

Disaster Recovery Centers: Catchment Analysis

Technical Report

Federal Emergency Management Agency
Recovery Directorate

DHS STEM Internship Program
Summer 2015

Heidi Hurst

heidi.hurst@associates.fema.dhs.gov

hurst@college.harvard.edu

misshurst@gmail.com

Table of Contents

Abstract.....	3
Intended Audience	3
Introduction	4
Research Project.....	4
Research Questions	4
Methodology	5
Setup & Execution	5
Data Sources & Inputs.....	5
Step 1: Catchment Areas.....	7
Step 2: Social Vulnerability	8
Step 3: Population Analysis	9
Step 4: Visualization	11
Case Study: DR-4223-TX	12
Caveats & Weaknesses	16
Applications.....	17
Next Steps	17
Skills & Knowledge Gained	17
Personal & Professional Impact	18
Conclusion	18
Acknowledgements	19

Abstract

This paper considers the application of geospatial analysis to DRC location decisions and explains a baseline methodology for analyzing populations served by one or more DRCs. This research supports Action Item #2 from FEMAStat Memorandum dated December 30, 2013 requiring that ORR-Recovery “articulate the data that FEMA will measure to assess whether DRCs are serving their intended purpose”¹.

The primary result of this project is a script written in Python 2.7 for use in ArcMap Desktop 10.2.2. This code takes inputs of DRC locations, impacted counties, and external data and provides outputs in the form of shapefiles and Excel spreadsheets, suitable for use by GIS specialists, analysts, and decision makers.

This research has the potential to help drive decisions about where to allocate DRCs. It could also support the operations of individual DRCs and enable data-driven collaboration between the DRC workgroup and the DSA Teams.

Spatial analysis – as this document hopefully makes clear – is about more than just maps. It is about generating useful insights, exploring data, facilitating decisions, and increasing FEMA’s ability to support survivors.

Intended Audience

This paper is intended for a technical audience with some background in disaster recovery operations. It is designed to accompany the source code for this project and provide detailed steps to recreating a similar analysis.

For those less interested in the technical details, please see the included Case Study for DR-4223-TX and the DRC Catchment Presentation.

¹ Response to FEMAStat Memorandum dated December 30, 2013 (Action Item #2)

Introduction

Disaster Recovery Centers (DRCs) provide important services for individuals in impacted counties. For those who survive tornados, hurricanes, floods, or other disasters, these pop-up shops provide information about FEMA programs, access to resources from volunteer organizations and other agencies, and a source of hope. FEMA opens anywhere between 200 and 600 DRCs per year, depending on the number of Individual Assistance disaster declarations². From 2008 to 2013, 1,674 DRCs were opened, providing service to 1,050,629 visitors³.

Location choice is vital in ensuring that DRCs are accessible to the target population and meet the goals of the coordinated disaster response. Infrastructure damage or lack of access to transportation may prevent survivors from reaching facilities within a reasonable amount of time, hindering effective recovery. Geographic information about local road networks and the distribution of vulnerable populations must be taken into account when selecting location choice to maximize both DRC impact and overall recovery potential.

Despite the clear importance of these geographic considerations, comparatively few resources have been dedicated to developing a rigorous methodology for determining DRC location. This research creates a baseline methodology for analyzing populations served by one or more DRCs. This spatial analysis is an important first step towards developing a rigorous, data-driven process for selecting DRC locations.

Research Project

Throughout the course of the summer, I developed a methodology to evaluate the relationship between social vulnerability and DRC location. This research supports Action Item #2 from FEMASat Memorandum dated December 30, 2013 requiring that ORR-Recovery “articulate the data that FEMA will measure to assess whether DRCs are serving their intended purpose”⁴.

The primary result of this project is a script written in Python 2.7 for use in ArcMap Desktop 10.2.2. This code takes inputs of DRC locations, impacted counties, and external data and provides outputs in the form of shapefiles and Excel spreadsheets, suitable for use by GIS specialists, analysts, and decision makers.

Research Questions

This project addresses the following questions:

How can FEMA leverage spatial analysis to evaluate DRC locations? Specifically, a) how many people can reach a DRC and b) what are those people like?

² Response to FEMASat Memorandum dated December 30, 2013 (Action Item #1)

³ Response to FEMASat Memorandum dated December 30, 2013 (Action Item #1)

⁴ Response to FEMASat Memorandum dated December 30, 2013 (Action Item #2)

These questions are examined on their own as well as through case studies of socially vulnerable populations in Individual Assistance designated counties for Texas DR-4223.

Methodology

This project was scripted in Python 2.7 for use with ArcMap Desktop 10.2.2. The methodology section will walk through each component of the analysis.

The methodology can be broken into four steps:

- 1) Determine catchment areas of DRCs (also called service areas or drive time areas)
- 2) Categorize population characteristics (here, SVI is used to categorize populations as high, medium, or low vulnerability)
- 3) Analyze populations served by DRCs
- 4) Visualize results

Although many of these can be completed without using Python, the first three steps can easily be combined into one cohesive code. This allows for more efficient analysis as well as standardization of products, allowing any analyst to rapidly run a rigorous analysis.

Setup & Execution

This code is written to operate on a specific ArcMap created following the standard formatting of other GIS products within the Agency; however, any map with an activated data frame named “Event Location” can be used. If this data frame is not activated and other frames are present, this analysis may not execute as expected and intermediate layers may be created in other data frames.

This script can be executed from an open map as described above by loading the entire code into the Python window of an open map or by “copy and paste” of one segment at a time. For debugging purposes, it is often most convenient to execute the code in the predetermined chunks to identify bugs and glitches. It may be easiest to modify the code in IDLE or a similar code editor before importing into ArcMap.

