scientific reports

OPEN

Limitations and considerations of using composite indicators to measure vulnerability to natural hazards

Celine Wehbe & Hiba Baroud

Vulnerability assessment plays a critical role in disaster management and requires the consideration of multiple dimensions that involve both the built and social environments. A common approach to address this problem is the use of composite indicators, which offer a simplified method to combine information across different dimensions and facilitate decision making. However, composite indicators present limitations in the context of hazard vulnerability. This study investigates the source of these limitations and provides ways to overcome shortcomings in the interpretation of composite vulnerability indicators. To conduct this investigation, a composite indicator is developed to assess the vulnerability of power and transportation infrastructure, while considering social vulnerability, to capture community hazard vulnerability. Using a case study of Harris County in Texas, we investigate the disparities in outcomes resulting from different calculation methods, such as sub-indicator weighting. The case study shows that the value of the indicator is not consistent across different calculation methods. Additionally, weighting the sub-indicators plays an important role in the value of the indicator. Combining infrastructure and social factors is found to be misleading in the interpretation of hazard vulnerability, and the use of bivariate maps is proposed to better distinguish between infrastructure and social vulnerabilities.

Different stability concepts have been used to analyze system behavior and characteristics under disruptions. These concepts include robustness, vulnerability, resilience, and risk. Although these measures are often mistakenly used interchangeably, they reflect distinct aspects of system behavior under disruption. Robustness refers to the ability of a system to maintain its performance in the face of disruptions[1], whereas resilience includes its ability to respond to and recover from these disruptions[2]. Risk considers the probability of a disruption occurring as well as its impact[3]. Vulnerability is defined as the susceptibility to incidents which results in considerable reduction in serviceability[4]. This study focuses on the evaluation of community vulnerability to natural hazards. The vulnerability of a community is considered to be a function of infrastructure and social vulnerability. It is worth noting that there are two types of social vulnerabilities, the first type relates to the provision of infrastructure services (e.g., limited access to critical services due to flooded roads), and the second type relates to the circumstances and social conditions of people that might affect their capacity to respond to and recover from disasters (e.g., poverty, disabilities, and older age, among others). This study considers the social vulnerability associated with the service provision of infrastructure systems to be embedded in the infrastructure component of community vulnerability. As such, this study uses the term *social vulnerability* in reference to the second type, which measures the social conditions and demographics of people influencing their capacity to respond to disasters.

Given that the vulnerability of communities to natural hazards is linked to both infrastructure and social factors, it is crucial to address both aspects in disaster preparedness[5]. Public infrastructures, such as power and transportation, are highly vulnerable to natural hazards. For example, heavy rain can accumulate on impermeable surfaces and cause floods, resulting in road closures[6]. During hurricanes, strong winds can cause trees to topple over electric lines, which results in power outages[7]. These disturbances interfere with emergency response decisions and actions. For example, the access of emergency responders to individuals in need of assistance and to critical facilities becomes limited when mobility is disrupted. Moreover, power interruptions restrict the communication of important information and evacuation strategies. As such, vulnerability of critical infrastructure must be quantified and communicated to decision makers to support disaster preparedness. Understanding the

Department of Civil and Environmental Engineering, Vanderbilt University, Nashville, TN 37235, USA. email: hiba.baroud@vanderbilt.edu

social context in which critical infrastructure operates (i.e., demographic and economic[8]) is equally important to identify how infrastructure vulnerability impacts different groups of the population, particularly disadvantaged groups. For example, highly dense areas require careful planning to evacuate. Elderly people or individuals with disabilities need more time and assistance. Individuals who fall below the poverty line lack resources to respond to disasters[9]. Social factors heavily influence the vulnerability of communities when faced with disruptions. They play a crucial role in developing effective resilience strategies, which is why community vulnerability should be measured using both, infrastructure systems and social systems. Given its multiple dimensions, measuring community vulnerability to natural hazards is challenging. The literature has traditionally approached vulnerability assessment from either an infrastructure-centric or a social-centric perspective, rather than considering both aspects simultaneously[10]. For example, infrastructure vulnerability has been widely studied using different methods. One common way to assess infrastructure vulnerability is qualitatively through experts' knowledge[11]. Such approaches are time-consuming, and specific to the infrastructure being considered, which means that results are difficult to scale and reproduce across infrastructure sectors and geographical regions[12]. Quantitative methods, such as physics-based models and data-driven models, have also been used. Physics-based models, similar to the work by Yang et al.[13], are driven by processes often described by a set of mathematical equations[14]. They produce consistent results due to the nature of the underlying physics, but they are computationally intensive. They also cannot provide assessments on models from different sectors and are difficult to generalize across different hazards. On the other hand, data-driven models use information from observed data to identify the characteristics and patterns of a system behavior or phenomenon without necessarily accounting for the underlying physical processes. These methods require high dimensional data to obtain accurate results[15]. In the case of transportation systems, several studies have developed data-driven methods to evaluate transportation infrastructure vulnerability, with a strong focus on flood events[16]. In the case of power systems, data-driven methods for power infrastructure vulnerability have focused on predictive modeling of power outages that are founded in statistical learning to predict the number of customers without power[17]. Among data-driven approaches, geostatistical analyses of vulnerability have been explored to capture spatial trends in vulnerability assessment using correlation analyses and spatial clustering. Examples of such studies select variables related to the phenomena to be measured and use statistical tools like Global Moran's I and Getis Ord Gi Hot Spot Analysis to analyze the degree of spatial autocorrelation between variables[18]. While effective in illustrating spatial trends, such methodologies lack the capacity for temporal analysis necessary to monitor vulnerability over time. Moreover, these approaches become more complex and less interpretable as more variables are considered. In the context of social vulnerability, data-driven analyses are primarily used to examine how socioeconomic factors and community characteristics affect a population's vulnerability to risks and disasters. For example, Chakraborty et al.[19] examine spatial variability in evacuation assistance needs during natural hazards. They develop a geophysical risk index, as well as a social vulnerability index using census information and flood data.

While existing research has made significant progress in assessing vulnerability to natural hazards, several gaps remain to be addressed. These research gaps include (1) the need for methods that scale vulnerability assessment across multiple systems and large geographic regions, (2) the integration of infrastructure and social dimensions of community vulnerability to natural hazards, and (3) the critical analysis of the considerations and limitations in the calculation and interpretation of these composite indicators. To address these gaps, this study develops a composite indicator that combines infrastructure (power and transportation) and social factors (e.g., poverty, age, disability), and conducts a critical analysis of different indicator construction methods to assess considerations and limitations of using composite indicators in the context of hazard vulnerability.

Hazard vulnerability assessment requires the integration of multiple factors which are often provided in different units and at different spatial scales. For example, demographic data is provided at the census tract scale whereas the National Land Cover Database (NLCD), offers data on factors that influence infrastructure vulnerability such as vegetation density or impervious surface density at a 30-m resolution. For the vulnerability assessment to be useful and help inform disaster planning, methods are needed that unify spatial resolutions as well as the units of different infrastructure and social vulnerability factors, to allow the composite indicator to combine information across multiple systems and scale up to measure vulnerability at a regional scale. A holistic and scalable approach that integrates infrastructure and social dimensions of hazard vulnerability across sectors and regions not only results in more efficient planning by providing a comprehensive understanding of potential risks and weaknesses, it also facilitates long-term planning and encourages inter-regional cooperation, promoting concerted action in disaster risk reduction initiatives.

The use of one composite indicator simplifies the interpretation of the results by condensing the information into a single figure, rather than comparing multiple variables with different scales and units. Composite indicators are formed by combining individual sub-indicators that describe different features of a system and integrating them into a single indicator[20]. The sub-indicators can be weighted to account for different levels of contribution to the overall composite indicator. In the field of vulnerability, composite indicators are more commonly studied in the context of social vulnerability[21] and economic vulnerability[22], but much less in infrastructure vulnerability. In this study, we specifically examine transportation and power infrastructure. In the case of transportation systems, El Rashidy et al.[23] develop a composite index to measure the resilience of road transport networks under disruptive events. The authors use scenario implementation to obtain the sub-indicators, instead of collecting observed data which may limit the generalization of the approach. There are no studies that use indicators to evaluate the vulnerability of power systems. In contrast, the application of composite indicators to social vulnerability is more common. The most widely known is the Social Vulnerability Index by Cutter et al.[8], where the authors used a factor analytic approach to compute a social vulnerability summary score. When considering the integration of both infrastructure and social factors in hazard vulnerability assessment, Müller et al.[24] were among the first to integrate social vulnerability with infrastructure vulnerability using an indicator-based approach. They combine physical and social vulnerability by taking a subjective approach. The authors use a questionnaire to determine sub-indicators weights, using inputs from both, households and experts, which provides a comprehensive assessment of weights but limited scalability and generalization. Other studies consider an objective approach by either assuming equal weights for all sub-indicators[9] or considering weights to be proportional to the inverse of the variance of sub-indicators[25]. While very few studies explored the integration of infrastructure and social factors in developing hazard vulnerability indicators, no prior studies investigated the considerations and limitations of constructing such indicators and interpreting them.

