Journal of Transport & Health 31 (2023) 101627

Contents lists available at ScienceDirect

Journal of Transport & Health

journal homepage: www.elsevier.com/locate/jth

Connecting people to food: A network approach to alleviating food deserts

Anna Sisk[a,*], Kristen Rappazzo[b], Tom Luben[b], Nina Fefferman[a,c]

[a] Department of Mathematics, University of Tennessee, 1403 Circle Dr, 227 Ayres Hall, Knoxville, TN, 37916, USA
[b] U.S. Environmental Protection Agency, Research Triangle Park, NC, USA
[c] Department of Ecology & Evolutionary Biology, University of Tennessee, 569 Dabney Hall, Knoxville, TN, 37996-1610, USA

A R T I C L E I N F O

Keywords:
Public transportation
Centrality measures
Food access
Urban planning
Food desert

A B S T R A C T

Introduction: In 2020, 13.8 million people in the United States struggled with food security. This means they were uncertain whether their food needs would be met. Where someone lives can influence struggles with food security. Food deserts are census tracts that experience high rates of poverty (20 percent of residents at/below poverty thresholds) and low access to grocery stores with nutritious foods. Food deserts and insecurity disproportionately affect disadvantaged communities and may contribute to health issues like diabetes, high blood pressure, and obesity. Public policies can be utilized to lessen the impact of food deserts and one way this can be achieved is through public transit.
Methods: We characterized the role public transportation plays in connecting food desert residents with food by formulating network models from data on food deserts, grocery stores, and public transportation systems for five representative locations: Brown Deer, WI; Lawrence, KS; Albuquerque, NM; Charlotte, NC, and Raleigh, NC. We analyzed these networks by looking at centrality measures, specifically degree and closeness. These centrality measures provide insight on the situation regarding grocery store access for food deserts.
Results: Results of the degree centrality measure varied across study sites; one site (Lawrence) had at least 1 bus stop within 0.25 miles (0.40 km) of the representative address for each food desert. Conversely, two sites (Charlotte and Raleigh) each had 2 representative addresses with 0 bus stops within 0.75 miles (1.21 km). When using the closeness centrality measure, 2 food deserts in Albuquerque had the highest number of grocery stores within 30 min (22 and 9) while 44% of food deserts in Raleigh had 0 grocery stores within 30 min.
Conclusions: Using these results, we identify how public transportation could better connect people with food and offer suggestions to city leaders as a way to help eradicate food deserts.

1. Introduction

In 2020, 13.8 million people in the United States struggled with food security, meaning they were unable to ensure or uncertain their food needs would be met (Coleman-Jensen et al., 2021). While struggling with food security has negative physical and psychological impacts on people of all ages, such as increased mental health issues and poor health outcomes, children face the most

* Corresponding author. Department of Mathematics, University of Tennessee Knoxville, 1403 Circle Dr, 227 Ayres Hall, Knoxville, TN, 37916, USA.

E-mail address: asisk9@alum.utk.edu (A. Sisk).

https://doi.org/10.1016/j.jth.2023.101627
Received 10 December 2022; Received in revised form 20 March 2023; Accepted 11 May 2023
Available online 2 June 2023

severe consequences (Gundersen and Ziliak, 2015). Children living or born into food insecure homes are at an increased risk of birth defects, anemia, cognitive issues, aggression, anxiety, asthma, depression, suicide ideation, and overall poorer physical and mental health outcomes than their food secure counterparts (Gundersen and Ziliak, 2015). Struggles with food security can be influenced by both where a person lives and if affordable and nutritious food is accessible. Food deserts are census tracts that are both low income and low access. The United States Department of Agriculture (USDA) defines a census tract as low income if it has a poverty rate of 20 percent or greater, or a median family income at or below 80 percent of the statewide or metropolitan area median family income. An urban/suburban census tract is defined to be low access if at least 500 persons and/or at least 33 percent of the population is living more than 1 mile (1.61 km) from a supermarket or large grocery store (for rural census tracts it is a distance of 10 miles) (Ploeg et al., 2011). Distances between a census tract and grocery store are calculated using the method described in (Transportation and N). While the existence of food deserts does not always imply its inhabitants are food insecure, it is often a contributing factor. Further, whether or not a person is food insecure, living in a food desert is associated with increases in the likelihood of health issues such as diabetes, high blood pressure, high cholesterol and obesity (National Research, 2009) due to decreased access to healthy and nutritious foods. Thus, addressing and eliminating food deserts is an important public health and social welfare objective.

While the ideal solution to ending food deserts is to increase the number of grocery stores near food deserts and/or lift communities out of poverty, this is rarely a realistic option. According to El-Geneidy, A. et al. (El-Geneidy et al., 2014) grocery stores and supermarkets are unlikely to be voluntarily located in urban food deserts due to economic reasons such as: higher urban land, labor, and utility costs, lower profit margins on perishable food items, and increased theft problems. Since city leaders and local governments do not have control over where grocery stores are located and it can be economically prohibitive to incentivize them to open in low income and urban areas, they must come up with alternative solutions to eliminate food deserts. Some of the most effective ways to do this are through policy interventions. Strategies such as revisiting city zoning codes to allow for grocery stores to open in previously unavailable areas (Development C.o.M.D.o.C. C.o.M.D.o.C. Development, 2019), streamlining and prioritizing food licenses (Development C.o.M.D.o.C. C.o.M.D.o.C. Development, 2019), and providing financial backing for fresh food initiatives (Loosemore, 2019) can help reduce the amount and effects of food deserts. While city governments cannot always bring grocery stores to people, they can help bring people to grocery stores. Offering shuttles that can transport food desert residents directly to and from grocery stores and working with rideshare and taxi companies to offer reduced fares to and from grocery stores can help increase access to food (Development C.o.M.D.o.C. C.o.M.D.o.C. Development, 2019; Ahangari et al.). In some urban and suburban areas, the public transit system can play an important role in connecting people to food (Ver Ploeg et al., 2015). Since food deserts, by definition, are low-income communities, public transportation is often the most cost-effective and sometimes the only way for people to travel around their cities. In addition, city leaders and departments have more control over their transit systems than they do the location of grocery stores. According to the Center Centers for Disease Control and Prevention (CDC) more affordable and accessible public transit will help increase access to healthier grocery options and decrease negative health outcomes associated with food insecurity (Control and C.f.D. and Prevention, 2014), which is crucial as the USDA reports that 31% of food insecure households do not use a personal vehicle to do grocery shopping (Ver Ploeg et al., 2015). Thus, analyzing the public transit system can provide important information for local decision makers working to eradicate food deserts.

