Models and Metrics to Assess Humanitarian Response Capacity

Jason Acimovic
Smeal College of Business, The Pennsylvania State University, acimovic@psu.edu

Jarrod Goentzel
Center for Transportation and Logistics, The Massachusetts Institute of Technology, goentzel@mit.edu

Working Paper - January 7, 2016

The race to meet vital needs following sudden onset disasters leads response organizations to establish stockpiles of inventory that can be deployed immediately. These government or non-government organizations dynamically make stockpile decisions independently. Even though the value of one organization's stock deployment is contingent on others' decisions, decision makers lack evidence regarding sector capacity to assess the marginal contribution (positive or negative) of their action. To our knowledge, there exist no metrics describing the system capacity across many agents to respond to disasters. To address this gap, our analytical approach yields new humanitarian logistics metrics based on stochastic optimization models. Our study incorporates empirical data on inventory stored by various organizations in United Nations facilities and in their own to offer practical insights regarding the current humanitarian response capabilities and strategies. By repositioning inventory already deployed, the system could respond to disasters in the same expected time with 7.3% and 19.7% lower cost for the items in our sample.

Key words: Humanitarian logistics, inventory pre-positioning, stockpiling, metrics

1. Introduction

Capability to rapidly deploy life-saving commodities in response to natural disasters is vital. Preparedness work to build humanitarian response capacity spans various actors and actions, which are often not coordinated. A common strategy government and non-government organizations use to improve response capacity is procuring stockpiles of critical commodities and pre-positioning them in various locations prior to disaster events. While increased inventory certainly improves response capacity, the incremental impact of continual stockpile deployments by various organizations is difficult to assess.

The reason is a complex system: dozens of organizations manage hundreds of line items in dozens of warehouses globally in order to respond to events that are random in location, type, and size. Thus, there are no sector metrics for humanitarian stockpile capacity, even though combined efforts across organizations determine the extent to which needs are met following a disaster. Moreover, there is no mechanism to guide numerous organizations' decisions toward system-wide stockpile optimization given inherent coordination challenges in the sector.

To more easily assess this complex landscape and guide decisions toward system improvement our analytical approach yields new humanitarian logistics metrics based on stochastic optimization models presented in the literature. We assess the quality of the humanitarian system with numerical values that relate to different objectives (e.g., cost, time, need met, etc.). Organizations can thus understand how their isolated inventory decisions affect the response capacity for the system as a whole. Such evidence enables decisions that effectively weigh internal objectives (e.g., procurement and warehousing costs, organizational mission, etc.) with contribution to system capacity.

We evaluate our approach with a combination of proprietary data from the United Nations (UN), publicly available data from various sources, and expert opinion on parameter values with no empirical basis. We show that the current allocation of inventory among the warehouses for which we have data can be improved significantly. By repositioning inventory already deployed, the system could respond to disasters in the same expected time with 7.3% and 19.7% lower cost for the items in our sample. Such efficiency gains translate directly into more humanitarian services for the same donation budget.

We show that collaboration is important because the distribution of disasters is not uniform with respect to geography, type, magnitude, and time of year. Organizations acting in isolation might optimally place inventory in the same location rather than deploy stock to serve more regions in a coordinated system. With 50 hypothetical organizations having equal shares of blankets, for instance, the system cost of not collaborating is 6.3% higher than optimal. Our approach does not require explicit collaboration since incremental decisions based on sector wide metrics of marginal change would improve system performance even if made independently.

We conduct sensitivity analysis to consider robustness to empirical data quality and to parameter value assumptions where empirical data are lacking. Metrics and optimal decisions are robust for most parameter value assumptions. They are moderately sensitive to the empirical data used for the risk portfolio. The sensitivity analysis guides further efforts to collect data and calibrate the model for sector use.

The approach proposed in this paper addresses important gaps in the growing literature on performance measurement for humanitarian logistics. It summarizes dynamic empirical data from a complex system with a few intuitive metrics. The metrics are based on accepted modeling approaches yet extend beyond myopic outcomes of independent organizations to measure sector wide capacity. And most importantly, they guide system improvement without the need for explicit coordination.

2. Research design

We consider the practical context of disaster response to motivate our capacity measurement approach and the empirical research we use to assess its potential. The research extends established modeling approaches to fill gaps in the humanitarian metrics literature.

2.1. Context and motivation

Immediately following a disaster that outpaces community coping mechanisms, various outside organizations rush to provide life-saving commodities to meet health, water, food, shelter, or other needs for the affected population. As noted earlier, the response is expedited by inventory prepositioned by these organizations, which could include government (local, regional, national, or foreign), non-government (NGO), military, or private sectors. The stock for this initial deployment could be centralized or deployed across several locations. For large-scale and/or urgent crises, organizations may choose to utilize several stock locations and incur the additional cost of shipping farther to meet needs. In most cases, this initial push is intended to meet human needs within the first few days, followed by replenishment from strategic suppliers based on assessments of need in the affected community. Hence, the initial push is typically transported by air unless ground transportation offers better transit time from a nearby stocking point; sea shipping is used for replenishment. Our model considers both air and land modes to optimize distribution cost and/or time for the initial response to a sudden onset disaster.

Poor response to a widely publicized event pushes organizations to take tangible actions; often this results in increasing the size and/or number of locations for critical commodity stockpiles (typically skewed toward the nature of a recent event and not a broader risk portfolio). On the other hand, constrained fundraising and/or expiration of stockpiled items pushes organizations to reduce stock. As a result, numerous organizations are continually adjusting stockpile deployment.

These dynamic decisions are made independently for the same population of potential beneficiaries. Organizations do not have explicit incentives to collaborate; this can be exacerbated by lack of centralized data and visibility into needs and roles of the actors (Tomasini and Van Wassenhove 2009). The result is a potentially chaotic response (Van Wassenhove 2005). Furthermore, there is no central authority to enforce compliance to coordinated solutions, such as optimization of sector wide stock.

Fragmented decision-making and limited transparency about response capacity make it difficult to assess, much less optimize, the combined level of preparedness for a region. As one manager at a large organization told us, a big unanswered question is "If we had one million dollars from a donor, what would we buy and where would we put it?"

2.2. Research questions and empirical study

To overcome coordination issues and improve system performance, we propose sector wide metrics based on analytical models for disaster response capacity to inform and evaluate the dynamic, independent decisions of numerous organizations. We emphasize that the scope of these metrics and our analysis involves only *immediate response* to sudden-onset disasters. We do not consider the cost of procurement or warehousing. Using these metrics, organizations may weigh their own procurement and storage costs with the impact on the system's ability to respond efficiently. We do not consider slow-onset disasters or 'steady-state' response. Items needed by beneficiaries in these situations would presumably be shipped by slow modes of travel possibly from suppliers directly instead of being shipped quickly from strategic warehouses.

We suggest that metrics could fill two gaps limiting collaborative decisions on stockpile capacity: (1) organizations lack evidence to understand the value of systemic rather than independent decision making, and (2) organizations lack guidance to operationalize system improvement. To close these gaps, our approach seeks to address two research questions:

1. What is the quality of current inventory positions across stockpile depots? (system assessment)
2. What is the value, positive or negative, of incremental change to the combined system? (decision support)

Dynamic answers to the first question provide effective evidence to motivate systemic actions by an individual organization and/or a coordinated group of decision makers. They also form the basis for more general insights regarding the value of stockpile capacity. Actions taken primarily focus on three decisions: (i) which items to buy; (ii) where to put these items; and (iii) stock transfers between depots. Proper application of answers to the second question will guide incremental change resulting from such decisions toward system improvement without the need for explicit coordination. In section 3 we describe the model and derived metrics aimed to answer these research questions.

We use empirical data to assess the potential for these metrics to answer the questions posed. Despite the challenges of coordination, we found two data sources for stockpile inventory spanning multiple organizations in the humanitarian sector. First, inventory quantities and owners for the six United Nations Humanitarian Response Depots (UNHRDs) around the world, which offer space for organizations at no-cost or on a cost-recovery basis, are published dynamically online (United Nations 2014). Second, several years ago the UN Office for the Coordination of Humanitarian Affairs (OCHA) conducted a "Global Mapping of Emergency Stockpiles" to track stock levels in various organizations' warehouses; participation was voluntary and each organization provided its own data (UN Office for the Coordination of Humanitarian Affairs 2014). To our knowledge, our study is the only analytical assessment of those data. Our study considers the combined sector capacity in these datasets across a broad risk portfolio. As our research does not focus on disaster forecasts, we leverage a widely used historical database for our scenarios.

2.3. Relevant literature

Our work is mainly related to two streams of literature. The first stream utilizes mathematical optimization to pre-position disaster stockpiles. Many researchers have used two-stage stochastic optimization models to position inventory in the first stage in order to optimize distribution to beneficiaries in the second stage, with demand often represented by a set of disaster scenarios. de Brito Junior et al. (2013) and Klibi et al. (2013) provide excellent reviews related to disaster response. Duran et al. (2011) is the most similar to ours; the authors worked with CARE international to optimize inventory placement in order to respond quickly to global disasters. We utilize similar two-stage stochastic linear programs as found in these papers as well as the existing stochastic optimization body of literature (Shapiro et al. 2014). Our contribution is to use such models to develop new metrics for the humanitarian sector, in a similar manner as Acimovic and Graves (2015) use a normalized (non-stochastic) linear program to develop an inventory balance metric for an online retailer.

The other relevant stream of literature consists of papers that define metrics for humanitarian logistics. Abidi et al. (2014) provide a thorough review of this literature, most of which emerged after the 2004 Indian Ocean tsunami. They conclude that the literature focuses on theory and models, with limited application to actual humanitarian supply chains, and that further work is needed in applying mathematical and stochastic programming for performance measurement. Haavisto and Goentzel (2015) offer a more recent review and identify the challenges aligning goals within an organization. The difficulty of aligning goals is one reason that the humanitarian logistics literature focuses on performance measurement within a organization rather than across the sector. Our effort applying stochastic programming using actual data to develop sector wide metrics fills these gaps in the humanitarian logistics measurement literature.

