UvA-DARE (Digital Academic Repository)

Optimal depot locations for humanitarian logistics service providers using robust optimization

Stienen, V.F.; Wagenaar, J.C.; den Hertog, D.; Fleuren, H.A.

DOI

[10.1016/j.omega.2021.102494](https://doi.org/10.1016/j.omega.2021.102494)

Publication date

2021

Document Version

Final published version

Published in

Omega

License

CC BY

[Link to publication](#)

Citation for published version (APA):

Stienen, V. F., Wagenaar, J. C., den Hertog, D., & Fleuren, H. A. (2021). Optimal depot locations for humanitarian logistics service providers using robust optimization. *Omega*, 104, [102494]. <https://doi.org/10.1016/j.omega.2021.102494>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

Optimal depot locations for humanitarian logistics service providers using robust optimization[☆]

V.F. Stienen^{a,*}, J.C. Wagenaar^a, D. den Hertog^b, H.A. Fleuren^a

^a Department of Econometrics and Operations Research, Tilburg University, Tilburg, The Netherlands

^b University of Amsterdam, Amsterdam Business School, Amsterdam, The Netherlands

ARTICLE INFO

Article history:

Received 3 February 2020

Accepted 17 May 2021

Available online 21 May 2021

Keywords:

Humanitarian logistics

Pre-positioning

Preparedness

Facility location

Transportation

Uncertainty

Robust optimization

ABSTRACT

When a disaster strikes an area, often international assistance is requested to help responding to and recovering from the disaster. The response is often characterized by the dispatch of relief items to the affected regions via depots from Humanitarian Logistics Service Providers (HLSPs). One of the challenges an HLSP faces is how to organize its network of depots. Important questions are how many depots should be opened and what should their locations be? We develop a method that uses historical data to determine depot locations that minimize the cost of transportation and minimize the maximum response time to a disaster. We examine the trade-off between the two objectives using the Pareto front. Furthermore, we use robust optimization to find solutions that are robust against uncertainty in the location and scale of future disasters. We provide a case study of the United Nations Humanitarian Response Depot (UNHRD), a large globally operating HLSP. Among other things, we compute potential cost savings that could be obtained when expanding the network with additional depots. Also, we show the need for using robust optimization when looking for solutions that are robust against the location and scale of future disasters. We generalize the approach by applying it to the independent disaster database EM-DAT, in which we focus on examining the added value of using a robust optimization framework. The first result holds for all Uncapacitated Facility Location Problems (UFLP) with uncertain demand, which are used to find the optimal number of depots. We conclude that incorporating the likely advantage of larger networks having less uncertainty in the expected cost of transportation, enhances the decision-making regarding the optimal number of depots. Even if at some point increasing the number of depots allowed to be opened may not decrease the expected cost of transportation significantly, it is likely to reduce the uncertainty in this expected cost of transportation. For the second result, we consider UFLP used for finding optimal locations of a given number of depots. We conclude that solutions based on nominal and robust optimization may perform similar on individual years that are not included in the optimization, but robust solutions guarantee a smaller cost of transportation for worst case scenarios. This uncertainty reduction is especially valuable when the solution has to be robust against extreme scenarios.

© 2021 The Authors. Published by Elsevier Ltd.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

1. Introduction

Earthquakes, floods, landslides and storms are examples of disaster types which affect many people across the world. When a disaster strikes, many of the affected regions call for international assistance to help responding to and recovering from the disaster. This response can, for instance, be characterized by the dispatch of

relief items (e.g., tents, tarpaulins, food, hygiene kits, etc.) to the affected regions by international relief organizations. To enhance response capabilities, humanitarian organizations often make use of pre-positioning inventory. This may significantly reduce the cost and time it takes to deliver emergency items to the affected people. However, for a relief organization it is often not easy to store goods anywhere in the world. For instance, opening a depot for a small amount of pre-positioned goods might be too expensive. As an alternative, they can use the services of Humanitarian Logistics Service Providers (HLSPs). As the name already suggests, an HLSP offers services to humanitarian organizations that improve the efficiency of the supply chain of these organizations. They try to in-

[☆] Area: Optimization Analytics. This manuscript was processed by Associate Editor Thomas Sharkey.

* Corresponding author.

E-mail address: v.f.stienen@tilburguniversity.edu (V.F. Stienen).

crease efficiency and effectiveness of the coordination of multiple relief agencies. For instance, an HLSP may offer storage space in specific places of the world. One of the largest HLSPs is the United Nations Humanitarian Response Depot (UNHRD), managed by the UN World Food Programme (WFP). UNHRD offers storage space to humanitarian organizations (for free) in multiple locations across the globe such that these organizations can consider these locations for pre-positioning their inventory. Besides offering storage space, UNHRD can assist in transporting stored relief items to the affected regions on behalf of the humanitarian organizations. Unlike the storage, this is done on a cost-recovery basis. It typically results in lower transportation costs due to cargo consolidation effects and pooling of resources.

The research described in this paper is done in cooperation with UNHRD. One of the aims is to determine efficient (new) depot locations for the UNHRD network. To accomplish this, we develop a deterministic optimization approach that uses historical data to determine optimal depot locations. We refer to this model as the *nominal* model. The historical data used for this case contains the shipments of UNHRD between January 2017 and April 2019. The two optimality criteria we consider are the cost of transportation and the maximum response time. The proposed approach can be used by any HLSP. We generalize the approach by applying it to the independent disaster database EM-DAT, in which we focus on examining the added value of using a *robust* optimization framework. This database consists of records of historical disasters, containing information on, for instance, the total number of people affected by a disaster. Using this database, we obtain solutions that are robust against uncertainty in the location and scale (e.g., the number of people affected by a disaster) of future disasters.

2. Literature review

Recently, [1] analyzed potential cost benefits for the UNHRD network by adding a *regional* distribution center in Kampala (Uganda), to better respond to humanitarian disasters in East Africa. They develop a model to optimize the inventory position at this distribution center. To do this, they compare the cost of shipping containers with and without a depot in Kampala. They evaluate their model on 5,000 demand scenarios. Their conclusion is that adding a distribution center in Kampala yields a mean cost improvement of 21% for the products considered. An important remark is that the location of the new distribution center (Kampala) is not determined by optimization. It is selected by UNHRD and the WFP office in Nairobi, mainly because it represents a *strategic* center supporting WFP logistics in this region. In our research, we describe an optimization approach for determining optimal depot locations. The subsequent decisions about *what stock should be pre-positioned where* could be evaluated using the approach described by [1].

Another paper that examines optimal depot locations for a large organization similar to an HLSP is written by [2]. They study the case of the international humanitarian organization CARE. They assume a given set of depot locations at which CARE can store relief items. This set consists of the depots of CARE and the depots of UNHRD. An integer programming model is formulated to minimize the average response time by pre-positioning inventory at these depots. Multiple cases are considered in which there is a maximum number of depots that can be used. They use data from the EM-DAT database of previous disasters between 1997 and 2006. When there is a maximum number of depots that can be opened, they implicitly select optimal depot locations out of a given set of depots. This is done in such a way that the network minimizes the average response time to the previous disasters used from the EM-DAT database. In our research, we include the determination of the set of potential depot locations. In [3], the findings are enhanced

by considering a larger time frame of 30 years of historical data. Furthermore, the effects of natural disaster trends on the optimal pre-positioning strategies are analyzed. For a recent overview of operations research papers applied to humanitarian operations, we refer to [4].

In a broader sense, our work is part of *humanitarian logistics*. In the field of humanitarian logistics, people often use concepts related to disaster preparedness, response and reconstruction. These terms are first introduced by [5] and [6]. Both papers are often considered as benchmark papers towards understanding the logistics operations during disasters. Many papers still use the concepts defined in these papers. [7] is a continuation of the work of [6]. The purpose of this paper is to evaluate how the research in disaster operations management has evolved and to what extent the gaps identified by [6] have been covered. For a more recent overview of the work done in humanitarian supply chain management, we refer to [8].

[9] review papers specifically using optimization modeling in humanitarian logistics. They distinguish models based upon their use before or after a disaster strikes an area. In particular, for the pre-disaster phase (the location problem studied in our research corresponds to this phase), they focus on papers that examine short-notice evacuation, facility location, and stock pre-positioning.

Determining optimal depot locations can be characterized as a *facility location problem*, in which a number of possible locations should be opened (or activated) to be able to deliver relief items. For an extensive literature overview of facility location problems, we refer to the work of [10]. More specifically, we focus on facility location problems with uncertainty in the demand for items from each facility. Taking uncertainty into account can be done in several ways, we refer to [11] for an overview of the literature on facility location problems under uncertainty. They categorize the literature in three classes; robust, stochastic and chance-constrained facility location problems.

[12] are one of the first researchers who analyze a facility location problem (or distribution network design problem) in a humanitarian relief setting. The model they develop determines, given a list of candidate distribution center (DC) locations, where to set up DCs and what amount of relief supplies should be stocked at these DCs. To cope with uncertainty in the location and impact of future disasters, they work with (demand) scenarios. A demand scenario represents a combination of a disaster location and impact. The probability of a demand scenario occurring is based upon historical data. The model then maximizes the expected coverage of a network of DCs, with respect to the demand scenarios. For an extensive literature review of facility location problems in a humanitarian setting, we refer to [13]. They distinguish four different facility location problems in a humanitarian setting: deterministic, dynamic, stochastic and robust facility location problems. The robust facility location problems discussed by [13] all use a robust programming approach as developed by [14]. They use scenarios, defined as a possible future states, to determine optimal decisions. Then, they optimize for the worst case scenario(s). For details, we refer to [14]. A disadvantage of this approach is that it can be very difficult to define a discrete set of scenarios that captures the uncertainty in, for instance, the demand parameter. Besides, it might be necessary to optimize for a large number of discrete scenarios, which is likely to end up in computationally intractable problems. We propose to use robust optimization as described by [15]. We use continuous sets of scenarios (uncertainty sets) that capture the future states we would like to safeguard ourselves against. At the same time, we keep the problem computationally tractable. For an overview of the literature of robust optimization applied to facility location problems, we again refer to [11].

