Evaluation and optimization of humanitarian aid using data and network sciences: case-study of Afghanistan in 2015

Peter De Ford*, Javier Cuervo†, Farooq Khan†, Samuel Johnson‡

Abstract
We propose techniques to evaluate and optimise the performance of humanitarian organizations at a country scale using open online data. To evaluate performance, initially three data based indicators are proposed, which measure: how humanitarian organizations in a country supply aid proportionally to the demands; how able is an organization to reach the provinces more in need; and how well is aid being coordinated by regions of neighbouring provinces. Then, we propose other four indicators based mainly on network centrality measures, which are calculated from a network of partnerships between organizations, that for many countries can be inferred from a popular type of dataset. These four indicators measure: how big organizations are taking advantage of their size to make partnerships and influence others; how influential organizations tend to have partnerships among them and less with not so influential organizations; the role of international organizations as connectors between organizations; and the participation of organizations in doing partnerships. To optimise performance, we propose an algorithm that redistributes aid geographically in such a way that optimises some of the data and network based indicators proposed. The results were promising when we tested our ideas in the Afghanistan humanitarian scenario of 2015. We also compared some indicators calculated for Afghanistan with other seven countries, and there were good insights obtained.

Keywords
Data science — Network science — Optimization algorithms — Humanitarian aid indicators

**Student from: Mathematics of Real-World Systems Centre for Doctoral Training, University of Warwick, Coventry CV4 7AL, United Kingdom*
†*Supervisor from: Polymaths Consulting, Birmingham Research Park, Birmingham B15 2SQ, United Kingdom*
‡*Supervisor from: Centre for Complexity Science, University of Warwick, Coventry CV4 7AL, United Kingdom*

1. Introduction

Humanitarian aid is the material and logistic assistance to people in need due to natural disasters or man-made disasters, whose main purpose is to save lives, alleviate suffering and maintain human dignity. It is rooted in the moral principle of humanitarianism, which promotes kindness, benevolence and sympathy towards all human beings. The origins of organized international humanitarian aid trace back to the 19th century, when China received donations and assistance from foreign countries during the Northern Chinese Famine of 1876-79. Since then, the humanitarian aid spirit has been present in natural disasters like the 2004 Asian tsunami or the 2010 Pakistani Floods, and in conflicts like the Palestinian-Israeli one. There is a clear distinction with developmental aid, which seeks to build a country's infrastructure, institutions and economy to promote sustainable development.

There are huge organizations like the United Nations and the International Red Cross and Red Crescent Movement who lead the planning of humanitarian aid in most emergencies, but still the humanitarian system is very decentralized due to thousands of organizations involved with different goals and agendas. Disorganization among aid agencies and self-interested aid have led the humanitarian system effectiveness to be criticised [1, 2], with a perfect example being the 2010 Haitian earthquake. One of the reasons for ineffectiveness is that most aid organizations have to compete for donations to survive. In this race for money, they face the 'Humanitarian Dilemma' of managing to balance their altruistic intentions with their own survival interests and donors' agenda. This causes lack of cooperation among agencies, competition, duplication of efforts, sacrifice of quality for quantity, and interests of local and national stakeholders being ignored or minimised. A Sri Lankan government official said after the 2004 Asian tsunami: 'I don't know which was worse, the first wave of water or the second wave of aid' [3]. In order to improve the organization and effectiveness of humanitarian aid, new scientific and technological approaches are becoming more common.

The most extensive scientific research in humanitarian aid has been made in the optimisation of supply chain logistics, using for example multi-criteria optimisation models and stochastic optimization [4, 5]. But recently, new open data available in the Internet have given rise to novel indicators of aid perfor-

mance as well as new approaches for analysing coordination efforts using network theory [6, 7, 8, 9, 10, 11, 12, 13]. Coscia [12] proposed indicators to measure systematically developmental aid performance at a global scale from online data. Seybolt [10] developed a systemic network theory of humanitarian assistance, and argues that the humanitarian system is becoming more effective by comparing the humanitarian operations of Rwanda in 1994 and Afghanistan in 2001. Moore [7], for a flood in Mozambique, determined that there appeared to be a positive association between higher on average organization network centrality scores and higher on average beneficiary numbers, specifically during the emergency period immediately following the flood. Gillmann [13] proposed four network topologies of coordination among organizations: loose–alliance, orchestrated–alliance, lead–partnership and lead–agency (see Figure 1); then, after studying the Sudanese Darfur crisis and the 2004 Asian tsunami, determined that the lead–partnership topology was the best for both disasters in terms of cooperation among organizations. Bisri [11] confirmed Gillmann's result by analysing the 2009 West Java earthquake in Indonesia. However, there is still need for indicators of how humanitarian organizations are performing individually and as a whole in a given country, that have the possibility of being calculated from open online data. The existence of these indicators would lead to a rapid evaluation and optimisation of aid.

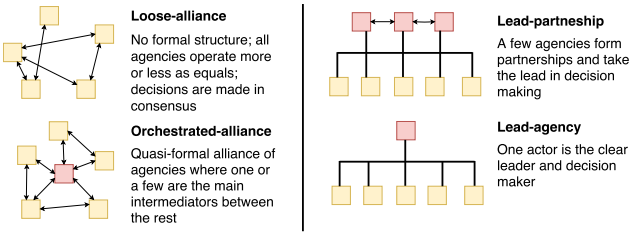

Figure 1. Network topologies of coordination among organizations defined by Gillmann [13].

Polymaths Consulting is a partner of the Warwick MathSys CDT who is very interested in tackling the humanitarian aid ineffectiveness from a complex systems perspective. In fact, they are willing to develop a software interactive tool that can help humanitarian coordinators and policy-makers to evaluate and optimize aid. As a first milestone towards it, this research project was created. We make use of open online data to design data based and network based indicators of how humanitarian organizations are performing individually and as a whole in a given country, and create an algorithm to redistribute the aid throughout the country in an optimal way based on the indicators proposed. We test the indicators and algorithm proposed in the Afghanistan 2015 humanitarian situation (not 2016 due to data limitations). This country was chosen due to its harsh socio-political situation that does not seem to end.

The report is structured as follows: in Section 2 we comment on the humanitarian situation in Afghanistan as of 2015; in Section 3 we describe the open online data obtained and how we prepared it for analysis; in Section 4 we propose three data based indicators to evaluate humanitarian aid; in Section 5 we explain how to infer a network of organizations partnerships from popular humanitarian aid datasets, and then propose four network based indicators to evaluate humanitarian aid from the network inferred; in Section 6 an algorithm that redistributes aid to optimize data and network based indicators is described; in Section 7 limitations are mentioned; finally, in Section 8 concluding remarks are made.

2. Humanitarian situation in Afghanistan

Afghanistan has been a great focus of conflict during the last four decades. Sadly, throughout this time the country became devastated due to the Soviet invasion (1979-1989), the Civil War (1989-1992), the Taliban Regime (1996-2001) and other conflicts. And even though in December 2001 the Taliban government was overthrown and a new Afghan government was formed, a huge power struggle between state and non-state armed actors is still causing conflict and affecting Afghanistan's development. Meanwhile, the international community, either adding fuel or not to the power struggle, have made an enormous aid intervention since 2002. An estimated of 50.7 billion USD have been given as official development assistance between 2002 and 2012, including 6.7 billion USD in humanitarian assistance [14]. Aid have brought important improvements in some basic development indicators; nevertheless, poverty, lack of education, insecurity, dependency on international aid, inequality and other important problems remain. This is reflected in the fact that Afghanistan ranks in the position 171 of 188 countries in the Human Development Index 2015 – last world position among non-African countries [15].

Afghanistan has a permanent structure of humanitarian aid. As of 2015, it was divided in five main areas – called **clusters**: 1. **Emergency, Shelter and Non-Food Items** (**ESNFI**), 2. **Food Security** (**FSAC**), 3. **Health**, 4. **Nutrition**, and 5. **Water/Sanitation** (**WASH**). There is an Inter-Cluster Coordination Team that assure coherence in achieving common objectives among clusters, avoiding duplication and ensuring areas of need are prioritized. Also, a Humanitarian Country Team (HCT) is in charge of guiding humanitarian action, which is composed mainly by representatives of United Nations agencies and representatives of non-governmental organizations (NGO). There were 149 humanitarian aid organizations reporting to the HCT in 2015, whose presence comprises most of the 398 districts of the 34 provinces of Afghanistan.

There are many failures in the structure of humanitarian aid, which are mentioned by the UNOCHA (United Nations Office for the Coordination of Humanitarian Affairs) in a report called *Afghanistan Coordination Architecture Review* [16] (see Appendix A for a summary). Also, a good overview of the situation per each of the five clusters is given in the

Humanitarian Needs Overview 2015 [17] (see Appendix A for a summary).

3. Data

In this section, we describe the datasets used in this research. Every dataset was always categorized as either demand or supply data. Demand data is any data that quantifies the humanitarian needs per province and cluster, while supply data is any data that quantifies the presence of humanitarian organizations per province and cluster. Some of the data required plenty of time to clean and prepare for analysis, while other required just to download.