3. Analytical approach

Our approach to developing sector metrics is grounded in analytical assessment of response capacity for the humanitarian system. A natural way to do this is to compare outputs of the current system with an optimal system; the metrics then measure normalized distance between the current and optimal state. We use a stochastic linear program to determine the optimal state, which is the technique adopted by much of the pre-positioning community as outlined in the literature review.

Within the subset of stochastic programming methods (see section 2.3), authors utilize different methods and objective functions: maximize throughput, minimize time-to-respond, minimize cost with penalty for unmet demand, minimize casualties, optimize a mix of objectives, and so on (Gralla et al. 2014). Our objective value considers time and cost, assuming that organizations will want to deliver goods as fast as possible while still being fiscally responsible. Implicitly, the model maximizes the amount of commodities delivered, since unmet demand is served by a dummy node with a significant cost penalty. We also find the efficient frontier between time and cost, so that organizations can make pareto optimal decisions according to their particular balance of speed and cost in their objectives (Tomasini and Van Wassenhove 2009).

The context and our objectives motivate slightly different application of stochastic optimization models than previous work, which focused on optimal facility location and capacity for one organization. First, we do not emphasize the facility location decision. We study stockpile humanitarian warehouse facilities that are already established, and any potential sites not yet established would be in prescribed locations. Thus, we focus on allocation of stock across a set of given locations. Second, we do not seek to prescribe strategic, optimal capacity (e.g. warehouse space, inventory) for an organization to implement, but rather to dynamically measure the quality of combined tactical capacity across organizations. We list other modeling assumptions/decisions as well as assumptions related to input parameters below:

1. The warehouse locations are given. We are focusing on tactical decisions, not strategic decisions such as where to place new buildings.

2. There are no warehouse capacity constraints. First, we assume organizations can acquire additional warehouse space in the vicinity if a primary warehouse were at capacity. Taking into account capacities would require a strategic decision including fixed costs for one organization. However, our model does provide the marginal value of inventory changes, which could be compared with the variable cost of temporary leased space or could prompt further study incorporating the fixed costs to deploy new warehouse space. Second, taking into account capacity would require an optimization model that considered all items simultaneously. We believe simplicity outweighs a more accurate but complicated model.

3. There are no transportation capacity constraints. We assume sufficient carriers are available for hire. Transportation may be difficult to procure depending on the market and context at

the time and, thus, is not easily generalized. The fixed time parameter could model origins with generally longer transportation procurement.

4. The model considers a single commodity at time. Results are easily combined with no bundle constraints across commodities. Bundling items together would require the analysis of all items in the depots, and would additionally significantly complicate the analysis reducing the interpretability of the results.

5. Each country has some level of internal capacity. A country can meet the needs for a certain number of affected people, beyond which outside assistance provides support. This number varies by country. The model itself is flexible enough to accept any value for a country's capacity to respond to a domestic disaster. Thus, when calibrated, the model will analyze capacity of the international community to respond to disasters only when they exceed a country's own ability to satisfy the needs of the beneficiaries.

6. All inventory is available for any global disaster. We assume that organizations' stocks earmarked for a specific region or purpose would not be included in the master inventory database. If stock count data are included, then we assume the owning organization is willing to use the items for an international response. Inventory earmarked for a specific region could be incorporated as an adjustment to a country's internal capacity.

7. The SLP is scenario based. Each scenario is an equally likely disaster from the past. We assume a good proxy for a forecast of future disasters consists of selecting a disaster that actually occurred after 1989 with probability $1/|K|$, where K is the set of all disasters. In consultation with analytics experts, the insurance sector uses a similar "equally likely scenario" approach, albeit with more sophisticated scenario generation. The weaknesses of this assumption are that probability distribution of disasters may change over time and that very large anomalous disasters may be weighted higher than they should be.

8. True demand for items is linear in persons affected. There is room to develop more sophisticated models that translate attributes of a given disaster into the amount of an item required by the outside aid community. However, creating such a model is beyond the scope of this paper. A simple linear relationship, we believe, captures most of the dynamics of the connection between persons affected and need. We note that the model does not require the relationship to be linear. Item need could be calculated off-line and fed into the SLP.

9. Relief stock is transported to the country's capital. The practical reason is that the data we use aggregates data at the country level. The capital is a good proxy since it is often the primary port of entry to a country for air transport. Inland capacity for final distribution is often a bottleneck in humanitarian response, but this is very contextual by event and cannot be generalized for our model. Customs is a common decoupling point for international humanitarian response supply chains.

10. Transportation cost depends on the weight of a given item. Further, the total cost for a specific item is linear in the number of units being shipped and in distance.

11. The model focuses on *transportation* cost or time. Like warehouse capacity, procurement cost for stockpile capacity would be considered in follow-on investment analysis, informed

by our response metrics. Warehouse processing cost or time could easily be incorporated in arc costs, though our study considers it negligible compared with transportation.

3.1. Model

The analytical foundation for our metrics is a scenario-based stochastic linear program (SLP) with two stages. The first stage uses the current inventory allocation for capacity measurement; it determines where to place inventory for the optimal capacity benchmark. The second stage is a transportation problem that allocates supply to demand, where the demand occurs at a single disaster node for each scenario in the risk portfolio. The SLP minimizes total expected time or cost to deliver the items from supply nodes to disaster nodes in the portfolio.

We first define parameters and variables for the SLP. A dummy supply node is employed in order to satisfy demand for disasters when the need exceeds combined stockpile inventory.

$I \ni i$ – Set of all depots and warehouses except the dummy node
i^W – The dummy supply node
$I^W \equiv I \cup i^W$: Set of all depots including dummy node
$K \ni k$ – Set of possible disaster scenarios
$J \ni j$ – Set of possible disaster locations
$M \ni m$ – Set of disaster types (e.g., storm, earthquake, etc.)
$R \ni r$ – Set of transportation modes
c_{ijr} – Cost in dollars to transport a single item from i to j via mode r
τ_{ijr} – Time to ship a single item from i to j via mode r (item independent)
$c_W(\tau_W)$ – The cost (time) from the dummy supply node to a disaster
p^k – Probability of scenario k occurring $\left(\sum p^k = 1\right)$
S_{jmt} – Domestic capacity to respond to a disaster in location j of type m at period t.
$\chi \in \mathbb{N}$ – Starting inventory in the system as a whole, not including the dummy node
TAP^k – Total Affected Population in scenario k
β_{jmt} – Units of item demanded per person at location j at time t for disaster type m
j^k – Location of disaster k
$d^k \equiv max\left(TAP^k \beta_{j^k,m,t} - S_{j^k,m,t}, 0\right)$: Actual demand for an item for disaster k
y^k_{ir} – Decision variable: how much to send from i to the disaster scenario k via mode r.
$\mathbf{X}$ – The $|I|$ dimensional vector of starting inventory in each supply node. Its elements are X_i. (May be decision variable or parameter.)

Note that several of the parameters may differ among line items: a bar of soap is much cheaper to ship than an entire kitchen set; a latrine plate can serve dozens of people while a blanket may serve only one or a fraction of a person. The SLP formulation for minimizing time is (note: the SLP that minimizes cost is analogous):

$$V^W(\mathbf{X}) \equiv \quad \min_y \sum_k p^k \sum_{i \in I^W, r} \tau_{i,j^k,r} y^k_{ir} \tag{1a}$$

$$s.t. \quad \sum_{i\in I^W,r} y_{ir}^k = d^k \qquad \forall k \tag{1b}$$

$$\sum_r y_{ir}^k \le X_i \qquad \forall i \in I, k \tag{1c}$$

$$y_{ir}^k \ge 0 \qquad \forall i,k,r \tag{1d}$$

The objective function (1a) minimizes the expected time to respond: the time to respond to any specific scenario is weighted by the probability of that scenario occurring. Constraints (1b) ensure that in each scenario, the demand is met. The model enforces the fact that the supply being shipped from any single depot must not be greater than the supply at that depot with constraints (1c). Constraints (1d) ensure flow must be non-negative. This is a simplified version of the model in Duran et al. (2011). We did test a more complicated model with delivery deadline cutoff times (see the appendix) but do not present those results here, in part because the optimal allocation of inventory is very sensitive to the choice of cutoff time.

A SLP is not required to calculate these metrics for current inventory (though it is easier). Additionally, some of the metrics we propose below do not require an SLP. However, using a linear program allows us to calculate dual variables and maintains similar structure as the model for our optimal benchmark. The formulation that optimizes inventory allocation makes $\mathbf{X}$ a decision variable and adds a constraint to ensure the sum of the supply is equal to χ.

$$V^{OPT,W}(\chi) \equiv \min_{\mathbf{X},y} \left\{ V^W(\mathbf{X}) : \sum_{i\in I} X_i = \chi, \quad X_i \ge 0 \;\; \forall i \in I \right\} \tag{2}$$

We define the objective values with the dummy costs subtracted as $V() \equiv V^W() - \sum_k p^k \sum_r \tau_{i^W,j^k,r} y_{i^W,r}^k$, where the $y_{i^W,r}^k$'s are fixed as the respective solutions to the above SLPs.

3.2. Metrics

By solving this stochastic linear program and manipulating the output, we create metrics aimed to answer the questions we posed in section 2. We first list the metrics and consider how they align with our research questions. We then describe each metric in detail.

$\Delta \equiv \dfrac{V(\mathbf{X})}{V^{OPT}(\chi)}$	Balance metric
$\mu \equiv \sum_k p^k d^k$	Weighted average of demand
$\mu' \equiv \sum_k p^k \;\; min\left(d^k, \chi\right)$	Average demand *met*
$\gamma \equiv \dfrac{\mu'}{\mu}$	Fraction of demand served
$\delta \equiv \sum_{k:d^k\le\chi} p^k$	Weighted fraction of disasters completely served (where the weights are the probabilities.)