All outputs of this analysis will be stored in a new GeoDatabase created in the “Input Validation” source of the code. This GeoDatabase is named “DR_XXXX_DRC.gdb” by default, where XXXX corresponds to the Declaration number. If the code is re-run, the old database must either be emptied or renamed. An improvement to this code would allow for overwrite privileges or create dynamic naming for the GeoDatabase.

Data Sources & Inputs

This code is written to take, as inputs, the following:

- a) a current disaster shapefile indicating the designation of each county

These shapefiles can be found in the [\\hqmac3f1\products\2015\Disaster\DRs](#) folder. It is assumed that shapefiles have a column titled “Designate” that specifies the county

designation; this analysis queries for “IA” or “IAPA” entries. Due to restructuring of product folders, this link may change.

If running an analysis without a declaration (e.g. pre-declaration or scenario), create a shapefile of selected counties within the GDB and assign the variable “IA_counties” to the name of the layer.

- b) a link to state Social Vulnerability Index data on the tract level

This analysis relies on data created and formatted by the CDC and downloaded from <http://svi.cdc.gov/>. It is formatted as an MDB file with geometry, allowing for direct rasterization without requiring a join to extant tract geometry. This analysis uses the “R_PL_THEMES” composite percentile ranking. In addition, the user must manually specify the breaks for raster categorization.

This code could be modified to utilize other data; some preliminary, inoperable structure for using census data has been included.

- c) DRC locations from FEMA’s REST service

This part of the code is the most likely to cause an error, as pulling data from REST services has proved tricky and unreliable. The code generates a URL with an embedded PJSON query to access the records. This PJSON table is then loaded into an empty Feature Set and copied to a new layer. This operation may fail for a number of reasons; the most frequent errors include empty query returns and server inaccessibility.

To ensure that the query returns values, manually paste the generated URL (“fsURL”) into a browser to confirm the url. The FEMA REST service can also be accessed online through <http://gis.fema.gov/REST/services/FEMA/DRC/FeatureServer/> and queried manually.

If data from the REST service are unavailable, data can be obtained through the DRC manager, news reports, or other internal or external resources. For more information on DRC locations, the DRC workgroup (Steve Michaels, Free Palmer, and others) is invaluable.

In addition, this process utilizes the following static data:

- d) Road network data (from the Homeland Security Infrastructure Program; found at [\\hqmac3f1\Static\GISdata\HSIP_DATA](#))
- e) Landscan data (conus-night; found at [\\hqmac3f1\Static\GISdata\Landscan](#))

Note that this Landscan dataset is restricted to the Continental US. Night estimates were utilized based on the unverified assumption that individuals are most likely to visit DRC locations close to their homes.

Other types of inputs (such as census data, ONA or IHP grant allocation, or registrations) can be utilized following the basic premises outlined in this document with some modifications to the code.

Step 1: Catchment Areas

The first step of this analysis is determining the catchment areas (also called service areas or drive time catchments) of DRC facilities. This process takes actual road networks and determines areas of maximum drive time to a given facility. For example, evaluating a 20 minute catchment area around a given facility would yield a polygon covering all the locations from which one could reach that facility in 20 minutes or less. Evaluating a number of different drive times (e.g. 20, 40, 60 minutes) provides a nuanced picture of population distribution.

These catchment areas are computed with ArcMap Network Analyst extension using HSIP Road Network data. By providing a link to the dataset within the code, the entire dataset was not loaded into the map document, thus reducing processing time. Any number of drive times may be used, although using more than three drive times in a single analysis is not recommended as it will increase processing time dramatically.

Areas can be computed using historical, current, planned, or hypothetical DRC locations; this code is set to execute with user specified types of DRCs for a specific declaration (e.g. all active DRCs, etc). This analysis uses average traffic time as an impedance, which may not accurately reflect traffic in post-disaster scenarios. Note that there are no excluded sources (e.g. dirt roads, carpool lanes), as excluding sources may result in unlocated facilities, particularly in rural areas.

Crucially, damaged infrastructure can be included as “impedances”, allowing the analysis to account for driving delays due to collapsed bridges, inundated roads, or other “off limits” roads. By using updated information as it becomes available, both analysts and decision makers can understand what areas have access to DRC facilities. This code does not include explicit support for including infrastructure damage; this would need to be added before the Solve function is executed near the end of the Network Analysis section.

These catchments are computed as polygons (rings). The DRC ID is added as the field “FacilityID” to ensure a valid join field for later analysis. Note that this assumes uniqueness of DRC ID number; if this assumption is invalid the remainder of the analysis may not execute as expected.

In the section “Service Area Polygons Rasters”, two rasters are created from the “Service Area/Polygons” layer. The first raster, “SA_PolygonBreak”, encodes the catchment time in minutes (e.g. 20, 40, 60) for each polygon. Importantly, it sorts with the priority field “PRIORITY”, equal to the value of the drive time over 100. If the drive time (“ToBreak”) exceeds 100, this expression may need to be modified to ensure that the smallest catchments are accurately converted.

The second raster, “SA_PolygonID”, encodes the DRC ID (“FacilityID”) for each facility. These two rasters are used to calculate into a third, “SA_PolygonToRaster”, that encodes both “FacilityID” and “ToBreak” values as follows:

$$\text{SA_PolygonToRaster} = \text{SA_PolygonBreak} + \text{SA_PolygonID} * 100$$

This resultant value is unique for each combination of ID and break time; thus the value 2085440 corresponds to a 40 minute catchment area (last two digits) for DRC ID 20854, Leonel Castillo Community Center in Houston, TX. (Again, this assumes drive time values less than 100.)