First demand dataset: Demand Table

The UNOCHA (United Nations Office for the Coordination of Humanitarian Affairs) launched the website called Humanitarian Response as an online tool to support humanitarian operations globally. They publish in it the reports called Humanitarian Needs Overview (HNO) – one per year and country (just fragile countries). In the Afghanistan 2015 HNO report [17], a table called Overall Needs and Vulnerability Index (ONVI) is included, which shows 20 indicators per province (each related to one of the five humanitarian clusters or to insecurity). For instance, some indicators present are Under-5 mortality, Poor Access to Safe Water and Vaccination Coverage Deficit. We associated each of the indicators in the ONVI table (see list of indicators in Appendix C) to one of the five humanitarian clusters (except the indicators associated with insecurity). Then, for each cluster, we calculated an overall indicator per province by averaging all indicators associated with the cluster using weights provided in the ONVI table (a higher weight meant the indicator was more critical). After this, we multiplied the cluster overall indicator of each province by the population of the province. In this way, a **Demand Table** that shows the demand of each province–cluster pair was made. Finally, the Demand Table was scaled for convenience in such a way that the demand from all provinces in each cluster sum 100 (table included in Appendix D). Basically, a province demands more humanitarian aid from a cluster if the overall indicator is more critical and if it has more population. For example, Badakhshan province has an overall nutrition cluster indicator more critical than Baghlan (see Appendix C), and has more population than Baghlan, therefore it demands more humanitarian aid in the nutrition cluster (see Appendix D). The population data of provinces and districts used to build the Demand Table was obtained from the official webpage of the Central Statistics Organization of Afghanistan.

Second demand dataset: World Bank dataset

A **World Bank dataset** with 38 development indicators per province was used as a complement of the ONVI table in Section 4 (see list of indicators in Appendix E). The dataset was extracted from a report made by the World Bank jointly with the Afghanistan Ministry of Economy called *Afghanistan Provincial Briefs* [18], released in 2013 with data updated up to 2011.

First supply dataset: 3W dataset

The UNOCHA's Humanitarian Response website also publishes periodically the **3W datasets** (Who does What Where?) for many countries, which contain the presence of humanitarian aid organizations per cluster and per province/district. A small excerpt of a 3W dataset is shown in Appendix F. We call each row of the dataset a **project**, which means the presence of an organization in some district doing activities related to a given cluster. The Afghanistan 3W dataset was downloaded for the first trimester of 2015 since, as we will see, is necessary to make it match in time with the Demand Table. It contains 2669 projects from 149 aid organizations (names of organizations in Appendix G). The 3W datasets for seven other countries (Haiti, Liberia, Mali, Niger, Sierra Leone, South Sudan and Sudan) in the same period were also downloaded.

Second supply dataset: Supply Table

A **Supply Table** showing the humanitarian supply per province-cluster pair was built from the 3W dataset of Afghanistan. The supply for each province-cluster pair was defined to be equal to the number of projects associated with the given province and cluster. For instance, if the province of Kabul has just two organizations working in the nutrition cluster, one present in Bagrami district and the other present in Bagrami and Kalakan districts, then the supply of nutrition aid in Kabul province is equal to 3 projects. The Supply Table was also scaled for convenience in such a way that the supply to all provinces in each cluster sum 100 (table included in Appendix D). The reason behind the Demand Table and Supply Table scaling is, as we will see later, to make a supply–demand zero-sum system from which we can analyse the effectiveness of aid distribution in an easier way. By zero-sum we mean that the sum of the supply to all provinces minus the demand from all provinces is equal to zero (per cluster).

4. Data based indicators

In this section we propose three indicators of the performance of humanitarian organizations individually and as a whole using the data described in Section 3. Indicators of the state of a system in the development and humanitarian aid sectors tend to be very simplistic (commonly in the form of percentages). There are exemptions like the Gini-index used to measure income inequality, which is a more complex indicator calculated in a non-trivial way. But curiously, complex indicators are usually more accurate in summarizing the state of a system. For example, Hidalgo's indicator of economic complexity outperforms GDP and other traditional indicators as measures of prosperity in a country [9]. The following are our three indicators proposed, which are more complex than simplistic.

Indicator: Humanitarian Presence Inequality (HPI)

The first indicator we created, and the less complex of the three, is called Humanitarian Presence Inequality (HPI). It measures at a country scale how is the supply of humanitarian aid proportional to the demand, by making an aggregation of supplies and demands over all province-cluster pairs. Before defining it, we introduce a surplus metric for each province-cluster pair:

$$Surplus_{i,j} = Supply_{i,j} - Demand_{i,j} \tag{1}$$

where i is a province belonging to the set P of all provinces, j is a cluster belonging to the set C of all clusters, and where $Demand_{i,j}$ and $Supply_{i,j}$ are taken from the Demand Table and Supply Table, respectively. Now, we define the Humanitarian Presence Inequality indicator as:

$$HPI = \sqrt{\frac{\sum_{i,j}(Surplus_{i,j} - \overline{Surplus})^2}{I * J}} \tag{2}$$

where $\overline{Surplus}$ is the arithmetic mean of $Surplus_{i,j}$ when averaged over all i and all j, I is the total number of provinces in P, and J the total number of clusters in C. Note that HPI is the sample standard deviation of the $Surplus_{i,j}$ over all i and j, and punishes severely provinces with very positive surplus or very negative surplus (deficit). This means that the smaller the magnitude of HPI, the better. Disaggregated versions of HPI can be obtained per province or per cluster.

We calculated first the HPI of Afghanistan, which was of 2.8097. In Section 6 we optimise this value using the algorithm proposed. Then, we calculated the HPI per of each of the 34 provinces of Afghanistan; table 1 shows the top 5 best and top 5 worst provinces.

Province	Population*	HPI province	Ranking
Panjsher	151000	0.32	1
Uruzgan	380500	0.42	2
Badghis	487800	0.64	3
Nimroz	162100	0.68	4
Farah	498900	0.68	5
Ghazni	1208600	3.27	30
Balkh	1298300	3.54	31
Wardak	586600	3.74	32
Kabul	4227200	4.27	33
Nangarhar	1489800	7.14	34

*in number of inhabitants

Table 1. HPI per province top 5 best and top 5 worst provinces.

We noticed that provinces with more population tend to have a bigger HPI per province (sample correlation coefficient between population and HPI per province for all provinces was 0.53). Additionally, we found that provinces with higher conflict profiles (data taken from the ONVI table insecurity indicators) have a slight correlation with HPI per province (sample correlation coefficient of 0.34). Then, we calculated HPI per cluster, shown in table 2 in descending order.

Cluster	HPI cluster
ESNFI	3.77
WASH	3.40
Nutrition	2.89
FSAC	2.08
Health	1.42

Table 2. HPI per cluster.

ESNFI and WASH clusters had the worst HPI per cluster, which coincides with disorganization reported for exclusively these two clusters [16]. We will later see that HPI per cluster is related to one of the network based indicators proposed.

Indicator: Organization Reachability (OR)

The OECD (Organization for Economic Co–operation and Development) says that the performance of a single organization cannot be evaluated in isolation from what others are doing [19]. In line with this, the second indicator proposed is an evaluation of an organization's performance which takes into account what all the others are doing. The indicator is called Organization Reachability (OR), and measures the capacity of an aid organization to reach the provinces/districts with the most deficit of humanitarian presence. The OR can be calculated for a given cluster, or for all clusters in which an organization is involved. For the first case, the OR is defined as:

$$OR_{j,k} = \frac{\sum_i(-Surplus_{i,j} * Supply_{i,j,k})}{Supply_{j,k}} \tag{3}$$

where i is a province; j is a cluster; k is an organization; $Supply_{j,k}$ is the number of projects in different districts where the kth organization is working in the jth cluster (no need of scaling); $Supply_{i,j,k}$ is the number of projects in different districts of the ith province where the kth organization is working in the jth cluster (no need of scaling neither); and $Surplus_{i,j}$ is the surplus of the jth cluster in the ith province (as previously defined). The more positive $OR_{j,k}$ is, the better reachability an organization has in a given cluster. The explanation to this is that if the organization has projects in districts of a province with deficit (negative $Surplus_{i,j}$), then the term in the summation will be positive, but if the organization is present in a province with positive $Surplus_{i,j}$ it will get a negative value in the summation term. The OR of an organization for all

clusters involved, OR_k, is defined as the arithmetic mean of all its $OR_{j,k}$ of the clusters in which is involved. It is worth to mention that if the projects are redistributed proportionally to the provinces' demands per cluster (lower the HPI), then all organizations OR_k will lean towards values closer to zero since the absolute values of surplus inside the summation will become smaller.

We calculated the OR_k for each of the 149 organizations in Afghanistan. A table in Appendix H shows the organizations with the top 10 best and top 10 ten worst OR_k values. We noticed that international organizations ranked much better than nationals. In fact, 8 international organizations are in the top 10 best and just 3 are in the top 10 worst. This suggest that international organizations in Afghanistan are being more able than the nationals in reaching the districts that need most aid.

Indicator: Regional Coordination Index (RCI)

The third indicator we present is called Regional Coordination Index (RCI). It measures how neighbouring provinces that have similar cluster overall indicators (Appendix C) are likely to share the same organizations in common. The ideal behind this is that regions of neighbouring provinces with very similar cluster overall indicators must have similar organizations working in them. The inspiration to create this indicator came when we realised that the Afghan provinces of Daykundy and Ghor are the neighbouring provinces with the most similar cluster overall indicators and do not share any organization in common. Before defining the RCI, we introduce two metrics called *distance* and *relevance*, which are used to compare neighbouring provinces.

