OPTIMAL DEPOT LOCATIONS FOR HUMANITARIAN LOGISTICS SERVICE PROVIDERS USING ROBUST OPTIMIZATION

By

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

12 October 2020

ISSN 0924-7815
ISSN 2213-9532
Optimal depot locations for humanitarian logistics service providers using robust optimization

V.F. Stienen¹, J.C. Wagenaar¹, D. den Hertog¹, H.A. Fleuren¹

¹ Tilburg University, Department of Econometrics and Operations Research, Tilburg, The Netherlands.

September, 2020

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. One of the conclusions is that expanding the current UNHRD network with one additional depot can result in a transportation cost reduction of up to 33%. 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.

Keywords Humanitarian Logistics · Pre-positioning · Preparedness · Facility Location · Transportation · Uncertainty · Robust Optimization
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 time it takes to deliver emergency items to the affected people. For a single 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 may be too expensive to realize. To increase the efficiency and effectiveness of the coordination of multiple relief agencies, Humanitarian Logistics Service Providers (HLSP) are established. As the name already suggests, an HLSP offers services to humanitarian organizations that improve the efficiency of the supply chain of these organizations. 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. Therefore, we also determine optimal depot locations based on open source data from the EM-DAT database. This database consists of records of historical disasters, containing information on, for instance, the total number of people affected by a disaster. Furthermore, we develop a robust optimization model to obtain solutions that are robust against uncertainty in the location and scale of future disasters. The scale of a disaster is represented by the total number of people affected by a disaster.

2 Literature review

Dufour et al. (2018) also examine the UNHRD network. They analyze 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. An important remark is that the location of this new distribution center (Kampala) is not determined by optimization. Instead, they optimize the inventory position at this chosen depot. To do this, they compare the cost of shipping containers

\footnote{It is selected by UNHRD and the WFP office in Nairobi, mainly because it represents a strategic center supporting WFP logistics in this region.}
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. In our research, we describe an optimization approach for determining optimal depot locations. The secondary decisions about what stock should be pre-positioned where could subsequently be evaluated using the approach described by Dufour et al. (2018).

Another paper that examines optimal depot locations for a large organization similar to an HLSP is written by Duran et al. (2011). They study the case of the international humanitarian organization CARE. They assume a given set of depot locations at which CARE can store its 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 open source historical data from the EM-DAT database of previous disasters between 1997 and 2006. Note that when there is a maximum number of depots that can be opened, they implicitly select optimal depot locations out of the given set of depots. This is done in such a way that the network minimizes the average response time to the historical disasters used from the EM-DAT database. In our research, we include the determination of the set of potential depot locations. In Bozkurt and Duran (2012), 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 Besiou et al. (2018).

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 Van Wassenhove (2006) and Altay and Green III (2006). 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. Galindo and Batta (2013) is a continuation of the work of Altay and Green III (2006). 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 Altay and Green III (2006) have been covered. For a more recent overview of the work done in humanitarian supply chain management, we refer to Behl and Dutta (2018).

Caunhye et al. (2012) 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 Laporte et al. (2019). 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 Correia and Saldanha-da Gama (2019) 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.

Balcik and Beamon (2008) 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 Boonmee et al. (2017). 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 Boonmee et al. (2017) all use a robust programming approach as developed by Mulvey et al. (1995). 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 Mulvey et al. (1995). 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 Ben-Tal et al. (2009). 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 Correia and Saldanha-da Gama (2019).

Zokaee et al. (2016) 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 Bertsimas and Thiele (2006). 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, which incorporate the uncertain environment of humanitarian logistics. Graß and Fischer (2016) review the literature concerning the pre-positioning of relief items combined with decisions related to facility location problems (e.g., distribution centers). They dis-
tistinguish 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., [Boonmee et al. (2017)]). 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 runtime 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 runtime. Based on this robust optimization framework, we show that when dealing with an uncapacitated facility location problem with uncertain demand in general, allowing more facilities to be opened does not only reduce the cost of transportation but it also likely reduces 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. Finally, we provide a case study with 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. We also 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 HLSPs. Finally, we describe 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 previously. 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 Figure 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 (UNHRD 2019a).