Finally, another raster calculation fills in the empty space in impacted counties with the value 9999, ensuring that the final service area raster layer (“SA_Raster”) covers the entire impacted area. The intermediate raster layers (“SA_PolygonBreak”, “SA_PolygonID”, and “SA_PolygonToRaster”) are deleted at the end of the code in the section entitled “Tidy Map Presentation”.

Step 2: Social Vulnerability

The second step of this analysis is importing and classifying additional data about the study area. These additional data could take a number of forms – disaster severity, targeted demographic information from the American Community Survey (ACS), mortgage rates, etc. This analysis utilizes the Social Vulnerability Index (SVI) compiled by the Center for Disease Control as one way of characterizing populations.

The SVI determines vulnerability based on 14 key Census variables grouped into four categories: Socioeconomic, Household Composition, Minority Status/Language, and Housing/Transportation. This is used to generate percentile ranking for all census tracts within a state, allowing comparisons of relative vulnerability. This index was chosen because of the rigorous academic work done by the CDC⁵ and the University of South Carolina⁶, the granularity of the data, and the availability of complete census tract level data for all states. (As a note, these data are unavailable for territories or Tribal Nations.)

These data are imported as a raster (“Eval_Full”) based on overall theme ranking (“R_PL_THEMES”) to facilitate further analysis. It is clipped to county outlines, creating raster “Eval_Raster”. The layer is then reclassified into three different categories: low vulnerability (below the 50th percentile), medium vulnerability (50th – 75th percentile), and high vulnerability (above the 75th percentile). This produces the final evaluation criterion raster “Eval_categ”, to be used in the next section of the code.

⁵Flanagan, Barry E.; Gregory, Edward W.; Hallisey, Elaine J.; Heitgerd, Janet L.; and Lewis, Brian (2011) "A Social Vulnerability Index for Disaster Management," *Journal of Homeland Security and Emergency Management*: Vol. 8: Iss. 1, Article 3.
<http://svi.cdc.gov/Documents/Data/A%20Social%20Vulnerability%20Index%20for%20Disaster%20Management.pdf>

⁶ See <http://sovius.org/>

The breakpoints for these classification were chosen out of convenience; further research may consider using the relationship between SVI and registration data, for example, to determine more appropriate break points. Break points must be specified by the text variable “census_class” in the “User Inputs” section at the top of the code. Further improvements to this code could consider creating dynamic or automatic breaks.

Step 3: Population Analysis

The third step, *Population Analysis*, constitutes the bulk of this analysis.

The catchment ID raster (“SA_Raster”) is combined with the evaluation criterion raster (“Eval_categ”) using the following formula (see code section “Zonal Classification”):

$$\text{catchment_raster} = \text{Eval_categ} + \text{SA_Raster} * 10$$

This simply appends an integer value (here, 1, 2, or 3) to the existing gridcode value. Thus, extending the previous example the gridcode 20854403 would indicate a high vulnerability area (3) in the 40 minute catchment zone of DRC 20854 (Leonel Castillo Community Center).

This process yield a number of “zones” or unique gridcode integers equal to:

$$(\# \text{ of drive times} + 1) \times (\# \text{ of SVI categories}) \times (\# \text{ of DRCs})$$

Once this information is created as a master raster layer, it is converted back to a polygon (see code section “Zonal Classification”). Note that as a consequence of multiple conversions between raster and polygon data types, the catchment polygon layer may not match precisely with the original county boundaries, SVI breaks, or network analyst service areas. However as the margin of error is typically small this imprecision is worth the added functionality.

The following section, “Add Fields to Catchment Polygon”, simply calculates fields for DRC ID, Drive Time, and Evaluation Criteria based on the gridcode. This makes visualization easier if information is desired in a map format. In order to receive tabular data this section is unnecessary.

Population counts are generated the section “Zonal Statistics Table Output”, which uses Landscan data and the catchment polygons to generate estimates of the number of people in each area. In order to do this, the Zonal Statistics as Table tool is utilized using the catchment polygon layer, the “gridcode” variable, and the “landscan_pop” raster. Only the mean is computed here; other statistics are unnecessary for the scope of this analysis.

A number of other fields are added to this table to ensure ease of analysis and interpretation. Population is estimated by multiplying the “MEAN” value by the “COUNT”. This multiplies the average people per pixel in each “zone” (unique gridcode value) by the total number of pixels, resulting in an estimate of the total zonal population. Note that since polygons do not overlap, there are no double counted individuals.

The sample excerpt from the catchment analysis table, below, shows the 5 important data fields available for each category. “DRC_ID” is the unique identification number associated with the closest facility; “DR_TIME” is the upper limit of the drive time catchment; “EVAL_CRIT” is the numeric equivalent of the population evaluation characteristic (here, SVI); and “TOT_POP” is the estimated number of individuals living in this type of area.

gridcode	DRC_ID	DR_TIME	EVAL_CRIT	TOT_POP
20854403	20854	40	3	196511

This excerpt, for example, should be interpreted as follows:

“An estimated 196,511 people live in high vulnerability areas between 20-40 minutes of DRC 20854.” It is assumed that users will know the lower limit of the drive time; the “DR_TIME” value is the upper limit of the catchment.

Before exporting this table to Excel, it is joined with the DRC data tables by DRC ID. (In order to ensure a seamless join, the DRC layers are combined to prevent duplication of fields.) This provides more information on the location, hours, closing date, etc. of the facility. The full entry for gridcode 20854403 identifies this DRC as Leonel Castillo Community Center in Houston, Texas, used for DR-4223-TX, opened on 2015-06-15.

For areas outside of a DRC catchment, the DRC ID and drive time are both listed as 99.