For the distance metric, we have that each province has 5 cluster overall indicators (see Appendix C), from which a 5–dimensional point in Euclidean Space can be created. The *distance* between two neighbouring provinces is the Euclidean distance between their 5–dimensional points. The smaller the distance means the two neighbouring provinces are more similar to each other regarding the five cluster overall indicators.

Relevance is a ratio between an observed and a randomly expected value, known as relative risk in statistics or lift in computer science. We use relevance as a measure of how similar are neighbouring provinces regarding the organizations that are present in each. To define it, consider N_{pro} as the number of existing organization-cluster pairs (from the 3W dataset), $N_{pro,a}$ as the number of organization–cluster pairs associated to province a, $N_{pro,b}$ as the number of organization–cluster pairs associated to province b, and $N_{pro,a,b}$ as the number of organization–cluster pairs associated to both provinces. Then, the relevance between provinces a and b is defined as:

$$R_{pro,a,b} = \frac{observed}{expected} = \frac{\frac{N_{pro,a,b}}{N_{pro}}}{\frac{N_{pro,a}}{N_{pro}} * \frac{N_{pro,b}}{N_{pro}}} \tag{4}$$

which can be simplified to:

$$R_{pro,a,b} = \frac{N_{pro,a,b} * N_{pro}}{N_{pro,a} * N_{pro,b}} \tag{5}$$

Ideally, the less distance between neighbouring provinces the more relevance they should have. To obtain the RCI, two separate rankings from best to worst must be made for distance and relevance. In the best case, all pairs of neighbouring provinces will rank in the same position in both rankings. The Regional Coordination Index (RCI) is designed to capture the differences in ranking positions, and is defined to be equal to one for the best case and zero for the worst case (when the differences in ranking position are maximized). The mathematical definition is as follows, consider representing each pair of neighbouring provinces as p, the difference in their ranking position as $diff_p$, and the total number of pairs of neighbouring provinces as M, then RCI is defined as:

$$RCI = 1 - \frac{\sum_p diff_p}{M^2} \tag{6}$$

if M is odd, and:

$$RCI = 1 - \frac{\sum_p diff_p}{M^2 + M} \tag{7}$$

if M is even. The denominator in the fraction is a scaling term that represents the maximum possible sum of differences, which is not equal for odd and even M. In general, the RCI indicator should reflect the good efforts of organizations and national/regional humanitarian coordination teams to organize the aid by regions of similar neighbouring provinces.

In Afghanistan, there are 80 pairs of neighbouring provinces in the set P of 34 provinces. For each pair, the distance and relevance metrics were calculated. Figure 2 displays a map of Afghanistan with the magnitudes of distance and relevance as colored links between each pair of neighbouring provinces. The neighbouring provinces of Daykundy and Ghor are highlighted in red because they had the biggest contrast between distance and relevance, ranking 1st on the distance ranking and 74th on the relevance ranking (see Figure 2). The RCI of Afghanistan in the first trimester of 2015 was 0.268, which is clearly from Figure 2 far away from optimality. In Section 6 we present an optimization algorithm that will increase the value of RCI.

As an addition to our research, we decided to find the regions of very similar neighbouring provinces based on the distance metric in order to inform about these to the Humanitarian Coordination Team (HCT) in Afghanistan - so they can improve the RCI. For this purpose, a hierarchical clustering approach [20] was used to cluster the 34 provinces by means of the *distance* metric. To be more sure about our results, we obtained the Euclidean distances between provinces not just

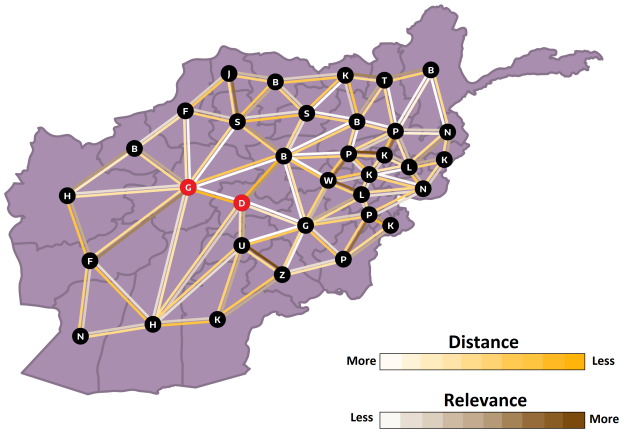

Figure 2. Map of Afghanistan with the magnitudes of distance and relevance as colored links between each pair of neighbouring provinces (map shows just the first letter of each province). The provinces of Daykundy and Ghor are highlighted in red, since they have the biggest contrast between distance and relevance (the less distance the more relevance expected).

using the humanitarian cluster overall indicators (Appendix C), but also using the developmental indicators from World Bank dataset described in Section 3 and Appendix E. The hierarchical clustering algorithm was applied for each of the two sets of distances. From the two dendrograms obtained, we chose heuristically in both cases to cluster the provinces in seven groups. The result was that we found five regions of neighbouring provinces that belonged to the same cluster in both cases, which are shown in Figure 3 differentiated by color. These regions are very similar regarding humanitarian and developmental indicators, so the HCT might consider to treat them as single units.

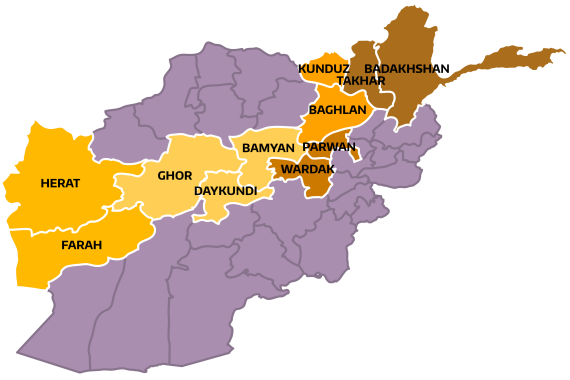

Figure 3. Five groups of neighbouring provinces (differentiated by color) that are very similar regarding humanitarian and developmental indicators.

5. Network based indicators

Coordination among organizations is critical to optimise the flow of resources and to increase aid effectiveness [21]. One form of coordination is through partnerships between agencies, where they exchange information, workers, resources and even money. Galaskiewicz [22] argues that one of the reasons why agencies establish partnerships is to guarantee organizational survival, even though they can lose some autonomy and independence. From all the partnerships among organizations in a country, there is a network of organizations that emerges, whose topological properties are related to the effectiveness of aid coordination [7]. In this section, we propose four indicators of humanitarian aid coordination based on network centrality measures.

To calculate the indicators we need a network of partnerships among organizations, which for many countries can be inferred from a 3W dataset. To infer the network, we turn up again to relevance, and assume that if two organizations working in a given cluster coincide in districts much more that what is randomly expected (assuming uniform distribution), then there is some partnership between both. In the implementation of the relevance metric to build the network, we decided to take into account the population of districts. The reason is that if two organizations coincide mainly in districts of big population the likelihood of having a partnership is less than if they coincide mainly in districts of small population. Thus, we define relevance mathematically in the following way; consider N_{org} to be:

$$N_{org} = \sum_{h \in S} \frac{D}{pop_h} \tag{8}$$

where h is a district–cluster pair, S is the set of all the district–cluster pairs in the 3W dataset (the universal set), D is the maximum district population (a scaling variable for computational purposes), and pop_h is the population of the district associated with the district–cluster pair h. Then, consider $N_{org,a}$, $N_{org,b}$, and $N_{org,a,b}$ to be respectively:

$$N_{org,a} = \sum_{h \in A} \frac{D}{pop_h} \tag{9}$$

$$N_{org,b} = \sum_{h \in B} \frac{D}{pop_h} \tag{10}$$

$$N_{org,a,b} = \sum_{h \in A \cap B} \frac{D}{pop_h} \tag{11}$$

where A is the set of all district–cluster pairs in which organization a is working, and B the set of all district–cluster pairs

in which organization b is working. Using these definitions, we calculate relevance in the same way as in Equation 4:

$$R_{org,a,b} = \frac{N_{org,a,b} * N_{org}}{N_{org,a} * N_{org,b}} \tag{12}$$

Using relevance, an unweighted undirected link between two organizations a and b is made if $R_{org,a,b} \geq 2$ – when they coincide working in the same district–clusters twice or more the expected value. We excluded from the network the organizations having less than three projects (i.e. working in less than three district-cluster pairs), since they make easy connections to others that might not hold in real life. For instance, an organization working just in Kabul–Nutrition would get connected to all other organizations in this district–cluster pair since the overlap with each other is very significant, but that might not be the case in the reality.