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. In this way economies of scale can be obtained. In this paper, we use the UNHRD network as a case study for our research. We evaluate the 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 distribution of depots. The first one is the cost of transportation, which summarizes the costs of transporting relief items from a depot to a disaster site. Secondly, we consider the maximum response time of a distribution of depots. This quantity is an upper bound for the time it takes to provide a disaster site with goods. 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. Among other things, we analyze the specific problem of expanding the UNHRD network with additional depots, given the existing configuration.

In practice, the amount of relief items stored in a depot largely depends on the decisions of individual organizations (partners). Partners decide for themselves where to store what amount of which items, based on the current network. Therefore, at 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. The capacity of a depot location is usually not the bottleneck. So, in this research, we focus on the optimal *locations* and as a secondary step, the optimal capacity per depot can be determined. In other words, we do not incorporate the capacity of depots. If a disaster strikes, any depot is able to deliver all relief items needed. This implicitly means that relief items during a disaster are always supplied through a *single* depot location. The one for which delivering the relief items has either the lowest cost of transportation, or a minimum response time. Furthermore, it is important to note that 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 upon historical data of previous disasters. This means that we find the optimal distribution of depots for the disasters that took place in the past. This is an important assumption in selecting the locations of the depots. 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 decision of where to establish/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: cost of transportation and 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 therefore 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. When 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 supplied through their nearest open depot. The only way in which this will not be the case in
practice, is when an organization wants to deliver its relief items to a specific location that is closest
to a hub in which this organization does not have any items stored. Therefore, this is more related
to optimizing a pre-positioning strategy for an organization \(^2\) instead of finding good locations for
depots of an HLSP. Note that this problem also depends on the scope/budget of an organization.
Finally, relief items are supplied through their nearest harbor and nearest airport. Note that it could
happen that these locations are destroyed by the disaster, in case another airport/harbor should be
used. As this will still be an airport/harbor that is close to its original location, incorporating such
events will likely not change the results of this study and is considered to be beyond the scope of
this research.

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 about 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 the 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 is obliged to open. 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
through which (open) depots relief items are supplied to each (historical) disaster. In other words,
the model assigns disasters to (open) depots. Because we assume that relief items are supplied
through their nearest open depot and through their nearest harbor/airport, 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

\(^2\)The challenge of optimizing where to store how much of which items.
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 combination to (open) depots.

We define groups of disasters that share their nearest airport and harbor as airport/harbor com-
binations (AH). So, instead of assigning disasters to open depots, we can assign airport/harbor combinations (AH) to open depots without loss of optimality. To do this, let $c \in C$ be the set of airport/harbor combinations. In case $AH \ c \in C$ is assigned to depot location $h \in H$, it means that all relief items necessary for the disasters in cluster $c$ are being transported from depot $h$ to $AH \ c$.

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 through their nearest open depot. For instance, when taking depot capacity into account, it might happen that affected areas of disasters in one group are provided with relief items through 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.

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 $AH \ c \in C$ through depot location $h \in H$ by $c_{hc}$. Furthermore, let $a_c$ represent the total number of people affected in $AH \ c$. Let $t_{hc}$ be the minimum response time in which people affected in $AH \ c$ can be provided with relief items through 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 $AH \ c$. 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 $AH \ c$ are provided with relief items through (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 AH (no capacity restrictions on the depots). Therefore, in the optimum, people affected in one AH are always provided with relief items through a single (open) depot, the one for which the cost of delivering to this AH 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.