In fact, combining information from multiple systems and sources into one value to describe the overall hazard vulnerability may lead to challenges in the interpretation. First, composite indicators are highly dependent on the weights assigned to sub-indicators. The literature presents multiple ways to combine sub-indicators into one composite indicator[26]. Different weighting methods will produce different results, which affects the final value of the composite indicator and significantly influences decision making[27]. Second, prior work on the inclusion of both infrastructure and social factors in the computation of a single indicator presents advantages in addressing the multi-dimensionality of vulnerability[25] but does not investigate how the combined information can be used in evaluating the risk of disasters. For example, if equal weights are considered, a high infrastructure vulnerability with a low social vulnerability could be treated similarly to a low infrastructure vulnerability and a high social vulnerability due to similarities in the numerical value of the composite indicator. As such, there is a need to investigate the considerations and limitations of using a composite indicator approach to quantify hazard vulnerability.

This study develops an indicator for hazard vulnerability to analyze the limitations and considerations of using composite indicators in the context of vulnerability assessment. The study area is in Harris County in Texas, and the composite indicator combines infrastructure dimensions (power and transportation) and social dimensions (e.g., poverty level, age, disability) to measure hazard vulnerability. The effect of weighting is analyzed by examining and comparing different weighting methods for the infrastructure vulnerability indicator. Then, the composite indicator value is analyzed and compared when various social factors are included. The outcome of this research can shed light on how different weighting methods influence the value of the composite indicator. It also emphasizes the importance of careful consideration when including multiple dimensions of vulnerability, specifically infrastructure and social factors. The outcome of the study provides a recommendation for alternative methods such as bivariate maps to distinguish these dimensions without compromising the value of combining them in the vulnerability assessment.

Results

The development and analysis of the composite indicator is demonstrated using a case study in Harris County, Texas with a focus on power and transportation infrastructure. Although the analysis shown focuses on a particular study area, our methodology is general and uses public data sources which makes it possible to reproduce in other areas.

Study area

As an area with a flat terrain that barely rises above sea level, Harris County is vulnerable to floods. Its environmental conditions such as humid climate and clay soils, in addition to its growing population and aging infrastructure, make it a target for climate-driven disasters including flooding and hurricanes[28]. In 2005, Houston, the main city in Harris County, was struck by Hurricane Rita, which caused extensive power interruptions all over the county that lasted for 6 days[29]. In 2008, Hurricane Ike swept through the city with high winds and strong rainfalls. The storm surge levels averaged near the 100-year levels for Harris County[30]. In 2017, Houston was hit by the costliest and most damaging tropical cyclone yet, Hurricane Harvey. It caused catastrophic flooding, leading to multiple deaths and widespread infrastructure failures[31]. In addition to the environmental stressors that impact infrastructure vulnerabilities, several social vulnerabilities are noted in this area. In 2018, among the three most densely populated counties in Texas, Harris County's Social Vulnerability Index (SVI) stood at 0.72, the highest among these counties. Nearly 3 million people, equivalent to 64% of the county population, reside in a census tract that exhibits greater vulnerability than half of all census tracts across the country[32]. Moreover, the highest deficit in shelters is noted in Harris County[33]. Data from the Federal Emergency Management Agency Resilience Analysis and Planning Tool show that the highest social vulnerability factors, such as the percentage of elderly, disabled, poor, and those with limited English proficiency, are all concentrated in the center of the area. With the increasing frequency and intensity of natural hazards combined with the inherent social and infrastructure vulnerability, there is a critical need to better understand the vulnerability of this community to natural hazards. This case study focuses on the power and transportation infrastructure given historical trends of power outages and roads flooding during disasters. In Texas specifically, the state has its own electric grid, making it impossible to access power from other states and increasing the importance of understanding its vulnerability and impact on communities[34]. Power and transportation infrastructure play a critical role in the response to and recovery from disasters by providing essential services to vulnerable communities. While the focus of the case study is in Harris County, TX, the analysis can be generalized and applied to other regions. The study is conducted at the census tract level using publicly available data, to provide a sufficiently high resolution for emergency management and communication. Harris County, which includes the city of Houston and a number of neighboring communities, is divided into 786 census tracts.

Development and analysis of the composite indicator

The steps involved in the construction of a composite indicator include (1) the selection of sub-indicators that represent different vulnerability dimensions, (2) the choice of a weighting method to allocate weights for each sub-indicator, and (3) the choice of an aggregation method to aggregate all weighted sub-indicators into a composite indicator. The selected sub-indicators are standardized to a common scale using normalization. Each sub-indicator is then assigned equal weights, or weighted using subjective methods, such as Budget Allocation Processes (BAP), or objective methods, such as Principal Component Analysis (PCA)[26]. Finally, the composite indicator is obtained by aggregating the weighted indicators, often through simple linear and geometric aggregation, or through more complex non-compensatory methods.

The first objective of this study is to assess how various weighting methods influence the final value of the composite indicator. As such, the composite indicator is first constructed by aggregating the infrastructure vulnerability sub-indicators using different weighting methods (details on the approach are provided in the "Methods" section). Then, the integration of the social dimension is studied to address the second objective of the study, which is to investigate considerations and limitations of using composite indicators in the context of hazard vulnerability. To address this objective, two approaches are employed and compared. In the first approach, the social factors are included as sub-indicators in the calculation of the overall composite indicator. In the second approach, bivariate maps are used to examine the infrastructure and social dimension of hazard vulnerability. A summary of the results is provided in the following subsections. The complete set of results can be found in the Supplementary Information.

Composite indicator for infrastructure vulnerability

To construct the composite indicator for infrastructure vulnerability, the sub-indicators are selected based on environmental factors that influence power and transportation vulnerability to natural hazards. The rationale behind the sub-indicators selection is explained in the "Methods" section. Sub-indicators for transportation infrastructure vulnerability were informed by prior studies in the literature that developed transportation vulnerability indicators[9,35,36]. Given that no prior work developed an indicator for power vulnerability, the selection of sub-indicators for the power infrastructure vulnerability were introduced by the authors and inspired by prior predictive models of power outages[37]. For example, in the case of above-ground power lines, sub-indicators for power infrastructure vulnerability include soil properties related to tree stability (e.g., soil moisture), and sub-indicators for transportation infrastructure vulnerability include surface properties related to permeability (e.g., impervious density). The data associated with these sub-indicators are collected for Harris County, TX. The details of data preparation can be found in the "Methods" section and the Supplementary Information. This section reports on the effect of weight assignment on the composite indicator which significantly influences the value of the composite indicator. The weighting methods used consider (1) equal weights, and (2) PCA-generated weights, obtained by using the loadings generated from PCA (more details can be found in the "Methods" section).

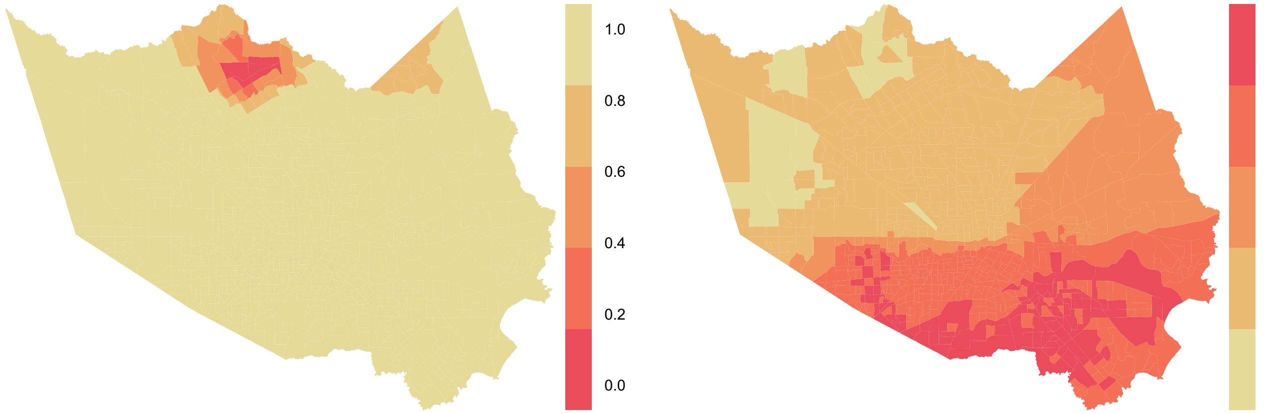
We can see from the weights in Table 1 that PCA in the power dimension assigns higher weights to sub-indicators of soil moisture and Soil Potential Index (SPI), and almost zero weight to canopy and root zone. For transportation, the PCA weights are more evenly distributed, with the lowest weight being assigned to wetland density, and the highest to impervious density.