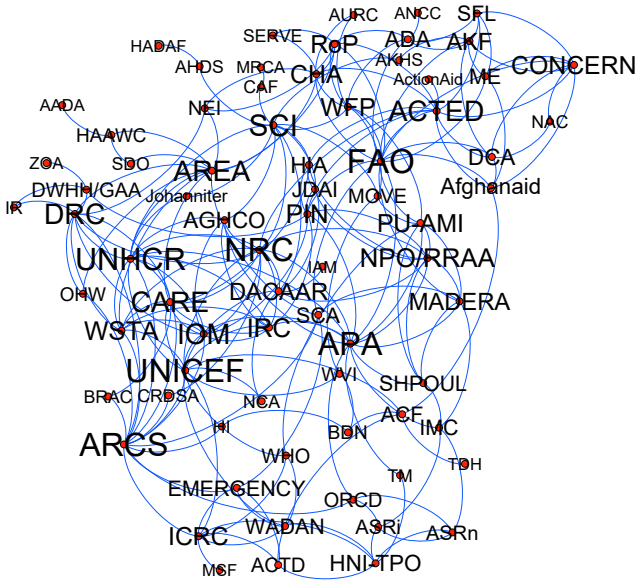

Figure 4. Network of partnerships between organizations inferred from the 3W dataset. Nodes are organizations and edges partnerships.

Figure 4 shows the network obtained for Afghanistan, composed of 75 nodes (organizations) and 206 edges (partnerships). The sizes of the organizations letters are proportional to their degree. At first sight, one can see the acronyms of some of the most popular aid organizations (i.e. UNICEF, UNHCR, FAO, WHO, ICRC) having high degree, which means many partnerships. Even though not all the organizations publish their partnerships in the Internet, we were able to find that many links/partnerships exist in real life. For instance: Afghan Red Crescent Society (ARCS) with the International Committee of the Red Cross (ICRC); World Food Programme (WFP) with Food and Agriculture Organization of the United Nations (FAO); or ACTED with ActionAid connected through the UK Department for International Aid in several projects. This network is mostly similar to the lead-partnership topology of Figure 1, which was determined to be the best in terms of cooperation [13].

The indicators of aid we are about to propose are based mostly on the following network centrality metrics: degree, eigenvector centrality and betweennness. Degree measures the quantity of partnerships an organization has, which might be proportional to how influential it is. Eigenvector centrality is the least popular of the three; basically, the vector of eigenvector centralities for all nodes in the network is equal to the left-hand eigenvector of the network's adjacency matrix associated with the eigenvalue of largest magnitude. This metric measures how is the organization connected to other important and influential organizations. Betweenness measures the extent to which an organization can exert greater influence in the paths between organizations, in the form of flow of information, resources, etc; also, it measures how organizations are important connectors between different sectors of the network.

Indicators: degree–size and degree–centrality

The first network based indicator proposed is called *degree–size*, and is defined as the sample correlation coefficient between the degree of organizations and their size. This indicator is meant to capture how big organizations are taking advantage of their size to make partnerships and influence others. The organization size is defined as the number of projects (or district–cluster pairs) in which the organization is present. For instance, if UNICEF is just present in Kabul–Nutrition, Kabul–Health, and Bagrami–WASH, its size is 3.

The degree–size of Afghanistan was 0.52. In order to get an idea of how high or low this was, we calculated the degree–size indicator for seven other countries from their 3W datasets in 2015. Table 3 shows the results. Most countries (including Afghanistan) had values between 0.24 and 0.67, showing that big organizations in these countries are very influential. Sierra Leone and South Sudan had a negative degree-size, evidencing that small organizations tend to have an important influence.

The second indicator proposed is called *degree–centrality*, defined as the sample correlation coefficient between the degrees of organizations and their eigenvector centrality. This indicator captures how influential organizations tend to have partnerships among them, and less with not so influential organizations. Table 3 contains the degree-centrality for Afghanistan and the seven other countries. All countries lie in the high value narrow range from 0.73 to 0.92, suggesting that at a global scale countries might all be close to this range (further research needed).

Indicator: betweenness–int

The third indicator proposed is called *betweenness–int*, defined as the number of international organizations present in the top 20 ranking of organizations with highest betweenness. The indicator captures the importance of international organizations in connecting together the different sectors of the network. In Afghanistan, the betweenness–int was 16, a

Country	Degree-size	Degree-centrality	Orgs.
Afghanistan	0.52	0.85	75
Haiti	0.49	0.92	42
Liberia	0.24	0.90	21
Mali	0.60	0.87	39
Niger	0.59	0.89	66
Sierra Leone	-0.30	0.92	13
South Sudan	-0.01	0.73	129
Sudan	0.67	0.91	85

Table 3. Degree-size and degree-centrality indicators for Afghanistan and seven other countries. The column called Orgs contains the number of organizations in the network.

high value that indicates that if international organizations go away (which is actually happening [17]) might leave a vacuum of power difficult to replenish [7]. One possible explanation to this high value is that international organizations (besides having more resources) have a greater capacity of remaining neutral during conflict than national organizations. This allows them to easily connect different sectors of the network, and therefore get more betweenness. Moreover, a language barrier that complicates the integration of national organizations into the network might be one of the reasons why nationals do not have high betweenness.

Indicator: network-participation

The last indicator is called *network-participation*, and is defined as the percentage of organizations connected to the network ($degree \geq 1$). We think that this indicator can explain low HPI values, since low participation would probably yield more disorganization. To test this using the Afghanistan data, we calculated the network-participation for networks inferred for each cluster (5 networks in total). It resulted that the clusters of Nutrition, FSAC and Health had HPI per cluster values smaller than 3 while having a network-participation greater than 0.7; in the other hand, ESNFI and WASH had HPI per cluster values greater than 3 while having a network-participation smaller than 0.7. This leads us to believe that low network-participation is truly related with a low HPI indicator, due to the reported disorganization in the ESNFI and WASH clusters that we mentioned before, but research with more data of many countries is needed to get conclusive results.

6. Optimization algorithm

In this section we describe an algorithm for optimising humanitarian aid (the psudo-code of the algorithm is presented in Appendix I). The algorithm optimises directly the HPI and the RCI indicators, and implicitly the network-participation indicator. What the algorithm does is to translate several organization's projects from their present districts to other districts. Commanding an organization project to move from one place to another might seem rude in the first instance; however, humanitarian aid organizations can very flexible due to the nature of humanitarian aid. For instance, in the case of Afghanistan, the Afghanistan Coordination Architecture Review [16] asserts that the majority of agencies are willing to change the location of activities to reduce gaps and duplication. The algorithm tries to maintain as much as possible the existing partnerships by moving especially projects of organizations with few or no partnerships (low degree organizations). To be more pragmatic, the algorithm tries to minimise HPI and RCI by moving projects just small distances. The distance between any two provinces is the network hop distance in a network of provinces where each province is connected only to all its geographical neighbours. The algorithm works in two phases: first, it minimises the HPI, and second, it minimises the RCI without altering the value of the HPI, and in both phases tries to increase network-participation by making organizations coincide in districts, which increases the likelihood of joint cooperation.

The first phase minimises the HPI by taking all surplus of projects in every province-cluster pair with surplus and translating them to province-cluster pairs with deficit. The algorithm works at a district level, and just takes out projects from district-cluster pairs with surplus, and assigns them to district-cluster pairs with deficit in other provinces. The demand (supply) for each district-cluster pair is estimated from the Demand (Supply) Table, by multiplying the population of the district times the demand (supply) of the province-cluster in question, and then dividing by the population of the province (an arithmetic 'rule of three'). Algorithm 1 in Appendix I describes this first phase in detail.

The second phase is concerned about the maximisation of RCI, and requires that provinces swap projects of same clusters. Since a swap does not represent an increase or decrease of supply per cluster for any province, then the HPI is unaltered. The algorithm iteratively finds the pair of neighbouring provinces whose relevance ranking is more below its distance ranking, and swaps some of their projects with other projects from their neighbours in order to correct the difference in ranking positions. The algorithm stops when no more projects can be possibly swapped. Algorithm 2 in Appendix I describes the second phase in detail.

After running the algorithm for our Afghanistan case, we obtained the following results. In the first phase, a total of 653 projects out of the 2669 were translated. The country HPI was reduced from 2.810 to 0.411. In the second phase, by translating 522 projects in regional areas, the RCI was increased from 0.268 to 0.729. Throughout both phases, the network–participation was increased from 0.50 to 0.62.

7. Limitations

The data based indicators proposed work under assumptions that might not totally hold under some circumstances. For instance, the demand of a given province-cluster pair was calculated as the multiplication of the cluster overall indi-

cator times the population of the province, but problems in the estimation of indicators like *sampling bias* might lead to inaccurate values of demand. The supply for a given province-cluster pair is assumed to be proportional to the number of projects belonging to that pair; however, project sizes might vary especially between big and small organizations. The HPI indicator assumes that it is a humanitarian coordination failure if supply is not proportional to the demand, but there might be political issues beyond humanitarian coordination that might make this not possible. Weak values in the OR indicator obtained by an organization (like UNHABITAT in our Afghanistan case) might be not because the organization is not reaching the districts with most deficit, but because they arrived first to those districts and then other organizations copied them in going to those districts. A low RCI value obtained might not reflect lack of regional coordination, but the fact that some pairs of neighbouring provinces are controlled each by different parties (i.e. government and insurgency) or belong each to different political regions of the country, which might hinder same organizations to have projects in the two provinces of these pairs.

The network based indicators are calculated from a network of partnerships inferred from a statistical significance metric. In real life some of these partnerships might not exist. Also, it is assumed that the presence of organizations throughout the districts is uniformly distributed; however, security issues would make organizations to prefer some districts over others.