[16] is a recent paper that applies robust optimization with continuous sets of scenarios to find cost optimal solutions for the problem of opening distribution centers and defining flows of (relief) goods from suppliers to distribution centers. They assume that most of the input parameters are uncertain. For the cost of transportation, they assume that there are certain lower and upper bounds. For the demand and supply parameters, they use the method introduced by [17]. This means that they consider a common *conservatism parameter* for each disaster region. This parameter controls the number of parameters that may simultaneously happen to deviate from their original nominal value. Furthermore, the uncertainty set for the demand parameter is determined by a *variability parameter*. This variability parameter indicates by how much percent the nominal values of the demand may change. For instance, if this parameter equals 0.05 (i.e., 5%), it means that the demand in a region may deviate by 5% from its nominal value. Given the conservatism degree, they select an optimal set of DCs on a regional scale. They include a case study in which there are 13 disaster regions, 6 suppliers and six potential DC locations. Finally, they also include a sensitivity analysis regarding the conservatism degree and the probability of constraint violations given these conservatism degrees. There are three main differences with our research. The first difference is related to the uncertainty sets used. We focus on uncertainty in the demand parameter, for which we use a more practical approach for defining the uncertainty involved. More generally, we base all parameter calibrations on historical data instead of using a variability and conservatism parameter. The second difference is that we consider multiple objectives instead of a single (cost) objective. Finally, we optimize on a global scale, transporting goods via sea and air, instead of a regional scale.

There are several other papers that incorporate the uncertain environment of humanitarian logistics. [18] review the literature concerning the pre-positioning of relief items combined with decisions related to facility location problems (e.g., distribution centers) in an uncertain environment. They distinguish two ways for incorporating uncertainty into the modeling. First, they discuss scenario-based approaches, in which each scenario contains information on demand, potential damages, transportation links, etc. The second way covers scenario-free approaches, which mainly consist of probabilistic parameters or chance-constrained methods. The first main disadvantage of these scenario-free methods is that parameters often do not follow a specific probability distribution. Even when they follow a probability distribution, it is often not known which probability distribution it follows. Much data needs to be available in order to estimate the probability distribution. Secondly, the applicability of the discussed scenario-free methods is limited due to problems being computationally intractable. We want to expand the class of scenario-free methods with a robust optimization approach that does not have these disadvantages. Based on historical data we define uncertainty sets that cover demand scenarios.

The main contributions of this paper can be summarized as follows. First, we develop a robust optimization framework that determines optimal depot locations for a humanitarian logistics service provider, while incorporating the uncertainty in the location and scale of future disasters. A common approach in the humanitarian aid literature is to use a discrete set of scenarios to capture the uncertainty in the parameters (see e.g., [13]). However, in order to capture the uncertainty of future disasters accurately, a large number of scenarios is necessary, resulting in a computationally intractable problem. Therefore, we propose to use a continuous set of scenarios in order to keep the formulation computationally tractable. For this approach, we develop specific model elements that significantly reduce the computation time of the mathematical model. Among other things, we develop a clustering algorithm that guarantees finding the same solution(s) as when solving the

model without the clustering algorithm, but with a significantly reduced computation time. Using data from the independent disaster database [19], we show that, when dealing with an uncapacitated facility location problem with uncertain demand in general, larger networks do not only reduce the cost of transportation, but they also likely reduce the uncertainty in the future cost of transportation. Incorporating these results enhances the decision-making regarding the optimal number of depots. Furthermore, we consider the optimal locations of a given number of depots. Although solutions based on nominal and robust optimization perform similar on individual years that are not included in the optimization, robust solutions still have the advantage of performing better in worst case scenarios. This uncertainty reduction can especially be valuable when the solution has to be robust against extreme scenarios. Next to these results based on data from the EM-DAT database, we examine a case study of UNHRD, a large, globally operating, humanitarian logistics service provider for which we perform, among other things, analyses about optimal expansion strategies.

The outline of this paper is as follows. In Section 3, we start by introducing UNHRD. We describe the current practices of delivering emergency aid to areas hit by a disaster, and we discuss the performance measures used to assess the quality of a geographical distribution of emergency response depots for UNHRD. Note that throughout this paper, the same performance measures are also used for other HLSs. We end this section by describing the optimization problems discussed in this paper. Then, in Section 4, we develop both the nominal and robust optimization model to solve the challenges introduced before. In Section 5, we describe the data that is used in the model. Furthermore, we also elaborate on pre-processing steps that are taken upfront in order to be able to solve the mathematical models efficiently. In Section 6, we provide numerical results and a case study of UNHRD. Finally, in Section 7, we conclude and draw potential lines of further research.

3. The UNHRD pre-positioning network

UNHRD is a network of six humanitarian support depots located in Italy, United Arab Emirates, Malaysia, Ghana, Spain and Panama (see Fig. 1), managed by WFP. In 2000 the first storage location was established in Brindisi, Italy, which was then managed by the Office of Coordination of Humanitarian Affairs (OCHA). Thereafter, the Inter Agency Standing Committee (IASC) transferred the mandate to provide rapid and accessible stockpiling and demobilization services to partners inside and outside the United Nations from OCHA to WFP ([20]).

The depot in Brindisi could be used by humanitarian organizations as a pre-positioning location. In 2006, WFP set up additional depots in Africa, the Middle East, South East Asia and Latin America. Each of these locations had been selected to provide easy access to the main transportation methods (airport, port and road systems). Furthermore, they had been determined in such a way that UNHRD was able to deliver to any location on earth within 24–48 hours after a request was made. Next to the free of charge services (e.g., storage space), UNHRD also offers multiple services for which the humanitarian organizations need to pay. For instance, UNHRD offers to transport relief items to the desired regions on a cost-recovery basis. Multiple items from different relief organizations can be transported together, which can lead to economies of scale. It is important to note that the amount of relief items stored in a depot depends on the decisions of the individual organizations (partners). Partners decide for themselves where to store what amount of which items. They also decide for themselves when and where to ship their items to. In this paper, we use the UNHRD network as a case study for our research. We evaluate the

Fig. 1. Current UNHRD network of depots.

performance of this network and we discuss scenarios that can be of interest to UNHRD.

We consider two objectives that determine the quality of a network of depots. The first one is the *cost of transportation*, which summarizes the costs of transporting relief items from depots to disaster sites. When discussing the cost of transportation, we mean the cost to deliver relief items to disaster sites from the depot that results in the least amount of transportation costs. Note that organizations may have (specific) items only pre-positioned in one depot, when wanting to respond to a disaster further away. This could lead to larger transportation costs than necessary based on the network. In Section 6, we discuss the potential cost savings when all relief items are delivered from the depot that results in the least amount of transportation costs. This, for instance, may require organizations to work together when pre-positioning their inventory. The second objective we consider is the *maximum response time* of a network of depots. This quantity is an upper bound for the time it takes to deliver relief items to a disaster site (from one of the depots in the network). We want to minimize both these objectives simultaneously.

Furthermore, we consider two different problems:

- **OPTIMAL:** We consider a greenfield situation in which we are allowed to open a given number of depots anywhere in the world.
- **ADDSOME:** We consider expansion strategies for an HLSP, in which we want to find the best choice for opening new depots. We analyze, for instance, potential cost/time savings when expanding the current UNHRD network with additional depots.

In this research, we focus on the optimal *locations* of depots. We do not incorporate the capacity of depots. First, when it is not known how much items partners are going to store in which depot, good locations need to be determined. Based on these locations possible capacity expansions can be examined when information becomes available about the use of a depot. Moreover, the capacity of a depot location is usually not the bottleneck. This means that if a disaster strikes, any depot is able to deliver all relief items needed. This implicitly means that relief items during a disaster always come from a *single* depot location. Furthermore, we also do not incorporate the cost of opening new depots. This relates to the previous assumption, as these costs largely depend on the size and location of a depot. In short, we examine the optimal *locations* of a fixed number of depots.

4. Mathematical approach

In this section, we discuss two mathematical approaches for solving the ADDSOME and OPTIMAL problem. The main question we try to answer is how to geographically distribute humanitarian response depots of an HLSP across the globe. Before elaborating on the approach, we first list the modeling assumptions and decisions. Then we discuss the first approach, referred to as nominal optimization. This approach determines HLSP depot locations based on data of previous disasters. This means that we find the optimal network of depots for the disasters that took place in the past. However, in general, we do not know where the next disaster will strike or how much people will be affected. This assumption is relaxed in the second approach, by using a *robust optimization* framework.

4.1. Modeling assumptions

We study the depot locations of an HLSP. Thus, we focus on the decisions of where to establish depots, and how to expand a network of depots. We do not seek to find an optimal pre-positioning strategy that includes recommendations of where to pre-position what items.

We want to find a geographical distribution of depots that minimizes two objectives simultaneously: the cost of transportation and the maximum response time. There might be a trade-off between the two objectives, which we want to visualize using a Pareto front. The Pareto front contains Pareto optimal solutions, which consist of all solutions in which one or more objectives can not be improved without deteriorating one or more of the other objectives. In this paper, we exclude so-called *weak* Pareto optimal solutions, solutions for which any change in an objective value will make at least one of the objective values no better off, while no objective value may be deteriorated. After constructing the Pareto front, the best solution depends on the preferences of the decision makers.

There are three main assumptions used throughout the modeling. First, the amount of relief items that an affected person needs does not depend on the disaster type. We assume that all essential relief items are needed every time. For instance, whether you are hit by an earthquake or a flood, you will always need basic relief items such as tarps, blankets, water containers, etc. Secondly, relief items are delivered to disaster areas via their nearest harbor and nearest airport. Note that it could happen that these locations are destroyed by the disaster, in case another airport/harbor

is used. As this will still be an airport/harbor that is close to the location of the disaster, incorporating such events will likely not change the results of this study and is considered to be beyond the scope of this research. Finally, relief items are delivered from the (open) depot that results in the least amount of transportation costs, or that has the smallest response time. Note that, in practice, organizations may want to deliver its relief items to a specific location that is closest to a depot in which this organization does not have any items stored. However, this is more related to optimizing a pre-positioning strategy for an organization than finding good locations for depots of an HLSP. Note that this problem also depends on the scope/budget of an organization.

4.2. Nominal optimization

We develop a single deterministic optimization model that is able to find a set of depot locations that minimizes the total cost of transportation, given an upper bound on the maximum response time of the set of depot locations. We start by explaining the details of this model. Thereafter, we explain how this model is used to obtain the Pareto front of the problem.

We frame the problem as a specific version of a facility location problem. Let $\mathcal{H}$ be the set of potential depot locations. Let $\mathcal{H}^f \subseteq \mathcal{H}$ be the set of depot locations that *must* be opened. For instance, consider the ADDSOME problem in which we want to expand the current UNHRD network with additional depots. This means that we put all the current depot locations in $\mathcal{H}^f$. Let n be the maximum amount of depots that can be opened. This means that we are allowed to select n depot locations from $\mathcal{H}$. The selected depots must, however, include all depots in $\mathcal{H}^f$.