Mathematical modeling is a useful tool for analyzing both public transit systems and food deserts. Agent-based models of food deserts and food access have allowed for the assessment of different interventions to increase food access, including adjustments to the public transit system (Widener et al., 2013; Abel and Faust, 2020) and statistical analysis has been used to analysis grocery shopping behavior, specifically considering if shopping occurs at the nearest store, frequency of shopping trips, and the role of vehicle ownership plays in these behaviors (Ahangari et al.). In addition to statistical analysis, GIS analysis has also be used to study how grocery shopping behavior is effected by grocery store proximity and travel time cost (including public transit) (Chavis et al., 2020). Network models are commonly used for studying public transit (Derrible and Kennedy, 2009; Fortin et al., 2016; Saidi et al., 2017). A network, sometimes referred to as a graph, is a collection of nodes and edges, where the nodes are connected to each other by the edges. This framework lends itself easily to public transportation systems as transit stops are connected to each other by transit lines. Thus, using a network model is both a time- and cost-effective way to study the relationship between public transit and food deserts. The results of such network models can provide valuable information to local decision makers working to determine how public transit can be better utilized to help alleviate issues stemming from food deserts.

In this paper, we characterize the role public transportation plays in connecting urban/suburban food desert residents with food by formulating network models from data on food deserts, grocery stores, and public transportation systems, where the public transportation lines serve as the edges and the food deserts and grocery stores are the nodes. We then analyze the networks by performing centrality measure analysis, specifically, degree and closeness. Finally, we offer some suggestions on how city leaders can use these results to help relieve problems caused by food deserts.

2. Methods

2.1. Model formulation

We identified urban and suburban locations as potential study sites and then used USDA's Food Access Research Atlas (USDA, 2021a) to construct a preliminary list of cities/towns which contained food deserts. To simplify the analysis, and for consistency across study sites, we decided to only use sites where the public transportation system is mainly a bus system. In our final selection, we wanted to ensure heterogeneity in the study sites, so we selected both urban and suburban locations that were distributed geographically across the country. The locations we studied are Brown Deer, WI; Lawrence, KS; Albuquerque, NM; Charlotte, NC; and Raleigh, NC.

Across the five study sites there are 61 food deserts distributed as follows: Brown Deer: 6; Lawrence: 5; Albuquerque: 15; Charlotte: 19; Raleigh: 16.

The conceptual framework for the network models is the food deserts and grocery stores are the nodes, while the transit lines are the edges. We implemented the network models and analyzed them in ArcGIS Pro. We primarily utilized the Network Analyst tool, Create Buffer tool, and Summarize Within tool. Each of the maps presented here and in supplemental material was created in ArcGIS Pro. Since some of the tools in ArcGIS Pro require a specific address or point, we chose one address per food desert to represent that food desert in the network. We refer to these addresses as the representative addresses. Candidates for the representative addresses were selected by looking at the food desert locations on Google Maps. We selected addresses that have both high population densities, like apartment buildings or locations in the middle of a community (e.g., a neighborhood park or school), and that are near the center of the food desert. Selecting representative addresses in this way allows for a better understanding of the situations of real people living in food deserts. Further, it helps ensure the calculations and analysis are done where people actually live in that food desert. For example, some food deserts have non-residential areas and thus do not have permanent residents living in them. Hence, analysis done with a representative address from non-residential areas won't be as accurate for those living in that food desert. The street addresses for the representative addresses and the food desert they represent for each study site can be found in the supplemental materials. Note that the food deserts are referred to by the last 3 digit of their census tract number.

2.2. Data

Information and graphical representation of the food deserts were obtained from the USDA's ArcGIS REST Services Directory, specifically using the foodaccess 2019 (MapServer) (foodaccess 2019 (MapServer), 2019). To represent the edges in the network model we created a public transportation network in ArcGIS Pro using General Transit Feed Specification (GTFS) public transit data and street centerline data. The GTFS data weas obtained from Open Mobility Data (OpenMobilityData.) and a local city's website (KU-LTS-GTFS. Available from) and the street centerline data was obtained from each city's open data bases (Office and M, 2018; Lawrence and C, 2021; Albuquerque and C; Transportation and N; County and W, 2021).

The USDA defines grocery stores to be establishments generally known as supermarkets and smaller grocery stores primarily engaged in retailing a general line of food, such as canned and frozen foods; fresh fruits and vegetables; and fresh and prepared meats, fish, and poultry. Convenience stores, with or without gasoline sales, are excluded (USDA, 2021b). We were unable to access the proprietary list of grocery stores used by the USDA, so we created our own grocery store list, based on the USDA's definition, for each city using the Reference USA/Data Axle Reference Solutions database. We sorted the data by Standard Industrial Classification (SIC) codes and focused on the 5411-Grocery Stores category. We included businesses where the codes in Table 1 are the primary SIC code, since many non-grocery stores (e.g., liquor stores, dollar stores) will have a secondary SIC code that labels them as grocery stores. Further, we included big box retailers or superstores (e.g., Walmart, Target) which are classified as department stores (SIC code 5311), but also include a grocery department. Wholesale clubs (SIC code 5311–10), for example Costco and Sam's Club, are not included as they generally require a paid membership to shop and thus are not accessible to all residents of a food desert. We also explicitly excluded SIC codes 5411–03: Convenience Stores and 5541–01: Service Stations-Gasoline and Oil from the search.

In order to verify the accuracy of the grocery store lists, we conducted a spot check on each city's list to ensure that only grocery stores were included. This check revealed many 'non-grocery' stores were included due to misclassified SIC codes. Thus, to increase the accuracy of the lists, we reviewed each entry on the lists by hand to determine if the store contributed to a positive food environment (offered fresh and nutritious foods in all the major food groups as stated in the USDA's definition given above). The initial process was conducted using Google to search the name and/or address of the business and look for evidence that the store sold fresh produce and meat, in addition to dry and/or packaged foods. This evidence could be in the form of pictures, customer reviews, social media, and/or websites. Businesses labeled as 'permanently closed' on Google were removed. Further, businesses that appeared (by their pictures or reviews) to be corner stores or neighborhood stores were also removed. These corner stores showed up most often in the Brown Deer list. These types of shops were not counted as grocery stores because although they may sell some grocery items, they act more as convenience stores, and are unlikely to be full-service grocery stores.

Businesses that were unable to be verified through the initial Google search were separated into two groups: one where the addresses from the list matched the address from a Google search and one where they did not match. For the entries where the addresses on the list and the Google search did not match, we checked the addresses on Google Maps. If there was no evidence a grocery store was

Table 1
SIC codes included in the grocery store lists.

SIC Codes	Descriptions
5411–05	Grocers-Retail
5411–06	Markets-Kosher
5411–07	Grocers-Ethnic Foods
5411–08	Grocers-Health Foods
5411–09	Grocers-Take-Out-Foods
5311–04209	Retail Shops (Walmart)
5311–02209	Department Stores (Walmart)
5311–02148	Department Stores (Walmart SuperCenter)
5311–02200	Department Stores (Target)

located at the street address on Google Maps, the entry was removed from the list. If there was evidence the businesses had either moved locations or changed names, the list was updated. For the businesses where the addresses did match but were unable to be verified in the initial search, a more detailed search was conducted; if there still wasn't enough evidence to definitively exclude a business, it remained on the list but was flagged to determine the effect of including or excluding these businesses on the results. Across all five study sites, there were initially 620 businesses labeled as grocery stores, but after the verification process the number of grocery stores was reduced to 408. Of the 212 businesses removed, 13 were labeled as permanently closed, 129 were food related businesses mislabeled as grocery stores in the database (e.g., corner stores, convenience stores, restaurants, liquor stores, etc.), and 70 were removed for miscellaneous reasons (addresses that did not have grocery stores at the location, businesses unrelated to food being mislabeled, duplicate entries, etc.). We conducted a sensitivity analysis between the initial grocery store lists and the lists after the verification to determine the effects of the inaccurate lists on our analysis.