The optimization algorithm estimates demand per district under the assumption that all districts in a province are equal regarding the five cluster overall indicators, but this is usually not the case. Additionally, the optimization algorithm assumes that organizations are willing to change the location of their projects, which might not be true in all cases.

8. Concluding remarks

In this research, we have proposed data and network based indicators of humanitarian aid that are useful to rapidly evaluate humanitarian organizations performance in a given country, and then an algorithm to optimise some of the indicators proposed. We chose Afghanistan as a case-study to test our ideas due to its complicated humanitarian situation and reported disorganization [16].

The data based indicators proposed were: Humanitarian Presence Inequality indicator (HPI), Organization Reachability (OR) and Regional Coordination Index (RCI). The HPI was designed to measure how is the supply of humanitarian aid proportional to the demands, which can be calculated at a country or province scale, or per cluster. We found that high HPI per province might be related to high population and insecurity. Also, we encountered that high HPI per cluster might be related to disorganization inside the cluster. We proposed the OR indicator to evaluate how are organizations able to reach the districts that are most in need. We found that in Afghanistan international organizations ranked much better than national ones. The RCI was proposed to capture how neighbouring provinces that have similar cluster overall indicators are likely to share the same organizations in common. The RCI for Afghanistan was far from optimality. The data based indicators work better for long term humanitarian situations like the one in Afghanistan. This is because time allows for census and estimation of indicators used to build the Demand Table.

Taking ideas from Scocia [12], we established a method to infer a network of partnerships that can be applied usign 3W datasets along with data of population per district. A network was built for Afghanistan, and also for seven other countries for comparison purposes. We proposed four network based indicators that work using the networks inferred. The degree-size indicator was proposed to capture how big organizations are taking advantage of their size to make partnerships and influence others. We found that Afghanistan and most of the other countries had a high degree-size indicator. The degree-centrality indicator was designed to capture how influential organizations tend to have partnerships among them. It resulted that all countries had high values inside a narrow range. We think that a high value of this indicator can reflect a lead–partnership network topology. We proposed the betweenness–int indicator to measure the importance of international aid organizations in the country, which for Afghanistan was really high. Finally, we created the network-participation indicator to measure the percentage of organizations participating in the network of partnerships. We found that low values of it might be related to low HPI.

In the last part, we created an algorithm that optimises the HPI, RCI and network–participation indicators. It achieves it by translating projects from districts with surplus to districts with deficit. When applied to Afghanistan, the three indicators were improved. We think that the optimisation algorithm is useful just as an informative tool, from which just some of the projects translated by the algorithm can be translated in real life.

This research was just the initial step of Polymaths Consulting to tackle humanitarian aid from a complex systems perspective. We consider that the implementation of the indicators or the algorithm proposed in a given humanitarian emergency of any country must be done hand in hand with the teams involved in the coordination of humanitarian affairs.

Our further work will focus on the following three areas. The first is finding ways to overcome the limitations mentioned in order to implement the techniques described here in real life scenarios. The second is to study in more detail how the network based indicators are related to the network topologies from Figure 1 described by Gillmann [13]. We think that high network-participation and a high degree-size might be the signature of the lead-partnership topology. The third is to propose a way in which agencies can self-organize to be more effective than centralized organization.

Acknowledgements

I would like to thank Farooq Khan and Javier Cuervo, supervisors from Polymaths Consulting, for giving me the opportunity of participating in this project and for the invaluable discussions we had. I am also thankful to Samuel Johnson, supervisor from The University of Warwick, for providing good ideas to the project. Finally, I extend my gratefulness to Kabir Isakhel (Senior Advisor Office of the President's Special Representative for Reform and Good Governance, Afghanistan) for the information provided. The financial support is gratefully acknowledged to the Foreign and Commonwealth Office of the United Kingdom under the **Chevening Scholarship** programme; also to the **MICITT** (Ministry of Science, Technology and Telecommunications of Costa Rica) under the Incentive Contract FI-135B-14.

References

[1] Dambisa F Moyo and Niall Ferguson. *Dead aid: Why aid is not working and how there is a better way for Africa*. Farrar, Straus and Giroux, New York, 03 2009.

[2] Timothy Nunan. *Humanitarian invasion global development in cold war Afghanistan*. Cambridge University Press, 2016.

[3] Ben Ramalingam and Michael Barnett. The humanitarian's dilemma: Collective action or inaction in international relief? Technical report, 2010.

[4] Paulo Gonçalves and Hugo Yoshizaki. Stochastic optimization for humanitarian aid supply and distribution of world food programme (wfp) in ethiopia. Technical report.

[5] Begoña Vitoriano, M. Teresa Ortuño, Gregorio Tirado, and Javier Montero. A multi-criteria optimization model for humanitarian aid distribution. *Journal of Global Optimization*, 51(2):189–208, 09 2010.

[6] Max Stephenson. Toward a descriptive model of humanitarian assistance coordination. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, 17(1):40–56, 03 2006.

[7] Spencer Moore, Eugenia Eng, and Mark Daniel. International ngos and the role of network centrality in humanitarian aid operations: A case study of coordination during the 2000 mozambique floods. *Disasters*, 27(4):305–318, 12 2003.

[8] Aurélie Charles, Matthieu Lauras, and Rolando Tomasini. Collaboration networks involving humanitarian organisations – particular problems for a particular sector. pages 157 – 165. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[9] Sebastián Bustos Michele Coscia. Sarah Chung Juan Jimenez Alexander Simoes Ricardo Hausmann, César A. Hidalgo and Muhammed A. Yıldirim. *The atlas of economic complexity : mapping paths to prosperity*. Updated edition. Cambridge, MA : The MIT Press, [2013].

[10] Taylor B. Seybolt. Harmonizing the humanitarian aid network: Adaptive change in a complex system. *International Studies Quarterly*, 53(4):1027–1050, 12 2009.

[11] Mizan B.F. Bisri. Examining inter-organizational network during emergency response of west java earthquake 2009, indonesia. *Procedia Environmental Sciences*, 17:889–898, 2013.

[12] Michele Coscia, Ricardo Hausmann, and César A. Hidalgo. The structure and dynamics of international development assistance. *Journal of Globalization and Development*, 3(2):1–42, 2013.

[13] Nina Gillmann. *Interagency Coordination during Disaster Strategic Choices for the UN, NGOs, and other Humanitarian Actors in the Field*. Baden-baden: Nomos, 2009.

[14] Lydia Poole. Aid and the transformation decade. Technical report, 2014.

[15] Humanitarian development report 2015. Technical report, 2015.

[16] Afghanistan coordination architecture review. Technical report, 2015.

[17] Humanitarian needs overview 2015. Technical report, 2015.

[18] Afghanistan provincial briefs. *World Bank*, 11 2014.

[19] *Guidance for Evaluating Humanitarian Assistance in Complex Emergencies*. OECD (Paris), 1999.

[20] Stephen C. Johnson. Hierarchical clustering schemes. *Psychometrika*, 32(3):241–254, 09 1967.

[21] Francisco Rey. The complex nature of actors in humanitarian action and the challenge of coordination. *Reflections on Humanitarian Action: Principles, Ethics and Contradictions. TNI/Pluto Press with Humanitarian Studies Unit and ECHO (European Commission Humaritarian Office), London*, 2001.

[22] J Galaskiewicz. Interorganizational relations. *Annual Review of Sociology*, 11(1):281–304, 01 1985.

Appendix A

Further details of the humanitarian aid situation in Afghanistan

The UNOCHA (United Nations Office for the Coordination of Humanitarian Affairs) produced a report called Afghanistan Coordination Architecture Review [16] which mentions the following facts about the humanitarian coordination structure. Humanitarian coordination is still required in Afghanistan; serious gaps remain in terms of establishing a centralised body to coordinate the aid. The existence of multiple different coordination mechanisms at one time in one place have led to duplication of roles and responsibilities. There is an obvious disconnect between national and sub–national levels coordination. Cluster leading NGOs (Non–Governmental Organizations) demonstrate little enthusiasm to share leadership responsibilities or support the sub–national coordination mechanisms without incentive. Resourcing across clusters varies substantially, from the relatively well–staffed and funded FSAC cluster to the not so well coordinated WASH and ESNFI clusters. An absence of sector/cluster standards and disagreement on technical approaches or agreed assistance packages has caused significant delays in delivery of humanitarian assistance. Most clusters have neglected the responsibility to ensure that clusters continue to operate only while they are strictly needed. The majority of partners continue to view coordination as information sharing. The majority of partners also wish to maintain a fairly high degree of organisational distinctiveness and autonomy but are willing to accept common guidance, or change the nature or location of activities to reduce gaps and duplication. The availability of recent, comparable national assessment data in Afghanistan remains a fundamental weakness to accurately identify priority needs and inform gap analysis. The increasing complexity and potential security implications for organisations when attempting to deliver assistance in new areas is also becoming an increasing impediment to providing assistance where it is most greatly needed. In the coming years, the conflict and displacement trends seen in 2015 are likely to remain consistent, if not increase.

