Objective function (1) minimizes the expected total service time for all the vessels. Objective function (2) minimizes the service start and finish time risk for all the vessels. Constraint set (3) ensures that each vessel will be served once, while constraint set (4) ensures that each vessel will either be served first or be preceded by another vessel. In a similar manner, constraint set (5) ensures that each vessel

INSERT Mutation

	Berth	1						2					
Before	Vessel	2	4	5	0	0	0	1	3	6	0	0	0
	Berth	1						2					
After	Vessel	2	3	4	5	0	0	1	6	0	0	0	0

SCRAMBLE Mutation

	Berth	1						2					
Before	Vessel	2	4	5	0	0	0	1	3	6	0	0	0
	Berth	1						2					
After	Vessel	4	3	1	0	0	0	6	5	2	0	0	0

Swap Mutation

	Berth	1						2					
Before	Vessel	2	4	5	0	0	0	1	3	6	0	0	0
	Berth	1						2					
After	Vessel	3	4	5	0	0	0	1	2	6	0	0	0

INVERT Mutation

	Berth	1						2					
Before	Vessel	2	4	5	0	0	0	1	3	6	0	0	0
	Berth	1						2					
After	Vessel	5	4	2	0	0	0	6	3	1	0	0	0

Fig. 4 Schematic illustration of the mutation operations

Table 1 Computational examples parameter values

Parameter	Value
Base problem instances	40
Berths	5
Planning horizon	1 week
Berth availability for the first time	Uniform (0, 10) hours
Vessel inter-arrival	3 h
Expected vessel handling time	Uniform (6, 42) hours
Vessel handling time standard deviation (as % of the mean handling time)	σ={.1, .2, .3, ………., .9, 1}
PPF confidence level ξ variations	ξ={.8, .85, .9, .95, .97, .99}

will either be last or will precede another vessel. Constraint sets (6) and (7) ensure that only one vessel can be served first and last at each berth. Constraint set (8) ensures that a vessel can be served after another vessel only if both are served at the same berth. Constraint set (9) ensures that the vessel service start time will be greater than the vessel's arrival time. Constraint set (10) estimates the expected service start time while constraint set (11) estimates the expected finish service time of each vessel. Constraint sets (12) and (13) calculate the upper bounds of the start and finish time of vessel j. Constraint set (14) estimates the service start time risk with ξ percent probability for each vessel. Constraint set (15) estimates the vessels' service finish time risk.

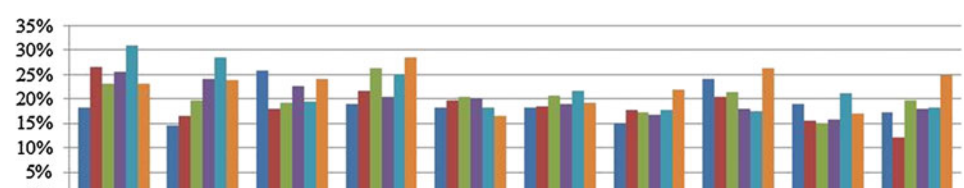
3 Solution algorithm

One of the main complexities of the RSBM lies in the definition of the PPFs. To the best of our knowledge, the published literature that formulated the berth allocation problem assuming deterministic or stochastic vessel handling times did not present any discussion on the distribution of the latter. For this reason, we turned our attention to published literature that applied simulation in order to evaluate the operations at marine container terminals (MCTs). For a detailed review of simulation applications in MCTs, we refer to Cartenì and de Luca [2]. In their research, it is reported that very few papers show estimated parameter values for their distribution functions and only a handful report values for QCs or vessel turnaround times. We would like to note that if a distribution function of the QC productivity is known, then the handling time of a vessel can be replicated as a linear combination of the QCs distributions. The distributions that replicated the productivity of QCs, as found in the literature, were Poisson, uniform, truncated normal, and gamma. For the vessel operations, the Erlang distribution was the most prominent; perhaps due to its direct relationship to the exponential distribution [17]. We would like to note that the loading and unloading process for the same vessel may follow different probability distributions. In this case, and if they cannot be approximated by a single probability function, the proposed approach cannot be applied as is. We leave this aspect of the problem as future research. Out of the reported distributions, only the linear combinations of the normal and Poisson distributions produce PDFs with closed form expressions, which are as follows:

Poisson If X_n~Poisson(λ_n) are n independent Poisson distributions, then (ΣX_n)~Poisson$(\Sigma\lambda_n)$ and $\sigma = \sqrt{\sum_n \lambda_n}$.