These results can be explained by examining the individual maps of the spatial distribution of the sub-indicators. All the figures in the manuscript and the Supplementary Information were developed by the authors using the ggplot2 library in R, version 4.3.0 (https://www.r-project.org/). Figure 1 shows the spatial distribution for the sub-indicators representing root zone and soil moisture around Harris County. The root zone is the extent to which tree roots extend in the soil, and the soil moisture is the moisture content of that soil. PCA is a method that estimates factor loadings based on the maximum variability. When examining these sub-indicators, it can be seen that the root zone has little geographic variability compared to the soil moisture. For this reason, not much information can be extracted from the sub-indicator describing root zone, and hence, PCA assigns a lower weight for that sub-indicator. For the transportation sub-indicators, there is more variability across all sub-indicators, with the least one being for wetland density, which is assigned the lower weight.

Once the sub-indicators have been weighted, they are linearly aggregated to obtain the final composite indicator value for each census tract within Harris County. The composite indicator values are displayed using cartographic representation to visualize the results in a geographic context, as this facilitates the recognition of spatial patterns. Equal intervals are chosen to display the different values in order to show the differences between method outcomes. Power vulnerability maps show different patterns for each method, as shown in Fig. 2. The complete set of maps can be found in the Supplementary Information.

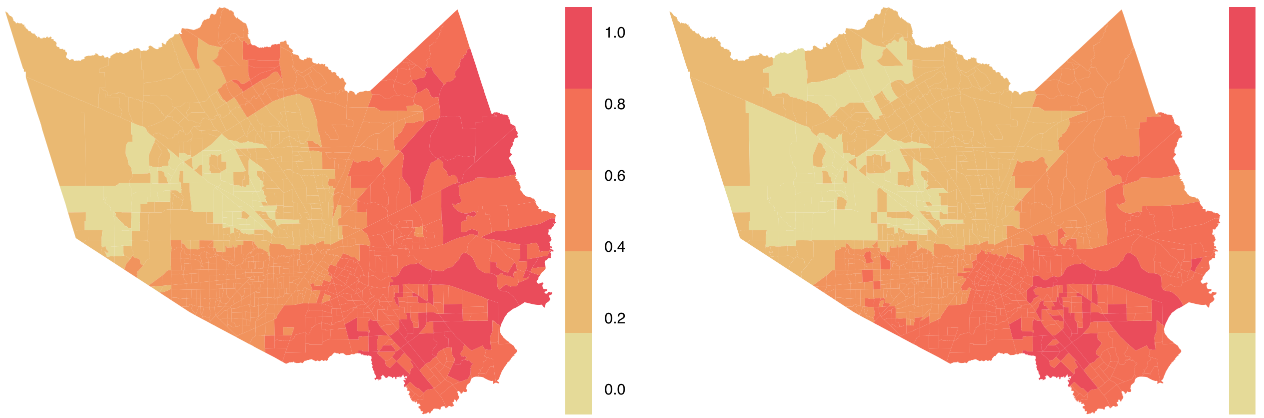
Power vulnerability sub-indicators				
	Canopy density	**Soil moisture**	**SPI**	**Root zone**
Equal	0.25	0.25	0.25	0.25
PCA	0.083	0.586	0.280	0.051
Transportation vulnerability sub-indicators				
	Wetland density	**Vegetation density**	**Impervious density**	**Closed roads**
Equal	0.25	0.25	0.25	0.25
PCA	0.174	0.271	0.307	0.248

Table 1. Calculated weights for equal and PCA weighting methods.



(a) Root Zone **(b)** Soil Moisture

Figure 1. (**a**) Root zone and (**b**) soil moisture have varying spatial distribution across Harris County, TX. Low vulnerability is indicated in yellow, high vulnerability is indicated in red.



(a) Power vulnerability (Equal weights) **(b)** Power vulnerability (PCA weights)

Figure 2. Power vulnerability indicator is different using (**a**) equal weights, and (**b**) PCA weights.

Equal weighting results in higher power vulnerability in East Harris County, while the vulnerability is shown to be higher in the South under the PCA method. These differences are due to the varying weights the methods offer, in combination with variations of the sub-indicator values geographically across the study area. For example, the power vulnerability using PCA weights follows a very similar pattern of vulnerability as soil moisture, which exhibits the most geographic variability between the sub-indicators. On the other hand, all the sub-indicators in the power vulnerability using equal weights contribute equally to the output. The composite indicator's value is highly dependent on its sub-indicators and varies greatly depending on the weighting methods applied. Therefore, using different weighting methodologies can lead to divergent results, introducing inconsistencies in prioritizing emergency preparedness and response. These discrepancies can mislead decision-makers and hinder the optimal allocation of resources and the implementation of effective emergency plans. To overcome these ambiguities, subjective weighting methods can be used to incorporate preferences from stakeholders, and guide results based on their prioritized needs. Weighting the sub-indicators therefore has a large effect on the final composite indicator, as the method of weighting treats data differently.