The Humanitarian Needs Overview 2015 [17] provide the following insights about the humanitarian situation for the five clusters. The Emergency, Shelter and Non-Food Items (ESNFI) cluster is critical for assisting the natural disaster displaced population. The country has a subarctic mountain climate with dry and cold severe winters, during which temperatures can fall to -20 degrees Celsius at higher altitudes and cause thousands of displaced people. The cluster must also assist more than a hundred thousand displaced people due to conflict. The main goal of the Food Security and Agriculture (FSAC) cluster is to support the severely food insecure population living with less than 1500 kilocalories/day. Some of the causes of food insecurity in the country are poverty, displacement due to conflict or natural disasters, unemployment and food price rises. The Health cluster is critical, since the intensified conflict has reduced the ability of rural populations to access health services due to checkpoints, military action and lack of transport. In some cases, health facilities have been damaged as a result of fighting. There are on average only 3 health workers per 10,000 habitants which is substantially below a minimum standard of 22. There are alarming levels of malnutrition and diseases across Afghanistan. The Nutrition cluster has to deal with Severe Acute Malnutrition (SAM) problems across all the country. There is a close association of insecurity and high prevalence of malnutrition. Poor access to safe water is also correlated with malnutrition. The Water/Sanitation (WASH) cluster have to deal with poor hygiene habits and high levels of diarrhoea. The data received by the cluster might be impacted by a sampling bias. For example, diarrhoea data is based on the number of people treated at health clinics, but clinics are not distributed proportionally to population density.

Appendix B

List of indicators in the ONVI table (by cluster):

Cluster	Indicator description	Cluster	Indicator description
ESNFI	Unmet natural disaster caseload, number of people needing life-saving assisstance	Nutrition	% of under-5 children with global acute malnutrition (GAM)
ESNFI	Integrated natural disaster analysis score (WFP)	Nutrition	Percentage of people below 1500 kilocalories per day
FSAC	Food consumption score (composite indicator)	WASH	Poor access to safe water, percentage without access to an improved source
FSAC	Household hunger scale (composite indicator)	WASH	Poor hygiene practices, percentage of households without soap available
Health	Deaths per 1000 births	Insecurity	Number of civilian casualties killed by conflict
Health	Acute diarrhoeal disease, cases per 1000 people (previous three year average)	Insecurity	Number of civilian casualties injured by conflict
Health	Measles prevalence, cases per 1000 people (previous 3 years average)	Insecurity	Number of civilian casualties killed by mines/UXO
Health	Measles number of outbreaks (previous 3 years average)	Insecurity	Number of civilian casualties injured by mines/UXO
Health	Acute respiratory infections (ARI), cases per 1000 people (previous 3 year average)	Insecurity	Number of insecurity incidents (3 year average)
Health	Vaccination coverage deficit, percentage of people without coverage for Penta3 (valid; before 1 year)	Insecurity	Number of people living within 500m of mine/UXO hazards
Health	Deliveries without a Skilled Birth Attendant, percentage deliveries without an SBA	Insecurity	Total people displaced (IDPs) in the last 3 years
Nutrition	% of under-5 children with severe acute malnutrition (SAM)	Insecurity	Total people displaced (IDPs) in the last year

Table 4. List of indicators in the ONVI table (by cluster).

Appendix C

The following table shows the cluster overall indicators per province, each calculated as a weighted average of the indicators related to the cluster provided in the ONVI table (weights were provided in the table). These indicators vary from 0 to 5, being 5 the most critical value.

Province	ESNFI	FSAC	HEALTH	NUTRITION	WASH
Badakhshan	2.5	4.5	4	4.7	3
Badghis	1.5	3	3.4	4.7	4
Baghlan	3	3	2.5	2.8	3.5
Balkh	2.5	2.5	3	2.3	3
Bamyan	1.5	3.5	3.3	2.8	3.5
Daykundy	1.5	3.5	3.4	2.3	4
Farah	2.5	2	2.5	2.7	3
Faryab	3	1.5	2.4	2.7	3
Ghazni	2	1.5	2.6	4.7	3.5
Ghor	2	3.5	3.4	2.3	4.5
Helmand	2	1.5	3	3.5	3
Herat	2	1.5	1.8	2.3	3
Jawzjan	4	1.5	3.3	2.8	3
Kabul	1.5	2	1.9	3.2	1.5
Kandahar	1.5	2	2.3	3.5	2.5
Kapisa	1.5	1	3.3	2.8	3
Khost	1.5	1	2.5	4.3	2
Kunarha	2	2	4.1	3.7	3.5
Kunduz	2	2.5	2.5	3.2	3.5
Laghman	2	3	3.9	4.7	2.5
Logar	1.5	2	3.5	2.3	3.5
Nangarhar	1.5	3	3.4	3.7	1.5
Nimroz	2	3.5	2.9	3.3	4
Nooristan	2	3.5	4.1	4.7	5
Paktika	2	1.5	3.1	3.3	4
Paktya	1.5	1	2.9	3.7	2
Panjsher	2.5	1	3	2.8	2.5
Parwan	1	2.5	2.4	2.8	2
Samangan	2.5	2	3.5	4.3	3.5
Sar-e-pul	2	2	4	3.5	3
Takhar	3	3	3.5	4.2	2
Urozgan	1	2	3	4.7	4
Wardak	1.5	2.5	3	3.7	2.5
Zabul	1	3	2.4	4	2

Table 5. Cluster overall indicators per province.

Appendix D

The following are the scaled Demand Table and scaled Supply Table (each column sums 100).

Province	ESNFI	FSAC	HEALTH	NUTRITION	WASH
Badakhshan	4.5	7	5.1	4.9	3.9
Badghis	1.4	2.4	2.2	2.6	2.7
Baghlan	5.2	4.5	3	2.8	4.3
Balkh	6.3	5.4	5.3	3.4	5.4
Bamyan	1.3	2.6	2	1.4	2.1
Daykundy	1.2	2.4	1.9	1.1	2.3
Farah	2.4	1.7	1.7	1.5	2.1
Faryab	5.7	2.5	3.2	3	4.1
Ghazni	4.7	3	4.3	6.4	5.8
Ghor	2.6	4	3.1	1.8	4.2
Helmand	3.5	2.3	3.7	3.6	3.8
Herat	7.1	4.6	4.5	4.8	7.7
Jawzjan	4.1	1.3	2.4	1.7	2.2
Kabul	12.2	14.1	10.9	15.2	8.8
Kandahar	3.5	4	3.7	4.7	4.1
Kapisa	1.3	0.7	1.9	1.4	1.8
Khost	1.6	0.9	1.9	2.7	1.6
Kunarha	1.7	1.5	2.5	1.8	2.1
Kunduz	3.8	4.1	3.4	3.6	4.8
Laghman	1.7	2.2	2.3	2.3	1.5
Logar	1.1	1.3	1.8	1	1.9
Nangarhar	4.3	7.5	6.9	6.2	3.1
Nimroz	0.6	0.9	0.6	0.6	0.9
Nooristan	0.6	0.9	0.8	0.8	1
Paktika	1.6	1.1	1.8	1.6	2.4
Paktya	1.6	0.9	2.1	2.3	1.5
Panjsher	0.7	0.3	0.6	0.5	0.5
Parwan	1.3	2.7	2.1	2.1	1.8
Samangan	1.8	1.3	1.8	1.8	1.8
Sar-e-pul	2.1	1.8	3	2.2	2.3
Takhar	5.6	4.8	4.6	4.6	2.7
Urozgan	0.7	1.3	1.5	2	2.1
Wardak	1.7	2.4	2.4	2.4	2
Zabul	0.6	1.5	1	1.3	0.8

Table 6. Scaled demand table.

Province	ESNFI	FSAC	HEALTH	NUTRITION	WASH
Badakhshan	1.3	8.3	9.3	8.1	8.1
Badghis	0.8	1.1	2.1	1.5	2.8
Baghlan	0.3	5.5	2.4	2.2	2
Balkh	1.9	6.2	3.4	3.3	10.6
Bamyan	0.3	0.3	0.7	3.1	0
Daykundy	0.2	0	0.3	5.9	0
Farah	0.5	0	1.5	0.2	0.8
Faryab	5.2	1.8	3.4	3.1	6.1
Ghazni	6.3	1.6	5.3	8.4	0
Ghor	0.6	0.9	2.2	2.9	2.4
Helmand	0.6	1.8	3.4	2	0
Herat	2.7	1.1	5	3.7	4.1
Jawzjan	1.4	1.7	2.1	2	2.4
Kabul	8	10.2	7.7	2	5.3
Kandahar	1.4	2.2	3.4	5.3	0.8
Kapisa	6.6	1.3	1.4	0	0
Khost	6	1.8	4.1	7.3	6.1
Kunarha	1.4	7.8	4.1	2.8	2
Kunduz	0.3	0.9	2.6	1.7	5.3
Laghman	3.1	3	1.2	1.1	2
Logar	5.3	2.9	1.7	0.7	2.8
Nangarhar	16.3	9.6	6.2	4.8	16.7
Nimroz	0.2	0.7	1.2	1.3	0
Nooristan	0.2	3.6	1.5	1.7	0
Paktika	4.5	2	2.6	4.6	0.8
Paktya	4.4	2.6	3.3	1.8	2
Panjsher	2.2	1.6	2.6	2	1.6
Parwan	5.8	2.6	4	2.2	2
Samangan	0.2	1.7	1.4	1.3	2.4
Sar-e-pul	0.8	1.6	1.2	2.6	0.8
Takhar	0.8	7	2.9	5.3	5.7
Urozgan	1.3	1.3	2.2	2.9	2
Wardak	9.1	4	1.5	0.2	2
Zabul	0.3	1.2	1.9	2	0

Table 7. Scaled supply table.