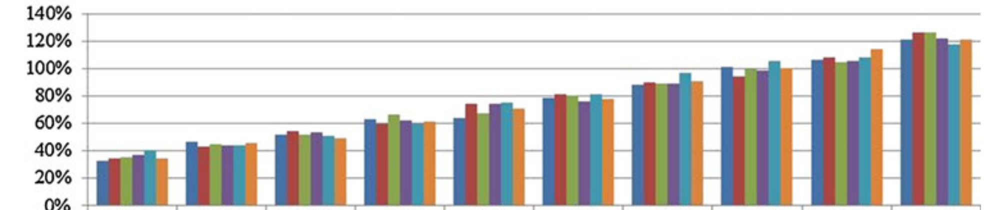
Normal If X_n~$N(\mu_n,\sigma_n)$ are n independent normal distributions, then $\sum_n X_n \sim N\left(\sum_n \mu_n, \sqrt{\sum_n \sigma_n^2}\right)$.

More complex distributions do not have closed form expressions and approximations have to be used. Even with the closed form expressions of linear combinations of the normal and Poisson distributions (which we can use to produce the PPFs), the RSBM still remains intractable (it is nonlinear). In order to tackle this issue and solve the RSBM within computationally reasonable times, we constructed an EAs-based heuristic [7], which is presented in the following subsection.



	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
ξ=80%	18%	15%	26%	19%	18%	18%	15%	24%	19%	17%
ξ=85%	27%	17%	18%	22%	20%	18%	18%	20%	15%	12%
ξ=90%	23%	20%	19%	26%	20%	21%	17%	21%	15%	20%
ξ=95%	26%	24%	23%	20%	20%	19%	17%	18%	16%	18%
ξ=97%	31%	28%	19%	25%	18%	22%	18%	17%	21%	18%
ξ=99%	23%	24%	24%	29%	17%	19%	22%	26%	17%	25%

Fig. 5 Average difference in MSC values (percentage) between the NPFS and the PPFS



	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
ξ=80%	32%	46%	52%	63%	64%	78%	89%	101%	106%	121%
ξ=85%	34%	43%	54%	60%	75%	81%	90%	94%	108%	126%
ξ=90%	35%	44%	51%	67%	67%	80%	89%	99%	104%	127%
ξ=95%	37%	44%	54%	62%	74%	76%	89%	98%	106%	122%
ξ=97%	40%	44%	51%	59%	75%	81%	97%	106%	109%	118%
ξ=99%	35%	46%	49%	61%	71%	78%	90%	101%	114%	121%

Fig. 6 Differences of MSC and EC values (in percent) for the PPFS

3.1 Evolutionary heuristic

The RSBM is a bi-objective minimization problem and this we adopt a multi-population EAs heuristic is adopted. The EAs heuristic proposed to solve the resulting problem consists of: (a) the chromosomal representation, (b) the chromosomal mutation, (c) the fitness function evaluation, and (d) the selection process. For scheduling problems, integer chromosomal representation is more adequate [5] and in this paper, we adopt an integer chromosomal representation that allows us to exploit the characteristics of the problem [6]. An illustration of the chromosome structure is given in Fig. 3 for a small instance of the problem with six vessels and two berths. As seen in Fig. 3, chromosome has 12 cells. The first six cells represent the six possible service orders at berth 1 and the last six cells the six possible service orders at berth 2. In the assignment illustrated in Fig. 3, vessels 2, 4, and 5 are served at berth 1 as the first, second, and third vessel, respectively, while vessels 1, 3, and 6 are served at berth 2 as the first, second, and third vessel, respectively.

The chromosomal mutation consisted of four different mutation types (insert, swap, inversion, and scramble) applied to the chromosomes of each generation. Each of the four types of mutation illustrated in Fig. 4 was based on the small example shown in Fig. 3, and has its own characteristics in terms of preserving the order and adjacency information.