The objective summarizes the total transportation costs of the assignment of AHS to depots. The first constraint ensures that people affected in each AH can only be supplied with relief items through open depots. The second constraint makes sure that each AH is assigned to a depot and the third constraint bounds the total number of open depots to $n$. The fourth constraint ensures that pre-
\begin{align*}
\text{minimize} & \quad \sum_{h \in H} \sum_{c \in C} a_c c_h x_{hc}, \\
\text{s.t.} & \quad \sum_{c \in C} x_{hc} \leq |C| y_h, \quad \forall h \in H, \quad (M.1) \\
& \quad \sum_{h \in H} x_{hc} = 1, \quad \forall c \in C, \quad (M.2) \\
& \quad \sum_{h \in H} y_h \leq n, \quad (M.3) \\
& \quad y_h = 1, \quad \forall h \in H^f, \quad (M.4) \\
& \quad \sum_{h \in H} t_{hc} x_{hc} \leq T, \quad \forall c \in C, \quad (M.5) \\
& \quad x_{hc} \in \mathbb{R}^+, \quad y_h \in \{0, 1\}, \quad \forall h \in H, \ c \in C. \quad (M.6)
\end{align*}

Model 1: The nominal depot location problem.

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. This is done by solving Model 1, in which we replace the objective by:

\begin{align*}
\text{minimize} & \quad \max_{c \in C} \left\{ \sum_{h \in H} t_{hc} x_{hc} \right\}.
\end{align*}

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 Ben-Tal et al. (2009).

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 AH.

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 AH \(c\), based on the historical data. In mathematical terms, we let \(\bar{a}_c\) represent the mean number of people affected in AH \(c\) 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., \(p = \sum_{c \in C} \bar{a}_c\). Now, the first assumption can be written as:

\[
\sum_{c \in C} a_c \leq p.
\]

The second assumption is that the number of people affected in any AH 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 AH, say \(l_c\) and \(u_c\) for all \(c \in C\). In mathematical terms, we say that \(a_c\) can be any value for which \(a = [a_1 a_2 \cdots a_{|C|}]\) is part of the following uncertainty set:

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

Note that this uncertainty set is also called a budget uncertainty set \((\text{Ben-Tal et al. (2009)})\). The method used to obtain the upper and lower bound on the number of people affected per AH 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 AH \(c'\). Moreover, the maximum number of affected people provided with relief items via AH \(c'\) in the considered years is 20,000. Then, the lower and upper bound for this AH 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 an AH experienced a disaster each year in the historical data used, the lower bound for that AH is zero.

Next, we describe how to find a solution to the optimization problem in which \(a \in U\). 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:

\[
\minimize_{a \in U} \ \max_{h \in H} \left\{ \sum_{c \in C} \sum_{h \in H} a_ch_cx_{hc} \right\}.
\]
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:

\[
\max_{a \in \mathbb{R}} \sum_{h \in H} \sum_{c \in C} a_c c_{hc} x_{hc}
\]

\[
\text{s.t. } \sum_{c \in C} a_c \leq p, \quad l_c \leq a_c \leq u_c, \quad \forall c \in C.
\]

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

\[
\min_{\lambda \in \mathbb{R}_+, \gamma, \beta \in \mathbb{R}_+^{\lvert C \rvert}} \lambda p + \sum_{c \in C} \gamma_c u_c - \sum_{c \in C} \beta_c l_c
\]

\[
\text{s.t. } -\sum_{h \in H} c_{hc} x_{hc} + \gamma_c - \beta_c + \lambda \geq 0, \quad \forall c \in C.
\]

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

\[
\text{minimize } \lambda p + \sum_{c \in C} \gamma_c u_c - \sum_{c \in C} \beta_c l_c,
\]

\[
\text{s.t. } -\sum_{h \in H} c_{hc} x_{hc} + \gamma_c - \beta_c + \lambda \geq 0, \quad \forall c \in C,
\]

\[
\lambda \in \mathbb{R}_+, \quad \gamma \in \mathbb{R}_+^{\lvert C \rvert}, \quad \beta \in \mathbb{R}_+^{\lvert C \rvert},
\]

\[(M.1) - (M.6).\]

Model 2: The robust optimization model for the optimal depot location problem.