2.3. Centrality measures

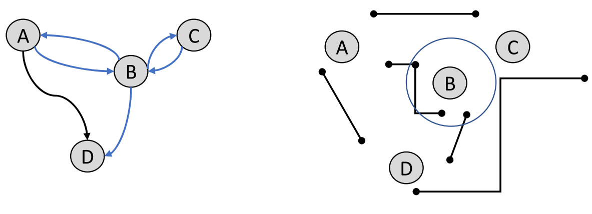
We analyzed the networks by looking at centrality measures, which quantify specific characteristics of the contribution to, and relative importance of, each component in the overall network organization. Centrality measures can focus on nodes or edges, and which is considered to be the most 'central' is defined by the measure. There are many types of centrality measures, common ones being degree, betweenness, closeness, and eigenvalue (Wan et al., 2021). Each centrality measure evaluates the importance of different aspects of a network in different ways. For our network models, we considered two types of centrality measures: degree and closeness. Both of these measures focus on which node is the most important to the network.

2.3.1. Degree

The traditional definition of the degree centrality is the number of edges that are connected to a node. Thus, the node with the highest number of edges connected to it, or the highest degree, is the most central to the network under this metric. We selected the degree centrality to analyze our network to show how connected a food desert is to the public transportation system. Note that the representative addresses (nodes) are not directly connected to the public transits routes (edges), so we have to slightly modify how we define and implement the degree centrality. Since public transit routes can only be accessed through bus stops, the number of bus stops accessible to a representative address provides an understanding of how connected that representative address is to the public transportation system as a whole. Thus, we define degree to be the number of bus stops within a particular radius of the representative address in each food desert. To calculate this we used the 'Create Buffer' tool in ArcGIS Pro to define a radius of a certain distance around the representative addresses. Then using the 'Summarize Within' tool we determined how many bus stops fell in that radius. We considered 3 different distances: 0.25, 0.5, 0.75 miles (0.40, 0.80, 1.21 km) based on (El-Geneidy et al., 2014). Fig. 1 provides a visual representation of both the traditional degree centrality and our modified version.

2.3.2. Closeness

Traditionally, the closeness centrality examines the path distance from one node to all the other nodes in the network. Thus, the lower the closeness the more central a node is to the network. This definition of closeness is not particularly useful for these analyses, as we are not interested in how close food deserts are to one and other or how close grocery stores are to one another; we are interested in identifying the distance between food deserts and grocery stores. Further, since we are exploring the role of public transportation, we are less interested in the distance between the two types of nodes as we are in the travel time between them (since public transit routes will rarely go directly from the starting point to the destination). Thus, our closeness centrality evaluates the number of stores that are within a certain total travel time from each food desert, where total travel time is defined as time spent walking and on public transit. Therefore, the higher the closeness, the more access a food desert has to grocery stores. This allows us to explore how connected food deserts currently are to grocery stores. We accomplished this using the Service Area layer, which is a part of the 'Network Analyst' tools. The Service Area layer shows the area reachable by public transit and walking within a given amount of time from a starting



a. b.

Fig. 1. Fig. 1 a is a simple example of a tradition network, where the nodes are represented by the gray circles labeled with capital letters and the edges are represented by the black and blue arrows. Following the traditional definition of the degree centrality, node B has a degree of five as shown by the blue colored edges. Fig. 1 b shows an abstracted example of our food desert/public transportation networks, where the gray circles labeled with capital letters are the representative address (nodes), the black lines are transit routes (edges), and the black circles at the ends of the transit routes are the bus stops. For Fig. 1 b the degree of B is three, since only three bus stops fell within a certain radius, shown by the blue circle, of the B. Note, a subtle difference between the two is the traditional degree centrality counts the edges, while our modified definition counts the number of ways to access the edges (not the edges themselves).

location on a particular day and time. Due to differences in the format of the GTFS data for the cities, the Service Area Layer was calculated for a general Wednesday at 7:00 p.m. for Lawrence, Albuquerque, Charlotte, and Raleigh and for the specific date of Wednesday, November 17th, 2021, at 7:00 p.m. for Brown Deer. We ran the simulation with three different cutoff times for total time traveled: 10, 20, and 30 minutes. Note that wait and transfer times for public transit are included when determining the total time traveled. After the creation of the service area maps for each time cutoff, we used the Summarized Within tool to determine how many stores are reachable within each service area.

In addition to defining closeness to be a centrality measure for the food desert nodes, we also defined a related centrality measure for the grocery store nodes. This centrality measure, called the overlap centrality, looks at how many food desert service areas a grocery store resides in. Although not a traditional centrality measure, we consider this to be informative because the higher a store's overlap the more important it is to the network for food access. We defined an overlap of 1 to mean a store is only reachable by one food desert, an overlap of 2 to mean a store is reachable by two food deserts, and so on. We calculate this centrality measure using the data generated from calculating closeness. In addition to determining the number of stores reachable in a service area, the Summarize Within tool can also generate the list of stores. We were then able to see which stores appeared in more than one food desert's list. Since the overlap centrality is particularly interesting for grocery stores with an overlap of 2 or more, we omitted all stores with an overlap of 1 from the results shown in the supplemental material.

2.4. Sensitivity analysis of representative addresses

While we purposefully selected representative addresses such that the address would be both in high population density areas and near the center of the food desert, there are many different methods that could be used for selecting representative addresses. Thus, we conducted a sensitivity analysis on representative address selection to determine the impact of these choices on the results of the degree and closeness analyses. We focused these sensitivity analyses on the example cities of Brown Deer and Lawrence, computing the degree and closeness centrality measures under two additional methods for selecting representative addresses. Then we employed a Wilcoxon Signed-Rank Test to determine if there was a statistically significant effect on the results depending on how the representative addresses were selected. The two additional methods we used to select representative addresses were: 1) randomly selecting a representative address from a list of high population density addresses (we will call this the 'random' method), and 2) selecting an address in the geographic center of the food deserts regardless of whether it is a residential area (we will call this the 'center' method). We will call the original method we used for selecting representative addresses the 'original' method. For the random method we used Reference USA/Data Axle Reference Solutions database to generate a list of apartments for each city. This was done by searching for the following apartment related SIC codes: 651304, 651303, 651307, and 651398. Since not all food deserts have apartments in them, we also included addresses of public K-12 schools, as these tend to be located in or near residential communities. After eliminating addresses not located in food deserts, the ones remaining were assigned a number. Using a pseudo-random number generator, one address was selected as the representative address for each food desert. If the randomly chosen address was an apartment complex the address was searched on Google to ensure it was accurate. If the address did not lead to an actual apartment, it was discarded and a new address was randomly selected using the same method as before. For the center method, we dropped a pin in the geographic center of the food desert (where street addresses exist) on ArcGIS Pro and those addresses were selected as the representative addresses. The addresses and results of the degree and closeness centrality analyses calculated with representative addresses from the random method and center method are provided in the supplemental materials.