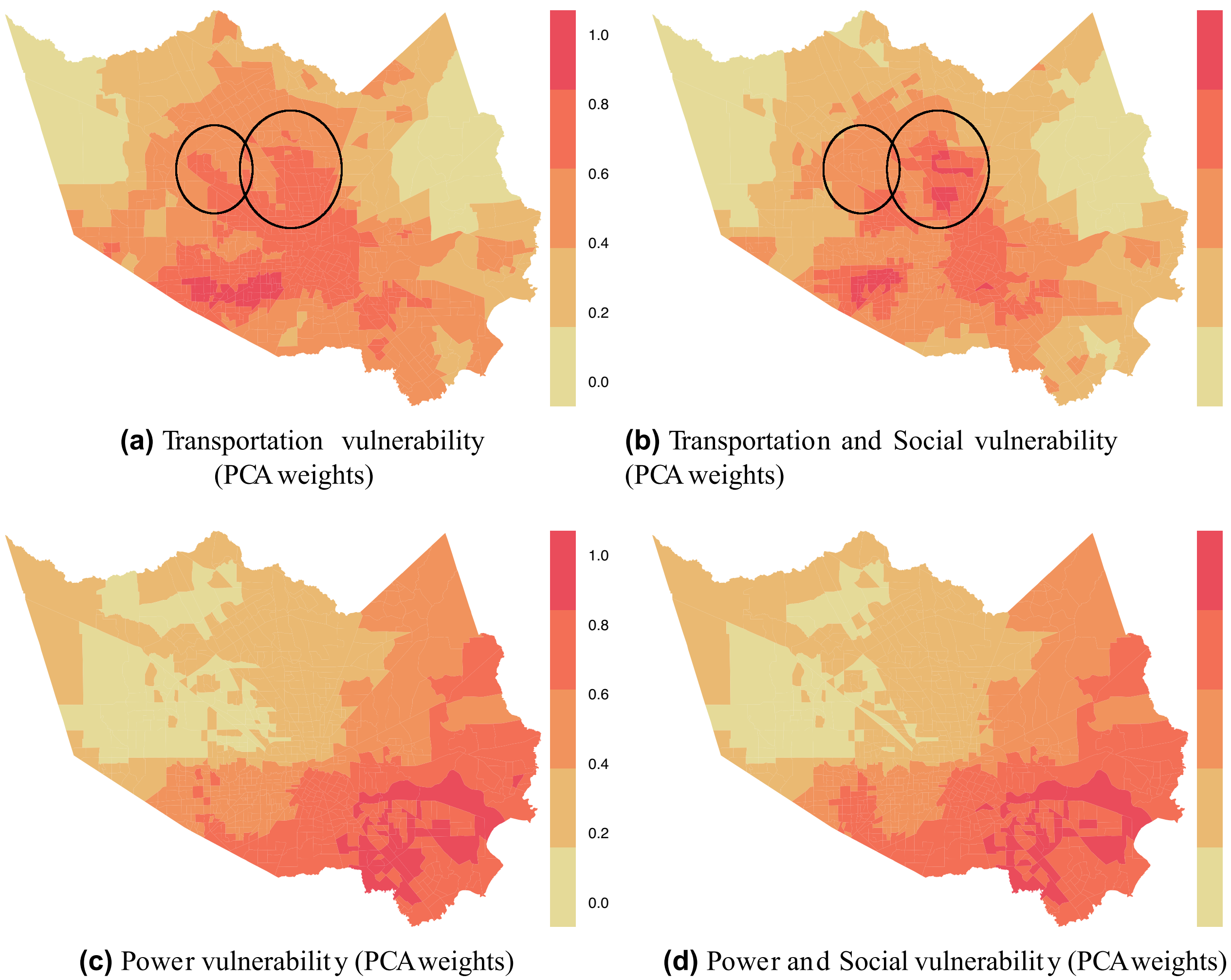
Integration of the social dimension

While composite indicators of infrastructure vulnerability can indicate where hazards are most likely to impact infrastructure, it is essential to recognize that vulnerability is not solely determined by physical infrastructure characteristics, and is also shaped by social factors. For example, the presence of marginalized communities in flood-prone areas may amplify vulnerability due to limited resources, inadequate evacuation plans, or reduced of access to information. These groups often lack the financial means to prepare for floods or evacuate effectively, compounded by inadequate transportation options and insufficient access to critical information. Without targeted support and inclusive planning, these disparities exacerbate the risks faced by marginalized populations during flood events. The social dimension highlights the importance of community awareness and preparedness in responding to natural hazards. This section presents the considerations and limitations of incorporating this dimension to infrastructure vulnerability. As a first step, the selected social factors are all included in the overall composite indicator as sub-indicators, along with the infrastructure sub-indicators, an approach that has been adopted in multiple studies in the literature[9,25]. These social factors include variables that describe populations that may be at higher risk of the adverse impacts of a natural hazard. The social factors added to the overall composite indicator consider people over 65 years old, people with disabilities, people with limited English proficiency, people without vehicles, people below the poverty line, and distance to the closest food hub. The same process described for constructing the infrastructure vulnerability indicator is used and adapted to include social sub-indicators. The complete set of results can be found in the Supplementary Information. The following section compares the infrastructure vulnerability to the overall vulnerability considering the social aspect, and examines the impact of this inclusion as well as the potential misinterpretation it may cause.

Heat maps of the overall vulnerability indicators, for power and transportation dimension, are shown in Fig. 3. Composite indicators considering only the infrastructure dimension are shown in Fig. 3 (a and c), and composite indicators considering both infrastructure and all social dimensions are shown in Fig. 3 (b and d).

For the transportation vulnerability, the general pattern of vulnerability assessment appears to be similar with and without social factors. However, when looking at a specific areas at a higher resolution, the addition of the social factors to the indicators resulted in differences in how overall hazard vulnerability is assessed. For example, the center of Harris County is characterized by a large number of the population below poverty line, population with disability, and population with limited English. This leads to an increase of the overall risk of transportation failure in this area which manifests in a higher overall vulnerability indicator. In contrast, the mid-west areas shows a lower overall vulnerability (after including social factors) even though the vulnerability of transportation infrastructure to failure is higher in this area, Fig. 3 (a). This outcome is due to residents having a low social vulnerability in these areas.

Incorporating the social dimension in a composite indicator along with the infrastructure dimension helps better assess the risk of natural hazards on communities. The aim is to capture the vulnerability of the population and its intersection with the vulnerability of critical infrastructure, which is a crucial factor in assessing the overall hazard vulnerability of a community. However, a closer inspection of the vulnerability maps reveal potentially misleading conclusions on the intersection between infrastructure and social vulnerability using



(a) Transportation vulnerability (PCA weights)

(b) Transportation and Social vulnerability (PCA weights)

(c) Power vulnerability (PCA weights)

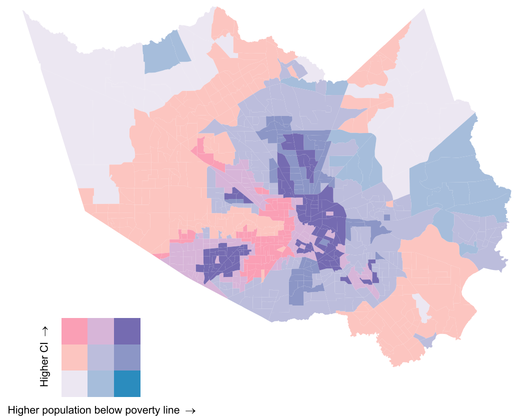
(d) Power and Social vulnerability (PCA weights)

Figure 3. Comparison of the composite vulnerability indicators considering the inclusion of social vulnerability indicators.

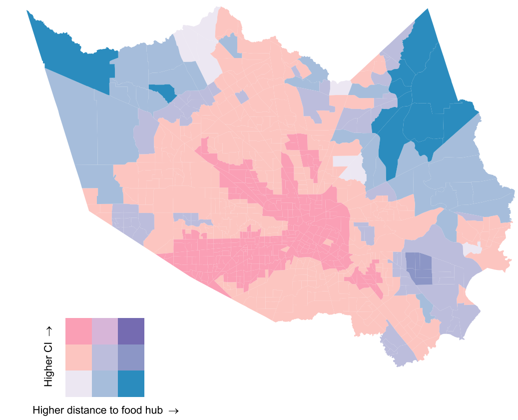
composite indicators. For example, a PCA-weighted power vulnerability composite indicator results in nearly identical vulnerability maps with and without social factors, which could potentially mislead risk managers into considering that social vulnerability has no influence on the overall hazard vulnerability, Fig. 3 (c and d). The rationale behind this outcome is due to the weighting method used. In this case, PCA assigns very small weights to the social indicators, making the changes insignificant. Likewise, a reduction of overall vulnerability after the inclusion of social factors does not necessarily mean that there is no social vulnerability, but could merely mean that the weight assignment shifted the contribution of the sub-indicator.

In order to investigate these variations, we propose the use of bivariate maps to visualize the patterns across different vulnerability dimensions. Bivariate maps help address challenges related to understanding relationships and patterns between two different variables within a geographic context. Figure 4 (a and b) shows the transportation vulnerability against the population below poverty line, and the population distant from food distribution hubs. Figure 4 (c and d) shows the power vulnerability against the population over 65 years old, and the population below poverty line. The complete set of bivariate maps can be found in the Supplementary Information.

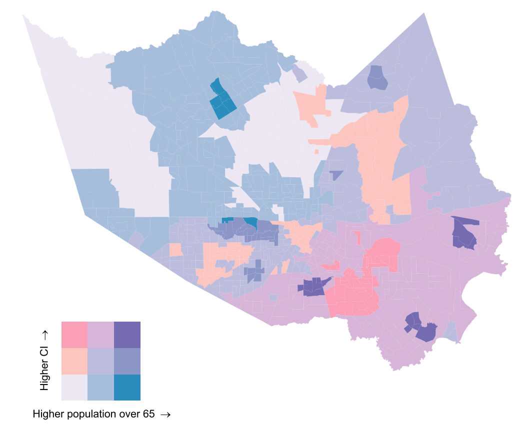
As opposed to what has been previously shown, the overall hazard vulnerability varies greatly depending on the type of social vulnerability considered. Decision makers have to prioritize areas where both the infrastructure and social vulnerability is critical. Bivariate maps help determine the specific needs based on the specific vulnerabilities where infrastructure might be failing in areas of high social vulnerability. These areas are shown in dark purple on the map. Darker blue areas represent areas with the highest social vulnerability, but where



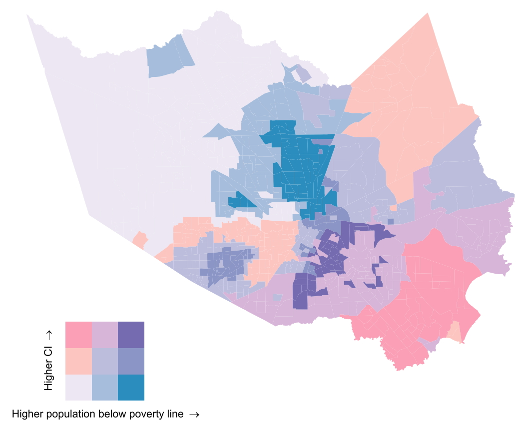
(a) Transportation vulnerability (PCA weights) vs. population below poverty line



(b) Transportation vulnerability (PCA weights) vs. closest distance to food hub



(c) Power vulnerability (PCA weights) vs. population over 65



(d) Power vulnerability (PCA weights) vs. population below poverty line

Figure 4. Bivariate maps can help identify specific areas of vulnerability, such as the areas that combine transportation vulnerability with (**a**) areas with population over 65 and (**b**) areas that are distant from food hubs, and the areas that combine power vulnerability with (**c**) areas with population over 65 and (**d**) areas with population below poverty line.

infrastructure is least likely to fail. Darker pink colors reflect areas where infrastructure is most vulnerable to failure and where social vulnerability is low. We classify three levels of vulnerability as low, moderate and high.