Definition 3 Let X be the feasible space of the RSBM and $x \in X$ be a feasible solution. Solution $a \in X$ dominates solution $b \in X$ if: $\{f_1(a) \leq f_1(b), f_2(a) f_2(b)\}$ or $\{f_1(a) f_1(b), f_2(a) \leq f_2(b)\}$. Any nondominated solutions form the Pareto front (PF). The set of PF solutions is denoted from now on as Ω.

[1] In this paper, we set the initial population equal to 100. Each chromosome is initialized based on the first come first served at the berth with the minimum service start time policy

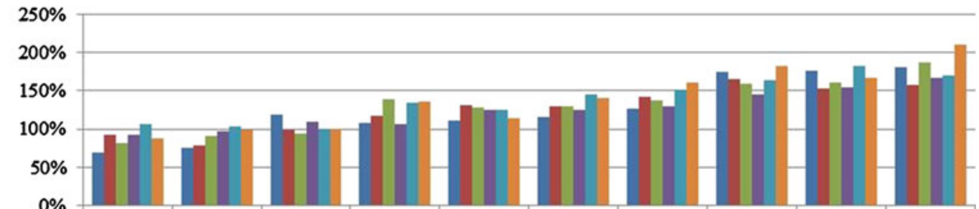
The RSBM is a bi-objective minimization problem; thus the smaller the values of each objective function, the higher the fitness value will be. At each iteration, out of all the available chromosomes, we select the ones that belong to the PF (of that iteration). If the number of chromosomes is less than the initial population we increase the population by randomly copying from the chromosomes in the current PF. The EAs heuristic can be summarized as follows:

Step 1: Initialize population of chromosomes[1] (i.e., parent chromosomes)
Step 2: Produce offspring chromosomes by mutating all parent chromosomes
Step 3: Select nondominated chromosomes from the offspring and parent chromosomes
Step 4: Replace parent chromosomes with the nondominated chromosomes from step 3
Step 4: If the convergence criterion[2] is met stop, else go to step 2.

3.2 Post-Pareto simulation

The algorithm described in the previous subsections will produce a number of nondominated solutions (i.e., berth schedules belonging to the PF). The next step will be to select one of these solutions as the schedule to be implemented. This follow-up step is known as post-Pareto analysis and can be quite a challenging task since, in the absence of subjective or judgmental information, none of the corresponding trade-offs can be said to be better than the others [8]. In the problem studied herein, we employ simulation as a means to select one schedule from the PF that will be implemented. The simulation entails the use of a simple Monte Carlo procedure that generates random instances of the vessels handling times and estimates an average of the total

[2] The EAs heuristic terminates if no new solutions enter the PF after 1,000 iterations or computational time exceeds 30 min



	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
ξ=80%	69%	76%	119%	107%	111%	115%	127%	175%	177%	181%
ξ=85%	92%	78%	98%	116%	130%	129%	142%	165%	153%	158%
ξ=90%	82%	91%	95%	139%	127%	129%	137%	159%	161%	187%
ξ=95%	92%	96%	109%	107%	124%	125%	130%	145%	154%	167%
ξ=97%	106%	104%	100%	134%	125%	145%	151%	163%	183%	169%
ξ=99%	87%	98%	100%	136%	114%	141%	161%	182%	166%	210%

Fig. 7 Differences of MSC and EC values (in percent) for the NPFS

service time over all the instances. The procedure can be described as follows. Let L be the total number of different handling time instances (i.e., realizations of the vessel's handling time) we wish to produce and $\mathrm{CPF}_{ij}(\cdot)$ the cumulative distribution function of vessel j's handling time at berth i. The procedure used in this paper to calculate the different vessel handling time instances at the different berths is as follows:

Monte Carlo procedure (MCP)

For $l=1{:}L$, $i=1{:}|I|$, $j=1{:}|J|$
Generate a number from the uniform distribution [0,1]: $u = U(0,1)$ and set $\mathrm{cs}_{ij}^{l} = \mathrm{CPF}_{ij}(u)$,
where:

CPF_{ij} is the cumulative distribution function of vessel j at berth i, and
cs_{ij}^{l} is the lth realization of vessel j's handling time at berth i
end

Definition 4 Let $x_n^{pf} \in \Omega \subset X$ be the nth PF solution of the RSBM.