$\phi \equiv \dfrac{V(\mathbf{X})}{\mu'}$	Average time (cost) per unit delivered
π_i^k	Dual variable for constraint (1c) for depot i and disaster k for *actual* allocation SLP
$\pi_i' \equiv \sum_k \pi_i^k + (1-\delta)\tau_W$	Adjusted dual variable for depot i over all scenarios

This set of metrics offers multiple dimensions for system assessment to address the research design questions: to assess the quality of the current inventory positions and to quantify the value of incremental change to the system while providing decision support. Current capacity is measured directly by fraction served metrics, γ and δ, which assess how sufficient system stock is to meet needs, and ϕ, which considers how quickly it can be deployed. The balance metric Δ measures the quality of current system stock allocation while also quantifying the impact of reallocation. Dual variables quantify the value of incremental change to the combined system stock. More specifically, the metrics' relationship to key decisions is such: (i) which items to buy - γ, δ; (ii) where to put these items - π_i'; (iii) should we transfer inventory - Δ.

3.2.1. The balance metric Δ The balance metric Δ assesses the allocation of current capacity. More specifically, it estimates *how far* out of balance the actual allocation of inventory is relative to the optimal. This metric is similar to the deterministic version reported in Acimovic and Graves (2015), which used data from an online retailer to show that an item's balance metric was highly correlated with the item's future actual incurred shipping costs.

We note a few properties of this balance metric:

1. It approximates the proportional increase in cost (time) to serve beneficiaries given that one's inventory is allocated as it *actually* is as opposed to being allocated *optimally*. Thus, it measures the value of change through reallocation.
2. The optimal value is 1. Anything greater than 1 is considered out-of-balance.
3. It is relatively robust to outliers in the risk portfolio. Exceptionally large disasters far exceed the on-hand supply for any item and could skew metrics whose calculations include all affected. The objective functions and solutions of the SLPs in section 3.1 depend only on those people who can be served by system inventory.
4. The metric is affected by which depots are considered. If a new depot is opened in a disaster hotspot and no inventory is placed there, then the balance metric will increase (because $V^{OPT}(\chi, n)$ will decrease). In this sense, operational managers can be alerted to the fact that the inventory is out-of-balance given the new depot.

3.2.2. Fraction served (γ) and weighted fraction of disasters covered (δ) γ represents the average weighted fraction of demand met, where the weights are derived from the risk portfolio probabilities. It does not depend on how the items are allocated among the depots . We note that it can be influenced by very large disasters. For example, including a disaster scenario affecting tens or hundreds of millions of people in the risk portfolio will have a significant impact on μ, the denominator in the fraction that defines γ. Thus, these values may appear low, and must be interpreted with this in mind.

δ, on the other hand, is relatively robust to outliers. If system stock contains only 1,000 items, then an unserved disaster affecting 1,001 or 10,000,000 people have the same impact on the calculation of δ. This robustness comes at the cost of not conveying a sense of the magnitude of the disasters that go unserved. δ provides *different* information from γ, and the two together can help operational managers understand the adequacy of the total inventory level.

3.2.3. Time and cost per unit delivered ϕ represents the average time (cost) to deliver one unit to a beneficiary from a depot. Even though this number should not be used to predict supply arrival times for operational planning in a real situation, where the context may not match model parameters, it can provide objective approximations of the *relative* time to deliver items to assess the quality of inventory strategies.

3.2.4. Dual variables The marginal increase in total expected time (cost) to serve more beneficiaries by adding a unit to depot i can be estimated by the *adjusted* dual variables (π'_i) for the SLP that utilizes actual inventory allocations (corresponding to constraints (1c)). We *adjust* the sum of the original dual variables $\sum_k \pi_i^k$; otherwise they would be dependent on the choice of the dummy value τ_W (c_W), which is rather arbitrary.

The value of π'_i may be positive or negative. In general, an organization adding a unit of inventory to the system might expect total costs to increase: if a disaster strikes, it might ship more items at a higher total cost, with the benefit of serving more beneficiaries. However, if the current stock allocation is particularly imbalanced, then adding inventory to the right place could actually decrease the expected cost, *while also* serving more beneficiaries.

4. Data

We use empirical data from several sources to assess how the proposed metrics address the issues raised and questions posed. While we discuss assumptions and overall details in subsections 4.1 to 4.4 more specific details regarding how we collected the data can be found in the appendix.

4.1. Disaster data

As mentioned earlier, we leverage historical data for our risk portfolio. Specifically, we assume that each disaster recorded between January 1990 and the summer of 2013 has an equal chance of occurring again, utilizing data from EM-DAT (Centre for Research on the Epidemiology of Disasters 2014). EM-DAT tracks the following: the month and year of the disaster, the country, the type of disaster, and the total affected population (as well as other fields we do not utilize). We include only sudden onset disasters: earthquakes, epidemics, floods, mass movement dry and wet, storm, volcano, and wildfire. We assume stockpiles are used for unexpected events rather than those that organizations could foresee (e.g., drought) and use procurement capacity, either directly or to immediately restore stockpiles. Of the records in this database for these disasters, 22% of the values for total affected population are null: we exclude these from our analysis. Despite the limitations of any database with the objective of recording details related to every disaster that occurs (Guha-Sapir and Below 2006), this database is recognized as the best of its kind and has been utilized by many other researchers working in similar domains (Peduzzi et al. 2009, Duran et al. 2011, 2013). We examine the potential bias of the null values in section 5.3.1.

4.2. Depot and inventory data

Governmental and non-governmental organizations that choose to stockpile disaster relief supplies may utilize their own depot or those offered by the government or other organizations. As mentioned earlier, we utilize two sources of data capturing depot information and inventory. The UNHRD website hosts a real-time stock report that details which organizations are housing which items where within the UNHRD network (United Nations 2014). The OCHA "Global Mapping of Emergency Stockpiles" database holds data voluntarily submitted by organizations regarding items in their stockpile: organization name, name of the city for the stockpile location, item name, and quantity (among some other fields) (UN Office for the Coordination of Humanitarian Affairs 2014). This database is proprietary and not open to the public. We combine the data from the UNHRD and OCHA databases to determine an overall estimate of organizations' inventory levels around the world. We merge duplicate records when one organization appears in both databases for the same item in the same city. For ease of interpretability of the results, we also merge the inventory of depots in the same country; otherwise, the model may dramatically shift inventory between Kuala Lumpur and Subang – which are about a 30 minute drive from each other – witih no real impact on overall cost. In this case, we put all of Malaysia's inventory in Subang when we run the model. This is one of seven similar instances. See the appendix for more details. In the end, we consider 25 depots in our analysis.

4.3. Time and cost data

We assume two modes of transportation are available: air and truck. We do not consider sea transportation as the study focuses on immediate deployment of the stockpile. As evident from the SLP formulations, we assume the cost of transporting goods is linear in the number of units being transported. Thus, for each warehouse-disaster-mode triplet, we calculate the time and the cost-per-metric-ton-km for the route. This, paired with information on the weights of each item, allows us to calculate the cost of shipping a single unit on a specific route.

For air, we assume the time and cost are based solely on distance along an arc of a great circle around the earth. For trucks we utilize Google's "Distance Matrix" API (application programming interface) to calculate the road distance and driving time between a warehouse and a disaster (Google 2014). If there is no way to drive on land between two points, the API may return results including a ferry trip over water. We believe this is reasonable because if a ferry route exists, organizations may use this route (or a similar one) to transfer goods via boat instead of using 100% air.

If no land driving route exists, then trucks cannot be used on a particular lane. Additionally, we assume trucks would not be used if the driving time were more than 100 hours. Of the 7175 warehouse-disaster arcs, about 31% are drivable according to the Google API. Of this 31% subset of drivable routes, 71% take 100 or fewer hours to traverse.

The time to move an item by air (truck) has a fixed component and a variable component dependent on the distance (driving time). We approximate the time and cost parameters as such: fixed cost per metric ton, air = 25 USD, truck = 10 USD; cost per kilometer per metric ton, air = 0.50 USD, truck = 0.1 USD; fixed time to procure, air = 6 hours, truck = 0 hours; speed, air = 600 km/hr, truck speed is irrelevant because driving times are acquired from the Google API. These values are based on data spanning numerous commercial and humanitarian projects conducted at the MIT Center for Transportation & Logistics and on conversations with humanitarian logisticians. As such, the parameters may vary among different organizations working in different contexts in different locations. Additionally, in section 5.3 we report on our sensitivity analysis in which we describe how the results change along with these input parameters.

4.4. Item specific data

We concentrate on seven items: blankets, buckets, jerry cans, kitchen sets, latrine plates, mosquito nets, and soap. These non-food items (NFI) are among the most common items sent to disasters by organizations and consequently have higher inventory levels at depots. Additionally, these items are different from each other in a way that necessitates modeling them differently.

Item weights were retrieved from International Federation of the Red Cross and Red Crescent Societies (2009). These weights are utilized to calculate transport cost, which is measured in cost per kilometer per metric ton. In section 3.1 we defined a parameter β that dictates for an item how many units are required per person affected. For the non-country and non-weather related items we assume the following: a bucket serves a family of five; two jerry cans serve a family of five; a kitchen set serves a family of five; a latrine plate serves 50 people; two mosquito nets serve a family of five, in countries where malaria is present; a bar of soap serves one person. We consulted the Sphere handbook (The Sphere Project 2014), historical appeals for funding that specify commodity requirements, and humanitarian experts (Bauman 2014) in order to estimate the items needed per person. To determine whether a country was at risk for malaria, we consulted the CDC website (Centers for Disease Control and Prevention 2014).

Blankets are one of the more complicated items because demand depends on the type of blanket and the climate by region and by month. To calculate how many blankets are required in a disaster, we interpolate the number of blankets per person required based on estimating the required number at different temperatures.