For example, when looking at the overall power and social vulnerability map in Fig. 3 (d), one can conclude that the overall area in Southeast Harris County is vulnerable to hazards when it comes to power infrastructure. However, when we investigate the bivariate maps, Fig. 4 (c and d), we can identify specific areas within the region where elderly and communities below the poverty are at most risk (dark purple) and where emergency preparedness and response efforts should focus. The remaining census tracts in Southeast Harris County are either areas in dark pink, which indicates high infrastructure vulnerability and low social vulnerability (lowest percentage of people below the poverty line) or areas in light purple which represent high infrastructure vulnerability and moderate social vulnerability (moderate percentage of people over 65). These areas are less critical than areas where both infrastructure and social vulnerability are high. Investigating vulnerability this way helps decision makers prioritize specific zones and allocate resources specific to the needs of the population.

Another example is shown in the transportation vulnerability bivariate maps. The overall vulnerability map for transportation and all social factors, Fig. 3 (b), shows a higher vulnerability in the center of Harris County. When looking at each social factor individually, such as the distance to food hubs, Fig. 4 (b), we notice that areas with the greatest concern in terms of accessibility to food hubs are areas with low transportation vulnerability (dark blue areas in northeast and northwest Harris County). Additionally, a few census tracts in southeastern part of the county exhibit a high vulnerability in terms of food hub accessibility and moderate vulnerability of the transportation infrastructure. When looking at the bivariate map of transportation and population below the poverty line, Fig. 4 (a), we notice a larger extent of the overall vulnerability in the middle of the county intersecting with a large number of the population below poverty line. The combined use of composite indicators with bivariate maps that distinguish between different social factors and different infrastructure sectors enables a more strategic disaster preparedness and response that targets specific needs of different group of marginalized communities under different types of disruptions (e.g., power outages, road closure).

Table 2 shows the percentages of census tracts that fall within the high vulnerable category (composite indicator value above 0.8) for overall hazard vulnerability maps, and within the high infrastructure-high social vulnerability bracket for each bivariate map. Based on the overall power vulnerability composite indicator map, approximately 8.4% of the Harris County census tracts are considered highly vulnerable. However, when looking at the corresponding bivariate map, this proportion is not uniform across all the social indicators. For example, none of the census tracts have a combined high power vulnerability and high percentage of people with limited English. However, almost 5% of census tracts fall within the bracket of high power vulnerability and high percentage of people below the poverty line. Similarly, 8% of the Harris County census tracts have high overall vulnerability for transportation and social factors, but none of the census tracts have both high transportation vulnerability and high social vulnerability with respect to distance to food hubs. On the other hand, 13.4% of the census tracts fall within the bracket of high transportation vulnerability and high percentage of people below the poverty line. This discrepancy and generalization of the vulnerability can be misleading to the decision makers and lead to the loss of crucial information. These bivariate maps show that the major social factor concern in areas with vulnerable infrastructure is the population below poverty line.

To summarize, the study reveals that vulnerability is not uniform across the county. While the overall vulnerability maps provide a broad understanding of vulnerable areas, the bivariate maps refine this knowledge, pinpointing specific communities facing increased risks. For the case of Harris County, it is clear that populations below poverty line emerge as particularly vulnerable groups, requiring targeted interventions. Disaster managers can use this information to prioritize resource allocation, directing efforts towards areas with the greatest need. For instance, to help low income households, emergency responders can focus on providing temporary power source such as backup generators for medical equipment during power outages or they can ensure that food is delivered to households when access to food hubs is disrupted due to flooded roads.

Combining sub-indicators of infrastructure and social factors is found to be misleading in the assessment of the vulnerability of a community. Including either one or all of the social factors in the overall composite indicator can mask intricate patterns and lead to the loss of information. When all the indicators are added together, there is a risk that planners might brush off certain areas in favor of others in their strategies. Composite indicators are designed to simplify complex information into a single value. Including social factors often requires reducing multifaceted social dynamics into a single metric, which can oversimplify the patterns of the social phenomena and overlook or misrepresent important factors. Bivariate maps are a valuable tool that enable the proper interpretation of composite indicators in the context of multidimensional hazard vulnerability.

	Bivariate maps						Overall maps
	Over 65	With disab.	Limited English	Without vehicle	Below pov. line	Hub dist.	–
Power	1.908	0.89	0	0.89	4.961	0.763	8.396
Transportation	1.399	1.017	1.145	3.435	13.486	0	8.015

Table 2. Percentages of census tracts that fall within the high vulnerability category.

Discussion

This study has shown the importance of carefully considering the weighting and aggregation methods for composite indicators generally, and the inclusion of social factors and infrastructure vulnerability specifically in the context of natural hazard vulnerability. The results emphasize the complex nature of composite indicators, given their inconsistency in outputs across different methods. Researchers, policymakers, and stakeholders must be cautious of the differences between methodologies and their underlying assumptions. Moreover, it is crucial to highlight that physical and social vulnerability are interconnected elements and require a holistic approach to their analysis. Determining the interaction between social and infrastructure vulnerability improves the allocation of resources by targeting high priority areas in a community. However, a direct incorporation of social factors in the overall composite indicator can be misleading of the overall hazard vulnerability. Overloading the indicator with several social factors makes it challenging to discern specific patterns to address specific needs of different vulnerable groups. The type of resources, efforts and plans depend on different types of vulnerabilities among a population group that is at risk. The policy implications of these findings extend beyond the immediate context of natural hazard vulnerability assessment. By highlighting the nuanced interplay between social factors and infrastructure vulnerability, this study demonstrates the importance of tailoring mitigation and adaptation strategies to the specific needs of communities.

Our results align with prior research that also detected spatial disparities in terms of the vulnerability of specific communities, such as economically underprivileged populations living within areas prone to flooding[9] as well as prior work that found social inequities in post-disaster assistance after hurricane Harvey[38]. In fact, recent work investigates how vulnerable groups are more likely to seek medical care in the aftermath of a disaster, emphasizing the need to differentiate between different aspects of infrastructure and social vulnerability[39]. More generally, prior work agrees with the fact that successful mitigation strategies require the consideration of all aspects of the multidimensional nature of vulnerability, such as the importance of including social sub-indicators along physical sub-indicators as they are interconnected components[9,40].

It is important to acknowledge and address limitations of our approach and analysis in future studies. One important consideration in composite indicators is that the correlation and interdependencies between different sub-indicators and social factors are not explicitly modeled in the aggregation approach. For example, percentage of population below poverty line may correlate with other social factors (e.g., percentage of people without a vehicle, percentage of people with disabilities). This limitation could be addressed by either removing highly correlated variables, or using weighting techniques that prevent the double counting of these variables. Other limitations are concerned with the time-dependent aspects of vulnerability. Composite indicators typically provide a snapshot assessment of the hazard vulnerability rather than a dynamic assessment in response to changing hazard conditions. Infrastructure vulnerability is dynamic and may change as a function of the forecast of natural hazards. While demographic factors may not change dynamically during a disaster, factors such as population shifts, economic fluctuations, or community resilience initiatives can all influence social vulnerability levels leading up to an event. To capture the evolving nature of the environmental conditions, a dynamic approach can be developed to adjust the weights or the choice of sub-indicators based on changing circumstances or the detection of anomalies. This adaptive approach can help maintain the relevance and accuracy of the indicator over time. This dynamic assessment can also account for the changes in human factors related to operators and emergency responders, as the expertise, training, and management of staff can significantly impact infrastructure performance. Finally, the dynamic approach can help resolve the issue of cascading failures across infrastructure sectors. The failure of power infrastructure can lead to cascading effects that impact the transportation system and vice versa. Therefore, composite indicators must be adapted to incorporate time-dependent variations.