In addition to computing this information at the individual DRC level, this code utilizes the Numpy scientific computing package to calculate summary statistics. The section “Numpy Manipulations” aggregates data to provide disaster-wide information.

gridcode	DRC_ID	DR_TIME	EVAL_CRIT	TOT_POP
201	0	20	1	901915
202	0	20	2	854743
203	0	20	3	1599449
401	0	40	1	1816396
402	0	40	2	839060
403	0	40	3	647468
601	0	60	1	1479252
602	0	60	2	441990
603	0	60	3	115713
991	99	99	1	1412227
992	99	99	2	915509
993	99	99	3	810778
1	0	0	1	5609790
2	0	0	2	3051302
3	0	0	3	3173408
200	0	20	0	3356107
400	0	40	0	3302924
600	0	60	0	2036955
990	99	99	0	3138514

The summary table (“catchment_summary”) includes information calculated for the entire disaster area with population broken out by drive time, by vulnerability, and by both. The table on the previous page shows catchment summary data for DR-4223-TX based on DRC locations on 7/27/2015.

In this context, DRC ID equal to 0 indicates all DRCs are included; DRC ID or drive time equal to 99 indicates areas outside of DRC catchment zones.

gridcode	DRC_ID	DR_TIME	EVAL_CRIT	TOT_POP
201	0	20	1	901915

The first row of this table, reproduced above, would be interpreted as follows:

“There are an estimated 901,915 people in low vulnerability census tracts within 20 minutes of any DRC for disaster 4223 (TX).”

These two tables (“catchment_analysis” and “catchment_summary”) are exported to Excel for visualization or further analysis.

The final lines of the code (“Tidy Map Presentation”) delete unnecessary layers, such as intermediate rasters. Additional code could be added here to format and export the map, increasing simplicity of use and encouraging automation.

Step 4: Visualization

The final step of the analysis, visualization, is not explicitly coded. It is up to the user to determine what visualization style is most effective for the given decision, task, or question at hand. Furthermore, there is a lot of space for creativity in these visualizations. For example, one could imagine an interactive visualization, perhaps in the JavaScript library D3, that allows users to dynamically select specific locations for which data will be displayed, or to toggle between DRC specific info and disaster wide information.

There are two areas of visualization: spatial and graphical.

Spatial visualization presents the shapefiles generated by this analysis in combination with other spatial data (point DRC locations, impacted counties, background layers) as a map. One challenge in presenting this visualization is encoding multiple variables. This case study used colors for catchment areas and hues for social vulnerability. Future visualizations may consider displaying only one vulnerability ranking and using hues to distinguish between catchment distances, for example.

Graphical visualization presents data detailed in either catchment analysis or catchment summary tables as charts and graphs for easier comprehension. This project relied on pie charts at the recommendation of other analysts in FEMA’s Recovery Office. In addition, this project explored cumulative population counts to show area under the curve graphs. This visualization provided a powerful example of differences between disasters and between population groups. For examples, see the Case Study of DR-4223-TX in the following section.

Case Study: DR-4223-TX

As one of many case studies, this project provided an analysis of DR 4223, Texas Severe Storms, Tornadoes, Straight-line Winds, and Flooding. This disaster was chosen because at the time of writing there were still a number of open DRC locations. Providing an analysis of active disaster facilities provided immediate relevance for the project as well as current data for the analysis.