We calculate the blanket needs per person by first calculating the thermal insulation needs per person according to temperature. Thermal insulation requirements by temperature are provided by the United Nations High Commisioner for Refugees (2012) (UNHCR). We assume beneficiaries' thermal insulation needs would be satisfied by a combination of their own very basic clothing (shirt and pants) and medium weight blankets – which have a consistent thermal insulation value and are the dominant blanket type – provided by the responding organizations. Using the above assumptions and UNHCR's thermal requirements, we regress number of medium blankets required in addition to basic clothing against temperature Fahrenheit to obtain our estimate of the blanket equation: $NumBlanketsPerPerson =$

$(3.34 - 0.044(NightlyLowTempInF))^+$, where $(a)^+ \equiv max(a, 0)$. To obtain $PeopleServedPerBlanket$ (the β parameter in our model), we take the inverse of $NumBlanketsPerPerson$. Note that $PeopleServedPerBlanket$ is *not* linear in temperature. If $NumBlanketsPerPerson = 0$, we assume that no blankets are needed, and set the demand to zero. We captured the average low temperature by month for each country in our database by querying the website World Weather Online (2014). Thus, if the nightly low were $32°F$, a beneficiary would require $(3.34 - 0.044(32))^+ = 1.93$ medium weight blankets, and each blanket would serve 0.52 people. See the appendix for more details on item-specific data and how blanket-needs and blanket-inventory are calculated.

5. Results

Having described the model in section 3.1, we utilize the data described in section 4 to calculate metrics outlined in section 3.2. We divide the results by the research question addressed as outlined in section 2: what is the quality of current inventory positions across depots (system assessment); what is the value of incremental change to the combined system (decision support). We conclude this section with a sensitivity analysis. Unless otherwise noted, demand refers to number of units required, not the number of people requiring the units.

5.1. System assessment results

Table 1 describes the system's ability to meet demand given the current on-hand inventory for all items in our study. Since all stock is deployed, these results are agnostic to the allocation of demand across the locations. The fraction of demand served (γ) tends to be small, even when almost one million blankets are kept in inventory, for example, and even when these blankets can cover all the demand in 96% of the disaster scenarios in our portfolio. This is driven by several disasters with reported TAPs in the hundreds of millions. As mentioned in section 3.2.2, the γ metric can be susceptible to outliers, but when considered along with δ describes a more complete picture of the system inventory adequacy.

Table 1 Summary metrics for items in actual depots (Ability to meet demand)

Item	Units	Demand (μ)	Demand met (μ')	Fraction of demand served (γ)	Fraction of disasters served (δ)
Blanket	852,563	561,746	75,150	0.13	0.96
Bucket	106,844	185,926	21,206	0.11	0.90
Jerry Can	437,530	371,852	61,300	0.16	0.93
Kitchen Set	126,143	185,926	23,105	0.12	0.91
Latrine Plate	4,650	18,593	1,321	0.07	0.83
Mosquito Net	395,588	428,591	64,386	0.15	0.91
Soap Bar	111,595	929,631	41,289	0.04	0.76

In order to assess the quality of system inventory allocation we present the resulting values of the Δ (balance) and ϕ (average time/cost to ship) metrics. Table 2 (table 3) shows the results when time (cost) is minimized in the SLP. This table also lists the corresponding cost (time), average distance traveled, and fraction of units by air.

Table 2 Time optimization for actual inventory allocation: summary metrics (the value being optimized is in bold)

Item	Balance Metric (Δ) (time)	Average time to ship an item (hrs) (ϕ)	Average cost to ship an item (USD) (ϕ)	Average distance traveled (km)	Fraction of units moved by air
Blanket	1.07	**15.6**	5.13	5,810	98%
Bucket	1.15	**16.9**	2.66	6,530	99%
Jerry Can	1.17	**16.4**	0.94	6,250	99%
Kitchen Set	1.13	**16.3**	15.60	6,190	99%
Latrine Plate	1.13	**17.5**	9.35	6,880	100%
Mosquito Net	1.14	**16.0**	1.34	5,990	100%
Soap Bar	1.16	**18.6**	0.39	7,580	99%

Table 3 Cost optimization for actual inventory allocation: summary metrics (the value being optimized is in bold)

Item	Balance Metric (Δ) (cost)	Average time to ship an item (hrs) (ϕ)	Average cost to ship an item (USD) (ϕ)	Average distance traveled (km)	Fraction of units moved by air
Blanket	1.15	22.6	**4.89**	5,870	81%
Bucket	1.35	23.5	**2.56**	6,540	84%
Jerry Can	1.37	25.6	**0.89**	6,280	78%
Kitchen Set	1.27	26.1	**14.63**	6,210	77%
Latrine Plate	1.25	25.9	**8.88**	6,880	83%
Mosquito Net	1.36	23.6	**1.28**	6,020	82%
Soap Bar	1.29	27.5	**0.37**	7,580	79%

From these tables, one can assess the quality of a given item deployment. For instance, blankets seem to be the best allocated item for both time and cost, whereas jerry cans seem to be the worst allocated. Items are more out of balance with respect to cost than time. The balance metric also quantifies the potential value of changing allocation. For these items, reallocation could improve response time between 7% and 17% or reduce cost between 15% and 37%. Quality is also defined by absolute time (cost) to respond. Note that soap bars are better balance than jerry cans with respect to time, but that it takes longer on average to deliver soap bars. This is in part due to the small number of soap bars in stock. This highlights the fact that an additional unit of inventory not only allows the system to serve more beneficiaries, but also may reduce the overall time or cost to respond.

Additionally, we can use the model to assess the right balance of cost and time by plotting the efficient frontier of the competing objectives. To do this, we minimize cost in the SLP while adding a constraint restricting the total time be less than one of ten equally spaced values between the shortest time possible (when time is minimized) and an upper bound on the best time (when cost is minimized). Figure 1 plots the actual inventory allocation for jerry cans against the efficient frontier, showing it is not pareto optimal. By reallocating items in the network, the system can maintain the same average time-to-respond while dramatically reducing costs. For instance, when cost is minimized, the resulting time-to-respond is double (100% more) the optimal. For a 5% increase in cost, the resulting increase in time-to-respond would be only 50% more than the optimal. For an 10% increase in cost, the resulting time-to-respond would be only about 15% more than the optimal. For a small increase in monetary budget, a much quicker response time can be realized.

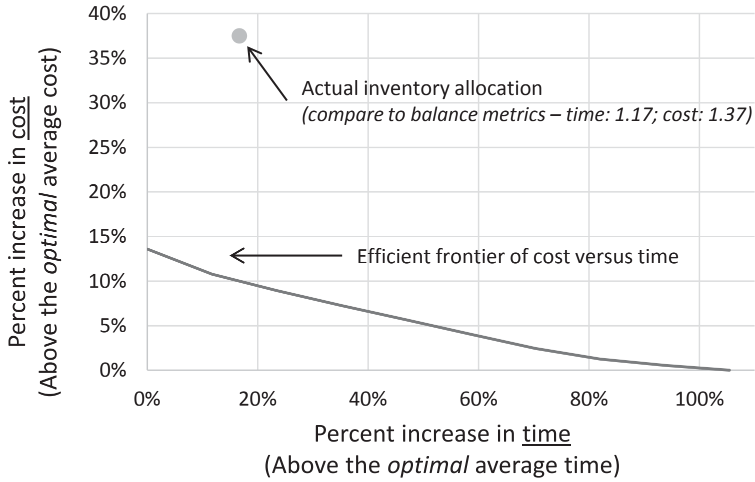

Figure 1 **Jerry cans: Efficient frontier of cost versus time.**

To gain further insight from the efficient frontier, we can calculate how much cost savings are achievable *while maintaining the current expected time-to-respond* through inventory reallocation. In graphical terms, this is equivalent to bringing the single point in figure 1 straight down onto the efficient frontier. In the jerry can example, the *increase* in cost-to-respond would drop from its current level of 37% to 10%, for an *overall* cost savings of $((0.37 - 0.10)/1.37 =)19.7\%$, while maintaining the same current expected time-to-respond of 17% worse than optimal. For all seven items, the achievable cost savings subject to no change in current time-to-respond are as follows: blankets -7.3%; buckets - 19.7%; jerry cans - 19.7%; kitchen sets - 13.9%; latrine plates - 14.7%; mosquito nets - 18.6%; soap bars - 18.7%. While these cost savings are less than those suggested by the balance metric column in table 3, they are - unlike the model solutions corresponding to table 3's values - attainable *without any degradation of time-to-respond* on average.

We examine how the optimal allocation of blankets varies by objective (time or cost). Figure 2 shows how the actual allocation of 852,563 blankets compares to the optimal allocation of blankets when both time and cost are optimized.

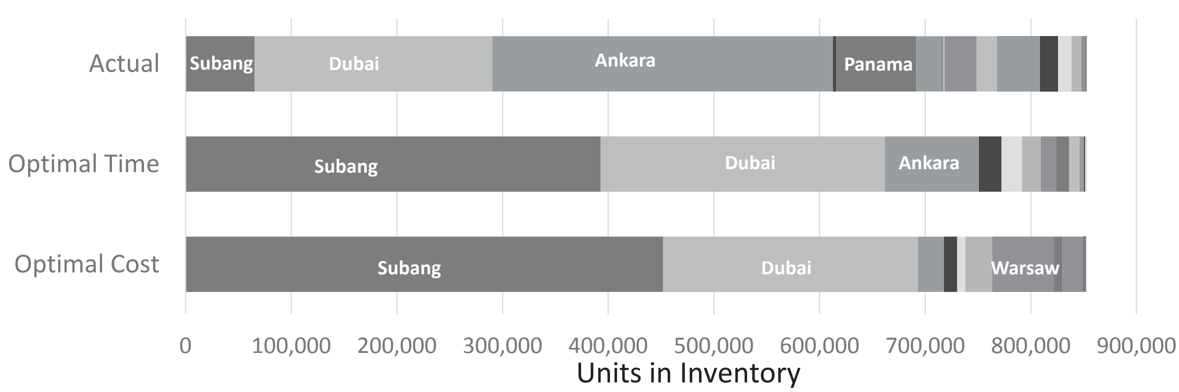

Figure 2 **Actual and optimal allocation of blankets across depots**

This figure suggests that it is better to have items in Ankara when minimizing time and in Warsaw when minimizing cost. Ankara is nearer to potential disaster locations via air transportation (based on our assumptions), while Warsaw is better connected to potential disaster locations via truck.

We present additional results regarding the impact if incremental inventory and delivery deadline cutoff times in the appendix.