5 Datasets and preparation

The main information used to determine the parameter values introduced in Section 4 consists of the locations and sizes of previous disasters. In this paper two different datasets are used. We use actual historical data provided by UNHRD and we use historical data from the independent database [EM-DAT](2019). UNHRD provided us with data about the outflow of goods from all 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. There could be multiple data points that ship relief items from Accra to Sierra Leone. We do not know whether these shipments are destined
for the response of the same disaster. Therefore, in the rest of this paper, we consider each country in this dataset as a disaster. The size of the disaster is represented by the sum of the weights of the incoming shipments to this country. In the end, this results in 113 different disasters in this dataset.

The second dataset we use is from EM-DAT (2019). EM-DAT 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 the 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. Each data point is structured similarly as in Table 4.

### Table 3: Data point of the UNHRD dataset.

<table>
<thead>
<tr>
<th>depot</th>
<th>Country</th>
<th>Transport Type</th>
<th>Gross Weight (mt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accra</td>
<td>Sierra Leone</td>
<td>Sea</td>
<td>2.710</td>
</tr>
</tbody>
</table>

Contrary to the UNHRD dataset, we know for the EM-DAT database that each data point represents a disaster. Another difference between the UNHRD dataset and the EM-DAT database is that the location in the EM-DAT database refers to a set of regions/places or provinces that are hit by the disaster in the corresponding country. Also, in the EM-DAT database, the mode of transportation is not included. Finally, in the EM-DAT database, we know the number of people affected instead of the weight that is transported to the affected region. When using the EM-DAT database, we specifically use disasters that occurred between 2000 and 2017. Furthermore, we only consider the natural sudden onset disasters: earthquakes, storms, landslides and floods. 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. This final set of disasters consists of 1558 disasters over 18 years, with an average of 87 disasters happening per year.

Next, we discuss the establishment of $\mathcal{H}$, $\mathcal{C}$, $a_c$, $t_{hc}$ and $c_{hc}$. First, for $\mathcal{H}$, 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 Virginia Economic Development Partnership (2014). They summarize the seaports responsible for 99% of total shipments worldwide into one overview. All of these harbors are considered to be large.

Secondly, we determine the elements of $\mathcal{C}$, the set of airport/harbor combinations. Note that these
are the combinations of an airport and a harbor that supply relief items during at least one of the disasters in the historical dataset. $C$ is obtained by clustering disasters for which the relief items are supplied through the same airport and harbor. To illustrate the clustering method, consider the situation in Figure 2a, 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 AH in which 30 people are affected. Similarly, AH $AB$ has 10 people affected and AH $BB$ has also 10 (5 + 5) people affected. The final (clustered) situation is shown in Figure 2b.

![Diagram](image)

(a) Original situation  
(b) Clustered situation

Figure 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 (AH) to depots. Note that the airports and harbors are supplied via a depot. However, for simplicity, the depots are omitted from this figure.

So, in the example, we now assign three airport/harbor combinations (AH) to (open) depots (i.e., $C = \{AA, AB, BB\}$). Recall that assigning these groups to depots results in the same solution as assigning individual disasters to depots. In this way, considering the EM-DAT database, we go from assigning 1558 disasters to assigning 363 AHs to open depots, while guaranteeing to obtain the same solution as when solving the model without the clustering procedure. By summing the total number of people affected by a disaster in each cluster (airport/harbor combination), we also obtain the total number of people affected in each AH, $a_c$, for all $c \in C$.

Thirdly, the minimum response time from potential depot location $h$ to any disaster in AH $c$, $t_{hc}$, consists of two time components. First, aircraft transport the relief items to the airport closest to the disaster site. Then, trucks transport the relief items to the disaster site (Figure 3).