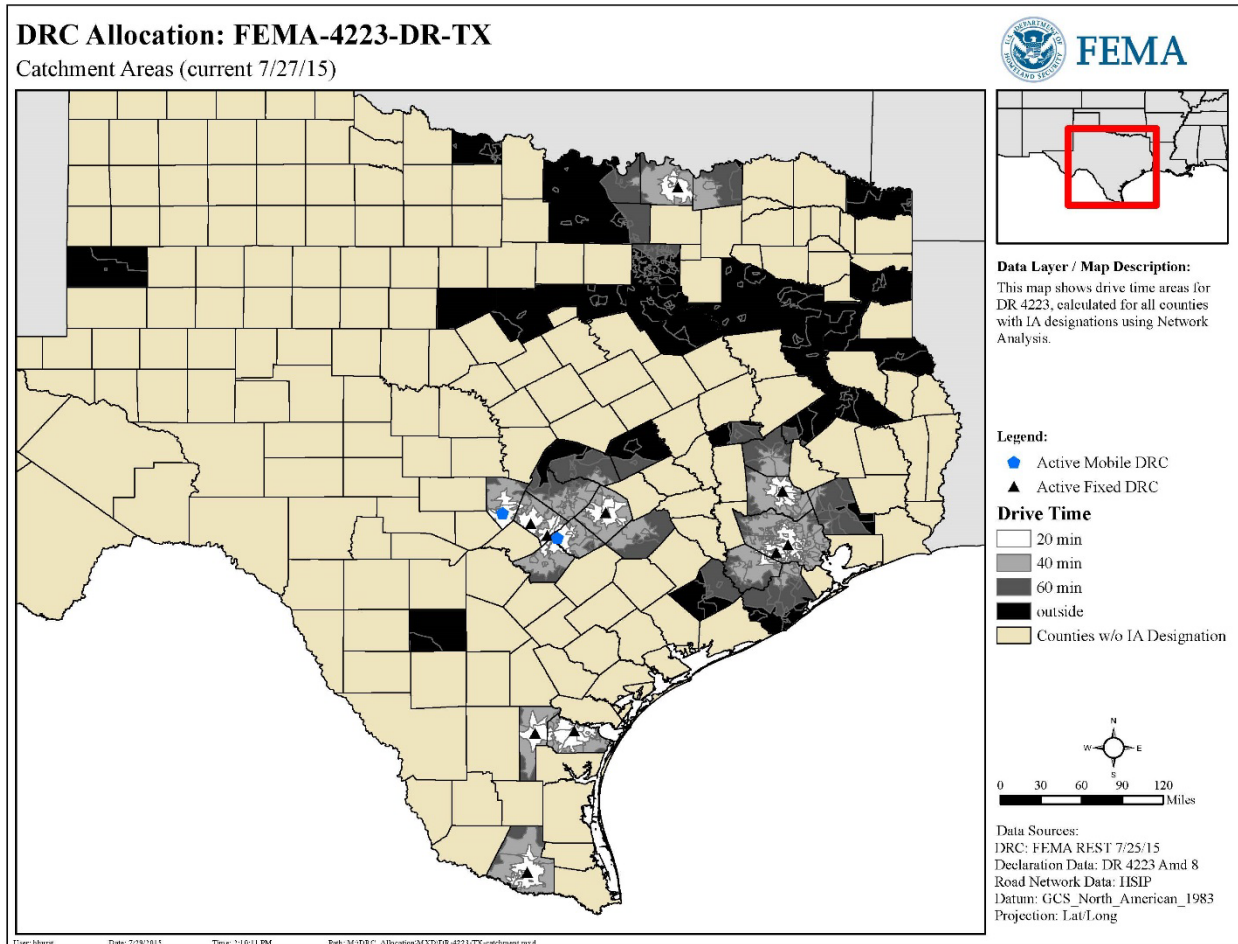


Figure 1: Catchment Areas

This disaster was declared May 29, 2015 for the incident period May 4, 2015 to June 22, 2015⁷. Data was pulled for this analysis on July 27, 2015. At that time, there were 12 active DRCs listed in the FEMA Rest Service. There were 10 counties with only IA designations and 36 with IAPA designations.

The first step in the analysis pulled data on the 2 active mobile facilities and 10 active fixed facilities from FEMA’s REST service. Then, a Service Area analysis was run to determine catchment areas (see Figure 1, above).

⁷ See <https://www.fema.gov/disaster/4223>

Next, census tracts were categorized based on their Social Vulnerability Index score (see Figure 2, below). Here, high vulnerability areas are represented in red.