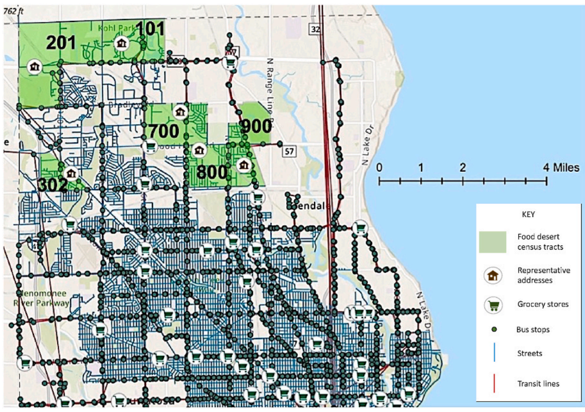


Fig. 2. Full network model for Brown Deer, WI. Note: The food desert census tracts are shaded in green and labeled by the last three digits of their census tract number. The representative addresses are symbolized by a white circle with a picture of a house inside and the grocery stores are symbolized by a white circle with a shopping cart inside. For the edges, the bus stops are marked by small green circles and the streets are the blue lines. The maroon lines represent the transit lines but note they do not always follow the streets or look like a typical map of the transit system. This is because they are representative of logical connections in the transit system, which is used by The Public Transit evaluator tool to determine travel times (Create and use a network).

To then determine how sensitive our results are to the selection method for representative addresses, we used the Wilcoxon Signed-Rank Test. This is a non-parametric test for two populations when the observations are paired, and it conducts a paired difference test of repeated measurements on a single sample to assess whether their population mean ranks differ. We conducted this test pairwise between the results for the original method's representative addresses and the random method's representative addresses and the original method's representative addresses and the center method's representative addresses, as we are interested in how the results in this paper could be affected. We ran the test for each distance/time cutoff. To do this analysis we used MATLAB's signrank function. We calculate the p, h , and stats values, where p is the p -value of a two-sided Wilcoxon Signed-Rank Test, h indicates a rejection of the null hypothesis or a failure to reject the null hypothesis at the 5% significance level, and stats returns the test statistic, W. Note that $h = 0$ represents failure to reject the null hypothesis and $h = 1$ represents rejection of the null hypothesis, where the null hypothesis is a zero median for the difference between paired samples.

3. Results

Fig. 2 is the full network map for Brown Deer, WI produced by ArcGIS Pro. The full network maps for the other study sites can be found in the supplemental information.

The map depicts the 6 census tracts identified as food deserts in the Northwest portion of the city and the majority of bus stops and grocery stores located in the central and Southeast portion of the city.

3.1. Degree

Fig. 3 shows the results of the degree analysis, and the maps of all five study sites with the food deserts, representative addresses, bus stops, and the 0.75-mile (1.21 km) buffer displayed. The full degree analysis results can be found in Tables 19–21 in the

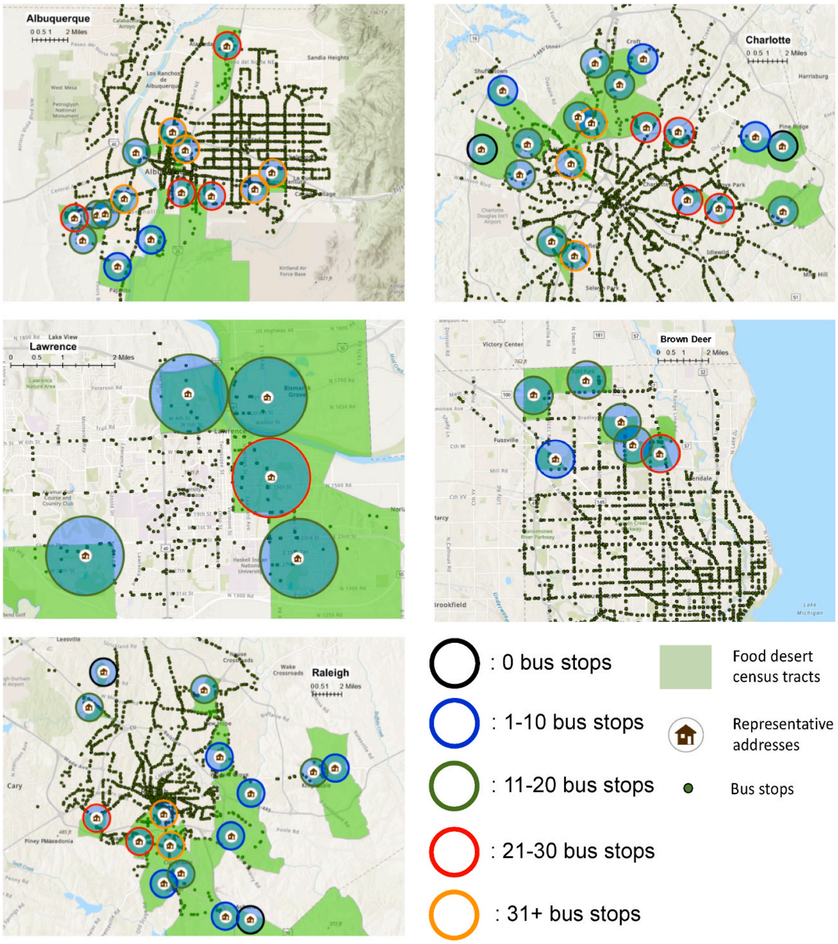


Fig. 3. Degree Results. Maps of all five study locations with food desert shaded in green, the representative apartments marked by a white circle with a house inside, bus stops represented by small dark green circles, and the 0.75 miles (1.21 km) radius circle shaded in light blue outlined with the color associated to the number of bus stops that fall inside it.

supplemental materials. The food desert with the highest degree within 0.75 miles (1.21 km) across all the study sites is Raleigh's 700 with 73 bus stops and the second highest is Albuquerque's 900 which has 60 bus stops. In Lawrence, all representative addresses have access to at least one bus stop within 0.25 miles (0.40 km), which is the smallest distance we considered. Conversely, 50% of representative addresses in Brown Deer have a degree of 0 within 0.25 miles (0.40 km), and the same is true for 47% of representative addresses in Albuquerque, 32% in Charlotte, and 44% in Raleigh. Across all the food deserts, Charlotte's 006 and 618 and Raleigh's 713 and 809 have the least access to their public transit systems as they have a degree of 0 within 0.75 miles (1.21 km). Thus, Raleigh has the largest disparity in public transit access for food deserts. On the other hand, Lawrence has the smallest disparity in public transit access as the largest degree within 0.75 miles (1.21 km) is 22 and the smallest degree is 14. Across all the study sites, 11–20 is the most common number of bus stops within a 0.75 miles (1.21 km) radius and this holds true on the city level for Brown Deer, Lawrence, and Charlotte. In Albuquerque 31+ is the most common number of bus stops within a 0.75 miles (1.21 km) radius, while in Raleigh 1–10 bus stops is the most common. Table 2 and Table 3 show the results of the Wilcoxon Signed-Rank Tests between the original method's results and the random method, and the original method and the center method for Brown Deer and Lawrence. Note, since $h = 0$ for each test, that we fail to reject the null hypothesis. Thus, there is zero median for the difference between paired samples.