The selection of depot locations is based on historical data. In the nominal optimization model, we try to find an optimal network of depots for a set of historical disasters. A historical disaster is characterized by a location and a number of people affected by this disaster. The model determines from which (open) depots relief items are delivered to each (historical) disaster. In other words, the model *assigns* disasters to (open) depots. Because we assume that relief items are delivered via the nearest airport/harbor, and from the (open) depot that results in the least amount of transportation costs, or that has the smallest response time, we know that:

- When minimizing the cost of transportation, we can ignore the cost of transporting relief items from the nearest airport/harbor to the disaster site. These costs are *always* made, regardless the geographical distribution of open depots. This, in turn, means that disasters that share their nearest airport and harbor are always *assigned* to the same (open) depot. Therefore, we do not need to assign *disasters* to depots. Instead, we can assign groups of disasters, that share their nearest airport/harbor, to (open) depots.
- When minimizing the maximum response time, we know that for each group of disasters that share their nearest airport and harbor, there is one disaster that has the longest response time. Therefore, if we set this response time as the response time of this whole group from its corresponding airport, we know that we can deliver all disasters in this group within this response time from the corresponding airport. In other words, we know that if a solution minimizes the maximum response time, all affected people of disasters within all groups can be provided with relief items within this minimized maximum response time. So, also in this case, we do not need to assign *disasters* to depots. Instead, we can assign groups of disasters, that share their nearest airport/harbor, to (open) depots.
- When minimizing the cost of transportation, while incorporating restrictions on the maximum response time, we know that

Model 1

The nominal depot location problem.

$$\begin{array}{ll}
 \text{minimize} & \sum_{h \in \mathcal{H}} \sum_{c \in \mathcal{C}} a_c c_{hc} x_{hc}, \\
 \text{s.t.} & \sum_{c \in \mathcal{C}} x_{hc} \leq |C| y_h, \quad \forall h \in \mathcal{H}, \quad (\text{M.1}) \\
 & \sum_{h \in \mathcal{H}} x_{hc} = 1, \quad \forall c \in \mathcal{C}, \quad (\text{M.2}) \\
 & \sum_{h \in \mathcal{H}} y_h \leq n, \quad (\text{M.3}) \\
 & y_h = 1, \quad \forall h \in \mathcal{H}^f, \quad (\text{M.4}) \\
 & \sum_{h \in \mathcal{H}} t_{hc} x_{hc} \leq T, \quad \forall c \in \mathcal{C}, \quad (\text{M.5}) \\
 & x_{hc} \in \mathbb{R}_+, \quad y_h \in \{0, 1\}, \quad \forall h \in \mathcal{H}, \quad c \in \mathcal{C}. \quad (\text{M.6})
 \end{array}$$

disasters that share their nearest airport and harbor, may be delivered from different depots, depending on their distance to the closest airport. In this case, we can only group disasters that occurred at the same location.

In the first two situations, we define groups of disasters that share their nearest airport and harbor as *airport/harbor combinations* (AHs). So, instead of assigning disasters to open depots, we assign AHs to open depots. In the third situation, we assign groups of disasters that occurred at the same location to open depots. Note that solving the grouped and the non-grouped problem are mathematically the same; both yield the same solution. The advantage of assigning groups of disasters is that it reduces the number of variables and constraints in the final optimization model, while still guaranteeing the same optimal solution as when solving the model without the clustering procedure. As a remark, suppose that affected areas are not necessarily provided with relief items from the depot that results in the least amount of transportation costs. For instance, when taking depot capacity into account, it might happen that affected areas of disasters in one group are provided with relief items from different depots. This means that we can not assign groups, in this way, to depots, while guaranteeing the same solution as when solving the model without the clustering procedure.

Let $c \in \mathcal{C}$ be the set of clustered disasters. In case cluster $c \in \mathcal{C}$ is assigned to depot location $h \in \mathcal{H}$, it means that all relief items necessary for the disasters in cluster c are being transported from depot location h to cluster c . Next, we define the cost parameters. Part of the transportation is done with aircraft (to the airport in c) and part via sea (to the harbor in c). We denote the cost of providing relief items to one person affected in cluster $c \in \mathcal{C}$ from depot location $h \in \mathcal{H}$ by c_{hc} . Furthermore, let a_c represent the total number of people affected in cluster c . Let t_{hc} be the response time in which people affected in cluster c can be provided with relief items from depot location h . Note that this is the time it takes to reach a disaster by aircraft and truck. Recall that the time it takes to deliver the goods from the airport to the disaster site is the time it takes to deliver the goods to any of the disasters in cluster c . Finally, let n be the maximum amount of depots that can be opened and let T be the maximum time allowed to respond (deliver relief items) to a disaster.

Thirdly, we define the decision variables of the mathematical model. Let y_h be a binary variable that indicates whether depot h is opened ($y_h = 1$), and let x_{hc} be a binary variable indicating whether people affected in cluster c are provided with relief items from (open) depot h . Note that in our case we are allowed to let x_{hc} be a nonnegative (continuous) variable instead. We assume that each depot is able to deliver relief items to any affected area in any cluster (no capacity restrictions on the depots). Therefore, in the optimum, people affected in one cluster are always provided with relief items from a single (open) depot, the one for which the cost of delivering to this cluster is the lowest, or the one for which the response time is the lowest. The optimal set of depots that should be opened when minimizing the cost of transportation can be found by solving Model 1 described below.

Model 2

The robust optimization model for the optimal depot location problem.

$$\begin{aligned} &\text{minimize} && \lambda p + \sum_{c \in \mathcal{C}} \gamma_c u_c - \sum_{c \in \mathcal{C}} \beta_c t_c, \\ &\text{s.t.} && -\sum_{h \in \mathcal{H}} c_{hc} x_{hc} + \gamma_c - \beta_c + \lambda \geq 0, \quad \forall c \in \mathcal{C}, \\ &&& \lambda \in \mathbb{R}_+, \gamma \in \mathbb{R}_+^{|\mathcal{C}|}, \beta \in \mathbb{R}_+^{|\mathcal{C}|}, \\ &&& (\text{M.1}) - (\text{M.6}). \end{aligned}$$

The objective summarizes the total transportation costs of the assignment of clusters to depots. The first constraint ensures that people affected in each cluster can only be supplied with relief items from *open* depots. The second constraint makes sure that each cluster is assigned to a depot and the third constraint bounds the total number of open depots to n . The fourth constraint ensures that pre-defined depot locations are opened. The fifth constraint bounds the maximum response time of the solution to T . Finally, constraints (M.6) define the domains of the decision variables.

Next, we explain how to find the Pareto front of the problem using Model 1. Recall that on the one hand we have the maximum response time and on the other hand we have the cost of transportation as objective. We start by discretizing the maximum response time. We assume that the accuracy of the maximum response time is one hour. In other words, for each integer hour only one Pareto optimal point may exist. Then, we determine the *minimum* maximum response time by solving Model 1, in which we replace the objective by:

$$\text{minimize} \quad \max_{c \in \mathcal{C}} \left\{ \sum_{h \in \mathcal{H}} t_{hc} x_{hc} \right\}.$$

Note that minimizing a max operator can easily be done by using the epigraph formulation of the objective function. Suppose that the minimized maximum response time turns out to be t^* . Then the first Pareto optimal solution can be found by solving Model 1, where T equals t^* . The next Pareto optimal solutions are found by repeatedly increasing (with one) the restriction on the maximum response time (T) and solving the corresponding Model 1. So we set T equal to $t^* + 1$, $t^* + 2$, etc. In other words, we minimize the total cost of transportation with a changing restriction on the maximum response time (Constraint (M.5)). This is done until we obtain a solution that only minimizes the cost of transportation (there is no upper bound on the maximum response time). Finally, note that both problems (OPTIMAL and ADDSOME) can be solved using Model 1.

4.3. Robust optimization

In the nominal model (i.e., Model 1), the optimal depot locations are based on known historical data. In other words, if we *would know* which disasters are going to happen, our procedure is able to provide us with optimal candidates. However, disasters are often not that predictable: there is uncertainty in the location and scale of future disasters. In this section, we incorporate this uncertainty into the mathematical model using robust optimization techniques as described in [15].

From now on, we assume that we want to find a set of depots that is optimal for an average unknown (future) year. This uncertainty in the location and scale of future disasters can be represented by the uncertainty in the number of people affected (a_c) in the average unknown future year in each cluster.

One straightforward way to optimize for this average unknown year is to set a_c equal to the average number of people affected (per year) in cluster c , based on the historical data. In mathematical terms, we let $\bar{a}_c$ represent the mean number of people affected in cluster c based on historical data. Then, we use $\bar{a}_c$ as value for

a_c in the nominal model to find the optimal depot locations for the average unknown year (which equals the average historical year).

The main disadvantage of this approach is that a disaster location in the historical dataset is expected to have an average amount of people affected in the average unknown year. However, disasters are typically uncertain in their size and this is most likely not represented by an average number. For instance, it is perfectly possible that a disaster strikes somewhere and affects a more than average number of people in that region. To find the depots for an average unknown year, we instead assume that a_c belongs to an *uncertainty set*. This is a continuous set of scenarios for the parameter a_c that might occur in the unknown year. We want to find solutions that guarantee a certain (maximum) cost of transportation/response time, independent of the occurred realization of a_c . This is done by using robust optimization. The set of scenarios is based on two assumptions. The first assumption is that, in the unknown year, a similar *total* amount of people is expected to be affected by a disaster as in previous years. To represent this assumption in mathematical terms, we define p as the sum of the mean number of people affected over all AHs, i.e., $p = \sum_{c \in \mathcal{C}} \bar{a}_c$. Now, the first assumption can be written as:

$$\sum_{c \in \mathcal{C}} a_c \leq p.$$

The second assumption is that the number of people affected in any cluster is bounded from below and above. Based on historical data we determine lower and upper bounds for the number of people affected in the average unknown year in each cluster, say l_c and u_c for all $c \in \mathcal{C}$. In mathematical terms, we say that a_c can be any value for which $a = [a_1 \ a_2 \ \dots \ a_{|\mathcal{C}|}]$ is part of the following uncertainty set:

$$\mathcal{U} = \left\{ a \in \mathbb{R}^{|\mathcal{C}|} : \sum_{c \in \mathcal{C}} a_c \leq p, \ l_c \leq a_c \leq u_c \quad \forall c \in \mathcal{C} \right\}.$$

Note that this uncertainty set is also called a *budget uncertainty set* ([15]). The method used to obtain the upper and lower bound on the number of people affected per cluster is called the Min-Max method. When we have several years of historical data available, we can use the Min-Max method to obtain l_c and u_c . Suppose, each year of the historical data there are at least 200 affected people provided with relief items via cluster c' . Moreover, the maximum number of affected people provided with relief items via cluster c' in the considered years is 20,000. Then, the lower and upper bound for this cluster in the average unknown year can be defined as 200, respectively 20,000. In mathematical terms, $l_{c'} = 200$ and $u_{c'} = 20,000$. Note that, unless a cluster experienced a disaster each year in the historical data used, the lower bound for that cluster is zero.