5.2. Decision support results

While charts reporting on the actual and optimal inventory allocation for various scenarios can provide insights, practitioners may need more specific evidence to support a decision. This is particularly the case when they may not be able to reallocate stock to the optimal site since their depots are in other locations. The dual variables provide an estimate of the value of an additional unit of inventory in each location. From this, one can estimate the value of transshipping between locations, or the cost of replenishing to a different depot.

Figure 3 shows the adjusted dual variables (π_i') returned by the optimization software. We do not consider or address degeneracy, multiple optimal dual variables, or the validity of the duals beyond infinitesimal perturbations. The adjusted duals are *estimates* of the increase in the time or cost objective function if an additional unit is placed at the specific depot. They are often positive because the total time or cost to serve increases when a unit is added to the system; this is the cost for the benefit of serving more people. Subang and Jakarta stand out because adding an item and serving more people actually *reduces* the cost and time. This is a product of significantly suboptimal actual inventory allocations.

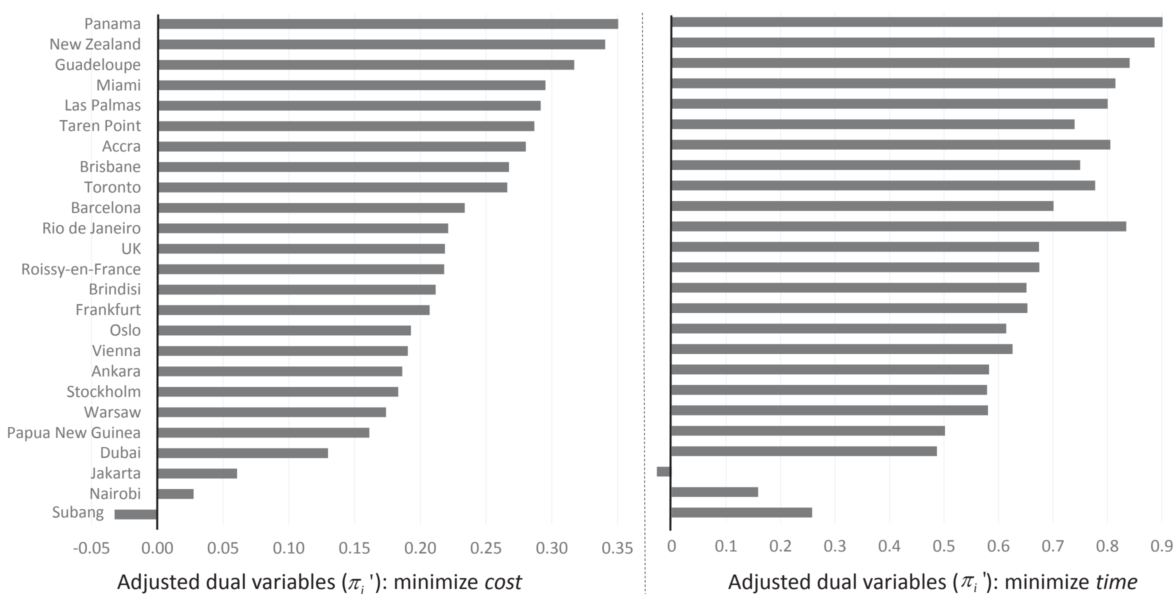

Figure 3 **Blankets: Dual variables for depots for actual allocation while minimizing time and cost**

5.3. Sensitivity analysis and discussion

Our model is based on data and assumptions that may be incomplete or inaccurate. We analyze the sensitivity of the model to data integrity, our choices of parameter values, and disaster risk profiles we utilized.

5.3.1. Data integrity In the appendix we analyze the distribution of null values in the disaster total affected population data. We find that it is not entirely uniform: some countries and disaster types have propensity for null values that are much higher or lower than the average (Pakistan has 47% null, while the Philippines has 7% null). We anticipate that this skewness of missing values may impact the optimal allocation of inventory, although we have not run any tests. For instance, if many values in Pakistan are null and if we ignore these corresponding disasters, it is possible the model may place too little inventory near Pakistan and too much inventory in Asia. Filling in these missing values using total killed and other attributes can be an avenue of future research.

5.3.2. Parameter values For many of our input parameters, the true values either lie in a range or are unknown. To explore how robust our study results are to these input data, we perform sensitivity analysis on several parameters, specifically reporting how changing input parameters within a range affects the balance metric and the optimal allocation of inventory. The detailed results of this sensitivity analysis can be found in the appendix. We summarize our findings here in table 4. For each parameter, we multiply the base case value by 2 (or by 10 for countries' abilities to respond) and observe the impact on the balance metric. If the balance metric went from 1.2 in the base case to 1.1 when we doubled the parameter value in question, we record this as a (negative) proportional increase of $((1.1/1.2) - 1 =) - 8.3\%$. The parameter values tested are in decreasing order of the absolute value of the impact on the proportional increase in the average value of the balance metric over the seven items. While the model's outputs change with the choice of these input parameters, the output varies in a way that can be easily explained. We believe the fact that the outputs change in explicable ways lends validity and credibility to our underlying assumptions. However, improvements can still be made and a next step is to calibrate the model with more accurate input parameter values using data from organizations themselves.

5.3.3. Disaster risk profile The results of our model are based on the disaster risk profile we built utilizing data from the EM-DAT database. In order to understand the model's sensitivity to the specific disaster risk profile we employed, we perform two experiments: one is on a subset of the data over 10 year rolling horizons and the other is on a dataset whose outliers are removed.

We first calculate each item's balance metric over a ten-year rolling horizon, beginning in 1980-1990 and ending in 2003-2013. Over this time, different disasters are included or excluded. We include details and figures in the appendix. The balance metric for different items changes depending on the subset of disasters being considered. We see that the absolute value of the balance metric does change over time. For instance, the balance metric for blankets (optimizing time) is about 1.04 in 1980, rises to 1.09 in 1986, and then stays at about 1.06 from 1991 onwards. In order to isolate how much of the variation is due to disaster scenarios affecting all items, and how much is due to disaster profiles that affect items differently, we normalize each item by the balance metric for buckets, which we set to 1.0. (We chose a normalizing item that was not affected by weather, time-of-year, or malarial prevalence within a country.) For the most part, blankets are the most balanced items – consistently at about 0.92 the normalized value of the buckets balance metric – while soap bars and jerry cans are some of the most imbalanced – mostly at

Table 4 Summary of sensitivity analysis to parameter values

Parameter *(objective type, multiplicative increase in parameter value)*	**Corresponding average proportional change in balance metric**	**Behavior of balance metric**	**Behavior of optimal allocation**
Fixed time to acquire airplane *(objective: time, increase: ×2)*	-3.3%	Decreases: as fixed time dominates total time, optimal and actual objective values are equally bad. Ratio converges to 1.	As air fixed time increases, allocation to Dubai and Subang (Ankara and Warsaw) decreases (increases). Ankara and Warsaw are more truck-friendly locations in our data.
Airspeed *(objective: time, increase: ×2)*	-3.3%	Decreases: see "fixed time."	As airspeed increases, allocation to Dubai and Subang (Ankara and Jakarta) decreases (increases).
Cost of air *(objective: cost, increase: ×2)*	1.3%	Increases: as air cost increases, penalty of not stocking enough in truck-friendly locations increases.	As long as air cost is more than truck, then allocation to Dubai (Warsaw) increases (decreases) with cost of air. As air cost dominates truck cost, allocation is decided by air cost alone, even though trucking is used slightly more often.
Capacity (number of beneficiaries) of countries' to respond to domestic disasters *(objective: time, increase: ×10)*	1.1%	Increases: as capacities increase, much of residual demand would be due to large disasters in Asia, which is not where inventory is now.	As capacities increase, more should be placed in Subang and less in Dubai and Ankara.

about 1.02 the normalized value of buckets. Rank order is often preserved, which suggests that the model is moderately robust: the value of the balance metric may change depending on the risk scenario, but if an item is imbalanced in one risk scenario, it will likely remain imbalanced for other risk scenarios. The most salient exception is that of mosquito nets. Mosquito nets' balance metric consistently declines from 1982 onwards. This reflects the trend that – based on the recorded disasters in the EM-DAT database – the optimal allocation of mosquito nets to African warehouses grew over the past few decades. Additionally, soap bars' and jerry cans' balance metrics – while consistent from 1980 until about 1997 – dropped in value after that and fell below the buckets' balance metric. This suggests the model is fairly useful in highlighting which items need special attention and further investigation. However, it can be somewhat sensitive to the disaster profile in certain situations. This points to the need to develop effective risk profiles for these types of models.

We also investigate the model's sensitivity to outliers: disasters scenarios with very large numbers of people affected. In our set of disaster scenarios, the 1% of the largest disasters (corresponding to 35 disasters) account for 66% of the total persons affected. These are truly outliers: the top five disasters occurred in China affecting on average (according to our data) 178,000,000 people per disaster. To understand the impact of these disasters, we rerun some of our experiments with the top 1% of disasters removed from the disaster risk profile. We do not propose that organizations should remove data when running the model and making decisions; rather, we remove these disasters as an academic exercise to understand model sensitivity to extremely large values. Of the 35 disasters removed, 24 occur in China, 10 in India, and 1 in Pakistan.

Table 5 shows the balance metric (optimizing cost), fraction of demand served, and fraction of disasters fully served for the base case (compare to tables 1 and 3) and for the case where we remove the top 1% of disasters and rerun the optimizations. The metrics do change somewhat as we might predict. As alluded to, fraction of disasters served is very robust to outliers and barely changes. Fraction of demand served is very susceptible to outliers, and changes significantly, increasing about three times on average. This makes sense as we are removing about two thirds of the demand (66%). The balance metric changes somewhat, depending on the item. The size of the shift in the balance metric will correspond to the inventory level. Jerry cans have one of the largest absolute shifts, and is also the item with the most units in inventory in terms of fraction of demand served. Soap bars can serve the least amount of demand and have one of the smallest shifts. Rank order is preserved: blankets are still the best balanced and jerry cans are still the worst. Considering that only 34% of the sum of demand remains in the experiment where we remove 1% of the biggest disasters, we believe that the model is relatively robust where we expect it to be so. The balance metric changes its values, but still highlights the appropriate 'best' and 'worst' balanced items.