3.2. *Closeness*

Fig. 4 shows the closeness results for all 61 food deserts across the different study sites. Tables 24–28 in the supplemental materials provide additional details for the information in Fig. 4 for each study site. Across all five cities, Albuquerque's 707 and 901 food deserts have the highest closeness with 22 and 9 grocery stores within 30 min, respectively. The food deserts with the highest closeness within 30 min for each city is as follows: Brown Deer's 900 with 7 grocery stores, Lawrence's 004 with 4 grocery stores, Charlotte's 603 with 5 grocery stores, and Raleigh's 707 with 6. Out of all the representative addresses across the 5 study sites, 20 (or about 33%) have a closeness of zero. Raleigh has the most, and highest proportion of, food deserts without grocery store access within 30 min at 44%, while Albuquerque and Lawrence have the lowest proportion with 20% each. Brown Deer and Charlotte fall in between with 33% and 37%, respectively. Four representative addresses have access to at least one store within 10 min (900 in Brown Deer, 201 in Charlotte, and 101, 301 in Raleigh) and out of the food deserts that have grocery store access, just over half must travel at least 20–30 min to access a grocery store. In addition, three of the five study sites -Brown Deer, Albuquerque, and Raleigh - have grocery stores with overlaps greater than 1, meaning the store is reachable by more than one food desert within 30 min. Albuquerque has the most with 10 grocery stores with an overlap greater than 1, while Brown Deer and Raleigh both have 2. Of the 14 stores all but 3 have an overlap of 2 food deserts; 3 stores in Albuquerque and 1 in Raleigh have an overlap of 3 food deserts. Table 4 and Table 5 show the results of the Wilcoxon Signed-Rank Tests between the original method's results and the random method and the original method and the center method for Brown Deer and Lawrence. Note, since $h = 0$ for each test, we fail to reject the null hypothesis. Thus, there is zero median for the difference between paired samples.

Five of the grocery stores across all of the study sites that appeared in the closeness analysis were flagged because we were unable to verify whether or not they provide fresh produce and meat (as indicated in the Methods section). One of the flagged stores was reachable by the 900 food desert in Brown Deer and one was reachable by the 901 food desert and another one by the 707 food desert in Albuquerque. In Charlotte the 305 food desert could reach a flagged store, and the same was true for the 018 food desert in Raleigh. Due to the low number of flagged stores appearing in each city's closeness analysis, the results would not be substantially different if the stores had been removed from the list. While including the flagged stores in the analysis had a relatively small impact, performing the closeness analysis on the unverified grocery store list has a larger impact. Across all the food deserts in each city, the sum of the closeness results for the number of grocery stores reachable within 30 min is increased by 73 when computing the closeness using the unverified list compared to the verified one. This breaks down to an increase in the closeness results across all the food deserts in Brown Deer by 12, in Lawrence by 5, in Albuquerque by 27, in Charlotte by 11, and in Raleigh by 18. The full results of the closeness analysis done with the unverified list is provided in the supplemental materials.

4. Discussion

We quantified the number of bus stops within different buffer distances (0.25, 0.5, 0.75 miles/0.40, 0.80, 1.21 km) of representative addresses in each food desert as a measure of their (representative addresses) degree centrality. While a high degree centrality does not guarantee reasonable access to fresh and nutritious food, it does give an idea which food desert nodes are most connected or

Table 2
Wilcoxon Signed-Rank Tests for Degree in Brown Deer. Results of the Wilcoxon Signed-Rank Tests between the original method and the random method and between the original method and the center method's degree analysis results in Brown Deer.

	Original vs. Random	Original vs. Center	Original vs. Random	Original vs. Center	Original vs. Random	Original vs. Center
Distance (miles/kilometers)	p	p	h	h	Stats	Stats
0.25/0.40	0.5625	0.8125	0	0	5	6
0.5/.80	0.0625	1	0	0	0	11
0.75/1.21	0.2812	0.4375	0	0	5	6

Table 3
Wilcoxon Signed-Rank Tests for Degree Lawrence. Results of the Wilcoxon Signed-Rank Tests between the original method and the random method and between the original method and the center method's degree analysis results in Lawrence.

	Original vs. Random	Original vs. Center	Original vs. Random	Original vs. Center	Original vs. Random	Original vs. Center
Distance (miles/ kilometers)	*p*	*p*	*h*	*h*	Stats	Stats
0.25/0.40	0.2500	0.2500	0	0	3	13
0.5/0.80	0.1250	0.1250	0	0	1	10
0.75/1.21	0.1220	0.1875	0	0	1	13

central to the network and thus more likely benefit from access to public transportation. This access to public transit could potentially lead to access to grocery stores that are further away. City leaders can use this centrality measure to decide where to add more bus stops or alter/create bus lines that will help connect low-degree census tracts to public transportation and improve access to food. This is especially important for food deserts that have no, or very low, access to public transit. Further consideration should be taken to ensure bus stops near the food deserts provide direct access to a grocery store (e. g., require no or minimal bus transfers) and that there are an adequate number of bus stops near grocery stores/business districts. Access to good public transit is important, especially in low-income neighborhoods, where other forms of reliable transportation may not be available. Lack of access to public transit can be especially concerning in census tracts that have low vehicle access (tracts in which more than 100 households have no access to a vehicle and are more than 1/2 mile (0.80 km) from the nearest supermarket) (USDA, 2021c), such as all of the food deserts in Brown Deer, Lawrence's 200, Albuquerque's 000, 200, 102, 901, and 707, Charlotte's 304, 305, 400, 401, 403, 404, 500, 603, 609, 610, 802, and Raleigh's 018, 101, 500, 803, and 900.

Determining what a high degree or adequate access to public transit means will vary across cities and depend on the city and their goals. For instance, in Lawrence, every food desert has a bus stop within 0.25 miles (0.40 km), but the most common number of bus stops within 0.75 miles (1.21 km) is 11–20, while in Albuquerque the most common number of bus stops within 0.75 miles (1.21 km) is 31+, even though 47% of Albuquerque's food deserts have no bus stops within 0.25 miles (0.40 km). An argument can be made for both locations for which city has a higher degree, and while the best option is to have the most bus stop access at the closest distance, that may not be immediately feasible. Thus, taking various factors into consideration such as existing infrastructure, budget, and population density, city leaders can decide on how best to increase the degree of their food deserts and thereby increase potential access to food.