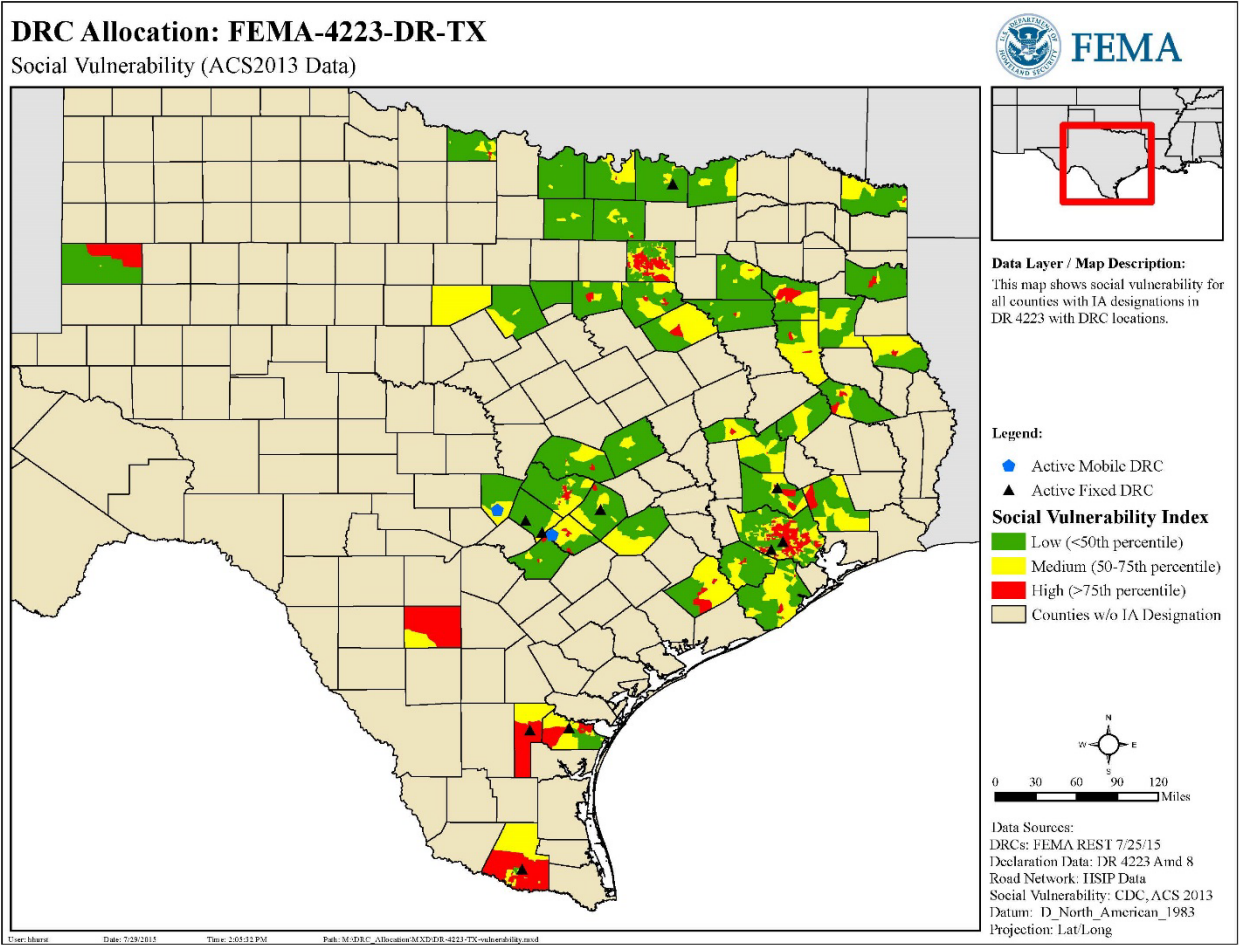


Figure 2: Vulnerability

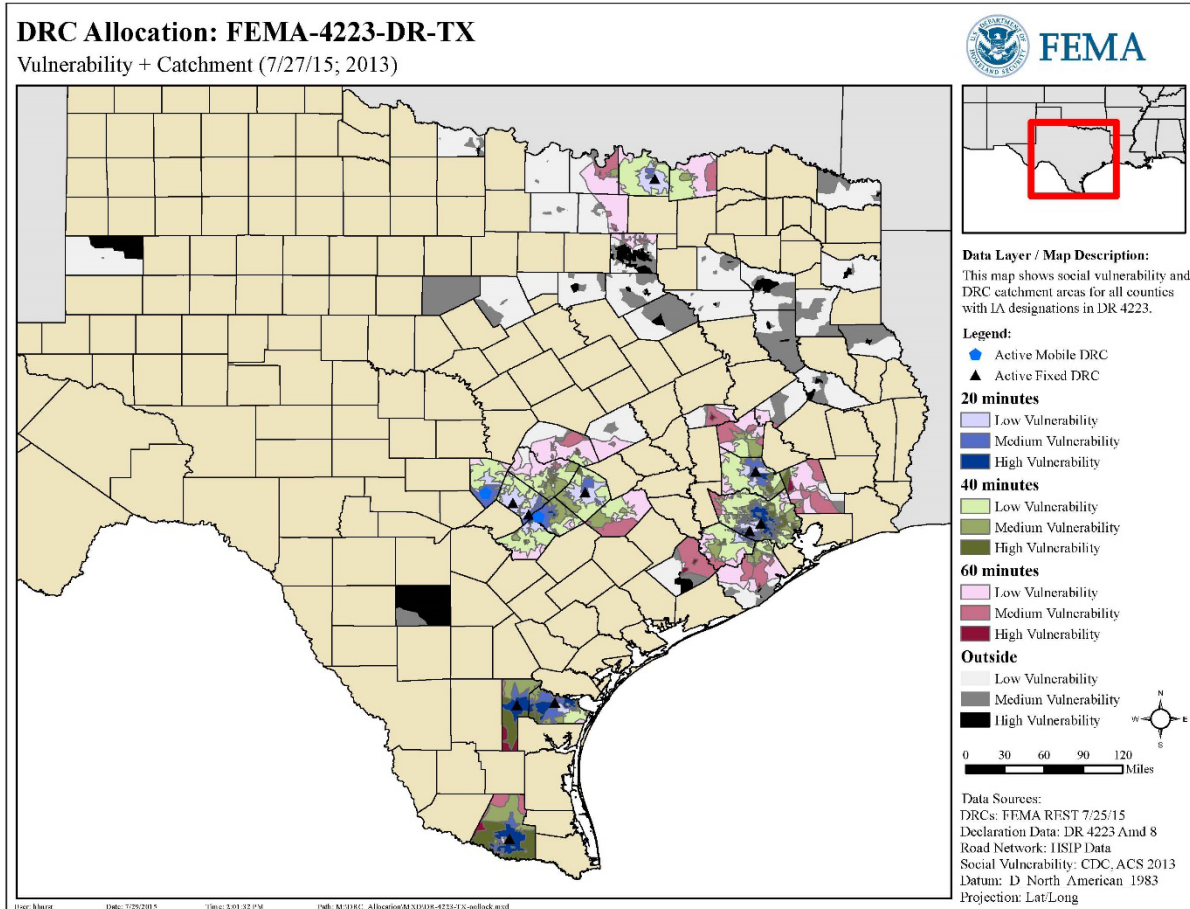


Figure 3: Combined Analysis

In the next step of the analysis, these two maps are combined to provide a comprehensive picture of the impacted communities (see Figure 3). The DRC locations are difficult to see against the polygons; a better choice of icon is needed.

This figure demonstrates the nuanced picture the data can paint of populations in this area. Consider, for example, the polygons represented in black. These are high vulnerability areas outside of an hour’s drive time. Individuals in these areas are more likely to need support and services than individuals in low vulnerability tracts. However, they are also less likely to be able to reach a DRC location. It is worth investigating these areas further. Do they have significant damage? Would they be good candidates for mobile DRC locations? Should DSA teams visit these communities?

In addition to providing maps, this type of analysis can be used to provide a number of graphs and charts. Consider the disaster-wide pie charts in Figure 4, showing breakdown within each vulnerability group. For example, one can see at a glance that the majority of individuals living in high vulnerability areas live within 20 minutes of a facility. However, there are still over 810,000 individuals in high vulnerability areas outside of an

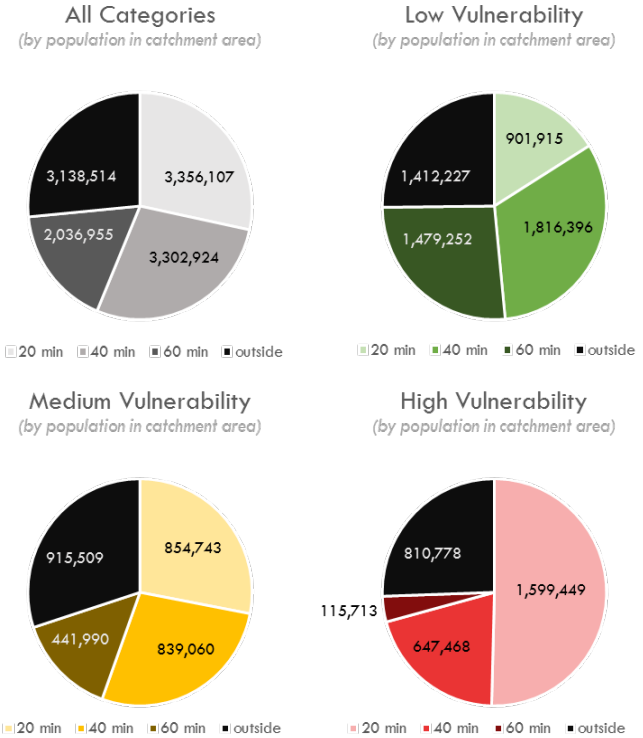
hour's drive. These charts allow comparison of resource access across groups and may allow decision makers to see if goals for DRC locations are being met.