Table 5 Metrics with outliers (top 1% of disasters) removed

Item	Balance Metric (Cost)		Fraction of demand served		Fraction of disasters completely served	
	Base Case	1% removed	Base Case	1% removed	Base Case	1% removed
Blanket	1.15	1.12	0.13	0.46	0.96	0.97
Bucket	1.35	1.33	0.11	0.33	0.90	0.91
Jerry Can	1.37	1.34	0.16	0.47	0.93	0.94
Kitchen Set	1.27	1.25	0.12	0.36	0.91	0.92
Latrine Plate	1.25	1.25	0.07	0.21	0.83	0.84
Mosquito Net	1.36	1.34	0.15	0.44	0.91	0.93
Soap Bar	1.29	1.28	0.04	0.13	0.76	0.77

In figure 4, we show the resulting optimal inventory allocations. As expected, less is stored in Subang which is near to the very large disasters in China we removed. However, Subang is still a significant warehouse: its allocation of inventory does not disappear when 66% of the demand data (most of this in Asia) is removed.

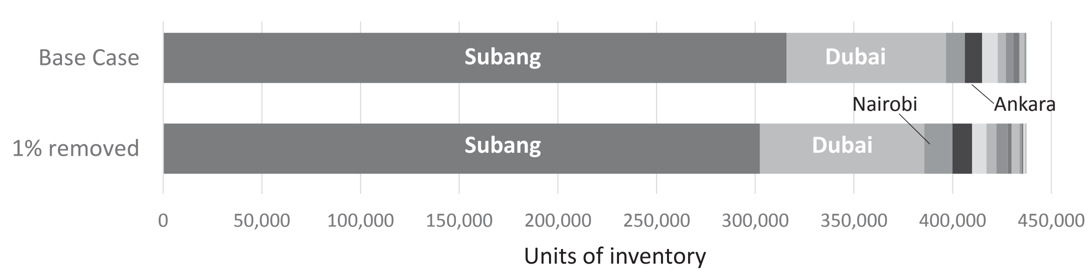

Figure 4 Optimal inventory for jerry cans (optimize cost) when 1% of values are removed

There is no doubt the model can be improved with better forecasts that better handle such outliers. However, we believe that we have shown that the model still predicts relatively consistent qualitative results in the presence of these super disasters. Thus, one might expect managers making decisions based on the model will be mostly correct.

6. Discussion

Working back from our questions, we now consider the context that motivated our approach. We noted two challenges that can lead to suboptimal system capacity. First, numerous humanitarian organizations are continually making adjustments to their stockpile inventory. Second, these organizations typically make these adjustments independently of others' actions. Metrics from the empirical study demonstrate that over time the system can easily become suboptimal for meeting disaster needs. For instance, figure 3 suggests that blankets be moved from Panama to Subang to save money and time. However, the inventory in Panama is owned not only by several organizations but also by organizations different from those that own inventory in Subang. To overcome these challenges, we proposed that metrics could fill two gaps that limit collaboration on stockpile capacity: (1) lack of evidence to understand the value of systemic rather than independent decision making, and (2) lack of guidance to operationalize system improvement.

Our metrics for system assessment help to close the first gap by quantifying the quality of the current system and the potential for improvement. The managerial use of metrics for decision support helps to close the second gap. We show that models and metrics outlined in this paper can help organizations explicitly estimate the time or dollar value of specific tactical decisions. Our approach provides information in a timely manner – in a real time dashboard we describe below, for example – as evidence to guide decisions that improve system performance.

In addition, the metrics can frame productive discussions among organizations and donors regarding further investments in capacity. Organizations have new evidence to demonstrate the systemic impact of their actions, and donors have evidence to determine where investment is most needed.

In this section, we discuss how the results can be used for system assessment, how the results can be used for decision support, and broader insights that can be derived from our model.

6.1. System assessment discussion

The fundamental system assessment questions focus on the quality of the current capacity and the value of incremental change. Regarding overall capacity, the γ and δ offer complementary metrics for quality assessment. While system stock serves a small fraction of the potential affected population, it is sufficient to meet needs for most disasters in the risk portfolio. The model offers the potential to refine the metric for fraction of disasters served with better data for domestic internal capacity.

Combined with ϕ, which considers how quickly system stock can be deployed, we can simultaneously quantify the impact on needs met and timeliness from changes in overall stock level. Organizations, and donors, can use this marginal analysis to prioritize investments.

The allocation of given system stock also has significant impact on system performance. The balance metric, Δ, succinctly addresses both research questions by assessing how well each item is allocated and by quantifying the potential for performance improvement through reallocation.

One useful tool we provide is to normalize the inventory data using our metrics. In table 1 we see that there are significantly more blankets than any other item. However, if we normalize the inventory by the number of people each unit is able to serve, we see that jerry cans actually have the highest capacity to serve beneficiaries in the "Fraction of demand served" column. We provide value my measuring inventory levels in terms of items' abilities to met demand as opposed to absolute numbers of units in warehouses.

We want to reemphasize that the numerical values derived from the model should not be taken literally. The model is already useful in allowing users to better understand the system, what is going right, and what is going wrong. However, it still needs to be calibrated using input from the humanitarian logistics community. Even when calibrated and based on better data, however, care must still be taken when interpreting the results. The model is most useful for making *relative comparisons*. For instance, if the model shows that average time to ship blankets is 15.6 hours and for buckets the time is 16.9 hours, in reality, buckets may not show up exactly 1 hour and 18 minutes later than blankets on average. Rather, for some reason that warrants further investigation, buckets are allocated less effectively than blankets. If buckets are a critical item, then perhaps more buckets should be added or they should be transferred.

This leads to a related question: does improvement of a few hours in response time really make a difference for humanitarian outcomes? First, recall that the savings are *on average*. If an item can be reallocated to be shipped on average in 14 hours instead of 16 hours, then this is a 12.5% reduction in system time-to-respond, and a 20% reduction in variable time, considering the 6 hour fixed time assumption. Second, a few hours of time could make a big difference within the 72 hour window commonly used to benchmark initial response efforts. For example, arriving a few hours earlier could result in a parking space at the airport that may not be available later. Finally, reducing flight times by a few hours could have important implications on transport capacity: being better positioned means that more round trips with a single airplane are possible within the 72 hour window. The UN and/or military can quickly mobilize an "air bridge" to take advantage of well positioned stock in this manner.

To complement the empirical study using historical data, we present a small case study using real time data to explore how a manager or coordinated group might actively monitor sector capacity. By compressing a lot of information into a simple dashboard, these metrics can alert managers or coordinated groups to changes that may have otherwise gone unnoticed. For this case, we monitor the publicly available UNHRD stockpile data daily and calculate relevant metrics. We also post our prototype dashboard publicly: `https://operations.shinyapps.io/HumanitarianLogisticsSummaryMetrics` (summary metrics); `https://operations.shinyapps.io/HumanitarianLogisticsDepotMetrics` (depot-specific metrics). We do not post this analysis, which is limited to sector inventory stored in UNHRDs, as definitive evidence for action. Rather, we hope it can facilitate valuable conversation among academics and practitioners as to the best way to implement, interpret, utilize, and validate such metrics.

A detailed screenshot of all the plots and metrics in the prototype dashboard is in the appendix. Here, we show in figure 5 detail of the balance metric and we describe the actions behind some metrics' changes. We see that in figure 5, "Organization A" added blankets and jerry cans to Accra, which – according to our analysis – might not be the optimal location due to the corresponding increase in the balance metric. On the other hand, the actions of "Organization B" adding kitchen sets *to* Dubai and "Organization A" deploying kitchen sets *from* Accra both helped to bring that line item into better balance.

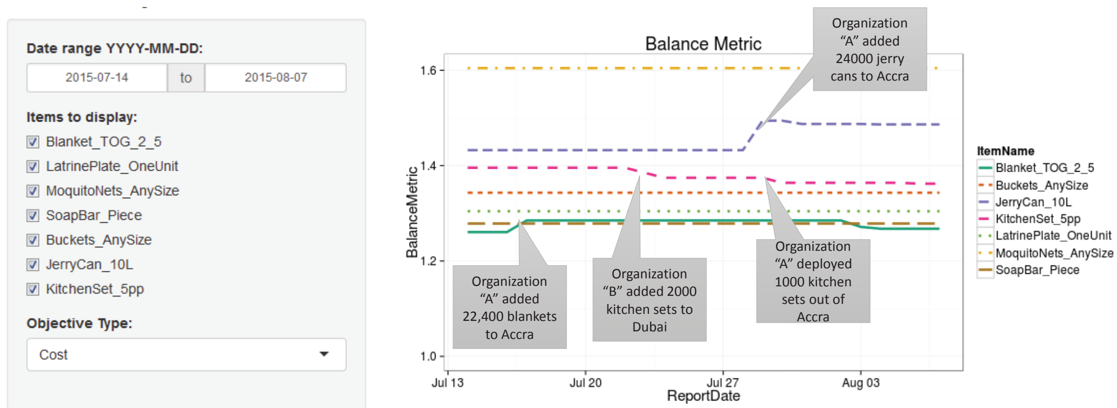

Figure 5 **Detail of balance metric from prototype dashboard noted with actions corresponding to metric changes.**

6.2. Decision support discussion

To consider how the metrics enable decision support, we walk through the decisions mentioned in the research design and consider how a manager or coordinated group could use the results from section 5 as evidence. While we focus on managerial decision-making, the discussion is equally relevant for donors who want their contributions to have the most impact on the system.

One strength of our approach is that a manager or coordinated group does not need to reoptimize the entire system frequently in response to every system change. The data will always have enough integrity issues that perfectly optimizing the system is a futile task anyway. In fact, we provide tools that allow managers to avoid running an optimization model at all, as long as some central entity calculates and publicizes the metrics regularly. Managers can monitor items' metrics, be alerted when the situation needs attention, and then make good (not necessarily perfect) incremental changes using the dual variables as a guide. The dual variables allow managers to understand the impact of taking an action that is not optimal, but that may be necessary due to context and which might capture much of the optimal solution's value.