Methods

This section describes the methods adopted to develop a composite indicator value for each census tract of Harris County. The process of constructing a composite indicator starts with the selection of sub-indicators that measure the phenomenon to be estimated, on the basis of their relevance, availability, and measurability. The sub-indicator values are collected for each census tract. Once data are acquired and missing values are imputed, normalization and orientation is applied to ensure all the sub-indicators have the same magnitude and direction. In our case study, sub-indicators for infrastructure and social vulnerability were collected for each census tract of Harris County. The data was complete, except for two census tracts which had missing root zone values. The missing values were imputed using the mode of the remaining values, since their quantity was small, and the majority of the data corresponded to the mode. All the sub-indicators were rescaled using a min-max normalization. Wetland density, vegetation density, root zone, and Standard Precipitation Index (SPI) sub-indicators were re-oriented to represent their contribution to the vulnerability (e.g., high wetland density decreases vulnerability). Following the data collection and processing, the sub-indicators are aggregated into one composite indicator. To assign different amounts of contribution for each sub-indicator to the overall composite indicator, a weight must be determined for each sub-indicator. Several subjective and objective methods have been proposed in the literature for weighting and aggregating sub-indicators. This study considers objective methods to facilitate the comparison across calculation and weighting approaches and the interpretation of the composite indicator. As such, equal weighting and PCA weighting methods are considered. The methodologies for each step are further described in the following sub-sections.

Selection of sub-indicators

The selection of sub-indicators is a crucial part of the development of a composite indicator, as different sub-indicators capture different properties of the system vulnerability, and hence produce different results. The selected sub-indicators must reflect the inherent vulnerability of the infrastructure systems and meet specific criteria.

Sub-indicators have to be relevant to the phenomenon being measured, they should be acquired from a trustful source, and ideally, they should be able to be measured consistently over time in order to monitor trends[41].

Composite indicators encompass several dimension of hazard vulnerability including infrastructure (power, transportation) and social (e.g., poverty level, age). Sub-indicators are selected for each dimension based on the criteria mentioned above, and data representing them are collected for each census tract of Harris County. The sub-indicators used in this study are summarized below.

Sub-indicators for power infrastructure vulnerability

In order to determine which sub-indicators are relevant to the power infrastructure vulnerability, we need to determine which factors, in the case of extreme weather events, make the area more vulnerable to power outages. During floods, heavy downpour can cause trees to fall down. Trees can also get overloaded by snow during snowstorm, causing the branches to break and fall. During hurricanes, strong winds tend to knock down trees and poles, which cuts the power lines, leaving households with no electricity. Hence, the sub-indicators are closely related to factors that cause tree or pole instability. The selected sub-indicators for the vulnerability of the power system are listed below.

- *Canopy density.* Densely forested areas are more likely to have their cables damaged during storms. Power lines can get entangled in branches due to strong wind in areas where there are many trees. High canopy density increases the vulnerability of its environment.
- *Soil moisture.* An increase in the moisture content of a soil typically results in the soil weakening and softening. Thus, when trees in saturated soils are subjected to strong wind, their roots are more easily pulled out, resulting in the trees getting blown over. This also decreases pole stability, leading to downed power lines. A high soil moisture content, therefore, increases the vulnerability of its environment. In the case of underground power lines, soil moisture is still a relevant sub-indicator whereby the thermal resistivity of cables in soil significantly decreases when its moisture increases[42]. This increases the risk of electrical faults or failures.
- *Standard precipitation index (SPI).* Drought conditions lead to stress in trees, making them more prone to falling. SPI is an indicator that compares the total amount of precipitation during a certain period of time in an area to the total long-term precipitation in that same area. Negative SPI values can be a sign of drought conditions, and hence low SPI values decrease the vulnerability of the environment.
- *Root zone.* The root zone is the extent to which tree roots spread. Deeper root zones are an indication of more stable trees, as it is harder to uproot them. A high root zone value therefore decreases the vulnerability of its environment.

Sub-indicators for transportation infrastructure vulnerability

In order to determine which sub-indicators are relevant to the transportation infrastructure vulnerability, we need to determine which factors, in the case of extreme weather event, make the area more vulnerable to road closures from flooding. Floods happen during events of heavy rain. When the precipitation is not absorbed, water accumulates on the pavement, causing the roads to overflow. Floods restrict the movement of the vehicles, and interfere with the road network, forcing some roads to close down. Moreover, during snowstorms, impervious pavements are more susceptible to becoming icy, causing accidents, which can also lead to road closure or increased traffic. Hence, the sub-indicators are closely related to factors relevant to the absorption of run-off. The selected sub-indicators for the vulnerability of the transportation system are listed below.

- *Wetland density.* Wetlands have the ability to store flood water from high runoff. They act as natural sponges that soak up the water, which reduces the severity of the flood. High wetland density decreases the vulnerability of its environment.
- *Vegetation density.* Dense vegetation slows down the flow of the water. Rainfall and runoff are captured by the vegetation, leaving less water available to flow over the surface. A high vegetation density therefore decreases the vulnerability of its environment.
- *Impervious surface density.* Impervious surfaces do not allow water to seep into the ground. They increase the rate and volume of runoff, and increase the risk of flood damage. High impervious surface density, therefore, increases the vulnerability of its environment.
- *Percentage of roads vulnerable to floods.* Knowing which roads are more likely to close down during an extreme weather event is an important step of disaster planning and management. It helps emergency response teams identify areas that will be less accessible following the hazard, and which alternative routes to take to get to their destination. The percentage of roads vulnerable to floods is determined by overlaying the 100-year floodplain with the road network of the census tract. A high percentage of roads vulnerable to floods increases the vulnerability of the environment.

Social factors of vulnerability

While it is important to know which areas have vulnerable infrastructure, it is also crucial to assess the specific needs of different groups of people living in those vulnerable areas, as a means to understand the ability of the residents to respond to the natural hazard and provide the necessary steps for proper disaster planning. The selected sub-indicators of population vulnerability are provided in the list below.

- *Percentage of population over 65 years old.* For the elderly, power outages can become a dangerous situation. Electricity is needed to power medical and assistive equipment, and to keep medications from going bad. Elderly also need more time and assistance for evacuation.
- *Percentage of population with disability.* Individuals with mental or physical disabilities required additional time and assistance during disasters. People with chronic illnesses and disabilities rely on electricity to power critical medical equipment, such as power wheelchairs, ventilators, dialysis or other mobility aid or treatments.
- *Percentage of population with limited English.* Individuals with limited English may face difficulties in understanding warning messages and evacuation instructions. Without access to power and internet, it becomes more difficult to receive news and instructions in their native language or to translate evacuation instructions.
- *Percentage of population without a vehicle.* During natural hazards, vehicles are needed for household members to evacuate and move to emergency shelters. Without access to a car, leaving to safer surroundings becomes challenging and more dangerous.
- *Percentage of population below poverty line.* Individuals living below the poverty line are typically located in areas where the infrastructure is not well maintained and prepared to withstand natural hazards, which increases their exposure to adverse effects of disasters. They also lack personal resources to respond to and recover from hazards and infrastructure failure.
- *Distance to closest food hub.* Proximity to food hubs can significantly influence the resilience and preparedness of communities, as shorter distances allow for quicker and more reliable access to food for people in need.

The data for all the sub-indicators were obtained from public sources which are summarized in Table 3.

Data preparation

Once the data of all the sub-indicators are acquired for each census tract, they first need to be processed before applying the weighting and aggregation to develop the composite indicator. Data processing is essential in order to identify outliers, make scale adjustments, and transform highly skewed sub-indicators[26].

Missing values

In many cases, data are only available for some areas, or only for certain sub-indicators. Imputation of the data is essential in order to obtain a complete data set. Depending on the number and pattern of the missing values, there are different methods used for data imputation. These methods fall under three categories: (1) case deletion, (2) single imputation or (3) multiple imputation. When the data are missing completely at random (MCAR), meaning that the cause of the missing data are unrelated to the data, and the amount of cases where the data are missing is small, a simple deletion of the cases is enough. However, this method is avoided as it means that certain census tracts will not be evaluated. Simple single imputation uses the mean, mode, or median of the remaining data to replace the missing values, while slightly more complex single imputation uses regression or clustering, which are more accurate methods. Multiple imputation averages the outcome across multiple imputed datasets[43].