![Diagram](image)

Figure 3: Response time of supplying an affected area of a disaster from a depot to the central location of the affected area.
The point at which the trucks deliver the relief items is called the central location of the disaster. For the UNHRD case, this is the center point of the country. The central location of a disaster in the EM-DAT database is the arithmetic mean of the locations of the regions where the disaster hits. Using this definition, we can write $t_{hc}$ in mathematical terms:

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

where $\alpha_{hc}$ is the time it takes to fly from depot location $h$ to the airport corresponding to AH $c$, and $\tau_{hc}$ is the time it takes to deliver relief items to the disaster site that is furthest away from the airport corresponding to AH $c$. To determine $\alpha_{hc}$, we first compute the distance between depot $h$ and the airport corresponding to AH $c$. This distance is computed using the great circle distance, using a radius of the circle equal to the radius of the earth (6373 km) plus 10 km (average altitude of an aircraft). Furthermore, we assume that an aircraft flies on average 850 km/h. Now, $\alpha_{hc}$ is computed by dividing the distance with this average speed.

The second time component, $\tau_{hc}$, is the time it takes to deliver relief items (by truck) to the disaster site that is furthest away from the airport corresponding to AH $c$. For this parameter, we determine the travel time between the airport corresponding the AH $c$ and all disasters in AH $c$. This is done by using an API from Mapbox (2020) that, among other things, includes the driving speed on roads when determining travel times. The travel time between the disasters in AH $c$ and the corresponding airport is defined by the maximum travel time from this airport to any AH $c$, $\tau_{hc}$. 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 the truck to deliver goods, but we fly the airplane to the central location (where relief items are dropped out). 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).

Finally, we determine values for $c_{hc}$, which summarizes the transportation costs of supplying relief items to one individual in AH $c$ from potential depot location $h$. We know that relief items are partly supplied via aircraft (in the immediate response phase) and partly via sea (in the recovery phase). Note that the UNHRD dataset contains information about the mode of transportation used. Based on the historical data provided by UNHRD, we conclude that 58.2% of demand is transported via sea and 41.8% is directly transported via air. Similar as for $t_{hc}$, $c_{hc}$ can hence be computed by considering the factors summarized in Figure 4. When determining values for $c_{hc}$, we take out the fixed costs, which are the costs made for driving from the nearest harbor/airport to the central location of the disaster. These costs are always made, regardless the proposed distribution of open depots, and hence do not influence the optimal solution.

![Figure 4: Cost of supplying relief items to a disaster area from a depot. The cost of transporting relief items with trucks is fixed (FC) and is not incorporated in the optimization.](image)

So, when minimizing the cost of transportation, we only consider the transportation costs from the
depot locations to the relevant airport/harbor (AH). In mathematical terms, we write $c_{hc}$ as follows:

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

where $\gamma_{hc} (\beta_{hc})$ is the distance, in kilometers, between depot location $h$ and the airport (harbor) corresponding to AH $c$, and $\kappa^{(a)} (\kappa^{(b)})$ is the cost per kilometer of transporting one metric tonne via air (sea). Moreover, $\delta$ represents the amount of metric tonnes needed for one individual at a disaster. Note that $\gamma_{hc}$ can be easily obtained from the computations for $t_{hc}$. To determine $\beta_{hc}$, 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 $180 \times 5$ rows and $360 \times 5$ 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 [WFP (2012)]. The results are tabulated in Table 5. Note that the cost of transporting relief items with trucks is not explicitly used in the model.

<table>
<thead>
<tr>
<th>Transport mode</th>
<th>Truck</th>
<th>Boat</th>
<th>Aircraft</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost (€/km/mt)</td>
<td>0.295</td>
<td>$(100/180) \times 0.295$</td>
<td>$(3500/180) \times 0.295$</td>
</tr>
</tbody>
</table>

Table 5: Cost per km per metric tonne over sea ($\kappa^{(b)}$) and by air ($\kappa^{(a)}$).