Appendix E

List of indicators per topic in the World Bank dataset described in Section 3:

Topic	Indicator description	Topic	Indicator description
Population	Percentage of the total population that lives in rural areas	Employment	Percentage of the labor force that is engaged in manufacturing and processing including handicrafts, food processing, tailoring, etc.
Population	Percentage of the total population that is female	Employment	Percentage of the labor force that is engaged in construction (e.g. roads, buildings, etc.)
Population	Total population	Employment	Percentage of the labor force that is engaged in transport and communication, retail and wholesale trade and other services
Population	Ratio of young dependents (population younger than 14) to the working age population (14 and above), expressed in percentage	Employment	Percentage of the labor force that is engaged in public sector jobs, including public services (health and education), UN/NGOs and other government services
Poverty	Percentage of the population whose expenditure on food and non-food items falls below the official poverty line	Education	Percentage of the total population aged 14 years and older that is literate, i.e. able to read and write
Poverty	Average shortfall of per capita consumption from the poverty line expressed as percentage of the poverty line. In calculating the depth of poverty, the shortfall of non-poor is treated as zero. This in dicator is also known as "poverty gap"	Education	Average years of schooling of the population aged 18 years old and above
Poverty	Average consumption of individuals consuming less than the poverty line (poor), expressed as a percentage of the poverty line	Education	Percentage of the population in primary school age [7,12] who are currently attending primary school (grade 1-6). The construction of this indicator does not include Islamic school
Poverty	Expenditure on food and non-food items (including durable goods and housing) consumed by households in a month, divided by household size. The per capita monthly consumption is expressed in Afgs at the 2011 prices	Education	Percentage of the population in secondary school age [13,18] who are currently attending secondary school (grade 7-12). The construction of this indicator does not include Islamic school
Poverty	Inequality measure that captures the deviation of the distribution of per capita consumption from a perfectly equal one. The value of the Gini index expressed in percentage ranges from 0 to 100, with 0 representing perfect equality and 100 complete inequality	Gender	Percentage of the female population aged 14 years and older that is literate, i.e. able to read and write
Food security	Percentage of the total population whose daily per capita intake is less than the minimum threshold of 2100 calories. No adjustment for age or gender is used in the estimation of the threshold	Gender	Ratio of girls' net attendance rate in primary school to boys' net attendance ratio in primary school, expressed in percentage
Food security	Percentage of the total population whose daily per capita intake is less than 1500 calories. No adjustment for age or gender is used in the estimation of the threshold	Gender	Ratio of girls' net attendance rate in secondary school to boys' net attendance ratio in secondary school, expressed in percentage
Food security	Percentage of the total population whose daily consumption of protein is less than 50 grams per day	Gender	Percentage of women aged 14 and over who, during the reference period of one week prior to the survey date, were economically active (employed/underemployed or unemployed)
Food security	Percentage of population whose food consumption scores are 42 and below. This measure aggregates the population that falls under the food consumption categories labeled 'border line' and 'poor'	Health	Percentage of women 49 years old and younger, ever married, who during their last pregnancy report at least one visit to a skilled provider (doctor, midwife, or nurse)
Labor market	Percentage of all persons aged 14 and over who, during the reference period of one week prior to the survey date, were economically active (employed/underemployed or unemployed)	Health	Percentage of women 49 years old and younger, ever married, who during the delivery of their last child were assissted by a skilled birth attendant (doctor, midwife, or nurse)
Labor market	Percentage of all persons aged 14 and over who, during the reference period of one week prior to the survey data, were in paid employment or self-employed and who worked at least eight hours	Health	Percentage of children aged 12 to 36 months that have received complete vaccines for BCG, DPT3, OPV3, and measles
Labor market	Percentage of the labor force (economically active individuals aged 14 and above) who, during the reference period of one week prior to the survey date, were either without a job and seeking for one, or working for less than 8 hours	Health	Percentage of children aged 12 to 36 months that have NOT received any of the scheduled vaccines for BCG, DPT3, OPV3, and measles
Labor market	Percentage of employed individuals aged 14 and over who, during the reference period of one week prior to the survey date, were working less than 40 hours, and reporting willingness and availability to work additional hours	Services	Percentage of the population whose main source of drinking water is one of the following: hand pump (in-compound or public); bored wells (hand pump or motorized); protected spring; pipe scheme (gravity or motorized); and piped water provided by the municipality
Labor market	Percentage of the labor force (economically active individuals aged 14 and above) that is literate	Services	Percentage of the population who has access to improved sanitation facility (a flush latrine or any other improved latrine)
Employment	Percentage of the labor force that is engaged in agriculture and livestock related activities	Services	Percentage of the population who has access to electricity from any of the providers (government, community or private) and sources (electricity grid, government generator, personal generator, community generator, solar, wind or battery)

Table 8. List of indicators per topic in the World Bank dataset described in Section 3.

Appendix F

The following table is an excerpt of 25 random rows (projects) from the 2669 rows (projects) in the 3W dataset.

PROVINCE	DISTRICT	ORG_ACRONYM	ORG_TYPE	CLUSTER
Kabul	Bagrami	ACF	International NGO	NUTRITION
Kabul	Chaharasyab	ARCS	ICRC	ESNFI
Kabul	Dehsabz	ACTED	International NGO	FSAC
Kabul	Estalef	ARCS	ICRC	ESNFI
Kabul	Kabul	HELVETAS	International NGO	FSAC
Panjsher	Onaba (Anawa)	ARCS	ICRC	HEALTH
Panjsher	Shutul	ARCS	ICRC	HEALTH
Parwan	Saydkhel	AREA	National NGO	FSAC
Parwan	Saydkhel	CARE	International NGO	ESNFI
Daykundi	Kajran	MOVE	International NGO	NUTRITION
Kunar	Barkunar	MADERA	International NGO	FSAC
Kunar	Barkunar	PU-AMI	International NGO	FSAC
Kunar	Khaskunar	AGHCO	National NGO	ESNFI
Laghman	Alingar	SCA	International NGO	NUTRITION
Nangarhar	Achin	SCI	International NGO	ESNFI
Nangarhar	Kot	MADERA	International NGO	FSAC
Nangarhar	Kuzkunar	APA	National NGO	ESNFI
Nangarhar	Kuzkunar	APA	National NGO	FSAC
Nuristan	Barg-e-Matal	IMC	International NGO	NUTRITION
Nuristan	Duab	APA	National NGO	FSAC
Ghazni	Waghaz	AADA	National NGO	HEALTH
Ghazni	Waghaz	AADA	National NGO	NUTRITION
Hirat	Hirat	UNICEF	United Nations	WASH
Hirat	Hirat	WFP	United Nations	FSAC
Hirat	Zindajan	HI	International NGO	HEALTH

Table 9. Excerpt of 25 random rows (projects) from the 2669 rows (projects) in the 3W dataset.

Appendix G

List of 149 organizations names and acronyms present in the 3W dataset described in Section 3:

ACRONYM	NAME	ACRONYM	NAME
AADA	Agency for Assistance and Development of Afghanistan	BDN	Bakhtar Development Network
AAID	Afghan Agency for Integrated Development	BRAC	Bangladesh Rural Advancement Committee Afghanistan
ABM	Afghan Blind Management	BVWO	Badakhshan Volunteers Women Organization
ACF	Action contre la Faim	CAF	Care of Afghan Families
ACTD	Afghanistan Center for Training and Development	CARD-F	Comprehencsive Agriculture Rural Developoment Facility
ACTED	Agency For Technical Cooperation & Development	CARE	Care International in Afghanistan
ActionAid	ActionAid	CARITAS-G	Caritas Germany
ADA	Afghan Development Association	CHA	Coordination of Humanitarian Assistance
ADRA	Adventist Development and Relief Agency	ChristianAid	Christian Aid
Afghanaid	Afghan Aid	CoAR	Coordination of Afghan Relief
AGHCO	Afghan General Help Coordination Office	CONCERN	CONCERN
AHDS	Afghan Health & Development Services	CRDSA	Coordination Of Rehabilitation Services For Afghanistan
AHEAD	Assitant for Health Education and Development	CRS	Catholic Relief Services
AIL	Afghan Institute of Learning	CWS	Church World Service-Pakistan/Afghanistan
AKF	Aga Khan Foundation	DAC	Danish Afghanistan Committee
AKHS	Aga Khan Health Service	DACAAR	Danish Committee for Aid to Afghan Refugees
ALO	Afghan Literacy Organization	DCA	Dutch Committee for Afghanistan
ANCC	Afghanistan National re-construction Co-ordination	DELTA	Development Education Learning & Training Association
APA	Afghan Planning Agency	DHSA	Development and Humanitarian Services for Afghanistan
APWO	Afghan Public Welfare Organization	DRC	Danish Refugee Council
ARAA	Ansari Rehabilitation Association for Afghanistan	DWHH/GAA	Deutsche Welthungerhilfe e. V. (German Agro Action)
ARCS	Afghan Red Crescent Society	EMERGENCY	Emergency Italian
AREA	Agency for Rehabilitation and Energy Conservation in Afghanistan	FAO	Food and Agriculture Organization
ARPD	Afghanistan Reconstruction And Planning Department	FGA	Future Generations Afghanistan
ASRi	Afghanistan Strategic Resolution	FOCUS	Focus Humanitarian Assistance
ASRn	Afghanistan Strategic Resourse	HAAWC	Humanitarian Assistance for Afghan Women and Children
AURC	Afghan United Rehabilitation Counsel	HADAF	Humanitarian Assistance for Development of Afghanistan
AVDA	Afghan Volunteer Doctors Association	HAPA	Humanitarian Action for People of Afghanistan
AWAKEN	Afghan Women and Kids Education and Necessities	HELP	Hilfe zur Selbsthilfe e.V.
AWARO	Afghanistan Water and Agriculture Research Organization	HELVETAS	Helvetas, Swiss Association for International Cooperation
AWRC	Afghan Women's Resource Center	HI	Handicap International
BARAN	Bu Ali Rehabilitation and Aid Network	HIA	Hungarian InterChurch Aid - Hungary

Table 10. Part 1 of the list of 149 organizations names and acronyms present in the 3W dataset described in Section 3.