Sub-indicator	Orientation	Data source
Power vulnerability sub-indicators		
Canopy density	(-)	NLCD 2016, Tree Canopy
Soil moisture	(-)	NASA, GES DISC
Standard precipitation index	(+)	NOAA, NIDIS
Root zone	(+)	USDA gSSURGO
Transportation vulnerability sub-indicators		
Wetland density	(+)	U.S. Fish & Wildlife Service, NWI
Vegetation density	(+)	NLCD 2016, Land Cover
Impervious surface density	(-)	NLCD 2016, Percentage Developed Imperviousness
% roads vulnerable to floods	(-)	U.S. Census TIGER/Line FEMA Floodmap
Social factors of vulnerability		
% population> 65 years old	(-)	FEMA RAPT
% population with disability	(-)	FEMA RAPT
% population with limited English	(-)	FEMA RAPT
% population without a vehicle	(-)	FEMA RAPT
% population below poverty line	(-)	FEMA RAPT
% distance to closest food hub	(-)	Houston Food Bank

Table 3. List of sub-indicators and corresponding sources. Acronyms are provided in the Supplementary Information. (+) indicates that a larger value contributes to lower vulnerability, (−) indicates that a larger value corresponds to higher vulnerability.

Normalization
Once a complete dataset is obtained for each sub-indicator, an important step of the construction of a composite indicator is the transformation of the data. Sub-indicators rarely have the same units or scale. Aggregating them in their raw format will allow the difference in magnitude to influence the overall value of the composite indicator. Normalizing the sub-indicator ensures all the sub-indicators have the same magnitude, refraining the composite indicator from favoring sub-indicators with higher values. Min-Max normalization, Eq. 1, is widely used in the field of composite indicators. It results in values that range from zero to one, by shifting the distribution, without changing it. This is a popular method due to its simplicity and ease of understanding. It also guarantees all features will have the exact same scale[44].

$$x_{min-max} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

Orientation
The orientation of the sub-indicators must reflect the correct direction of the composite indicators. Given that a high composite indicator value should indicate high vulnerability, it is expected that all sub-indicators must show high vulnerability at high values. Instead, some sub-indicators have high values that reflect low vulnerability. Thus, they need to be re-oriented[45]. To re-orient a sub-indicator, we use Eq. 2, which takes the complementary value of the sub-indicator.

$$x_{oriented} = 1 - x_{min-max} \tag{2}$$

Weights
One misconception about weights in a composite indicator is that those weights are a measure of the importance of the sub-indicator[46] whereas these weights are concerned with their contribution to the value of the composite indicator. The assigned weight has a big influence on the final indicator value, and this can be used to control the amount of contribution each sub-indicator has on the output. Sub-indicators with a small assigned weight have little impact on the end result, and vice-versa. This study considers two methods for weighting sub-indicators, equal weights and the Principal Component Analysis.

Equal weights
Equal weights, Eq. 3, is a popular method of weight assignment due to its simplicity. Applying equal weights could imply equal importance between the sub-indicators. If there are no empirical grounds or experts' knowledge available to assign importance to the sub-indicators, equal weights assumes all sub-indicators are equally important, and hence all sub-indicators are contributing equally. It could also indicate insufficient knowledge of the statistical relationship between the sub-indicators. It is used in the construction of many sub-indicators. For example, in the Human Development Index, equal weight is assigned to all three dimensions of the composite indicator, as well as to all the sub-indicators of the composite indicator. This is based on the assumption that all human beings value the dimensions equally[47]. It should be noted that equal weights could lead to oversimplification and potential misinterpretation of the weights, as well as make it more prone to double counting when sub-indicators are highly correlated.

$$w_{equal} = \frac{1}{\text{total number of sub-indicators}} \tag{3}$$

Principal component analysis (PCA)
Objective data-driven approaches have been developed to overcome the problem of subjective weighting. These approaches estimate which variables get assigned the highest weights and which get assigned the lowest weights. When presented with a large number of sub-indicators, Principal Component Analysis (PCA) is a popular method that is used, as it can reduce the size of the data, while keeping as much of the variation as possible. Wiréhn introduced a new weighting method that uses PCA to assign weights[48]. This method is not meant to be used for dimension reduction, rather the loadings of the principal components are used to compute the weights. The first step is to conduct PCA on the data set to obtain the eigenvectors and eigenvalues. Then, the loadings of each principal component are obtained. The absolute value of the loadings of the first principal component are applied to the sub-indicators. These loadings are divided by the sum of loadings of the first principal component to achieve a total weight of one, as shown in Eq. 4. Here, the factor loading of sub-indicator i in the first principal component is λ_{1i}.

$$w_{i,PCA} = \frac{\lambda_{1i}}{\sum_{i=1}^{m} \lambda_{1i}} \tag{4}$$

Aggregation
The final step in forming a composite indicator is the aggregation, which combines the values of the set of weighted sub-indicators into a single composite indicator. The two main aggregation methods used are compensatory and non-compensatory[49]. The former allows trade-offs while the latter does not. The existence of trade-offs refers to the fact that poor performance in some sub-indicators can be compensated by high values

of other sub-indicators. In other words, non-compensatory methods output a low score for the case where any sub-indicator has a low value, while compensatory methods can make up the presence of low scores with the presence of high scores. Linear aggregation, Eq. 5, is a compensatory method. It assumes constant trade-offs for all cases. Here, i refers to the index of the sub-indicator and j is the index of the census tract.

$$CI_{j,L} = \sum_{i=1}^{n} x_{ij} w_{ij} \tag{5}$$

In order to compare the results from each method, the output is further scaled using the min-max normalization after the aggregation to obtain a value ranging between 0 and 1, where 0 indicates least vulnerable and 1 indicates most vulnerable.

The final composite indicator is constructed twice. First, the infrastructure sub-indicators are linearly aggregated using equal and PCA weights to determine the effect of weighting, then the composite indicator is constructed again using both infrastructure and social sub-indicators to assess the interconnectedness of the infrastructure and social dimensions. In the second construction of the composite indicator, which includes both infrastructure and social factors, the PCA method was used for assigning weights.

Data availability

The data and the code that support the findings of this study are available from the corresponding authors upon request.