For each one of the solutions in the PF, we estimate the mean value of the total service time (from now on referred to as the mean simulated cost or MSC) over all the L realizations of the vessel handling times (obtained from the MCP) as: $\mathrm{MSC}\left(x_n^{\mathrm{pf}}\right) = \frac{\sum_{k=1}^{K} f_1\left(\mathrm{cs}_{ij}^{l}, x_n^{\mathrm{pf}}\right)}{L}$. The solutions with the minimum MSC over all the schedules in the PF (from now on referred to as the pruned PF solution or PPFS and denoted by x^{ppfs}) is selected as the schedule to be implemented. The algorithm was coded and implemented in Matlab 7.7.0[3] and the experiments were performed on an ASUS desktop personal computer (E5300@2.60 GHz) with 6 GB memory.

[3] www.mathworks.com

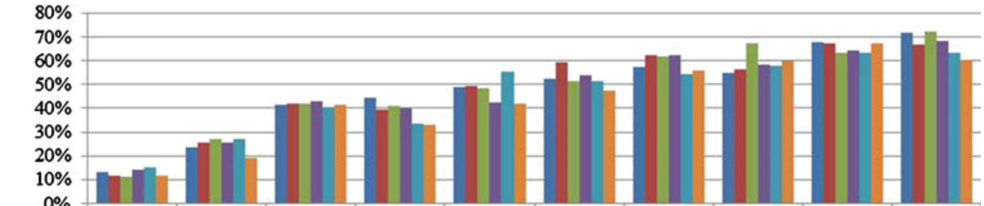
4 Computational examples

We developed 40 base problem instances, where vessels are served with various handling volumes at a multi-user container terminal with five berths and a planning horizon of 1 week. The range of variables and parameters considered herein were chosen from Golias et al. [8] and we report them herein for consistency purposes. Availability of berths for the first time in the beginning of the planning horizon (i.e., parameter S_i) was calculated using a uniform probability with a minimum of zero and a maximum of 10 h. Vessel inter-arrivals were calculated based on an exponential distribution with a mean of 3 h. Expected vessel handling times (loading and unloading) ranged from 6 to 42 h based on a uniform distribution. We assumed that over all the available berths one would be at the preferred berth (i.e., minimum expected handling time). The minimum expected handling time of a vessel at the other berths is obtained by adding a time penalty based on the distance of a particular berth from the preferred berth, with a maximum time penalty increase of 50 %. Assuming a normal distribution for the vessel handling times, we developed ten different variations for the standard deviation σ (where σ increases by 10 % from 10 to 100 % of the mean) and six different variations for the confidence level ξ (where ξ is equal to 80, 85, 90, 95, 97, and 99 %) for each dataset. In total, 2,400 problem instances were developed (40 different instances with the same handling times but different variations of the standard deviation σ and the probability for the confidence level ξ for each one of the 40 base problem instances). These data is shown in tabular form (Table 1) to assist the reader.

4.1 Evaluation of berth allocation policy

In this subsection, we evaluate the payoff of introducing the second objective function. For each one of the 2,400 problem instances previously described, we obtained the PF using the heuristic algorithm presented in Section 3. For each schedule in the PF, we calculated the MSC over a sample size of L=5,000. As discussed in the previous

Fig. 8 Nondominance of PPFS



	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
ξ=80%	13%	24%	41%	44%	49%	52%	57%	55%	68%	72%
ξ=85%	12%	26%	42%	40%	50%	59%	62%	56%	67%	67%
ξ=90%	11%	27%	42%	41%	49%	51%	62%	67%	63%	72%
ξ=95%	14%	25%	43%	40%	43%	54%	62%	58%	64%	68%
ξ=97%	15%	27%	40%	34%	56%	52%	54%	58%	63%	63%
ξ=99%	12%	19%	41%	33%	42%	48%	56%	60%	67%	60%