The final parameter we determine is the amount of metric tonnes needed for one individual at a disaster, $\delta$. Note that this is only needed for the EM-DAT database, since the UNHRD dataset already contains this information. For the EM-DAT database, we base this amount upon the total number of people affected. We assume a linear relationship between the demand at a disaster and the total number of people affected. Using nine operation update reports of [UNHRD (2019b)]\textsuperscript{3}, we find that we need to transport approximately 0.128 metric tonnes of relief items for every 1,000 people affected. Note that it is a constant that will not affect the optimal depot locations. It is included to obtain more realistic estimates for the cost reductions.

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.

\textsuperscript{3} 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.
6.1 UNHRD case

We start with the results of applying the nominal model for solving the ADDSOME and OPTIMAL problems on 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, because we do not have as much data as for the EM-DAT dataset to test and compare robust solutions with nominal solutions. The implications of using robust optimization can be better discussed for the EM-DAT dataset. To get an indication about how robust the solutions obtained in this section are, we do 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 dataset). It should be noted that all analyses may be applied to both datasets.

The accuracy of the response time for the Pareto front is set at one hour. The minimized 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. Assuming that all goods are also transported from their nearest depot to the disaster site, the total minimal cost of supplying relief items to the affected areas of the disasters in the UNHRD dataset (from January 2017 to April 2019) equals 86.3M.

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 Pareto front of this problem is given in Figure 5a. We observe that there are only two Pareto optimal solutions. Both are tabulated in Table 6, where the cost of transportation is summed over the whole period (i.e., January 2017 - April 2019).

<table>
<thead>
<tr>
<th>Proposed depot location</th>
<th>Total cost of transportation</th>
<th>Maximum response time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solution 1</td>
<td>Adelaide, Australia</td>
<td>85.7M</td>
</tr>
<tr>
<td>Solution 2</td>
<td>Mombasa, Kenya</td>
<td>57.4M</td>
</tr>
</tbody>
</table>

Table 6: Pareto optimal solutions of the ADDONE problem using the UNHRD dataset.

So, 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 49.3% higher. This choice depends on the preferences of the decision makers, in this case UNHRD. Note that the minimal maximum response time of the current network is 29h. Therefore, with establishing one additional depot a transportation cost reduction of 33.5% 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 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 an other country. We compare the countries on rank-based and percentile-based criteria as is done by Dufour et al. (2018). 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 are the Worldwide Government Indicators obtained from the World Bank (2019), listed in Table 7.
### 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.

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

It turns out that Solution 1 is the only solution that ensures a maximum response time of 24 hours. 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.

### 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 solution for a specific measure.

<table>
<thead>
<tr>
<th>Time</th>
<th>Cost</th>
<th>Location</th>
<th>EODB</th>
<th>CPI</th>
<th>LPI</th>
<th>PSnV</th>
<th>RQ</th>
<th>RoL</th>
<th>CoC</th>
</tr>
</thead>
<tbody>
<tr>
<td>24h</td>
<td>85.7M</td>
<td>Adelaide, Australia</td>
<td>18</td>
<td>13</td>
<td>18</td>
<td>78</td>
<td>98</td>
<td>93</td>
<td>93</td>
</tr>
<tr>
<td>26h</td>
<td>85.9M</td>
<td>Port Moresby, Papua New Guinea</td>
<td>108</td>
<td>138</td>
<td>148</td>
<td>26</td>
<td>28</td>
<td>21</td>
<td>17</td>
</tr>
<tr>
<td>26h</td>
<td>86.0M</td>
<td>Auckland, New Zealand</td>
<td>1</td>
<td>2</td>
<td>15</td>
<td>99</td>
<td>99</td>
<td>98</td>
<td>100</td>
</tr>
</tbody>
</table>

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 to not build a depot in Kenya. This is caused by the fact that Mombasa (Kenya) is the only potential depot location in eastern Africa. This is due to Jeddah and Durban being the 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 reductions obtained when we are allowed to open more than one additional depot. The results are shown in Figure 5B.
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 solution for a specific measure.