Next, we describe how to find a solution to the optimization problem in which $a \in \mathcal{U}$. In other words, how to find a solution that guarantees a minimal cost of transportation, independent of the occurred realization of a . To do this, we use the following (new) objective function:

$$\text{minimize} \quad \max_{a \in \mathcal{U}} \left\{ \sum_{h \in \mathcal{H}} \sum_{c \in \mathcal{C}} a_c c_{hc} x_{hc} \right\}.$$

Next, we want to get rid of the maximization problem in this objective, since this makes the optimization problem nonlinear. Note that we can write this maximization problem as follows:

$$\begin{aligned} &\max_{a \in \mathbb{R}^{|\mathcal{C}|}} && \sum_{h \in \mathcal{H}} \sum_{c \in \mathcal{C}} a_c c_{hc} x_{hc} \\ &\text{s.t.} && \sum_{c \in \mathcal{C}} a_c \leq p, \\ &&& l_c \leq a_c \leq u_c, \quad \forall c \in \mathcal{C}. \end{aligned}$$

Table 3
Data point of the UNHRD dataset.

Depot	Country	Transport Type	Gross Weight (mt)
Accra	Sierra Leone	Sea	2.71

The next step is to formulate the dual of this optimization problem. For that, we introduce several new variables, $\lambda \in \mathbb{R}_+$, $\gamma \in \mathbb{R}_+^{|\mathcal{C}|}$ and $\beta \in \mathbb{R}_+^{|\mathcal{C}|}$. The dual of the problem is then given by the following optimization problem:

$$\begin{aligned} \min_{\lambda \in \mathbb{R}_+, \gamma \in \mathbb{R}_+^{|\mathcal{C}|}, \beta \in \mathbb{R}_+^{|\mathcal{C}|}} \quad & \lambda p + \sum_{c \in \mathcal{C}} \gamma_c u_c - \sum_{c \in \mathcal{C}} \beta_c l_c \\ \text{s.t.} \quad & - \sum_{h \in \mathcal{H}} c_{hc} x_{hc} + \gamma_c - \beta_c + \lambda \geq 0, \quad \forall c \in \mathcal{C}. \end{aligned}$$

This means that the robust optimization model can be summarized as in Model 2.

5. Datasets and preparation

In this section, we describe how the elements of the sets $\mathcal{H}$ and $\mathcal{C}$, and the values of the parameters a_c , t_{hc} and c_{hc} , of the model described in Section 4 are determined. The elements of $\mathcal{C}$ and the values of the parameter a_c , are based on a so-called *disaster database*. This database contains information about the location and size (the amount of relief items needed in metric tonnes) of historical disasters. Before discussing the determination of all parameter values, we first describe the two disaster databases that we use.

5.1. Disaster database

In this paper two different case studies are distinguished. One that is based on historical data provided by UNHRD and one that is based on historical data from the [19] database. The data provided by UNHRD contains information about the outflow of goods from all their depots in the months between January 2017 and April 2019. The EM-DAT database is an independent database that also contains information about disasters that did not receive relief items via one of the UNHRD depots. Moreover, this database contains data from a larger time frame, and it contains, other than the UNHRD dataset, information about *when* a disaster happened (which year). These last two features specifically allow us to examine the added value of using a robust optimization framework. We discuss, for both case studies, the characteristics of the disaster database.

First, UNHRD provided us with data about the outflow of goods from all their depots in the months between January 2017 and April 2019. Each data point is structured in a similar way as in Table 3.

This example shows that in the 2.5 years between January 2017 and April 2019, there has been a shipment from the depot in Accra (Ghana) to Sierra Leone. The weight of this cargo was 2.71 metric tonnes (or 2,710 kilograms) and it was transported via sea. Note that there may have been multiple shipments from Accra to Sierra Leone in these 2.5 years. All these shipments are registered separately in the UNHRD dataset. We assume that the point to which the relief items are shipped is the center point of the receiving country. Since there is no time registration for a shipment (besides the information that it occurred between January 2017 and April 2019), we do not know whether shipments to a specific country are destined for the response of the same disaster. So, we either use all separate records in the UNHRD dataset as disasters, or we combine the records that ship relief items to the same country into one disaster. In the latter case, the size of the disaster is the total weight of incoming shipments to this country. Note that both

Table 4
Data point of the EM-DAT database.

Year	Country	Location	Disaster type	Total affected
2017	Bangladesh	Dinajpur, Kurigram,...	Flood	8,000,000

choices result in the same optimal solution(s). This is due to the assumption that all relief items are delivered via the nearest airport/harbor, and from the (open) depot that results in the least amount of transportation costs, or that has the smallest response time (see Section 4.2). When choosing to combine all records that ship relief items to the same country, we end up with a disaster database containing 113 different disaster locations (countries).

Secondly, we use data from [19]. The EM-DAT database contains data on the occurrence and effects of natural and technological disasters from 1900 to the present day. It is based upon various sources such as the UN agencies, the US Office of Foreign Disaster Assistance, national governments, the International Federation of Red Cross and Red Crescent Societies, NGOs, insurance companies, research institutes and the media. For a disaster to be entered into this database, either i) 10 or more people have been reported killed, or ii) more than 100 people have been reported affected, or iii) a declaration of a state of emergency has been released or iv) a call for international assistance has been made. We specifically use disasters that occurred between 2000 and 2017. Furthermore, we only consider the natural sudden onset disasters: earthquakes, storms, landslides and floods. These are the disasters for which we assume relief items are mostly delivered via HLSPs. Finally, we assume that only disasters that happened in a country that is not classified as a country with a very *high* Human Development Index (HDI) (above 0.8) are served from one of the depots. Countries with a very high HDI are expected to resolve the disaster situation without the help of HLSPs. Each data point of this EM-DAT database is structured similarly as in Table 4.

Every record in the EM-DAT database contains information about a specific disaster with a time registration. Moreover, the disaster location in this dataset refers to a set of regions/places or provinces that are hit by the disaster in the corresponding country. For this database, we assume that the point to which the relief items are shipped is the arithmetic mean of the GPS coordinates that represent the center points of the locations of the regions where the disaster hit. Finally, the size of a disaster in the EM-DAT database is reported as a total number of people affected. To be able to determine the shipment size in metric tonnes, we determine the average amount of relief items needed for one individual affected by a disaster. We assume a linear relationship between the amount of relief items needed and the total number of people affected. Using nine operation update reports of [21]², we find that we need to transport approximately 0.128 metric tonnes of relief items for every 1,000 people affected. Based on this number, we can now determine the amount of relief items needed in a disaster (in metric tonnes). One important remark is that this number is a constant that will not affect the optimal depot locations, it is included to obtain more realistic estimates for the cost reductions. In the end, this results in a disaster database containing 1,558 disasters over 18 years, with an average of 87 disasters each year.

5.2. Parameter value determination

Next, we discuss how we determine the elements of the sets $\mathcal{H}$ and $\mathcal{C}$, and the values of the parameters a_c , t_{hc} and c_{hc} .

² Operation update reports summarize the information of a specific operation performed by the UNHRD. Information such as the amount of goods transported from any depot to the disaster area.

Fig. 2. The clustering method. In the original situation (a), we assign disasters to depots, and in the clustered situation (b), we assign Airport/Harbor combinations (AHs) to depots. Note that the airports and harbors are supplied via a depot. However, for simplicity, the depots are omitted from this figure.

$\mathcal{H}$: For $\mathcal{H}$, the set of potential depot locations, we determine all locations which contain a large airport *and* a large harbor. We use community data from ourairports.com to find all locations/cities containing a large airport. A *large* airport is defined as an airport that i) schedules major airline services with millions of passengers per year or ii) is a major world military base. For the harbors, we use a dataset maintained by the [22]. They summarize the seaports responsible for 99% of total shipments worldwide into one overview. All of these harbors are considered to be *large*.

$\mathcal{C}, a_c$: We determine the elements of $\mathcal{C}$, the set of clusters of disasters that are assigned to the same depot. Note that when minimizing the cost of transportation, while incorporating restrictions on the maximum response time, we only cluster disasters that happen at the same location. When minimizing one of the objectives, the cost of transportation or the maximum response time, the clusters are the combinations of an airport and a harbor that supply relief items during at least one of the disasters in the disaster database. $\mathcal{C}$ is obtained by clustering disasters for which the relief items are delivered through the same airport and harbor. To illustrate the clustering method, consider the situation in Fig. 2(a), where the affected areas from disasters are supplied through their nearest harbor and airport. The numbers indicate the number of people affected by that disaster. The two affected areas from the two disasters on the left are both supplied via Airport A and Harbor A. This means that in total, 30 people are affected in the airport/harbor combination AA. We cluster the two disasters on the left (with 10 and 20 people affected) into one group in which 30 people are affected. Similarly, group AB has 10 people affected and group BB has also 10 (5 + 5) people affected. The final (clustered) situation is shown in Fig. 2(b).

So, in the example, we now assign three airport/harbor combinations (AHs) to (open) depots (i.e., $\mathcal{C} = \{AA, AB, BB\}$). Recall that assigning these clusters to depots results in the same solution as assigning individual disasters to depots. In this way, considering the EM-DAT disaster database, we go from assigning 1,558 disasters to assigning 363 groups of disasters to open depots, while guaranteeing to obtain the same solution as when solving the model without the clustering procedure. This clustering algorithm is thus not restrictive and does not affect the quality of the solutions. By summing the total number of people affected by a disaster in each cluster (AH), we also obtain the total number of people affected in each cluster, a_c , for all $c \in \mathcal{C}$.

t_{hc} : The response time from potential depot location h to any disaster in cluster c consists of two time components. First, aircraft transport the relief items to the airport corresponding to cluster c . Then, trucks transport the relief items to the disaster site (Fig. 3).