section, the schedule to be implemented will be the one with the minimum MSC. To evaluate the effectiveness of the proposed policy we initially compared the $\mathrm{MSC}\left(x^{\mathrm{ppfs}}\right)$ value to the MSC value of the solution in the PF with the minimum expected cost (EC) defined as: $\arg\min_{x\in X} f_1(x)$ (i.e., the solution we would obtain if we did not consider the risk function) and from now on will be referred to as the Nadir PF schedule or NPFS and denoted by x^{npfs}. To compare the berth throughput of the PPFS and the NPFS (i.e., $\mathrm{MSC}\left(x^{\mathrm{ppfs}}\right)$ and $\mathrm{MSC}\left(x^{\mathrm{npfs}}\right)$ values), we estimated the following: $\frac{\mathrm{MSC}\left(x^{\mathrm{npfs}}\right)-\mathrm{MSC}\left(x^{\mathrm{ppfs}}\right)}{\mathrm{MSC}\left(x^{\mathrm{npfs}}\right)}$ for each of the 2,400 problem instances.

Figure 5 shows the average values of the berth throughput difference over the 40 different base problem instances of each σ-ξ combination, answering the following question: "*On average, should we expect a gain in berth throughput if we choose the PPFS over the NPFS and by how much?*". Independent two-sample z tests performed on the simulated values of the objective function of the two schedules rejected the null hypothesis (that the samples have the same mean) at the 99 % confidence interval. We observe that the PPFS always produced a smaller MSC (i.e., positive percentages), thus a gain in berth throughput should be expected if the PPFS is chosen (as opposed to the NPFS). For example, for the first σ, the average improvement in the berth throughput of the PPFS as compared to the NPFS (over the 40 different base problem instances) is between 19 and 25 % (for the different ξ's). We further observe that the benefit from using the PPFS remains significant throughout the range of σ and ξ.

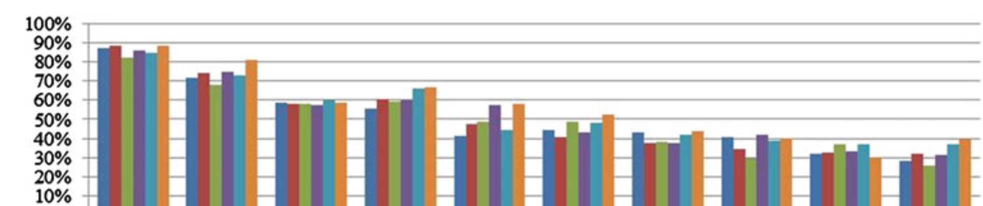
For each PPFS and NPFS of the 2,400 problem instances, we further calculated the differences between the values of their MSCs and ECs (i.e., the value of the first objective function in the PF using the expected vessel handling times) using the following formulas:

$$\frac{\mathrm{MSC}\left(x^{\mathrm{ppfs}}\right)-\mathrm{EC}\left(x^{\mathrm{ppfs}}\right)}{\mathrm{EC}\left(x^{\mathrm{ppfs}}\right)}, \text{where, } \mathrm{EC}\left(x^{\mathrm{ppfs}}\right)=f_1\left(E\left(c_{ij}\right),x^{\mathrm{ppfs}}\right) \text{and}$$
$$\mathrm{MSC}\left(x^{\mathrm{ppfs}}\right)=\frac{\sum_{l=1}^{L} f_1\left(\mathrm{cs}_{ij}^{l},x^{\mathrm{ppfs}}\right)}{L} \tag{17}$$

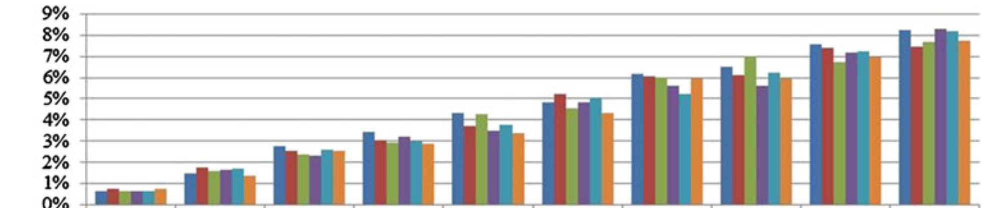
$$\frac{\mathrm{MSC}\left(x^{\mathrm{npfs}}\right)-\mathrm{EC}\left(x^{\mathrm{npfs}}\right)}{\mathrm{EC}\left(x^{\mathrm{npfs}}\right)}, \text{where, } \mathrm{EC}\left(x^{\mathrm{npfs}}\right)=f_1\left(E\left(c_{ij}\right),x^{\mathrm{npfs}}\right) \text{and}$$
$$\mathrm{MSC}\left(x^{\mathrm{npfs}}\right)=\frac{\sum_{l=1}^{L} f_1\left(cs_{ij}^{l},x^{\mathrm{npfs}}\right)}{L} \tag{18}$$