![Graph](image1.png)

(a) Pareto front ADDONE.

![Graph](image2.png)

(b) Minimal cost of transportation/maximum response time analysis ADDSOME.

Figure 5: (a) The Pareto front for the ADDONE problem, and (b) the minimal cost of transportation/maximum response time when adding more than one depot to the network (ADDSOME problem).

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 Figure 5a). 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 Figure 7a.
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 Figure 7a 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 34.2% 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 Figure 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 southern 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 Figure 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 Figure 7b, we show the results of optimizing for a different number of depots. First, we observe that a network of three depots already has the minimal maximum response time. 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 depots, a network can be formed that does not improve greatly when adding more depots.

![Graph showing the Pareto front for the OPTIMAL problem and minimal cost of transportation maximum response time analysis with different number of depots.](image)

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 cost of transportation. First, we compare the average case, in which the number of people affected in an AH equals the average number of people affected over all years between 2000 and 2017 for this AH. Secondly, we compare the worst case performance, in which the number of people affected in an AH is contained in the $\text{Min-Max}$ uncertainty set based on the EM-DAT dataset.

<table>
<thead>
<tr>
<th></th>
<th>Original UNHRD network</th>
<th>ADDONE solution</th>
<th>OPTIMAL solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>103.6</td>
<td>102.6</td>
<td>72.1</td>
</tr>
<tr>
<td>WC(Min-Max)</td>
<td>147.6</td>
<td>145.9</td>
<td>119.0</td>
</tr>
</tbody>
</table>

Table 10: Average expected yearly cost of transportation for the solutions based on the UNHRD dataset (in million dollars (M)).

In Table 10, we observe that adding one additional depot using nominal optimization does not result
in a more robust solution. The average and worst case are very similar. 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 use of robust optimization more thoroughly as we are then able to obtain solutions based on robust optimization.

### 6.2 EM-DAT case

In this section, we consider the **OPTIMAL** problem, using the EM-DAT dataset. 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. Note that we will not discuss the Pareto front, because this gives similar insights as for the UNHRD case. 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 AH and we assume that the total number of affected people per year over all AHs 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 number on historical data. In Figure 8, the total number of people affected by a disaster (as reported in the used EM-DAT dataset) 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 Figure 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 dataset of the years between 2000 and 2017. Both the nominal and robust solution are visualized in Figure 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
Figure 8: Total number of people affected by a disaster between 2000 and 2017 in million people (M).

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

<table>
<thead>
<tr>
<th></th>
<th>Nominal sol.</th>
<th>Min–Max sol.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>34.1</td>
<td>41.0</td>
</tr>
<tr>
<td>WC (Min–Max)</td>
<td>98.9</td>
<td>85.4</td>
</tr>
</tbody>
</table>


year. When we apply this solution to the average case (in which, for each AH, the number of people
affected in the unknown year equals the average number of people affected in that AH in the previous years), we 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 Figure 10.

![Figure 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.](image)

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.

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 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 AH is now bounded by the minimum number of people affected in the previous years in
that AH, and a factor times the maximum number of people affected in the previous years in that
AH. We examine two factors, 1.5 and 2. Note that originally this factor was equal to 1. The results
are shown in Figure 11.

![Figure 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.](image)

From these figures, we observe that the conclusions about reducing the uncertainty still hold. More-
over, we observe that the robust solution is less affected by the evaluation on the larger uncertainty
set. The expected average yearly cost of transportation stays approximately the same, whereas in-
creases when using the nominal solution. In Figure 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. Sec-
ondly, 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 Figure 12a.

![Figure 12a](image)

Figure 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.

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 if it guarantees a much better 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 AH, we now may expect all scenarios that are in the \textit{Min-Max} uncertainty set based on the years 2000 to 2014. In other words, any disaster that happened in the past for a particular AH, is also a possible disaster in the year 2015. 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 Figure 12b.