This chart shows cumulative population in 10 minute increments. This displays slightly different data and addresses a different question than the pie charts in Figure 4.

Figure 5 shows that 50 percent of individuals living in high vulnerability census tracts can reach a DRC in under 20 minutes, based on this estimation. However, for low vulnerability groups this jumps to nearly 42 minutes.

Note that the lines indicating the 75th and 90th percentiles are not visible on this chart. This corresponds to the large uptick of the graph between 60 minutes and "outside", indicating that a large percentage of individuals are located beyond an hour's drive from a facility. The shape of this curve is unique to this specific disaster, this population, and this configuration of DRC locations.

Figure 4: Disaster-Wide Pie Charts



Catchment Population By Drive Time

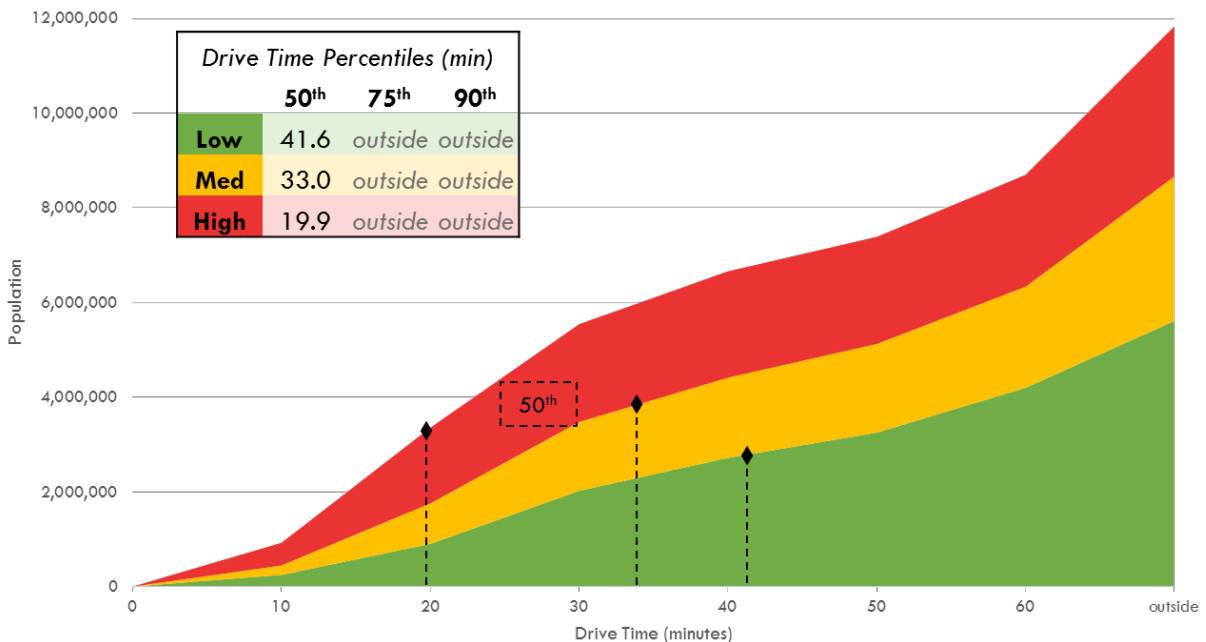


Figure 5: Disaster-Wide Catchment

In addition to disaster-wide metrics, this process allows for breakdowns of information at the individual DRC level. This information may be useful for DRC managers or other individuals directly involved in the field.

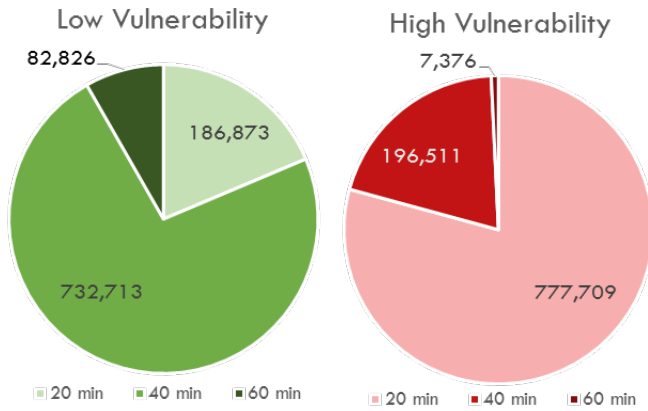


Figure 6: Vulnerability & Catchment for Leonel Castillo Community Center

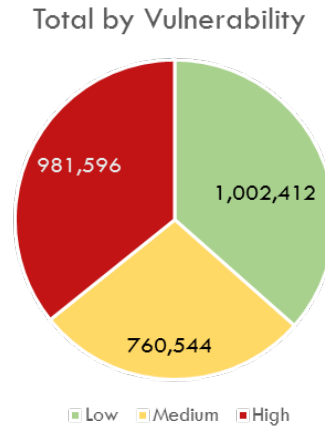


Figure 7: Vulnerability Breakdown for Leonel Castillo Community Center

Figure 6 shows breakout by time catchment for high vulnerability areas and low vulnerability areas served by Leonel Castillo Community Center in Houston, TX.