The average values over the 40 base problem instances for each $\sigma-\xi$ combination are reported in Figs. 6

Fig. 9 Nondominance of NPFS



	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
ξ=80%	87%	72%	59%	56%	41%	45%	43%	41%	32%	28%
ξ=85%	88%	74%	58%	60%	48%	41%	38%	35%	33%	32%
ξ=90%	82%	68%	58%	59%	49%	49%	38%	30%	37%	26%
ξ=95%	86%	75%	57%	60%	57%	43%	38%	42%	34%	32%
ξ=97%	85%	73%	60%	66%	44%	48%	42%	39%	37%	37%
ξ=99%	88%	81%	59%	67%	58%	52%	44%	40%	30%	40%



	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
ξ=80%	1%	1%	3%	3%	4%	5%	6%	7%	8%	8%
ξ=85%	1%	2%	3%	3%	4%	5%	6%	6%	7%	7%
ξ=90%	1%	2%	2%	3%	4%	5%	6%	7%	7%	8%
ξ=95%	1%	2%	2%	3%	3%	5%	6%	6%	7%	8%
ξ=97%	1%	2%	3%	3%	4%	5%	5%	6%	7%	8%
ξ=99%	1%	1%	3%	3%	3%	4%	6%	6%	7%	8%

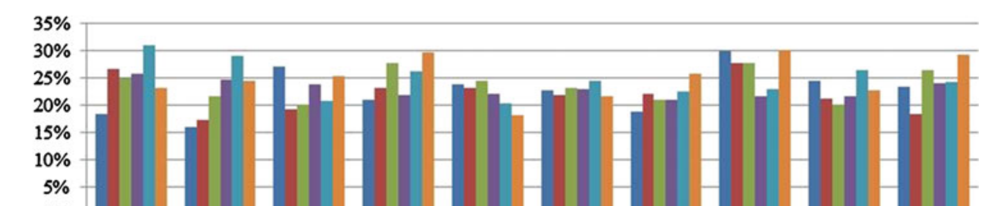
Fig. 10 Expected berth throughput loss under the PPFS

and 7 for the PPFS and the NPFS, respectively. The percentages reported in this table answer the following question: "*On average, should we expect a gain in berth throughput (either from the PPFS or the NPFS) as compared to their ECs and by how much?*" For example, for dataset 1, σ=10 % and ξ=80, the MSC$\left(x^{\text{ppfs}}\right)$would be 34 % larger than the EC $\left(x^{\text{ppfs}}\right)$. On the other hand, for the same dataset and $\sigma-\xi$ combination, the NPFS would result in the MSC$\left(x^{\text{npfs}}\right)$ value being 73 % larger than the EC$\left(x^{\text{npfs}}\right)$. In general, we observe that the PPFS produces more reliable schedules (i.e., smaller differences to the expected values). As expected, the difference between the EC (expected value) and the MSC (mean cost) increases (for both the NPFS and PPFS) as variability of the handling times (i.e., σ) increases but the PPFS shows a significantly smaller change compared to the NPFS. On the other hand these differences do not change significantly as we increase the level of confidence ξ (for the same σ) which increases the reliability of the PPFS.

4.2 PPFS and NPFS dominance

We further evaluated the dominance of the PPFS and NPFS in the PF by calculating the number of times, out of the L different vessel handling times realizations, where the PPFS or the NPFS did not have the maximum berth throughput. Average values over the 40 different base problem instances for each $\sigma-\xi$ combination are reported in Figs. 8 and 9 (for the PPFS and the NPFS, respectively), answering the following question: "*How likely is it that the PPFS or the NPFS do not provide the best berth throughput over all the solutions in the PF when a random realization of the vessel handling times is obtained?*".