6.2.1. Which items to buy Imagine a manager or coordinated group deciding whether or not to procure an item, and if so, which one. She may sort the items by δ and by γ using the data in table 1. From a system-wide perspective, soap is the item that serves the smallest fraction of demand and the smallest fraction of disasters covered. An additional soap bar will alleviate a unit of unserved demand in fraction $(1-\delta)=0.24$ of the time a disaster strikes.

The manager could further place a value on each item inversely proportional to its importance in a typical disaster, say v. Then, the manager may sort items by $v \cdot \gamma$ or $v \cdot \delta$. Those items with the smallest values would be targeted for procurement. Determining values for v's is an avenue for future research as part of a dialogue with input from the humanitarian logistics community.

6.2.2. Where to put these items Assume now that a donor provides funding for blanket procurement and a manager or coordinated group must decide where to place them. The manager can use the dual variables to determine which depot results in the smallest increase (or biggest decrease) in total time or cost to serve beneficiaries. Thus, he would add it to the depot with the smallest π'_i. According to figure 3, Subang would be the best place to minimize cost, and Jakarta would be the best place to minimize time. Adding a unit to Subang would *decrease* the expected total cost, *even though the system is also serving more people*. This is because the current inventory position is imbalanced. One can rank preferable warehouses by the dual values. If certain non-quantifiable factors prevented the manager from placing items in Subang, he might choose Nairobi instead, the location with the second lowest dual value. Thus, dual variables offer an objective assessment of the value of a location to consider with other factors such as the political climate, incentives, and risks in making decisions.

6.2.3. Stock transfers between depots One way a manager or coordinated group can highlight items that are out of balance is to sort items by balance metric Δ. Looking at tables 2 and 3, we see that jerry cans are the most imbalanced item with respect to cost and time. With this item prioritized, the manager could decide among a few alternatives to improve the situation:

1. Investigate potential systematic issues that could cause the imbalance. Is there one dominant organization that is not optimizing inventory or collaborating? Were jerry cans delivered to a suboptimal location due to vendor error or cheaper procurement costs? Are there location risks that outweigh cost or time advantages? Acimovic and Graves (2015) showed that one large online retailer addressed previously undetected systematic errors by implementing a simpler non-stochastic version of this balance metric.

2. Transfer jerry cans from one warehouse to another. The operations manager could move a unit of inventory from the depot with the largest $\overline{\pi}'_{i_1}$ to the depot with the smallest $\underline{\pi}'_{i_2}$. The estimated value of doing this is $\overline{\pi}'_{i_1} - \underline{\pi}'_{i_2}$. If the value of reallocation exceeds the shipment cost, then the manager might shift inventory from depot i_1 to depot i_2. For instance, in examining data similar to figure 3, but for jerry cans instead of blankets, we estimate the value of moving a jerry can from Panama (jerry can dual variable of 0.139) to Subang (jerry can dual variable of -0.011) would net about (0.139 - (-0.011) =)\$0.15 USD in cost savings.

3. Procure new jerry cans into the warehouse with the smallest π'_i.

4. In a non-emergency phase of a disaster, ship items to beneficiaries from the warehouse with the lowest *true cost*, defined as $\min_i c_{ij} - \pi'_i$, in an attempt to rebalance the system.

6.3. Deriving broader insights

Beyond the research questions posed, our model and metrics can be used to answer broader questions posed by the humanitarian logistics community. The examples below are illustrative of ways in which these model-based metrics can be used and also offer evidence that, when considered with the assumptions and sensitivity analysis, could already guide strategies in the humanitarian community.

6.3.1. Cost structure for rapid response The results in tables 2 and 3 provide insights regarding the potential to optimize cost or time in global disaster response. These objectives depend greatly on the transportation mode mix, since air is very fast but very expensive compared with truck. Global organizations focused on cost may attempt to strategically deploy stock in order to more effectively use ground transportation. Our results show this is not easy given the current network structure. Trucks are only useful for up to 2% of the stock when time is the objective, whereas they carry around 20% when cost is being minimized. We show in the appendix that for a wide range of inventory levels, trucks carry only between 15% and 25% of the freight when minimizing cost. Since the global road network is fairly disconnected (and often too long to drive when it is connected) organizations must still rely heavily on air, even when 25 depots around the world are considered. Since trucks cannot effectively reach many disasters from current depot locations, the potential to significantly reduce cost by shifting modes from air to truck is limited. Transportation for the immediate response is going to be costly.

However, given the stock allocation in our data, there is significant potential to reduce the high cost of transporting 80% of freight by air through more effective distributed location of stock. The balance metric, Δ, indicates that proper allocation can reduce cost by 15-37% for these items. The value of a global footprint for stockpile deployment is not only to respond quickly in saving lives but, perhaps as important, to lower the inherently high cost of transporting critical items within 72 hours of a disaster in order to leverage limited funding to reach more people.

6.3.2. Allocation strategies according to system inventory Figure 2 showed the optimal allocation given the *current* inventory level for blankets of 852,563. Figure 6 extends this to show the optimal allocation of blankets when time is minimized for several levels of system inventory.

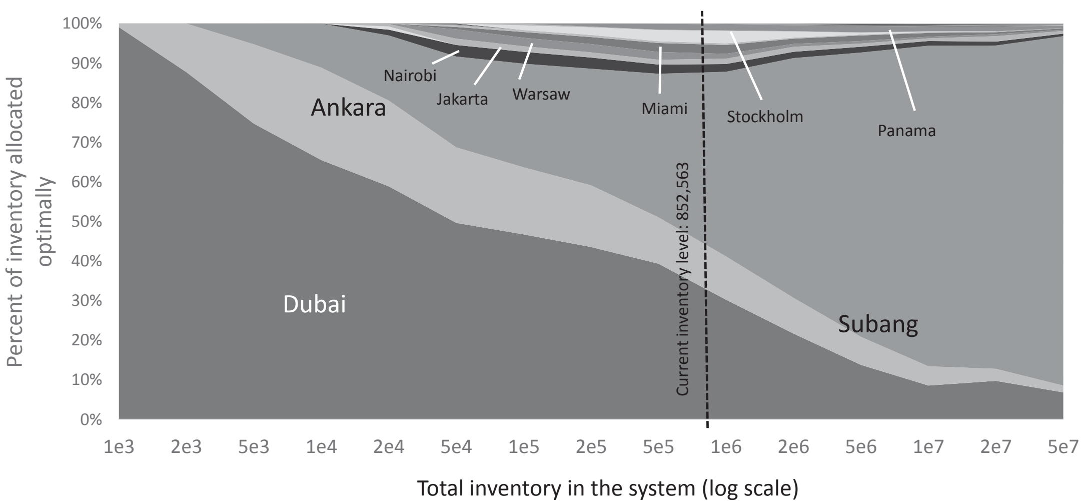

Figure 6 **Blankets: Optimal allocation across depots versus total inventory level (minimize time)**

From figure 6 we see that Dubai and Ankara are good locations if there is little inventory in the system. As inventory is added to the system, Subang becomes increasingly important because of the larger disasters that occur in Asia. The dotted line on this figure corresponds to the current total inventory level; as such the allocation along this dotted line matches the middle stacked bar in figure 2. It is interesting to note that when placing all 852,563 blankets optimally among only Dubai, Ankara, and Subang (the warehouses with the largest allocations of inventory) and forcing all other warehouses to hold zero units results in 3.6% additional expected deployment cost as opposed to using all warehouses as shown in figure 6. We note in figure 6 that one is not necessarily decreasing the amount of inventory in (for instance)

Ankara as the total inventory increases from about one million to fifty million. Rather, the *proportion* of inventory kept in Ankara reduces, even if the actual inventory stays constant or increases. Similar plots for jerry cans and mosquito nets are in the appendix.

It is interesting to note that Stockholm is a useful location to store blankets. This is due to the fact that flight paths on the spherical earth position Stockholm closer to certain disaster locations like China than other possible depots in the dataset. In the scenario when one million blankets are kept in the system, it is optimal to keep about 3% of these blankets in Stockholm, which serves disasters in China most often, followed by Mexico (as backup to limited stock in the Western Hemisphere), and then Russia.

6.3.3. Static and dynamic inventory strategies Different regions of the world have different disaster seasons in different months: flood season, hurricane season, monsoon season. Additionally, for blankets in particular, as the seasons change so do the requirements to keep people warm in a particular country. However impractical, if inventory *could* be moved among depots every month, what *would* the optimal month-by-month allocation be? If inventory is kept in a static (but optimal) allocation throughout the year, what is the impact on cost-to-respond? That is, how suboptimal (in terms of response only, ignoring transshipment costs) is a static inventory policy versus one that moves inventory twelve times throughout the year?

We discuss this in detail in the appendix and we display the results there in figures. In our analysis, we optimally reallocate one million blankets twelve times across twelve time blocks corresponding to the months. During the northern hemisphere's summer, more blankets are allocated to Asia - specifically Subang - as well as Stockholm. As hurricane season picks up in the Caribbean in the fall, more inventory is allocated to Panama and Miami. As winter approaches, inventory is placed in Africa and Turkey.

Rearranging this inventory every month can be costly, and may not be realistic for many organizations. If blankets were optimally allocated statically using data from disasters across the entire year, how 'bad' would this allocation be each month? During the northern hemisphere's summer, the system would incur costs about 1-2% more than the optimal month-specific allocation. From November to February, however, the cost-to-respond is 3-7% more expensive than the month-specific optimal inventory allocation. This is partly driven by the fact that large disasters in Asia during the summer heavily impact the static model's placement of inventory, at the expense of smaller disasters around the world at other times of the year.

We can examine each depot's adjusted dual variable on a month-to-month basis (assuming an actual static inventory allocation) to estimate the savings of mid-year inventory transfers. We examine the dual variables for Rio de Janeiro (Subang) because it has the lowest dual values among all depots in the winter (summmer) and highest in the summer (winter). The dual variable for Rio de Janeiro varies from -0.25 in December and January to 0.6 in June. Subang's dual variables range from from 0.1 in December and January to 0.2 in March through June. Interpreting the data from this figure, a manager – instead of moving all inventory throughout the system every month across all depots – may choose only to swap inventory between these two locations. For every blanket moved from Rio de Janeiro (Subang) to Subang (Rio de Janeiro) in summer (winter), approximately 0.40 (0.35) USD could be saved in cost-to-respond.