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 may be 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 \textit{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 describe a brief complexity analysis. Solving the model requires runtime that depends on the size of the input of the model. We examine the relation between the runtime and i) the number of training years used, and ii) the number of depots to be opened. In Figure 13 we show the runtime of the solver when solving the \textit{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).

![Figure 13: Runtime 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.](image)

First, we observe that the runtime increases approximately linearly with the number of training years. We also observe that optimizing a different-sized network does not affect the runtime much. So we
conclude that, using the EM-DAT database, the runtime of the nominal and robust optimization 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 Figure 14.

First, we perform the sensitivity analysis on the ADDSOME problem. We focus on using the $p^\text{max}$ in this case, as this increases the expected cost of transportation. We want to avoid having a much higher cost. So, we again run the robust optimization model, but now based on $p^\text{max}$. Note that the nominal solution does not depend on the choice for $p$. Figure 15 shows the results of applying both solutions to the nominal case and the worst case, which is represented by the Min-Max uncertainty set with a factor 2, i.e., the maximum number of people affected per AH is increased by this factor.

The main difference is that in this case, when using $p^\text{max}$, the robust solution does not reduce the uncertainty 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 $p^\text{max}$ 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 $p^\text{max}$ 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. Figure 14 shows 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 $p$. 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 increase 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 left 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 AH is contained in the Min-Max uncertainty set, where a total of 39M (= $p_{min}$) people are affected in the average unknown year. Notice that a nominal solution can not occur in a scenario in which $p_{min}$ is used, since the sum of the people in the nominal scenario equals the average number of total people affected, $p$.

<table>
<thead>
<tr>
<th></th>
<th>Min-Max sol.</th>
<th>$p_{min}$</th>
<th>$p_{max}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{min}$ Nominal</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>WC (Min-Max)</td>
<td>4.8%</td>
<td>-1.5%</td>
<td></td>
</tr>
<tr>
<td>$p$ Nominal</td>
<td>-0.3%</td>
<td>0.7%</td>
<td></td>
</tr>
<tr>
<td>WC (Min-Max)</td>
<td>-2.3%</td>
<td>-0.5%</td>
<td></td>
</tr>
<tr>
<td>$p_{max}$ Nominal</td>
<td>-0.3%</td>
<td>0.7%</td>
<td></td>
</tr>
<tr>
<td>WC (Min-Max)</td>
<td>-1.5%</td>
<td>0.1%</td>
<td></td>
</tr>
</tbody>
</table>

Table 12: Possible percentual transportation cost increases for using $p$ instead of $p_{min}$ (left column) or $p_{max}$ when solving the OPTIMAL problem.
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 equals $p_{\text{min}}$. 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 $p_{\text{min}}$, results in a cost increase of 0.1%. Similarly, in expectation, choosing the solution based on $p$ over the solution based on $p_{\text{max}}$ results in a cost decrease of 0.1%. 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 region consists of an airport/harbor combination (AH) from which relief items are sent to the disasters areas. These 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, when applying the obtained solutions (based on the nominal model) to unknown datasets, we observe that adding one additional depot using nominal optimization does not result in a more robust solution.

We generalize this approach by applying it to the independent database EM-DAT containing also information about historical disasters that may not be in the UNHRD database. 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 past years. To do this, we define, based on historical data, an uncertainty set that captures possible (future) disaster scenarios. In defining this uncertainty set, we assume that the total number of people affected in an AH may vary between an upper and lower bound. Furthermore, we assume that the total number of people across the earth 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 allowing for 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 may not decrease the expected transportation cost significantly, does reduce the uncertainty in this expected transportation cost. This result should be incorporated in any uncapacitated facility location problems in which there is uncertainty in the demand. Solving such models often leads to a decreasing marginal transportation cost reduction when allowing for larger networks (in which more depots can be 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.
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.

References