While the degree centrality focuses on how connected a food desert is to the public transit system, closeness examines how connected a food desert is to grocery stores through the public transit system. We chose to run the closeness analysis for Wednesday at 7:00 p.m. because this is after the traditional workday ends and is late enough that it will avoid major rush hour traffic that typically increase travel times. Note, that these cutoff times represent total time traveled, which includes walking plus public transit. This is important to consider because walking long distances present different challenges than riding a bus, such as harsh weather conditions, safety, accessibility, and carrying heavy groceries. Analyzing closeness is valuable for mitigating harm caused by food deserts because it shows which food deserts could be prioritized based on need. While ideally cities would be able to increase food access for everyone in food deserts, those that have a low closeness, and especially those without any access to grocery stores within 30 min, should be addressed first. This could look like enacting short-term solutions in those areas immediately, such as mobile food pantries or providing rides to grocery stores, while more permanent solutions are implemented. Further, by using a network model to address these issues, simulations can be run to help determine where routes could be modified or added to connect food deserts to previously unreachable grocery stores or make access to them quicker. Studying this via modeling is much more efficient for cities than making changes to their transit systems and studying the realized effects only after implementation. While ideally closeness should be used to increase food access, at the least it should be used to ensure access won't be made worse by changes to the public transit system. Thus, a closeness analysis should also be run when any adjustments are made to the public transit system, even if the changes aren't directly related to food access.

While city leaders don't have complete control over where grocery stores open, they can influence and incentivize businesses to open in certain locations. Thus, considering an overlap centrality can help maximize the impact city leaders have on grocery store locations. Note that currently not many stores have an overlap greater than 1, so most grocery stores are not in ideal locations for increasing food access. Thus, city leaders should try to incentivize stores to open in locations where they would have a higher overlap, so food access can be increased for multiple food deserts at once. Note, that spacing out grocery stores such that they have a low overlap may be a purposeful strategy by the grocery stores to maximize profit, so urban planners and economist may need to investigate this further. By knowing the overlap of current or proposed stores, it could help city leaders make decision about allocation and distribution of resources. For example, stores with a higher overlap could be more likely to receive government aid or be of higher priority for the necessary licensing and other bureaucratic procedures. Further, if any grocery stores with a higher overlap close, city leaders will be aware that the location is an important one for food access and could try to incentivize a new grocery store to open in a similar location.

When considering the relationship between degree and closeness, we would not necessarily expect to see any correlation between the two as a mathematical property of an arbitrary network. Although, when considering a relationship between them as a property of networks that describe human communities, it may be intuitive to think the higher a food desert's degree is the higher their closeness should be (since they have more access to public transit and therefore better access to grocery stores that are farther away); that is not

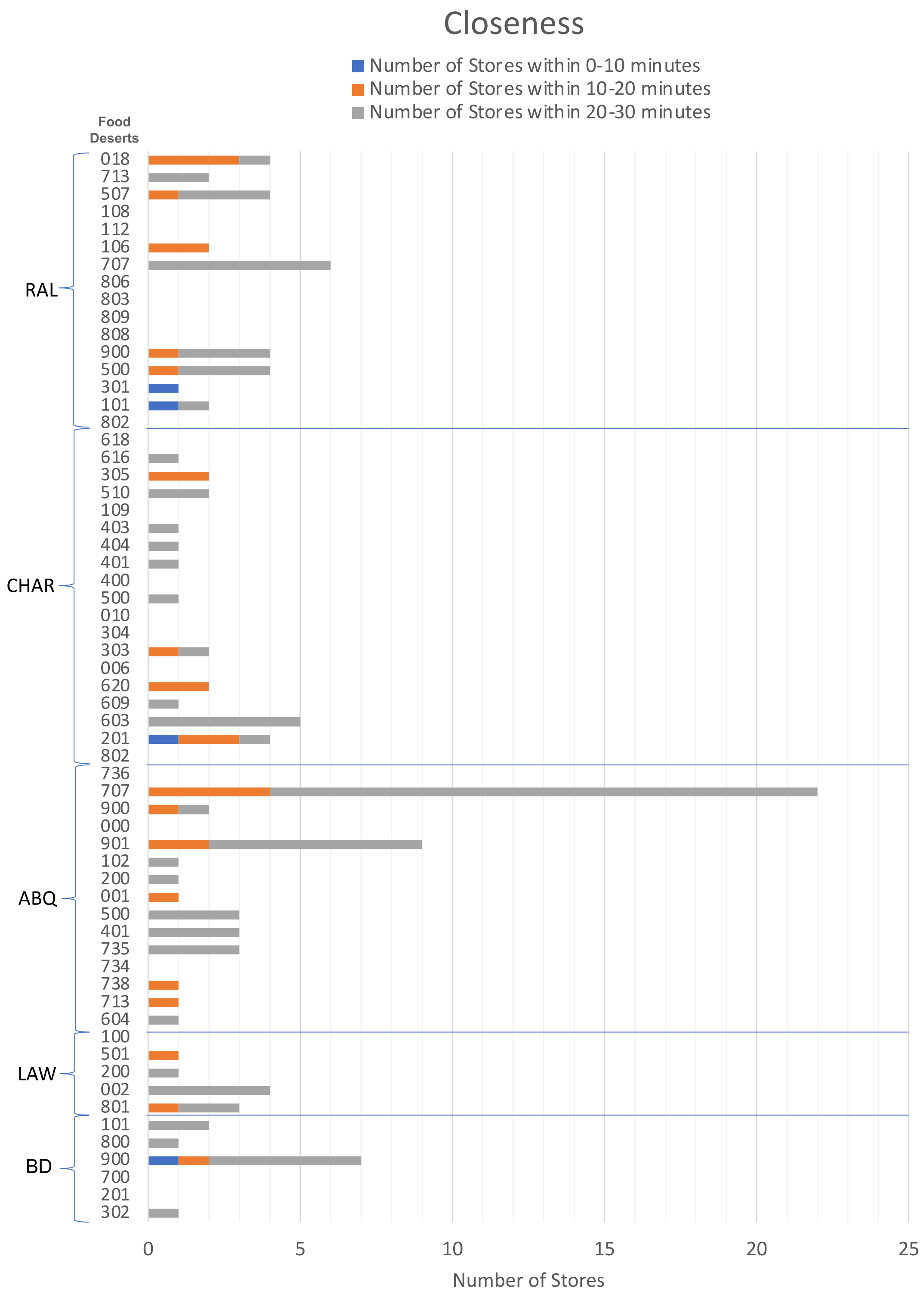


Fig. 4. Closeness Results. The number of grocery stores within 0-10, 10–20, 20–30 min of total travel time for all of the representative addresses across the study sites. BD-Brown Deer; LAW-Lawrence; ABQ- Albuquerque; CHAR-Charlotte; RAL-Raleigh.

necessarily the case (although, Brown Deer and Lawrence seem to follow this trend). On the other hand, in general, we do see if a food desert has a lower degree, they tend to also have a lower closeness. For example, of the 4 food deserts that have a degree of 0 within 0.75 miles (1.21 km), 3 don't have access to grocery stores within 30 min. One has access to two grocery stores within 20–30 min

Table 4
Wilcoxon Signed-Rank Tests for Degree in Brown Deer. Results of the Wilcoxon Signed-Rank Tests between the original method and the random method and between the original method and the center method's degree analysis results in Brown Deer.