Fig. 3. Response time of delivering relief items from a depot to the central location of an affected area.

Note that the trucks deliver the relief items to one location: the central location of the disaster. Using this definition, we can write t_{hc} , the response time to deliver relief items from depot location h to any disaster in cluster c , as follows:

$$t_{hc} = \alpha_{hc} + \tau_c,$$

where α_{hc} is the time it takes to transport relief items from depot location h to the airport corresponding to cluster c (using aircraft), and τ_c is the time it takes to deliver relief items from the airport corresponding to cluster c to the disaster site, that is furthest away (using trucks), in cluster c . To determine α_{hc} , we first compute the distance between depot h and the airport corresponding to cluster c . This distance is computed using the great circle distance, using a radius of the circle equal to the radius of the earth (6,373 km) plus 10 km (average altitude of an aircraft). Furthermore, we assume that an aircraft flies on average 850 km/h. Now, α_{hc} is computed by dividing the distance with this average speed.

The second time component, τ_c , is the time it takes to deliver relief items from the airport corresponding to cluster c to the disaster site that is furthest away (by truck). For this parameter, we determine the travel time between the airport corresponding to the cluster c and all disasters in cluster c . This is done by using an API from [23] that, among other things, includes the driving speed on roads when determining travel times. Then, τ_c is the maximum travel time from the airport corresponding to cluster c to a disaster in cluster c . One additional assumption is that we assume that the travel time to the nearest airport (from the central location) is less than 24h. If this is not the case, we do not use trucks to deliver goods. Instead, we only use aircraft to deliver the goods to the disaster area (e.g., by using airdrops containing relief items). In total, only about 8% of the disasters in the EM-DAT database are located at least 24 hours away from its nearest airport (given that at least one route could be found).

c_{hc} : The *transportation cost* of supplying relief items to one individual in cluster c from potential depot location h (c_{hc}), can be computed by considering the factors summarized in Fig. 4. In this figure, we also indicate which costs are *fixed* (FC). These costs are *always* made, regardless the proposed distribution of open depots, and hence do not influence the optimal solution. Therefore, we do not incorporate these costs into the optimization models, or use them when determining values for c_{hc} .

Fig. 4. Cost of delivering relief items from a depot to a disaster area. The cost of transporting relief items with trucks is fixed (FC) and is not incorporated in the optimization.

Table 5

Cost per km per metric tonne over sea ($\kappa^{(b)}$) and by air ($\kappa^{(a)}$).

Transport mode	Truck	Boat	Aircraft
Cost (€ /km/mt)	0.295	$(100/180) * 0.295$	$(3500/180) * 0.295$

So, when minimizing the cost of transportation, we only consider the transportation costs from the depot locations to the relevant airport/harbor. Note that the costs made for driving from the nearest harbor/airport to the central location of the disaster are fixed costs. In mathematical terms, we write c_{hc} as follows:

$$c_{hc} = \delta(\rho^{(a)}\kappa^{(a)}\gamma_{hc} + \rho^{(b)}\kappa^{(b)}\beta_{hc}),$$

where γ_{hc} (β_{hc}) is the distance, in kilometers, between depot location h and the airport (harbor) corresponding to cluster c , $\kappa^{(a)}$ ($\kappa^{(b)}$) is the cost per kilometer of transporting one metric tonne via air (sea), and $\rho^{(a)}$ ($\rho^{(b)}$) is the percentage of relief items delivered via air (sea). Moreover, δ represents the amount of metric tonnes needed for one individual at a disaster. As argued in Section 5.1, we need approximately 0.128 metric tonnes of relief items per thousand people affected. When determining γ_{hc} , we use the distance computations performed to determine t_{hc} . $\rho^{(a)}$ and $\rho^{(b)}$ are determined using the information about the mode of transportation in the UNHRD dataset. We conclude that 58.2% of demand is transported via sea and 41.8% is directly transported via air. To determine β_{hc} , we need to compute the shortest distance between potential depot locations and harbors when using a boat. To do this, we use an ESRI Shapefile of all land on earth. We then divide the earth into rectangles. We know that there are 360 degrees of longitude and 180 degrees of latitude. To be able to recognize all waters, we divide the earth in a grid of 180×5 rows and 360×5 columns. We classify each rectangle as land or water (based on the ESRI Shapefile). Then, we determine the shortest path from a possible depot location to a harbor, where we only allow for rectangles that are classified as water. Note that we take care of the shape of the earth over which we compute the shortest path.

To determine $\kappa^{(a)}$ and $\kappa^{(b)}$, we use information from year reports of [24]. The results are tabulated in Table 5. Recall that the cost of transporting relief items with trucks is not explicitly used in the model. When minimizing the cost of transportation, we only consider the transportation costs from the depot locations to the relevant airport/harbor.

6. Numerical results

In this section, we discuss the results of applying the nominal and robust optimization approach to the two different data sets. We discuss the results of the analyses on the UNHRD and EM-DAT database in Section 6.1 and 6.2, respectively. The calculations are done on a PC with a CPU that has a clock rate of 2.3 GHz and with 8GB RAM. Furthermore, the optimization model is implemented in Python and the model is solved using Gurobi 8.1.1.

6.1. UNHRD case

We start with the results of applying the nominal model to solve the ADDSOME and OPTIMAL problems using the UNHRD dataset. We specifically focus on the Pareto front, which typically provides insights in solutions before making decisions. The robust model will not be applied to this dataset. We do not have as much data as we have of the EM-DAT database to compare robust solutions with nominal solutions. The implications of using robust optimization can be better discussed for the EM-DAT database. To get an indication about how robust the solutions obtained in this section are, we perform a robustness analysis for some of the solutions. This means that we test the solutions on situations with unknown disasters (obtained from the EM-DAT database). It should be noted that all analyses may be applied to both datasets. Moreover, the accuracy of the response time for the Pareto front is set at one hour.

In this research, we assume that each disaster area receives relief items from the depot that results in the least amount of transportation costs, or from the depot that has the smallest response time. However, this may not always be the case. As discussed before, UNHRD is partner-driven regarding the logistics operations. Organizations that store items in one of the depots decide for themselves when and where to ship their items to. Organizations may have (specific) items only pre-positioned in one depot, when wanting to respond to a disaster further away. This may result in transportation flows that are not optimal from a cost perspective. Cost savings could be obtained when all relief items are delivered from the depot that results in the least amount of transportation costs. This, for instance, may require organizations to work together when pre-positioning their inventory. The total cost of transportation, from January 2017 to April 2019, when using the flows described in the UNHRD dataset, is approximately 130.1M. If we assume that all relief items come from the depot that leads to the lowest total cost of transportation, we would end up with a total cost of transportation of 86.3M, which is a 34% reduction in cost. In a similar way, the maximum response time of the current UNHRD network of six depots equals 29 hours, according to the definition of the response time made in this research. Since we assume that each disaster area receives relief items from the depot that results in the least amount of transportation costs, or receives items from the depot that has the smallest response time, the solutions obtained in this research will be compared with this minimal cost of transportation/response time.

We start with the ADDSOME problem, where we are allowed to open one additional depot. We refer to this problem as the ADDONE problem. We determine, given the current geographical distribution of depots, optimal locations for establishing one new depot. The suitability of any new location depends on the objectives: the maximum response time and the total cost of transportation. The Pareto front of this problem is summarized in Table 6, where the cost of transportation is summed over the whole period (i.e., January 2017 - April 2019).

We observe that there are only two Pareto optimal solutions. We either accept Solution 2 with a minimal maximum response time of 29 hours (and a total cost of 57.4M), or we go for a maximum response time of 24 hours (which is approximately 17% less) while the total cost of transportation is 49% higher. This choice depends on the preferences of the decision makers, in this case UNHRD. Note that the maximum response time of the current network is 29h. Therefore, with establishing one additional depot a transportation cost reduction of 33% may be attained, while keeping the maximum response time the same.

Note that both solutions require a depot in a specific country. For Solution 1, a new depot has to be established in Australia

Table 6

Pareto optimal solutions of the ADDONE problem using the UNHRD dataset.

	Proposed depot location	Total cost of transportation	Maximum response time
Solution 1	Adelaide, Australia	85.7M	24h
Solution 2	Mombasa, Kenya	57.4M	29h

Table 7

Rank-based and percentile-based criteria used to compare countries. The most right column indicates whether solutions become better when the value of the corresponding measure is higher/lower.

	Criterion	Abbreviation	Better solutions
Rank-based	Ease Of Doing Business	EODB	↓
	Corruption Perception Index	CPI	↓
	Logistics Performance Indicator	LPI	↓
Percentile-based	Political Stability and absence of Violence	PSnV	↑
	Regulatory Quality	RQ	↑
	Rule of Law	RoL	↑
	Corruption Control	CoC	↑

Table 8

Solutions of the ADDONE problem, all with a minimal maximum response time, compared based on the criteria listed in Table 7. The bold printed numbers represent the best choice among the solutions for a specific measure.

Time	Cost	Location	EODB	CPI	LPI	PSnV	RQ	RoL	CoC
24h	85.7M	Adelaide, Australia	18	13	18	78	98	93	93
26h	85.9M	Port Moresby, PNG	108	138	148	26	28	21	17
26h	86.0M	Auckland, New Zealand	1	2	15	99	99	98	100

and for Solution 2 in Kenya. However, countries have their own (dis)advantages in, for instance, political stability and the quality of available infrastructure. It can be of interest to compare our original solutions with alternative solutions that have the same maximum response time but which propose a new depot in a different country. We compare the countries on rank-based and percentile-based criteria as is done by [1]. Note that for the rank-based criteria, lower numbers are better, and for the percentile-based criteria, higher numbers are better. For example, a value that is greater than 75% of the values of other countries is said to be at the 75th percentile, where 75 is the percentile rank. The criteria we use are the Worldwide Government Indicators obtained from the [25], listed in Table 7.

We examine the optimal solution (with respect to the maximum response time), when we exclude all potential depot locations in Australia. This means that we end up with a solution that minimizes the maximum response time by adding a depot not in Australia. This solution has therefore other characteristics (EODB, CPI, etc.). Note that this solution is, in terms of transportation cost and maximum response time, dominated by Pareto optimal Solution 1 (adding a depot in Australia, Adelaide). After this step, we repeat the process to obtain another (dominated) solution, but again in a different country. The results are stated in Table 8.