Results reported in Figs. 8 and 9, show that the PPFS is not always the best schedule over all the L different realizations of the vessel handling times. Furthermore, as σ and ξ increase the probability that the PPFS will not be the best schedule (given a random instance of the vessel handling times) increases while the same probability for the NPFS decreases. This observation contradicts the initial claim that the PPFS should be selected as the schedule to be implemented under conditions of high handling time variability (i.e., high values of σ) as its dominance decreases with the increase of the handling time variability (i.e., σ). For example, given a random instance of the vessel handling times and for the case of ξ=80 % and σ=100 %, PPFS has a 72 % probability that another schedule (from the PF) will have a better throughput. For the same case, this probability for the NPFS is only 28 %. As these results do not report the loss of



	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
ξ=80%	18%	16%	27%	21%	24%	23%	19%	30%	25%	23%
ξ=85%	27%	17%	19%	23%	23%	22%	22%	28%	21%	18%
ξ=90%	25%	22%	20%	28%	24%	23%	21%	28%	20%	26%
ξ=95%	26%	25%	24%	22%	22%	23%	21%	22%	22%	24%
ξ=97%	31%	29%	21%	26%	20%	24%	23%	23%	26%	24%
ξ=99%	23%	24%	25%	30%	18%	22%	26%	30%	23%	29%

Fig. 11 Expected berth throughput loss under the NPFS

berth throughput when the PPFS or the NPFS are dominated an additional measure of performance was introduced: the expected berth throughput loss. To calculate the latter measure assume that for a test instance, the PPFS or the NPFS is not the best schedule K times out of the L different realizations of the vessel handling times. Then, the expected berth throughput loss for the PPFS and the NPFS can be calculated by Eqs. 19 and 20, respectively:

$$\sum_{k=1,\ldots.K} \frac{\max\left(0, \left(f_1\left(\mathrm{cs}_{ij}^{k}, x^{\mathrm{ppfs}}\right) - \min\limits_{x_n^{pf} \in \Omega} f_1\left(\mathrm{cs}_{ij}^{k}, x_n^{\mathrm{pf}}\right)\right)\right)}{f_1\left(\mathrm{cs}_{ij}^{k}, x^{\mathrm{npfs}}\right)} / K \tag{19}$$

$$\sum_{k=1,\ldots.K} \frac{\max\left(0, \left(f_1\left(\mathrm{cs}_{ij}^{k}, x^{\mathrm{npfs}}\right) - \min\limits_{x_n^{\mathrm{pf}} \in \Omega} f_1\left(cs_{ij}^{k}, x_n^{\mathrm{pf}}\right)\right)\right)}{f_1\left(\mathrm{cs}_{ij}^{k}, x^{\mathrm{npfs}}\right)} / K \tag{20}$$

In Eqs. 19 and 20, x_n^{pf} is the solution from the PF with the highest berth throughput for the nth vessel handling time realization. Results of the expected berth throughput loss over the 40 different base problem instances, and for each $\sigma-\xi$ combination, are reported in Figs. 10 and 11 (for the PPFS and NPFS, respectively). These percentages answer the following question: "*What is the expected loss in the total berth throughput as compared to the optimal schedule given a random realization of the vessel handling times?*". We observe that, on average, we should expect a loss in the berth throughput of less than 8 %, if for a given vessel handling time realization the PPFS is not the optimal solution. On the other hand, the same expected loss, for the NPFS, exceeds 20 %. Thus, even though the PPFS is less dominate as σ increases (Fig. 8) the loss in berth throughput is significantly lower than the loss of the NPFS. These results increase confidence that the PPFS will perform better (as compared to the NPFS) when a random instance of the vessel handling times is realized.