Within a 20 minute drive of this facility, there are approximately 186,873 people who live in low vulnerability areas and 777,709 people who live in high vulnerability areas. These charts may allow DRC managers to better anticipate what type of survivors they will serve.

Figure 7 shows the vulnerability breakdown of all individuals within 60 minutes drive of Leonel Castillo Community Center (note that this does not include individuals within 60 minutes who are closer to another facility).

Caveats & Weaknesses

There are a number of caveats and weaknesses to this model. This is not a final project but a first step, and there are many opportunities to improve upon this research.

The most glaring inaccuracy of this model is inconsistent datum across different data sources and layers. In order to ensure accuracy, the model and template need to be adjusted to ensure that all data have the same datum and projection.

There are a number of other caveats to this model and these data sources. This only works for the contiguous United States, as SVI is only calculated for states and Landscan is only for the contiguous US. This process includes no estimation of margin of error, an important component of ensuring rigor and usability.

The code has some bugs and glitches and is best run in chunks, allowing for troubleshooting. In particular, the process of pulling data from the REST service is clunky and doesn't provide adequate error checking.

Applications

There are many applications of this research to the DRC program, to field services, and to Recovery's efforts more broadly.

This research has the potential to help drive decisions about where to allocate DRCs. Before opening facilities, state and federal officials could analyze what populations would be served and assess whether that meets their goals. By determining what populations are currently served, officials could determine if another facility is warranted. This process allows for comparison of potential locations, providing firm analytic footing to prevent inefficient, politically motivated, or otherwise suboptimal DRC locations.

This research could also support the operations of individual DRCs. By obtaining detailed information on the population served by a given facility, DRC Managers could have a better idea of what resources are needed at a specific location.

In addition, this research could enable data-driven collaboration between the DRC workgroup and the DSA Teams. Identifying high vulnerability populations far from DRC catchment areas could serve as one way to identify locations for DSA teams to visit. This could prevent duplication of efforts and ensure easier access to resources.

Next Steps

This research is one small view into how spatial analysis can help support field services and recovery operations. In the future, this research should be extended to predict, select, and rank possible DRC locations in addition to evaluating current ones.

This work could be expanded to look at a number of different variables, such as IA registrations, ONA grant award amounts, IHP grants, or other internal data. This could be aggregated to 1km or census tract level to maintain confidentiality.

I hope that future DHS-STEM Interns, GIS Analysts, or Recovery Analytics & Budget professionals will continue to improve this analysis.

Skills & Knowledge Gained

During the course of this internship I gained a great deal of knowledge, skill, and patience. FEMA is a large, complex organization and I enjoyed the challenge of learning about every side of it, from program areas to staffing designations. As a citizen, I appreciated the opportunity to learn more about disaster preparedness; on a personal note I feel much more prepared for a natural disaster after my time at FEMA.

I have gained a number of skills during my time here. I have expanded my ability to use Python by coding extensively in ArcPy. I have become a much GIS analyst – if I can consider myself a GIS analyst at all – through my extensive use of ESRI software and the

Network Analyst extension. In addition, I have come to understand the challenges in data organization that FEMA and other similar agencies face.

Personal & Professional Impact

This internship has exposed me to a veritable wealth of opportunity. In June, I had the privilege of attending the Geospatial Intelligence Symposium in Washington, DC. This four day conference introduced me to applications of geospatial analysis from industry, intelligence agencies, contractors, and educational institutions. In addition to attending numerous keynote addresses and informational sessions, I was able to speak with industry leaders about the future of geospatial research. I look forward to maintaining my contact with the individuals I met there as I progress with my career.

I have also presented my research and findings to a number of groups across FEMA. This process was an invaluable learning opportunity and I benefited greatly from the opportunity to share my research with others and learn from their questions and feedback. I have had many wonderful, candid conversations with my FEMA colleagues. I appreciate their wisdom and advice and look forward to keeping in touch.

My experience here at FEMA has reaffirmed my desire to pursue a graduate degree in geospatial analysis. Moving forward, I hope to strengthen my quantitative and design skills through both educational and professional opportunities.

Conclusion

Spatial analysis – as this document hopefully makes clear – is about more than just maps. It is about generating useful insights, exploring data, facilitating decisions, and increasing FEMA’s ability to support survivors. In order to do the best we can, in order to promote recovery and aid survivors, we *must* combine empathy with analysis.

Above all else, it is this dynamic combination of skills that draws me to the work of disaster recovery. FEMA has an unparalleled opportunity to combine deep compassion with rigorous analytics, to use both brains and hearts to better the lives of survivors. As we gain access to deeper, richer data about the lives of Americans – from social media intake to sewage output – we face a choice.

Do we shy away from data, believing them the harbingers of anesthetized SOPs and soulless decision making? Or do we embrace these advances as mission central, let vital data unlock deeper empathy, and rise to the opportunity to truly “improve our capability to prepare for, protect against, respond to, recover from, and mitigate all hazards”?

Based on available data, I believe the choice is clear.

Acknowledgements

This project could not have been possible without the countless hours of meetings and patient support and encouragement from the wonderful DRC working group and the whole Recovery Directorate.

With special thanks to the hardworking DRC Group; Mick Brierley, Anna Krzystyniak-Sobiewska, David Warren, Greg Edgar, and Michelle Bonifas in the incredible Recovery GIS Team; Carter Hawley; Chris Vaughan and his team for getting me access to data; Tim Allen, statistical wizard; Chris Smith for his feedback and encouragement; my fellow intern Kennedy Akwuba; my mentor Julia Moline; and my (very patient) supervisor Carlos Davila.

This project was funded and made possible through the DHS HS-STEM Summer Internship Program administered by ORAU.

For questions or concerns, please contact Heidi Hurst – hurst@college.harvard.edu; misshurst@gmail.com.