We observe that having a depot in Australia is necessary to ensure a maximum response time of 24 hours. Based on the preferences of UNHRD decision makers, additional measures could be taken into account. Using this information, a choice can be made about where to establish a new depot and what costs are incurred for not choosing the optimal solution. A similar analysis is done for the solution with a maximum response time of 29 hours. The results are summarized in Table 9.

We observe that there is a large increase in the cost of transportation when we choose not to build a depot in Kenya. This is caused by the fact that Mombasa (Kenya) is the only potential depot location in eastern Africa, Jeddah and Durban are closest locations to Mombasa that are classified as potential depot locations.

So far, we looked at the situation in which we are allowed to add *one* additional depot. Next, we extend the analysis by comparing possible transportation cost and maximum response time

Fig. 5. The minimal cost of transportation/maximum response time when adding more than one depot to the network (ADD SOME problem).

reductions obtained when we are allowed to open more than one additional depot, the ADD SOME problem. The results are shown in Fig. 5.

Note that adding zero additional depots to the network represents the current situation (in which the total cost of transportation is 86.3M). Adding one additional depot may result in a total cost of transportation of 57.4M (by installing a depot in Mombasa, see Table 9). We observe that adding one additional depot to the network is likely to be the most efficient expansion in terms of transportation cost reduction. Adding more than three depots, of course, still reduces the cost of transportation, but not as much as the first expansions do. Regarding the maximum response time, we observe that by adding only two additional depots, the minimum maximum response time (23 hours) can be attained.

Secondly, we determine the Pareto front of the OPTIMAL problem. We first assume that we want to geographically distribute six depots across the globe that minimize both the total cost of transportation and the maximum response time simultaneously. The Pareto front of this problem is given in Fig. 7(a).

Table 9

Solutions of the ADDONE problem, all with a maximum response time of 29h, compared based on the criteria listed in Table 7. The bold printed numbers represent the best choice among the solutions for a specific measure.

Time	Cost	Location	EODB	CPI	LPI	PSnV	RQ	RoI	CoC
29h	57.4M	Mombasa, Kenya	61	144	68	13	44	38	15
29h	67.4M	Jeddah, Saudi Arabia	92	58	55	24	55	57	66
29h	77.5M	Durban, South Africa	82	73	33	36	63	52.4	57
29h	78.4M	Eilat, Israel	49	34	37	17	87	82	80
29h	81.4M	Yangon, Myanmar	171	132	137	13	19	17	32

Fig. 6. Pareto optimal solutions for the OPTIMAL problem on the UNHRD dataset. The green, red, black and yellow dots correspond to the solutions with a maximum response time of 23h, 24h, 29h and 31h, respectively.

As can be seen from this Pareto front, there are six Pareto optimal solutions. All solutions represent a specific trade off between the total cost of transportation and the maximum response time. We can compare these solutions with the current network of UNHRD, which is represented with a black dot. Note that the minimal maximum response time of the current network is 29 hours. From Fig. 7(a), we observe that, based on the assumptions made in this model, a maximum response time of 29 hours can be realized (with a network of six depots) with 56.8M, which would be 34% lower than the cost of transportation corresponding to the original UNHRD network. This strengthens the claim that expanding the current UNHRD network with for instance a single depot could significantly reduce the cost of transportation. In Fig. 6, four of the six Pareto optimal solutions are visualized on a map. These four solutions summarize the Pareto optimal solutions best.

We observe that there are four locations that are optimal in all four Pareto optimal solutions. The remaining two depots depend on the minimal maximum response time. The solutions that minimize the maximum response time (green/red dots) both open a depot in Australia. The number of people affected in this region is not very high, but, when minimizing the maximum response time, we also want to make sure that these more remote locations can be supplied within the minimized maximum response time. It will, however, not be an efficient idea when minimizing the cost of transportation. This can also be seen in Fig. 6, where the optimal locations for minimizing the cost of transportation are located in regions where more disasters happened.

Next, we also look at other sizes of networks. Until now, we composed networks of six depots. In Fig. 7(b), we show the results of optimizing for a different number of depots. First, we observe that a network of three depots may already result in the minimal maximum response time of 23h. Moreover, we see that the marginal value of having a network of seven or more depots becomes small. With the availability of between five and seven de-

pots, a network can be formed that does not improve greatly when adding more depots.

Finally, the solutions obtained in this section are based on nominal optimization. These solutions do not incorporate uncertainty in the location and scale of future disasters. To get an indication of how robust these solutions are against these uncertainties, we examine the performance of three solutions on uncertain disaster situations. The three solutions are the original UNHRD network, the solution of the ADDONE problem and the solution of the OPTIMAL problem (with six depots). The last two are based on the Pareto solution that minimizes the total cost of transportation. First, we compare the average case, in which the number of people affected in a cluster equals the average number of people affected per year in this cluster by disasters reported in the EM-DAT database between 2000 and 2017. Secondly, we compare the worst case performance, in which the number of people affected in a cluster is contained in the Min–Max uncertainty set based on the EM-DAT database. Note that the disasters in the EM-DAT database span the years between 2000 and 2017, where the UNHRD dataset covers only disasters that occurred between January 2017 and April 2019. The results are shown in Table 10.

We observe that the ADDONE solution does not result in a more robust solution compared to the original UNHRD network. The average and worst case cost are very similar. The ADDONE solution also does not reduce the average expected yearly cost as much as it did when using the UNHRD dataset. This means that when using a different dataset (in this case the EM-DAT database), results are likely to change due to differences in location and scale of incorporated disasters. We do observe that the OPTIMAL solution is expected to reduce the average expected yearly cost of transportation significantly. In the next section, we examine the added value of using robust optimization more thoroughly by obtaining solutions while incorporating the uncertainty in the location and scale of future disasters.

Fig. 7. The Pareto front for the OPTIMAL problem, and (b) the minimal cost of transportation/maximum response time when having different-sized networks (OPTIMAL problem). The black dot represents the current network of six depots of UNHRD.

Table 10

Average expected yearly cost of transportation (in million dollars (M)) of the solutions based on the UNHRD dataset, applied to disasters from the EM-DAT database (2000 - 2017).

	Original UNHRD network	ADDONE solution	OPTIMAL solution
Average	103.6	102.6	172.1
WC(Min-Max)	147.6	145.9	119.0

6.2. EM-DAT case

In this section, we specifically focus on the added value of solutions that are robust against the location and scale of future disasters. We do this by considering the OPTIMAL problem, using the EM-DAT database. For this case study, we do not discuss the Pareto front, because this gives similar insights as for the UNHRD case. We focus on minimizing the total cost of transportation, without incorporating the maximum response time. We start with the OPTIMAL problem in which we are allowed to open six depot locations across the globe. We solve the nominal model and verify whether the obtained solution is *robust* against uncertainty in the locations and scale of future disasters. Then, we extend the analysis with the OPTIMAL problem in which we are allowed to open other amounts of depot locations across the globe. Finally, we divide the data in a training and test set in order to examine the results of the optimization models on unknown data.

We start with the OPTIMAL problem in which we want to create a network of six depots. So, based on the data from EM-DAT, we want to find a set of depot locations that is optimal (minimal cost of transportation) for future years. Solving the nominal model results in a solution with an average yearly cost of transportation of 34.1M. We define this solution as the *nominal* solution. Note that this solution minimizes the cost of transportation in an average future (unknown) year, when such a year is equal to an average historical year. If this is not the case, other solutions might be preferred over this nominal solution. Therefore, we investigate whether this solution is robust against uncertainty in the scale and location of disasters. We do this by solving the robust optimization model introduced in Section 4.3. For this approach, we use upper and lower bounds on the number of affected people per airport/harbor combination (cluster), and we assume that the total number of affected people per year over all clusters stays approximately the same.

First, we examine the total number of people expected to be affected by a disaster in the average unknown year. We base this

Table 11

Robust optimization results (yearly average cost of transportation in million dollars (M)) using the EM-DAT database (2000 - 2017). Min-Max refers to the problem using the Min-Max uncertainty set.

	Nominal sol.	Min-Max sol.
Average	34.1	41.0
WC (Min-Max)	98.9	85.4

number on historical data. In Fig. 8, the total number of people affected by a disaster (as reported in the used EM-DAT database) is shown (for the years between 2000 and 2017).

Based on this figure, we can not conclude that there is a strong trend up/downwards. Therefore, in this case, we expect that in the average unknown year an average total number of people will be affected by a disaster. In mathematical terms, p is approximately 111M. At the end of this section, we perform a sensitivity analysis on this number. We compare results when using the minimum and maximum number of total people affected in a year (based on Fig. 8).

We start with determining the robust solution for finding a network of six depots that minimizes the cost of transportation. We do this based on the EM-DAT database of the years between 2000 and 2017. Both the nominal and robust solution are visualized in Fig. 9. We observe that the depots corresponding to the robust solution are more spread out over the world. The depots corresponding to the nominal solution are more clustered towards regions that had many people affected in disasters in the historical dataset. The cost of transportation for both solutions is shown in Table 11.

We observe that the robust solution guarantees a maximum cost of transportation of 85.4M per year. When we apply this solution to the average case (in which, for each cluster, the number of people affected in the unknown year equals the average number of people affected in that cluster from the previous years), we

Fig. 8. Total number of people affected by a disaster between 2000 and 2017 in million people (M).

Fig. 9. The nominal (green dots) and robust (red dots) solution for the OPTIMAL problem using historical data from 2000 till 2017.

obtain a cost of transportation of 41.0M. If we optimize on this average case, and obtain the *nominal* solution, the total cost of transportation would be 34.1M. On the contrary, if we apply the nominal solution to the case in which the number of people affected is in the Min–Max uncertainty set, we may end up with a transportation cost of 98.9M. Applying the robust solution is therefore expected to increase the average yearly cost of transportation with approximately 20% when it turns out that in the upcoming unknown years, the number of people affected is indeed equal to the mean of the previous years. However, if this is not the case, this solution may save approximately 14% per year in the cost of transportation. Note that this cost saving is in terms of a larger number. As an example, the 14% in our case corresponds to a cost reduction of 13.4M and the increase of 20% corresponds to 6.9M cost increase.

We also observe that the range of the cost of transportation between the average scenario and the worst case scenario is approximately 30% smaller for the robust solution. This suggests that the robust solution reduces the uncertainty in the expected cost of transportation in future years when optimizing a network of six depots. To check whether this claim also holds for other network sizes, we extend the analysis with experiments on the number of depots to be opened. The results are shown in Fig. 10.

As an example, for a network of six depots, the lower bound of the robust solution is 41.0M. This amount is incurred when it turns out that an average number of people is affected in the unknown year, when the robust solution has been implemented. The upper bound corresponds to the situation in which the worst case scenario of the Min–Max uncertainty set occurs.