ACRONYM	NAME	ACRONYM	NAME
HNI-TPO	Health Net International - Trans cultural Psychosocial Organization	NFUAJ	National Federation of UNISCO Association in Japan
HRDA	Human Resource Development Agency	NPO/RRAA	Rural Rehabilitation Association for Afghanistan
IAM	International Assistance Mission	NRC	Norwegian Refugee Council
ICARDA	International Center for Agricultural Research in the Dry Areas	OBA	Organization of Basic Advancement
ICRC	International Committee of the Red Cross	ODCG	Development and Care Group
IFRC	International Federation of Red Cross and Red Crescent Societies	OHW	Organization of Human Welfare
IMC	International Medical Corps	ORCD	Organization for Research and Community Development
INTERSOS	Intersos - Humanitarian Organization for Emergency	ORD	Organization for Relief and Development
IOM	International Organization for Migration	OXFAM	OXFAM
IR	Islamic Relief	PIN	People in Need
IRA	Islamic Relief for Aghanistan	PIPA	People in Peril Association
IRC	International Rescue Committee	PMS	Peshawar-kai Medical Service
JDAI	Joint Development Associates International	PRB	Partners in Revitalization and Building
JEN	Japan Emergency NGO	PU-AMI	Première Urgence-Aide Medicale Internationale
Johanniter	Johanniter International	RAADA	Rehabilitation Association and Agriculture Development for Afg.
JVC	Japan International Volunteer Center	RCDC	Rural Capacity Development Committee
MADERA	Mission d' Aide au Developpement des Economies Rurales en Afg.	RF	Rupani Foundation
MarieStopes	Marie Stopes International	RoP	Roots of Peace
MCA	Mehtherlam Coordination Assistance	RORA	Relief Organization for Rahabilitation of Afghanistan
MCDO	Marwa Cultural and Development Organization	RRDPO	Rural Rehabilitation and Development Organization
ME	Mission East	RSA	Reconstruction Service for Afghanistan
MEDAIR	MEDAIR	SAF	Solidarity for Afghan Families
MEDIOTHEK	Mediothek Afghanistan	SCA	Swedish Committee for Afghanistan
MercyMalaysi	Mercy Malaysia	SCI	Save the Children Federation International
MOVE	MOVE Welfare Organization	SDO	Sanayee Development Organization
MRCA	Medical Refresher Courses for Afghans	SERVE	Serving Emergency Relief and Vocational Enterprises
MSF	Medecins Sans Frontieres (Doctors Without Borders)	SFL	SFL International
MSI	Management Systems International, Inc.	SHA	Swiss Humanitarian Aid Unit
NAC	Norwegian Afghanistan Committee	SHPOUL	Sound Humanitarian Participatory and Organization Uplift
NCA	Norwegian Church Aid	SHRDO	Serve Health Relief and Development Organization
NCRO	New Constancy and Relief Organization	SI	Solidarites International
NEI	Nutrition & Education International	SIKA	Stability In Key Areas

Table 11. Part 2 of the list of 149 organizations names and acronyms present in the 3W dataset described in Section 3.

ACRONYM	NAME	ACRONYM	NAME
SOFAR	Salam Organization for Afghanistan Rehabilitation	UNMACCA	Mine Action Coordination Centre for Afghanistan
Solidarites	Solidarites	VARA	Voluntary Association for Rehabilitation of Afghanistan
STARS	Skills Training And Rehabilitation Society	WADAN	Welfare Association for Development of Afghanistan
TDH	Terre Des Hommes	WC-UK	War Child UK
Tearfund	Tearfund	WFP	World Food Program
TIKA	Turkish International Cooperation and Development Agency	WHO	World Health Organization
TLO	The Liaison Office	WoW	Women for Women
TM	Tamir-e-Millat	WSTA	Watan Social and Technical Services Association
UNHABITAT	United Nations Human Settlements Programme	WVI	World Vision International
UNHCR	United Nations High Commissioner for Refugees	ZOA	ZOA Refugee Care
UNICEF	United Nations Children's Fund		

Table 12. Part 3 of the list of 149 organizations names and acronyms present in the 3W dataset described in Section 3.

Appendix H

Organization Reachability (OR) top 10 best and top 10 worst organizations.

ACRONYM	NAME	ORG_TYPE	OR	RANK
IRA	Islamic Relief for Aghanistan	International NGO	13.20	1
BRAC	Bangladesh Rural Advancement Committee Afghanistan	International NGO	8.17	2
NEI	Nutrition & Education International	International NGO	6.93	3
TDH	Terre Des Hommes	International NGO	4.60	4
HELP	Hilfe zur Selbsthilfe e.V.	International NGO	4.47	5
IFRC	Int. Fed. of Red Cross and Red Crescent Societies	ICRC	4.47	6
AWRC	Afghan Women's Resource Center	National NGO	3.95	7
CARD-F	Comprehensive Agriculture Rural Developoment Facility	National NGO	3.95	8
CARITAS-G	Caritas Germany	International NGO	3.95	9
NFUAJ	National Federation of UNESCO Association in Japan	International NGO	3.95	10
WSTA	Watan Social and Technical Services Association	National NGO	-4.30	140
NPO/RRAA	Rural Rehabilitation Association for Afghanistan	International NGO	-4.30	141
APWO	Afghan Public Welfare Organization	National NGO	-4.32	142
TLO	The Liaison Office	National NGO	-4.43	143
AAID	Afghan Agency for Integrated Development	National NGO	-4.51	144
SI	Solidarites International	International NGO	-4.54	145
APA	Afghan Planning Agency	National NGO	-5.45	146
AGHCO	Afghan General Help Coordination Office	National NGO	-6.59	147
FGA	Future Generations Afghanistan	National NGO	-7.77	148
UNHABITAT	United Nations Human Settlements Programme	United Nations	-12.00	149

Table 13. Organization Reachability (OR) top 10 best and top 10 worst organizations.

Appendix I

The following are the pseudo-codes for the first and second phases of the optimization algorithm, respectively.

```
create an empty vector of projects and call it POOL;
# Choose which projects to put temporarily in the POOL
for every cluster do
    calculate surplus per province;
    for every province with a surplus greater than 0 do
        calculate surplus per district;
        for every district with surplus greater than zero in descending order do
            order organizations present in the given district and cluster by degree in ascending order;
            put as many projects from the organizations with the lowest degree temporarily in the POOL, without making
            the current district or current province be in deficit (negative surplus);
        end
    end
end
# Translate chosen projects from the POOL to new districts
for every cluster do
    calculate surplus of projects per province;
    for every province with a surplus smaller than 0 in ascending order do
        calculate surplus per district;
        for every district with surplus smaller than zero in ascending order do
            assign to the district as many projects (of the given cluster) from the POOL, without making the district or the
            province to have surplus in that cluster, and trying to assign the organizations from the POOL that have
            partnerships with the organizations already present in the current district;
        end
    end
end
```

Algorithm 1: Algorithm for the optimization of HPI.

```
rank each pair of neighbouring provinces in ascending order according to their distance metric, and in descending order
according to their relevance metric;
order the pairs of neighbouring provinces according to their differences in rankings (relevance ranking position minus
distance ranking position);
while projects available to be swapped do
    find the pair of neighbouring provinces (A and B) with the most positive difference in rankings;
    find the set of all provinces that are neighbours of A or B (except A or B) and call it Neighbour–Set;
    for every cluster do
        find a set of all organizations (working in the given cluster) present in province A and the Neighbour-Set at the
        same time and call it Set-A; do the same for province B and call it Set-B;
        make province A to swap (if possible) projects from the organizations with the lowest degree possible present in
        its districts by other projects from the provinces of the Neighbour-Set whose organizations are in Set-B;
        make province B to swap (if possible) projects from the organizations with the lowest degree possible present in
        its districts by other projects from the provinces of the Neighbour-Set whose organizations are in Set-A;
        flag each organization project swapped, and do not allow it to swap any more;
    end
end
```

Algorithm 2: Algorithm for the optimization of RCI.