5 Conclusions

In this paper, we formulated the discrete space and dynamic vessel arrival berth allocation problem as a bi-objective optimization problem with the objective to maximize the berth throughput and minimize the risk of the berth schedule, under the assumption that vessel handling times are stochastic parameters with known probability distributions. In order to maximize the reliability of the berth schedule, a risk measure dependent on the vessel-to-berth assignment was proposed. In order to solve the resulting problem, a combination of an EA-based heuristic and a simulation-based Pareto front-pruning heuristic were proposed. Based on computational results, it was concluded that considering the proposed risk can provide schedules with improved berth throughput and higher reliability. The schedules from the proposed approach were either the best schedule or marginally deviated from the best schedule (a deviation ranging from 1 to 8 %). The proposed model formulation is limited to cases where the linear combination of the vessel handling time distributions provide a new distribution with a closed form expression (e.g., all the individual vessel handling time distributions follow either a normal or Poisson distribution). Future research can focus on expanding the model and solution algorithm to incorporate more generic handling time distributions.

Acknowledgments This material has been partially supported by the Intermodal Freight Transportation Institute (IFTI), University of Memphis and the Kathikas Institute of Research & Technology (KIRT). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of IFTI or KIRT.

References

1. Briand C, La Trung H, Erschler J (2007) A robust approach for the single machine scheduling problem. J Scheduling 10:209–221
2. Cartenì A, de Luca S (2009) A micro simulation model of the Salerno container terminal. Technical paper, Dept. Of Civil Engineering, University of Salern. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.127.7218&rep=rep1&type=pdf
3. Cowling P, Johansson M (2002) Using real time information for effective dynamic scheduling. Eur J Oper Res 139:230–244
4. Du Y, Xu Y, Chen Q (2010) A feedback procedure for robust berth allocation with stochastic vessel delays. In: Proceedings of the World Congress on Intelligent Control and Automation, IEEE, Jinan, China, 2210–2215
5. Eiben AE, Smith JE (2003) Introduction to evolutionary computing. Springer, New York
6. Goldberg DE (1989) Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, Reading
7. Golias MM, Boilé M, Theofanis S (2007) The stochastic berth scheduling problem. Proceedings of the Second TRANSTEC Conference. Prague, Czech Republic
8. Golias MM, Boilé M, Theofanis S (2010) Discrete berth scheduling problem: towards a unified mathematical formulation. In: Journal of the Transportation Research Board: Transportation Research Record, vol 2168., pp 1–8
9. Imai A, Nagaiwa K, Tat CW (1997) Efficient planning of berth allocation for container terminals in Asia. J Adv Transp 31:75–94
10. Imai A, Nishimura E, Papadimitriou S (2001) The dynamic berth allocation problem for a container port. In: Transportation Research Part B: Methodological, vol 35., pp 401–417
11. Meisel F (2009) Seaside operations planning in container terminals. Physica, Berlin
12. Moorthy R, Teo CP (2006) Berth management in container terminal: the template design problem. OR Spectr 28(4):495–518
13. Pfeiffer A, Kádár B, Monostori L (2007) Stability-oriented evaluation of rescheduling strategies, by using simulation. Comput Ind 58(7):630–643

14. Saharidis GKD, Golias MM, Boilé M, Theofanis S, Ierapetritou M (2010) The berth scheduling problem with customer differentiation: a new methodological approach based on hierarchical optimization. Int J Adv Manuf Technol 46(1–4):377–393
15. Theofanis S, Boile M, Golias MM (2009) Container terminal berth planning: critical review of research approaches and practical challenges. Transp Res Rec J Trans Res Board Mar Transp Port Oper 2100:22–28
16. Wang B, Lai X, Xi L (2005) Single-machine partial rescheduling with bi-criterion based on genetic algorithm. Adv Intell Comput 3644:947–956
17. Winston W (2003) Operations research applications and algorithms, 4th edn. California, USA, Duxbury Press
18. Wu SD, Storer RH, Chang PC (1993) One-machine rescheduling heuristics with efficiency and stability as criteria. Comput Oper Res 20:1–14
19. Zhou PF, Kang HG (2008) Study on berth and quay crane allocation under stochastic environments in container terminal. Syst Eng Theory Pract 28(1):161–169
20. Zhou PF, Kang HG, Lin L (2006) A dynamic berth allocation model based on stochastic consideration. In: Proceedings of the 6th World Conference on Intelligent Control and Automation, IEEE. Dalian, China