Received: 1 March 2024; Accepted: 19 July 2024
Published online: 20 August 2024

References

1. Dong, S., Wang, H., Mostafavi, A. & Gao, J. Robust component: A robustness measure that incorporates access to critical facilities under disruptions. *J. R. Soc. Interface* **16**, 20190149 (2019).
2. Carlson, J. *et al. Resilience: Theory and Application* (Tech. Rep., Argonne National Lab.(ANL), Argonne, IL, 2012).
3. Ross, R. *Nist sp 800-30 rev. 1: Guide for Conducting Risk Assessments* (National Institute of Standards Technology, 2012).
4. Berdica, K. An introduction to road vulnerability: What has been done, is done and should be done. *Transp. Policy* **9**, 117–127 (2002).
5. Coleman, N., Esmalian, A. & Mostafavi, A. Anatomy of susceptibility for shelter-in-place households facing infrastructure service disruptions caused by natural hazards. *Int. J. Disaster Risk Reduct.* **50**, 101875 (2020).
6. Blum, A. G., Ferraro, P. J., Archfield, S. A. & Ryberg, K. R. Causal effect of impervious cover on annual flood magnitude for the united states. *Geophys. Res. Lett.*https://doi.org/10.1029/2019GL086480 *(2020).*
7. Davidson, R. A., Liu, H., Sarpong, I. K., Sparks, P. & Rosowsky, D. V. Electric power distribution system performance in Carolina hurricanes. *Nat. Hazard. Rev.* **4**, 36–45 (2003).
8. Cutter, S. L., Boruff, B. J. & Shirley, W. L. Social vulnerability to environmental hazards. *Soc. Sci. Q.* **84**, 242–261 (2003).
9. Chang, H. *et al.* Assessment of urban flood vulnerability using the social-ecological-technological systems framework in six us cities. *Sustain. Cities Soc.* **68**, 102786 (2021).
10. Atzl, A. & Keller, S. A systemic approach for the analysis of infrastructure-specific social vulnerability. *Publication Series of UNU-EHS No. 17/2012* 27 (2013).
11. Snedaker, S. & Rima, C. Chapter 4—risk assessment. In *Business Continuity and Disaster Recovery Planning for IT Professionals* 2nd edn (eds Snedaker, S. & Rima, C.) 151–224 (Syngress, Boston, 2014). https://doi.org/10.1016/B978-0-12-410526-3.00004-0.
12. Seppänen, H., Luokkala, P., Zhang, Z., Torkki, P. & Virrantaus, K. Critical infrastructure vulnerability-a method for identifying the infrastructure service failure interdependencies. *Int. J. Crit. Infrastruct. Prot.* **22**, 25–38 (2018).
13. Yang, Y., Ng, S. T., Zhou, S., Xu, F. J. & Li, H. A physics-based framework for analyzing the resilience of interdependent civil infrastructure systems: A climatic extreme event case in hong kong. *Sustain. Cities Soc.* **47**, 101485 (2019).
14. A tale of two approaches: Physics-based vs. data-driven models. (2021) https://jpt.spe.org/twa/author/hector klie. https://jpt.spe.org/a-tale-of-two-approaches-physics-based-vs-data-driven-models.
15. Klie, H. A tale of two approaches: Physics-based vs. data-driven models. (2021) https://jpt.spe.org/a-tale-of-two-approaches-physics-based-vs-data-driven-models.
16. Chen, X.-Z., Lu, Q.-C., Peng, Z.-R. & Ash, J. E. Analysis of transportation network vulnerability under flooding disasters. *Transp. Res. Rec.* **2532**, 37–44 (2015).
17. Nateghi, R., Guikema, S. D. & Quiring, S. M. Comparison and validation of statistical methods for predicting power outage durations in the event of hurricanes. *Risk Anal. Int. J.* **31**, 1897–1906 (2011).
18. García-Ayllón, S., Tomás, A. & Ródenas, J. L. The spatial perspective in post-earthquake evaluation to improve mitigation strategies: Geostatistical analysis of the seismic damage applied to a real case study. *Appl. Sci.* **9**, 3182 (2019).
19. Chakraborty, J., Tobin, G. A. & Montz, B. E. Population evacuation: Assessing spatial variability in geophysical risk and social vulnerability to natural hazards. *Nat. Hazard. Rev.* **6**, 23–33 (2005).
20. Nardo, M. *et al.* Tools for composite indicators building. *Euro. Com. Ispra* **15**, 19–20 (2005).
21. King, D. & MacGregor, C. Using social indicators to measure community vulnerability to natural hazards. *Aust. J. Emerg. Manag.* **15**, 52–57 (2000).
22. Noy, I. & Yonson, R. Economic vulnerability and resilience to natural hazards: A survey of concepts and measurements. *Sustainability* **10**, 2850 (2018).
23. El Rashidy, R. A.H. & Grant-Muller, S. A composite resilience index for road transport networks. In *Proceedings of the Institution of Civil Engineers-Transport*, vol. 172, 174–183 (organizationThomas Telford Ltd, 2019).
24. Müller, A., Reiter, J. & Weiland, U. Assessment of urban vulnerability towards floods using an indicator-based approach-a case study for Santiago de Chile. *Nat. Hazard.* **11**, 2107–2123 (2011).
25. Erena, S. H. & Worku, H. Urban flood vulnerability assessments: the case of dire Dawa city, Ethiopia. *Nat. Hazards* **97**, 495–516 (2019).
26. Commission, J.R.C.-E. *et al. Handbook on Constructing Composite Indicators: Methodology and User Guide* (OECD publishing, 2008).
27. Munda, G. & Nardo, M. Constructing consistent composite indicators: The issue of weights (2005).
28. Associated Press. Why does houston flood so often and so heavily? https://www.nbcnews.com/storyline/hurricane-harvey/why-does-houston-flood-so-often-so-heavily-n796446.

29. National Weather Service. Hurricane rita—post storm report hgx. https://www.weather.gov/hgx/climate_reviews_pshhgx_Rita05.
30. Harris County Flood Control District. Hurricane ike 2008. https://www.hcfcd.org/About/Harris-Countys-Flooding-History/Hurricane-Ike-2008.
31. Blake, E. & Zelinsky, D. *Hurricane Harvey* (Tech. Rep., National Hurricane Center, 2018).
32. Understanding Houston. Vulnerability to and impacts from disasters: Exploring unequal effects of natural disasters on houston communities. https://www.understandinghouston.org/topic/disasters/vulnerability-impacts#overview.
33. Karaye, I. M., Thompson, C. & Horney, J. A. Evacuation shelter deficits for socially vulnerable Texas residents during hurricane Harvey. *Health Serv. Res. Manag. Epidemiol.* **6**, 2333392819848885 (2019).
34. McInerny, C. Why is texas on its own electric grid? (2021) https://www.kut.org/energy-environment/2021-07-22/texas-electric-grid-february-blackouts-the-disconnect.
35. Cutter, S. L. *et al.* A place-based model for understanding community resilience to natural disasters. *Glob. Environ. Chang.* **18**, 598–606 (2008).
36. Beccari, B. A comparative analysis of disaster risk, vulnerability and resilience composite indicators. *PLoS Curr.*https://doi.org/10.1371/currents.dis.453df025e34b682e9737f95070f9b970 *(2016).*
37. Nateghi, R., Guikema, S. & Quiring, S. M. Power outage estimation for tropical cyclones: Improved accuracy with simpler models. *Risk Anal.* **34**, 1069–1078 (2014).
38. Griego, A. L., Flores, A. B., Collins, T. W. & Grineski, S. E. Social vulnerability, disaster assistance, and recovery: A population-based study of hurricane Harvey in greater Houston, Texas. *Int. J. Disaster Risk Reduct.* **51**, 101766 (2020).
39. Rickless, D. S. *et al.* Social vulnerability and access of local medical care during hurricane harvey: A spatial analysis. *Disaster Med. Public Health Prep.* **17**, e12 (2023).
40. Nickdoost, N., Shooshtari, M. J., Choi, J., Smith, D. & AbdelRazig, Y. A composite index framework for quantitative resilience assessment of road infrastructure systems. *Transp. Res. Part D Transp. Environ.* **131**, 104180 (2024).
41. New Zealand Ministry of Social Development. Criteria for selecting indicators—social report 2010. (2010) https://socialreport.msd.govt.nz/2010/introduction/criteria-for-selecting-indicators.html.
42. Czapp, S. & Ratkowski, F. Effect of soil moisture on current-carrying capacity of low-voltage power cables. *Przeglad Elektrotech.*https://doi.org/10.15199/48.2019.06.29 *(2019).*
43. Jadhav, A., Pramod, D. & Ramanathan, K. Comparison of performance of data imputation methods for numeric dataset. *Appl. Artif. Intell.* **33**, 913–933 (2019).
44. Feldmeyer, D., Wilden, D., Jamshed, A. & Birkmann, J. Regional climate resilience index: A novel multimethod comparative approach for indicator development, empirical validation and implementation. *Ecol. Ind.* **119**, 106861 (2020).
45. Hudrliková, L. *et al.* Composite indicators as a useful tool for international comparison: The Europe 2020 example. *Prague Econ. Pap.* **22**, 459–473 (2013).
46. Becker, W., Saisana, M., Paruolo, P. & Vandecasteele, I. Weights and importance in composite indicators: Closing the gap. *Ecol. Ind.* **80**, 12–22 (2017).
47. United Nations Development Programme. Frequently asked questions. https://www.undp.org/sites/g/files/zskgke326/files/migration/tr/UNDP-TR-HDI-IHDI-GDI-GII-MPI-FAQ-EN.pdf.
48. Wiréhn, L., Danielsson, Å. & Neset, T.-S.S. Assessment of composite index methods for agricultural vulnerability to climate change. *J. Environ. Manage.* **156**, 70–80 (2015).
49. Jeffreys, I. The use of compensatory and non-compensatory multi-criteria analysis for small-scale forestry. *Small-scale For. Econ. Manag. Policy* **3**, 99–117 (2004).

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grants No. 2133352 and 1944559.

Author contributions

H.B. conceived the idea and designed the study. C.W. developed the modeling approach and performed the data analysis. All authors contributed to reviewing the results and writing and editing the manuscript.

Competing Interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-68060-z.

Correspondence and requests for materials should be addressed to H.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.