	Original vs. Random	Original vs. Center	Original vs. Random	Original vs. Center	Original vs. Random	Original vs. Center
Distance (minutes)	*p*	*p*	*h*	*h*	Stats	Stats
0-10	1	1	0	0	1	1
10-20	1	1	0	0	1	1
20-30	0.2500	0.2500	0	0	6	13

Table 5
Wilcoxon Signed-Rank Tests for Degree in Lawrence. Results of the Wilcoxon Signed-Rank Tests between the original method and the random method and between the original method and the center method's degree analysis results in Lawrence.

	Original vs. Random	Original vs. Center	Original vs. Random	Original vs. Center	Original vs. Random	Original vs. Center
Distance (minutes)	*p*	*p*	*h*	*h*	Stats	Stats
0-10	1	1	0	0	0	0
10-20	1	1	0	0	1	1
20-30	0.7500	0.3750	0	0	4.5	8.5

mostly by walking (since they don't have access to a bus stop within 0.75 miles/1.21 km). However, this is not a bi-directional relationship. Thus, low degree usually implies lower closeness, but low closeness does not imply low degree. So, based on our study sites we can conclude that degree does not necessarily have a positive effect on grocery store access. This may imply cities are currently under utilizing their public transit systems as a way to increase food access. Many food deserts have decent access to public transit, but the public transit is not connecting them to grocery stores. For future public transit planning, officials should explicitly take food access into account. This could be as simple as adding bus stops near grocery stores on existing routes or more complicated such as adding new or modifying existing lines such that they reach grocery stores. To get a better understanding of the relationship between degree and closeness more city-specific factors may need to be considered and evaluated individually.

For the sensitivity analysis of the representative addresses for both Brown Deer and Lawrence, all the Wilcoxon Signed-Rank Tests show that there is no evidence that the choice of representative addresses, or the method used to select representative addresses, has a statistically significant impact on the overall structure of the results (e.g., highest degree/closeness to lowest degree/closeness). Thus, while the individual numerical results of the degree and closeness analysis may change depending on the method used for selecting representative addresses, insights gained from the results are unlikely to be impacted. Although, we do suggest that sensitivity be considered on a single-city basis when analysis from a network is being used to make actionable recommendations. (Note that we only performed this sensitivity analysis on Brown Deer and Lawrence and determined that, since there was no evidence of impact, these initial investigations could safely forgo running the same analysis on the other study sites at this stage. This presents an opportunity for future exploration to see if urban transportation networks provided consistency of sensitivity in centrality measures across scales of networks size. Abstract work on network centrality sensitivities on arbitrary networks (Murai and Yoshida, 2019) and particular network centralities on specifically urban networks (Lämmer et al., 2006) have already been investigated, though not in connection with food access.)

4.1. Limitations

The primary limitation of this study was issues we had with the database used for the grocery store data. Some of these issues were inaccurate classification of businesses (e.g., a music store being classified as a grocery store), repeated entries, business headquarters being labeled as a grocery store itself, etc. By performing the spot check described in the Methods section, we were able to identify these issues, but if the spot check had not been conducted these issues would not have been noted and subsequently corrected. As discussed in the same section, using the inaccurate lists inflates the number of groceries stores accessible to food deserts which conceals the severity of the issue of food access. This could lead to an inadequate response to address food insecurity in those communities. Overall, having access to a more complete data set would improve the accuracy of the models and results. While we know that the original grocery store lists contained businesses that did not satisfy our definition of a grocery store, we were not able to determine if businesses that would satisfy the definition were left out, or misclassified, by the database. Another limitation to our analysis is working with representative addresses. Although we showed that different methods for selecting representative addresses did not cause a statistically significant change in the average ranks of the food deserts results for the closeness and degree analysis within each study site, it can affect the numerical values of the results. Thus, if city leaders want to employ this model in their cities, it will be important to determine the representative addresses in a manner that is most appropriate to their efforts and desired outcomes.

In the Methods section we note this study is focused on cities with a bus system as the only or primary mode of public transportation, which limited the cities we could select as study sites. This was done to simplify the network and analysis, but in the future cities with more complex transit systems could be included in the analysis. Although consideration will still need to be taken to ensure comparison analysis is done between cities with similar modes of public transportation. In addition to considering multi-modal transportation systems, an explicit focus on city size may offer additional insight. In the Discussion section it is noted that Brown

Deer and Lawrence follow the trend that food deserts with higher degree also tend to have higher closeness, while the other cities did not. Brown Deer and Lawrence are the two smallest study sites selected (with populations of 12,507 and 94,934 in 2020 respectively), while Albuquerque, Charlotte, and Raleigh are considerably larger (with populations of 564,559, 874,579, and 467,665 in 2020 respectively) (Bureau U.S.C, 2020). By taking population and population density into account, we can also explore weighting the network by population served and the ridership capacity of the transit systems by looking at different centrality measures and incorporating other network analysis tools. Thus, in future multi-city studies considering population size may be useful to identify possible trends in transit and food deserts that appear in larger cities versus smaller.

There are different types of centrality measures that can be used to evaluate a network. We purposely selected the degree and closeness centrality measures for the information they provide about the connectivity of food deserts to both the public transit system and the grocery stores, and the realistic characterization for the current state of food access in food deserts. Our results and recommendations may have been different if different centrality measures had been implemented; our results are a direct consequence of the centrality measures we chose. Finally, this study focuses on access to the closest grocery stores, but this may not be realistic as people have grocery store preferences and will often travel further to reach a desired grocery store. So, in the future incorporating different grocery store characteristics, like pricing, product range/availability, and location safety into the model could offer better insight into the real distance people travel for groceries.

5. Conclusion

Analyzing a complex situation like food deserts with a network model and centrality measures can be an incredibly useful tool for city leaders. It allows for adjustments to be made in the model and simulations to be run, before they are implemented in the public transit systems, which is a time- and cost-effective way to consider issues surrounding food access and deserts. This lets city leaders tackle these issues dynamically. As changes are made, analysis can be run to offer adjustments before the next set of changes are implemented. This is especially important as individual behavior and grocery stores are likely to react in real time to changes. In this paper, we used network models to characterize the role public transportation has in connecting food deserts to grocery stores in 5 U.S. cities. Our network models focused on two centrality measures: degree and closeness. Results of the degree centrality measure varied across study sites; one study site (Lawrence, KS) had at least 1 bus stop within 0.25 miles (0.40 km) of the representative address for each food desert. Conversely, two study sites (Charlotte, NC and Raleigh, NC) each had two representative addresses with 0 bus stops within 0.75 miles (1.21 km). When using the closeness centrality measure, 2 food deserts in Albuquerque, NM had the highest number of grocery stores within 30 min (22 and 9), while 44% of food deserts in Raleigh, NC had 0 grocery stores within 30 min. These results will be useful to city leaders and planners in developing strategies to use public transportation to improve food access to residents of food deserts in their cities.