If the blankets can be moved via a slow mode of transport (such as water or rail) for less than this, than it may be cost effective to do some seasonal-specific inventory rearranging.

6.3.4. The value of collaboration We investigate and summarize here the cost of not collaborating and examine the resulting impact on inventory allocation. Further details are in the appendix. To quantify the cost of not collaborating, we assume that the current inventory of 852,563 blankets are evenly distributed among N identical organizations where $N \in \{1, 2, 5, 10, 20, 50, 100, 200, 500\}$. Each organization optimizes the placement of its share of the inventory $(852,563/N)$ in isolation. We compare the system cost-to-respond when these organizations act in isolation versus the system cost-to-respond if the organizations work together to optimize the placement of all of the inventory. The cost is higher when organizations work in isolation. The increase in cost-to-respond for the system is 0.3% when $N = 2$, 1.0% when $N = 5$, 2.0% when $N = 10$, 4.1% when $N = 20$, 6.3% when $N = 50$, 9.3% when $N = 100$, 13.1% when $N = 200$, and 17.8% when $N = 500$. When the number of organizations is small (or, as we show in the appendix, if the inventory is concentrated at a small number of organizations), there may not by much system degradation when working in isolation. However, as the number of organizations grows, it is increasingly important to cooperate and make decisions considering the *system* capacity in order to minimize the overall *system* cost.

Figure 6 lends insight as to why this degradation occurs. When there are very many organizations, each organization has very little inventory $(852,563/N)$. Thus, we see in figure 6 that when inventory levels are small, the model allocates inventory mostly to Dubai, which is – on average – near to most disasters of all sizes. When there are very few organizations, each organization holds a lot of inventory. Thus, according to figure 6, when inventory levels are high much inventory is placed in Subang to respond to the large disasters recorded in Asia. Thus, once the inventory in Dubai is at a high enough level to serve many of the small and medium sized disasters, our model places much of the additional inventory in Subang to serve beneficiaries in Asia's large disasters. Having lots of isolated organizations is suboptimal because too much inventory gets placed in the warehouse close to all disasters of all sizes (Dubai) and not enough is placed in the warehouse near the large disasters (Subang).

In the appendix, we show additional results relating to collaboration. We show on a stylized model with artificial arc costs that not collaborating can be arbitrarily bad. We also show that using actual organizations' actual inventory levels for blankets, the cost of not collaborating is only 2%. This is because organizations are in fact not identical, and 75% of the inventory of the 23 organizations holding blankets is concentrated at a only five organizations.

6.3.5. Proposed index Still, institutionalizing change in a sector can be daunting. We hypothesize that developing an index may facilitate faster adoption of common metrics. Translating a set of metrics into a composite value that is simple and dynamic may provide a useful gateway to encourage use of the evidence. As an initial step in this direction, we normalize the metrics in this paper and propose a sector-wide Response Capacity Index (RCI) that considers stock deployment decisions across organizations. Details are in the appendix.

This RCI encompasses information on a variety of aspects that policy analysts and decision makers might want to consider with a single composite number. Furthermore, it offers a simple measure to encourage transparency and accountability with the public regarding resources established for their benefit. Much like other indices or industry benchmarking efforts, we hope that the value of this dynamic information provides incentive for organizations to share data on a regular basis. The formal calibration of the RCI and assessment of its ability to create a virtuous cycle of data sharing and decision making, are left for further research.

7. Conclusion

Numerous humanitarian organizations deploy resources to increase logistics capacity in responding to natural disasters. Often these efforts are not coordinated and the combined capacity to meet needs is difficult to assess. Focusing on stockpile inventory, we propose new humanitarian logistics metrics that enable assessment and evaluation of response capacity while also providing evidence to guide dynamic, independent decisions toward system improvement.

Empirical studies using data from the United Nations demonstrate the potential for this approach. Results show that the combination of metrics, each focused on a distinct dimension of capacity, paints a new and effective common operating picture for the humanitarian sector. At the same time, we outline how the system metrics that quantify marginal change can guide managers toward decisions that improve system performance. Finally, the metrics offer quantifiable evidence in shaping and evaluating broader policies and interventions.

There are limitations to our approach that provide avenues for future work. First, we utilize past disaster data to forecast future needs. Our results are only moderately robust to the scenario selection from these historical data. The metrics should be based on better forecasting methods, such as those used by the insurance sector to forecast future risk to property. Second, we make several assumptions about parameters that can be improved by further engaging the humanitarian community for more extensive and more current data. These data might include: broader and more current inventory levels; better estimates of each country's internal capacity to respond; and better estimates of delivery costs and times. Third, the model ignores strategic capacity. It could be extended to consider such capacity, such as warehouse space or procurement budget, and thus become a multi-commodity model. Further consideration would be required as to how to mix strategic and tactical decisions, as well as fixed and variable costs, in assessing current capacity. Further expert opinion may be needed regarding the priority and importance of each commodity. Fourth, response capacity incorporates more than stockpile inventory. The structure of the SLP is easily extended to include supplier capacity to replenish depots and/or delivery directly to the affected community. For instance, if several NGOs have contracts with a single supplier whose capacity is below the sum of the contracted amounts of the NGOs, the model would highlight this. Fifth, we propose a prototype index, the RCI, which uses a simple average of the components. Further work could develop a more rigorous methodological basis for the index, incorporating expert opinion to scale and weight the components, consider further attributes, and leverage further empirical data for testing. Finally, we hypothesize that the value of information provides incentive for further information sharing. Empirical study regarding the interest and use of the proposed metrics could confirm or deny this virtuous cycle.

We aim to continue model-based development of response capacity metrics. In publicly sharing our research and our metrics, we hope to provide practical contribution while encouraging discourse to improve the methods.

8. References

Abidi, H., de Leeuw, S., and Klumpp, M. (2014). Humanitarian supply chain performance management: a systematic literature review. Supply Chain Management: An International Journal, 19(5/6):592–608.

Acimovic, J. and Graves, S. (2015). Making Better Fulfillment Decisions on the Fly in an Online Retail Environment. Manufacturing & Service Operations Management, 17(1):34–51.

Bauman, N. (2014). Assumptions for shelter NFI standards. Personal communication. Global Shelter Cluster. [email; 28-August-2014].

Centers for Disease Control and Prevention (2014). Malaria information and prophylaxis, by country. `http://www.cdc.gov/malaria/travelers/country_table/a.html`. [Online; accessed 01-September-2014].

Centre for Research on the Epidemiology of Disasters (2014). EM-DAT: The OFDA/CRED international disaster database. `http://www.emdat.be/`. Universit catholique de Louvain. Brussels, Belgium. [Online; accessed 28-June-2013].

de Brito Junior, I., Leiras, A., and Yoshizaki, H. T. Y. (2013). Stochastic optimization applied to the pre-positioning of disaster relief supplies decisions in brazil. Technical Report POMS 2013 paper.

Duran, S., Ergun, Ö., Keskinocak, P., and Swann, J. L. (2013). Humanitarian logistics: advanced purchasing and pre-positioning of relief items. In Handbook of Global Logistics, pages 447–462. Springer.

Duran, S., Gutierrez, M. A., and Keskinocak, P. (2011). Pre-positioning of emergency items for CARE international. Interfaces, 41(3):223–237.

Google (2014). The google distance matrix API. `http://developers.google.com/maps/documentation/distancematrix/`. [Online; accessed 17-October-2014].

Gralla, E., Goentzel, J., and Fine, C. (2014). Assessing trade-offs among multiple objectives for humanitarian aid delivery using expert preferences. Production and Operations Management, 23(6):978–989.

Guha-Sapir, D. and Below, R. (2006). Collecting data on disasters: easier said than done. Asian Disaster Management News, 12(2):9–10.

Haavisto, I. and Goentzel, J. (2015). Measuring humanitarian supply chain performance in a multi-goal context. Journal of Humanitarian Logistics and Supply Chain Management, 5(3):300–324.

International Federation of the Red Cross and Red Crescent Societies (2009). IFRC emergency items catalogue.

Klibi, W., Ichoua, S., and Martel, A. (2013). Prepositioning emergency supplies to support disaster relief: A stochastic programming approach. Technical Report CIRRELT-2013-19.

Peduzzi, P., Dao, H., Herold, C., and Mouton, F. (2009). Assessing global exposure and vulnerability towards natural hazards: the disaster risk index. Natural Hazards and Earth System Science, 9(4):1149–1159.

Shapiro, A., Dentcheva, D., and Ruszczyski, A. P. (2014). Lectures on Stochastic Programming: Modeling and Theory, Second Edition. SIAM.

The Sphere Project (2014). The sphere handbook. `http://www.spherehandbook.org/`. [Online; accessed 22-May-2014].

Tomasini, R. M. and Van Wassenhove, L. N. (2009). From preparedness to partnerships: case study research on humanitarian logistics. International Transactions in Operational Research, 16(5):549–559.

UN Office for the Coordination of Humanitarian Affairs (2014). Global mapping of emergency stockpiles. `http://www.humanitarianinfo.org/stockmap`. [Online; accessed 17-October-2014].

United Nations (2014). United nations humanitarian response depot. `http://www.unhrd.org/`. [Online; accessed 17-October-2014].

United Nations High Commisioner for Refugees (2012). Core relief items catalogue. Technical report, Geneva, Switzerland. `http://www.unhcr.org/cgi-bin/texis/vtx/home/opendocPDFViewer.html?docid=53fc545b9&query=core%20relief%20catalogue`.

Van Wassenhove, L. N. (2005). Humanitarian aid logistics: supply chain management in high gear. Journal of the Operational Research Society, 57(5):475–489.

World Weather Online (2014). World weather online. http://www.worldweatheronline.com/. [Online; accessed 29-October-2014].