Fig. 10. Average expected range of yearly cost of transportation in million dollars (M) for different-sized networks. The lower bound is the average situation and the upper bound is the worst case situation.

From this figure, we conclude that the robust solution in each case reduces the uncertainty in the expected cost of transportation. This may be beneficial for organizations that benefit from a more constant cost of transportation. As an example, if an organization relies on a constant flow of donations, it might be beneficial to allow a slightly higher average expected transportation cost, in return for less uncertainty in the expected transportation cost. Secondly, we observe that a larger network has the additional advantage of less uncertainty in the expected cost of transportation. We see that increasing the number of depots allowed to be opened to above four does not decrease the expected cost of transportation significantly. This suggests that any solution on the flat part of the graph is equally good. However, based on the result

Fig. 11. Average expected range of yearly cost of transportation in million dollars (M) for different-sized networks. The lower bound is the average situation and the upper bound is the worst case situation.

Fig. 12. Average yearly cost of transportation per person affected when using 2000-2013 as training set. Solutions are applied to (a) the (actual) data of the corresponding year, (b) the artificial uncertainty set resulting in a worst case scenario.

of robust optimization, we observe that, in this case, it is best to choose the solution that is furthest to right of the flat part. This solution is expected to have the smallest uncertainty in the expected cost of transportation. This result should be incorporated in general uncapacitated facility location problems with uncertain demand. Larger networks tend to have the additional advantage of less uncertainty in the expected cost of transportation. Even if at some point increasing the number of depots allowed to be opened may not decrease the expected transportation cost significantly, it does reduce the uncertainty in the expected transportation cost. This makes sense as when the number of depots increases, the average distance from a disaster site to a depot gets smaller.

Next, we examine the performance of both solutions when disasters occur that affect more people than disasters in that region did in the years of the historical data set used. The number of people affected per cluster is now bounded by the minimum number of people affected in the previous years in that cluster, and a *factor* times the maximum number of people affected in the previous years in that cluster. We examine two factors, 1.5 and 2. Note that originally this factor was equal to 1. The results are shown in Fig. 11.

From these figures, we observe that the conclusions about reducing the uncertainty still hold. Moreover, we observe that, especially for smaller networks, the robust solution is less affected by the evaluation on the larger uncertainty set. The expected average yearly cost of transportation does not change as much as when using the nominal solution. In Fig. 11, this can be concluded from the increasing size of the green area for the nominal solution.

In short, for this problem, there are two main conclusions. First, as expected, robust optimization can be used to find solutions that have less uncertainty in the expected cost of transportation. Secondly, when considering the OPTIMAL problem, adding more depots to a network has the additional advantage of uncertainty reduction in the expected cost of transportation. Even if at some point adding more depots may not decrease the expected cost of

transportation significantly, it does reduce the uncertainty in this expected cost.

Next, we focus on the number of years used in the optimization. We specifically consider the OPTIMAL problem in which we are allowed to establish six depots. We concluded that implementing a robust solution might be a better alternative than implementing the nominal solution. However, we did not check the results on a real test set of future (unknown) disasters. Therefore, we split the dataset in a *training* and *test* set. The training set is used for the optimization and the test set is used to check the performance of the solutions on unknown disasters. To do this, we determine optimal depot locations based on information from the years in the training set. We again compare the nominal and robust solution. Both solutions are tested on the years in the test set. Recall that when we consider only a few years in the training set, the model based on the Min-Max uncertainty set may not give reliable results. The upper and lower bounds do not represent an appropriate range. Also, in this case there may be many locations that did not encounter a disaster in the training set.

We consider a training set of 14 years, the years between 2000 and 2013. This means that we obtain the nominal and robust solution based upon the information from the years 2000 till 2013. These solutions are then evaluated on the years in the test set, 2014 to 2017. To be able to compare solutions, we focus on the cost of transportation per person affected instead of the total cost of transportation. Note that we know the number of people affected in these years. The results are shown in Fig. 12(a).

We observe that both solutions result in an approximately equal transportation cost per person for all test years. The nominal solution performs slightly better than the robust solution. A remark about these results is that we test on four different years. It is possible that no extreme events occurred in these years, which the robust solution incorporated in the uncertainty set. So if the performance of the robust solution is not very different from the nominal solution in the test years, but it guarantees a much better

Fig. 13. Computation time of the solver given the number of training years included (left figure), and given the size of the network to be optimized in number of depots (right figure). The blue (red) lines represent the solutions of the nominal (robust) optimization model.

solution in worst case scenarios (that apparently did not occur), then it is still a good/better alternative. Therefore, we also look at worst case scenarios. A worst case scenario, for instance for the year 2014, is defined as follows. Instead of only looking at the actual number of people affected per cluster, we now may expect all scenarios that are in the Min-Max uncertainty set based on the years 2000 to 2014. In other words, any disaster that happened in the past for a particular cluster, is also a possible disaster in the year 2014. We define the worst cases for 2015 to 2017 in a similar way. Note that the uncertainty set grows as the year becomes larger, since the minimum and maximum bound can only expand when more years are included. The results of applying the nominal and robust solution to the worst case scenarios are shown in Fig. 12(b).

We observe that, when applying the nominal and robust solution to the worst case scenario, the robust solution guarantees a much smaller transportation cost (approximately 15% smaller) per person than the nominal solution. Combining this with the result that the robust and nominal solution are comparable in the actual case, we conclude that the robust solution is a better alternative. This is especially the case when one is dealing with large uncertainty sets (i.e., when the minimum and maximum bound in the Min-Max uncertainty set are not close to each other). In this case, it is more beneficial to reduce the uncertainty in the expected cost of transportation. We may not expect all scenarios to happen often, but if we can safeguard ourselves to these scenarios without sacrificing much in the expected case, this is a better option than using the nominal solution.

Next, we analyze the practical computational behavior of the model. From theoretical (worst-case) complexity theory it is well-known that facility location problems are NP-hard. However, in practice, we are still able to solve our problem efficiently. We exploit, for instance, the fact that disasters can be clustered, in this case into airport/harbor combinations. Therefore, increasing the training years does not necessarily increase the amount of disaster locations that needs to be assigned to depots, as additional disasters may be absorbed into already existing airport/harbor clusters. We examine the relation between the computation time and i) the number of training years used, and ii) the number of depots to be opened. In Fig. 13, we show the computation time of the solver when solving the OPTIMAL problem, given both optimization model characteristics. When considering the number of training years (left figure), we optimize a network of six depots, and when considering the size of the network (right figure) we optimize using the full training set of 18 years (2000–2017).

First, we observe that the computation time increases approximately linearly with the number of training years. We also observe that optimizing a different-sized network does not affect the computation time much. So we conclude that, using the EM-DAT database, the computation time of the nominal and robust opti-

mization model is linear in the size of the possible inputs that we use (the number of training years and the size of the network).

We conclude this section by performing a sensitivity analysis on the expected total number of people affected by a disaster in an average unknown year (the parameter p in the robust optimization model). In this section, we used that in the average unknown year an *average* total number of people will be affected by a disaster. Next, we perform two sensitivity analyses on this parameter by considering the situation in which we expect a total number of people to be affected by a disaster equal to the total number of people affected in the year with the *highest* and *lowest* total number of people affected by a disaster. These numbers are shown in Fig. 14.

First, we perform the sensitivity analysis on the ADDSOME problem. We want to examine the possible uncertainty reduction in transportation costs when using robust optimization, as we did for the average p . To be able to compare the results with the original setting (using p), the situation in which an average number of people per cluster is affected should be in the uncertainty set of the robust optimization model. When using $p^{\min}$, this is not the case. The sum of the average number of people affected per cluster exceeds $p^{\min}$. Therefore, we focus on using $p^{\max}$ in this case. So, we again run the robust optimization model, but now based on $p^{\max}$. We also examine the situation in which we scale the possible number of people affected in a cluster by 2. Fig. 15 shows the results of applying both solutions to the average case and the worst case (for both situations). Note that the nominal solution in both situations does not depend on the choice for p .

The main difference is that, when using $p^{\max}$, the robust solution does not reduce the uncertainty in transportation costs for a given problem as much as it did when using the original p . This is likely to be the case because evaluating the solutions on such a large uncertainty set (in which $p^{\max}$ people are affected by a disaster somewhere in the world) means that many different locations may experience a disaster. Both solutions are simply not able to safeguard against all of these scenarios. Because of the large total number of people affected, an unfortunate (inefficient) situation can always be constructed. Therefore, based on this sensitivity analysis, we conclude that using $p^{\max}$ instead of p results in similar conclusions when determining the optimal number of depots by solving the OPTIMAL problem. Note that the total number of people affected does, as expected, affect the size of the observed uncertainty reduction of the robust solution.

Secondly, we perform the sensitivity analysis for determining the optimal depot locations by solving the OPTIMAL problem. We determine the changes in the expected yearly cost of transportation when implementing the original solution (with p) instead of the solution based on $p^{\min}$ or $p^{\max}$. The analysis is done for the whole training set of years between 2000 and 2017. Table 12 lists the possible cost *increases* due to using the solution based on p .

Fig. 14. Total number of people affected by a disaster between 2000 and 2017.

Fig. 15. Average expected range of yearly cost of transportation in million dollars (M) for different-sized networks using the uncertainty set based on $p^{\max}$. The lower bound is the average situation and the upper bound is the worst case situation.

Table 12

Possible percentual transportation cost increases when using p instead of $p^{\min}$ or $p^{\max}$ when solving the OPTIMAL problem.

		Min-Max sol.	
		$p^{\min}$	$p^{\max}$
WC (Min-Max)	$p^{\min}$	4.8%	-1.5%
	p	-2.3%	-0.5%
	$p^{\max}$	-1.5%	0.1%
Average	$p^{\min}$	-0.3%	0.7%
	p	-0.3%	0.7%

In this table the columns represent solutions of the optimization model. As before, the Min-Max solution refers to the solution of the robust optimization model when using the Min-Max uncertainty set. Additionally, now $p^{\min}$ and $p^{\max}$ represent the p used in the uncertainty set of this optimization model. For instance, the first column represents the possible cost increases that may be incurred when using p instead of $p^{\min}$ in the robust optimization framework. The rows represent the scenarios for which the solutions are tested. Here, the most second column refers to the p used in the scenario. As an example, the first row is the scenario in which the vector of people affected in each cluster is contained in the Min-Max uncertainty set, where a total of 39M ($= p^{\min}$) people are affected in the average unknown year.