While this study offers suggestions for city leaders for how to best utilize their public transit system to increase food access, the next step is to implement these suggestions in a city's model to see the impact it would have on the centrality measures. Further, we could offer more impactful suggestions for how to modify public transportation by conducting a sensitivity analysis between food deserts that have low vehicle access and those that do not. This would offer insight into where public transit plays a more important role in food access and which areas need to be prioritized for transit improvements, as people who have reliable vehicle access are less likely to rely on public transportation to reach grocery stores. More generally, this work could help develop a new definition of food deserts that takes public transportation into account. Right now, even if a food desert has adequate access to grocery stores through public transit, it is still considered a food desert by the original definition. Thus, by having a definition that takes public transit into consideration, cities will be able to more easily identify census tracts that are most isolated from food access and focus their resources and support where it is needed most. In addition to studying access to grocery stores, network models like this can be employed by city leaders to study the role public transportation plays in accessing other essential services and resources, such as employment, medical care, and education. While network modeling has often been applied to study public transportation and other model types (such as agent-based models) have been used to study individuals' behaviors around public transit use and food access, our research combines these two ideas and uses network modeling to explore food access through public transit. This novel approach to understanding the role public transportation plays in food access allows for suggestions to be made to city leaders on how to best utilize their public transit system to help alleviate the burden of food deserts.

Disclaimer

The views expressed in this article are those of the author(s) and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

Author statement

Anna Sisk: Conceptualization, Methodology, Formal analysis, Writing-original draft, Writing-review and editing; Kristen Rappazzo: Writing-review and editing; Tom Luben: Writing-review and editing; Nina Fefferman: Conceptualization, Methodology, Writing-review and editing, Supervision.

Financial disclosure

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

None.

Data availability

Data will be made available on request.

Acknowledgements

The authors would like to thank Drs Anne Weaver and Nichole Kulikowski for their review and comments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jth.2023.101627.

References

Abel, K.C., Faust, K.M., 2020. Modeling complex human systems: an adaptable framework of urban food deserts. Sustain. Cities Soc. 52, 101795.
Ahangari, S., et al., Understanding Access to Grocery Stores in Baltimore City.
Albuquerque, C.o., streets. https://www.cabq.gov/.
Bureau, U.S.C, 2020. QuickFacts. Available from: https://www.census.gov/quickfacts/fact/table/raleighcitynorthcarolina,charlottecitynorthcarolina,albuquerquecitynewmexico,browndeervillagewisconsin,lawrencecitykansas/PST045221.
Chavis, C., Jones, A., Center, E., 2020. Understanding Access to Grocery Stores in Food Deserts in Baltimore City.
Coleman-Jensen, A., et al., 2021. Household Food Security In the United States in 2020. Economic Research Service.
Control, C.f.D. and Prevention, 2014. Healthier Food Retail: an Action Guide for Public Health Practitioners. US Department of Health and Human Services.
County, W., streets in wake county, NC. https://data-ral.opendata.arcgis.com, 2021.
Create and use a network dataset with public transit data. Available from: https://pro.arcgis.com/en/pro-app/latest/help/analysis/networks/create-and-use-a-network-dataset-with-public-transit-data.htm#.
Derrible, S., Kennedy, C., 2009. Network analysis of world subway systems using updated graph theory. Transport. Res. Rec. 2112 (1), 17–25.
Development, C.o.M.D.o.C., 2019. In: C.o.M.D.o.C. Development (Ed.), Milwaukee Fresh Food Access Report.
El-Geneidy, A., et al., 2014. New evidence on walking distances to transit stops: identifying redundancies and gaps using variable service areas. Transportation 41 (1), 193–210.
foodaccess2019 (MapServer), 2019. ArcGIS REST services directory. Available from: https://gis.ers.usda.gov/arcgis/rest/services/foodaccess2019/MapServer.
Fortin, P., Morency, C., Trépanier, M., 2016. Innovative GTFS data application for transit network analysis using a graph-oriented method. Journal of Public Transportation 19 (4), 2.
Gundersen, C., Ziliak, J.P., 2015. Food insecurity and health outcomes. Health Aff. 34 (11), 1830–1839.
KU-LTS-GTFS. Available from: https://lawrencetransit.org/developer/.
Lämmer, S., Gehlsen, B., Helbing, D., 2006. Scaling laws in the spatial structure of urban road networks. Phys. Stat. Mech. Appl. 363 (1), 89–95.
Lawrence, C.o., streets. https://lawrenceks.org, 2021.
Loosemore, B., 2019. How can cities end food deserts? Here are 4 solutions that worked. in Courier Journal. https://www.courier-journal.com/story/news/local/2019/01/10/innovative-solutions-ending-food-deserts/2512232002/.
Murai, S., Yoshida, Y., 2019. Sensitivity analysis of centralities on unweighted networks. In: The World Wide Web Conference.
National Research, 2009. C., in the Public Health Effects of Food Deserts: Workshop Summary. National Academies Press (US) Copyright 2009 by the National Academy of Sciences, Washington (DC). All rights reserved.
Office, M.C.L.I., street centerlines. https://gis-mclio.opendata.arcgis.com, 2018.
OpenMobilityData.
Ploeg, M.V., Nulph, D., Williams, R., 2011. Mapping Food Deserts in the United States. Amber Waves.
Saidi, S., et al., 2017. A generalized framework for complex urban rail transit network analysis. Transportmetrica: Transport. Sci. 13 (10), 874–892.
Transportation, N.C.D.o., road data. https://connect.ncdot.gov.
USDA, 2021a. Go to the Atlas. 2021 tuesday. Available from: https://www.ers.usda.gov/data-products/food-access-research-atlas/go-to-the-atlas/.
USDA, 2021b. Documentation. Food environment Atlas friday. Available from: https://www.ers.usda.gov/data-products/food-environment-atlas/documentation/.
USDA, 2021c. Documentation. Food access research Atlas 2021 monday. Available from: https://www.ers.usda.gov/data-products/food-access-research-atlas/documentation/#definitions.
Ver Ploeg, M., et al., 2015. Where Do Americans Usually Shop for Food and How Do They Travel to Get There? Initial Findings from the National Household Food Acquisition and Purchase Survey.
Wan, Z., et al., 2021. A survey on centrality metrics and their network resilience analysis. IEEE Access 9, 104773–104819.
Widener, M.J., Metcalf, S.S., Bar-Yam, Y., 2013. Agent-based modeling of policies to improve urban food access for low-income populations. Appl. Geogr. 40, 1–10.