We observe that most of the cost increases are *negative*. This means that by using p , we end up with better results in most of the scenarios. The largest cost increase may be obtained when using the robust model with its uncertainty set based on p , and implement it in a scenario where the total number of people affected happens to be $p^{\min}$. In this case, we would end up with a maximum cost increase of approximately 4.8%. However, in expectation, choosing this solution (based on p) over the solution based on $p^{\min}$, results in a cost increase of 0.2%. Similarly, in expectation,

choosing the solution based on p over the solution based on $p^{\max}$ results in a cost *decrease* of 0.3%. Therefore, based on this sensitivity analysis, we conclude that using p is not a very restrictive constraint when solving the OPTIMAL problem.

7. Concluding remarks

When a disaster strikes an area, often international assistance is requested to help responding and recovering from the disaster. This response can be characterized by the dispatch of relief items to the affected regions. Often relief agencies make use of storage space provided by Humanitarian Logistics Service Providers (HLSPs). In this paper, we discuss the problem of finding optimal depot (storage) locations for an HLSP. An optimal network of depots aligns with two main objectives, the cost of transportation and the maximum response time to a disaster.

Our methodology starts by defining a list of candidate depot locations. These locations all contain a large harbor and are close to a large airport. Then, we compute the costs related to delivering relief items from a candidate depot location to a disaster region. More specifically, the disaster regions are clusters of disasters that receive relief items from the same open depot. The costs consist of the cost of delivering via aircraft (immediate response phase) and the cost of delivering with ships (recovery phase). Furthermore, we compute the response times of delivering goods to disaster sites from the candidate depot locations. Based on these parameters, we develop a nominal optimization model that uses historical data to determine optimal depot locations. We examine the trade-off between the two objectives using the Pareto front. Furthermore, we apply robust optimization in order to find solutions that are robust against uncertainty in the location and scale of future disasters.

We apply the nominal model to a case study from the United Nations Humanitarian Response Depot (UNHRD), a large, globally operating, HLSP. We start by finding Pareto optimal solutions using historical data. We observe that a fast response often leads to

a larger total cost of transportation and vice versa, which clearly shows the trade-off between the cost of transportation and the maximum response time. The best solution for UNHRD can be obtained by using its preferences for both objectives. Based upon these preferences, one of the Pareto optimal solutions qualifies as the *best* solution. Besides the preference for objectives, UNHRD also has a preference for solutions. For instance, it might be easier to establish a depot in certain countries. Therefore, we include country specific measures (e.g., the Ease of Doing Business measure (EODB) and the Corruption Perception Index (CPI)) that may influence decision-making for UNHRD. Based on the optimization results, we conclude that establishing a depot in Mombasa (Kenya) is the optimal expansion strategy for UNHRD. This candidate depot location is also part of all the Pareto optimal solutions when optimizing a network of six depots. We extend the analysis for UNHRD with examining the results when more depots can be opened. We conclude that adding more than three depots to the current network does not have that much added value to the network. Most of the added value can be obtained by adding one additional depot to the network. Moreover, we conclude that a network with between five and seven depots is likely to be optimal in a way that adding more depots is not of large added value and opening less depots has a large negative impact on the cost of transportation. Finally, we observe that when evaluating the solutions using a different dataset, results are likely to change due to differences in the location and scale of incorporated disasters. This shows the potential for robust solutions, solutions that are robust against the uncertainty in the location and scale of future disasters.

We generalize the approach by applying it to the independent database EM-DAT, that also contains information about historical disasters that may not be in the UNHRD dataset. Since this database also reports the year of a disaster, we use robust optimization to verify whether our solutions of the nominal model are robust against uncertainty in the location and scale of future disasters. In other words, we want to find solutions that are optimal for future years, instead of for previous years. To do this, we define, based on historical data, an uncertainty set that captures possible (future) disaster scenarios. When defining this uncertainty set, we assume that the total number of people affected in a cluster may vary between an upper and lower bound. Furthermore, we assume that the yearly *total* number of people affected by a disaster stays approximately the same.

We start with evaluating the robust optimization results when deciding about the number of depots that can be opened. We conclude that robust optimization can be used to reduce the uncertainty in the average expected yearly cost of transportation compared to using nominal optimization. Secondly, we conclude that a larger network of depots has the additional advantage of uncertainty reduction in the expected cost of transportation. Even if at some point adding more depots does not decrease the expected transportation cost significantly, it does reduce the uncertainty in this expected transportation cost. This result should be incorporated in any uncapacitated facility location problem in which there is uncertainty in the demand. Solving such models often leads to a decreasing marginal transportation cost reduction when having larger networks (in which more depots are opened). However, larger networks tend to have the additional advantage that they are likely to reduce the uncertainty in the expected cost of transportation.

Moreover, we test the results on unknown (to the optimization model) data. We consider the problem in which we are allowed to establish six depots. To evaluate results, we split the historical data in a training and a test set. The training set is used for the optimization and the test set is used to evaluate the solutions. We observe that both the nominal and robust solution obtained from the training data result in an approximately equal transportation

cost per person for all test years in the test data. We also observe that, when applying the nominal and robust solution to worst case scenarios, the robust solution guarantees a much smaller transportation cost per person than the nominal solution. This means that, especially when one is dealing with large uncertainty sets, the robust solution may be a better option. In this case, it is more beneficial to reduce the uncertainty in the expected cost of transportation. We may not expect all scenarios to happen often, but if we can safeguard ourselves to these scenarios without sacrificing much in the expected case, this is a better option than using the nominal solution. Note that the results of using robust optimization with the EM-DAT database are also applicable for the UNHRD case study.

Potential lines for further research include extending the current model with, for instance, disaster-specific demand. In the models developed, we used the same amount/type of relief items needed by an individual in any disaster type (e.g., earthquake, flood, landslide, etc.). Furthermore, one can extend the approach with a secondary step of finding the capacities of potential depot locations, provided that the data about the depots is available, as discussed before.

Declaration of Competing Interest

No.

CRediT authorship contribution statement

V.F. Stienen: Conceptualization, Methodology, Software, Writing - Original Draft. **J.C. Wagenaar:** Conceptualization, Supervision. **D. den Hertog:** Conceptualization, Supervision. **H.A. Fleuren:** Conceptualization, Supervision.

Acknowledgement

We would like to thank United Nations Humanitarian Response Depot (UNHRD), in particular Georgia Farley, for the collaboration and for providing us with the data/information for the case study. Furthermore, we thank Perry Heijne from the Zero Hunger Lab at Tilburg University for his valuable feedback and insights during this whole research.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.omega.2021.102494](https://doi.org/10.1016/j.omega.2021.102494)

References

- [1] Dufour É, Laporte G, Paquette J, Rancourt M-È. Logistics service network design for humanitarian response in East Africa. *Omega* 2018;74:1–14.
- [2] Duran S, Gutierrez MA, Keskinocak P. Pre-positioning of emergency items for care international. *Interfaces* 2011;41(3):223–37.
- [3] Bozkurt M, Duran S. Effects of natural disaster trends: a case study for expanding the pre-positioning network of CARE international. *Int J Environ Res Public Health* 2012;9(8):2863–74.
- [4] Besiou M, Pedraza-Martinez AJ, Van Wassenhove LN. OR applied to humanitarian operations. *Eur J Oper Res* 2018;269(2):397–405.
- [5] Van Wassenhove LN. Humanitarian aid logistics: supply chain management in high gear. *J Oper Res Soc* 2006;57(5):475–89.
- [6] Altay N, Green III WG. OR/MS research in disaster operations management. *Eur J Oper Res* 2006;175(1):475–93.
- [7] Galindo G, Batta R. Review of recent developments in or/ms research in disaster operations management. *Eur J Oper Res* 2013;230(2):201–11.
- [8] Behl A, Dutta P. Humanitarian supply chain management: a thematic literature review and future directions of research. *Ann Oper Res* 2018;283(1–2):10011044.
- [9] Caunhye AM, Nie X, Pokharel S. Optimization models in emergency logistics: A literature review. *Socio-Econ Plan Sci* 2012;46(1):4–13.
- [10] Laporte G, Nickel S, Saldanha-da Gama F. Location science. Springer; 2019.
- [11] Correia I, Saldanha-da Gama F. Facility location under uncertainty. In: *Location science*. Springer; 2019. p. 185–213.

- [12] Balci B, Beamon BM. Facility location in humanitarian relief. *Int J Logist* 2008;11(2):101–21.
- [13] Boonmee C, Arimura M, Asada T. Facility location optimization model for emergency humanitarian logistics. *Int J Disast Risk Reduct* 2017;24:485–98.
- [14] Mulvey JM, Vanderbei RJ, Zenios SA. Robust optimization of large-scale systems. *Oper Res* 1995;43(2):264–81.
- [15] Ben-Tal A, El Ghaoui L, Nemirovski A. Robust optimization, 28. Princeton University Press; 2009.
- [16] Zokaee S, Bozorgi-Amiri A, Sadjadi SJ. A robust optimization model for humanitarian relief chain design under uncertainty. *Appl Math Model* 2016;40(17–18):7996–8016.
- [17] Bertsimas D, Thiele A. A robust optimization approach to inventory theory. *Oper Res* 2006;54(1):150–68.
- [18] Groß E, Fischer K. Prepositioning of relief items under uncertainty: A classification of modeling and solution approaches for disaster management. In: *Logistics Management*. Springer; 2016. p. 189–202.
- [19] EM-DAT. 2019. The Emergency Events Database - Université Catholique de Louvain (UCL) - CRED, D. Guha-Sapir - www.emdat.be, Brussels, Belgium.
- [20] UNHRD. UNHRD background. 2019a. Accessed: 02-12-2019 from <https://unhrd.org/about-us>.
- [21] UNHRD. UNHRD publications. 2019b. Accessed: 14-12-2019 from <https://unhrd.org/publications>.
- [22] Virginia Economic Development Partnership. Seaports of the world. 2014. Accessed: 10-11-2019 from <http://exportvirginia.org/wp-content/uploads/2014/04/Seaports-of-the-World.pdf>.
- [23] Mapbox. Directions API mapbox. 2020. <https://www.mapbox.com>.
- [24] WFP. WFP logistics in 2012: Changing the way we deliver. 2012. <https://www.wfp.com>.
- [25] World Bank. Worldwide Governance Indicators. 2019. Accessed: 19-11-2019 from <https://info.worldbank.org/governance/wgi/>.