Title

Unseen Risks: Social Vulnerability and the Unsheltered Homeless Under Wildfire

Author

Zheng, Zhongqi

Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Santa Barbara

Unseen Risks: Social Vulnerability and the Unsheltered Homeless Under Wildfire

A Thesis submitted in partial satisfaction of the

requirements for the degree Master of Arts

in Geography

by

Zhongqi Zheng

Committee in charge:

Professor Alan Murray, Chair

Professor Susie Cassels

Professor Stuart Sweeney

December 2024

The thesis of Zhongqi Zheng is approved.

Susie Cassels

Stuart Sweeney

Alan Murray, Committee Chair

December 2024

Unseen Risks: Social Vulnerability and the Unsheltered Homeless Under Wildfire

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Professor Alan Murray, for encouraging me to explore the topics that spark my interests. Thanks for your patient, meticulous, always accessible and enlightening guidance that introduced me into a delightful research journey.

My appreciation also goes to my committee members, Professor Stuart Sweeney and Professor Susie Cassels, for your invaluable comments and suggestions throughout the way.

A heartfelt thank you to my family and friends for your unwavering trust and support during my demotivation moments. I'm also deeply grateful to my cohort and colleagues who accompanied me when I first started graduate school.

Finally, I want to acknowledge the assistance of the Community Services Department in the County of Santa Barbara Housing and Community Development Division, as well as the generous support of Yardi Systems through the Wildfire Resilience Initiative at the University of California, Santa Barbara.

ABSTRACT

Unseen Risks: Social Vulnerability and the Unsheltered Homeless Under Wildfire

by

Zhongqi Zheng

Demographic and socio-economic omissions in social vulnerability assessments are not uncommon. However, conclusions from incomplete population characterization contribute to inequity and marginalization of at-risk segments of society, individuals who require care, resource assistance, and tailored public policy. The unsheltered homeless are considered here as an underrepresented segment of socially vulnerable people in hazard assessment. The unsheltered suffer disproportionately higher levels of direct exposure to natural and human-caused disasters, yet their lack of stable housing and often frequent moves make them both hidden and invisible in traditional census information. An exploratory spatial data analysis framework is developed to combine geographic information systems (GIS), statistics, and geostatistics to assess the implication of this omission. This research treats the integration of annual government counts and continuously updated, community-contributed information as a means of capturing regional spatial patterns of unsheltered homelessness at a fine geographic level. Comparison is then made to approaches based on traditional social vulnerability measures to evaluate consistency and representativeness. The findings suggest a significant gap in existing approaches. This study sheds light on potential avenues for more inclusive and accurate assessments of social vulnerability in the context of hazards.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

Disasters are an unfortunate yet regular occurrence across the globe. O'Keefe et al. (1976) suggest that people are a central concern when a disaster strikes, indicating that socio-economic impacts are not uniform across different segments of society. Subsequent research recognizes that social vulnerability represents the capacity of a community to prepare for, respond to, and recover from the impacts of disaster (Cutter et al., 2003; Flanagan et al., 2011; Spielman et al., 2020). Social vulnerability encompasses a broad spectrum of issues, including economic loss, public health stress, and mortality. For ease of interpretation and application, researchers often transform a multidimensional characterization of vulnerability into a composite score, known as a social vulnerability index (SVI). This index serves as a simplified explanation of complex demographic factors, such as race/ethnicity, age, sex, and household composition, along with other socio-economic and neighborhood conditions that reflect aspects of vulnerability within society. The SVI approach is commonly used to evaluate potential or realized impacts of natural or human-caused hazards at different geographic levels, and drives public policy, mitigation planning, and resource allocation (Cutter et al., 2003; Flanagan et al., 2011; Rufat et al., 2019; Murray et al., 2023). Any misspecification, misinterpretation, or underrepresentation of individuals, however, may be problematic, particularly in high-risk areas. There is a wide range of hazards found in places like California, for example, where floods, drought, earthquakes, storm surges, wildfires, toxic releases, etc. are not uncommon. Of course, a significant concern is the exponential increase in frequency, intensity, and scale of these hazards, particularly wildfire, and robust social vulnerability assessment is lacking (Paveglio et al., 2018; Lambrou et al., 2023; Murray et al., 2023, Pludow, 2024).

Given the significance of social vulnerability and the need to measure and quantify it, derived indices are increasingly subject to questioning and criticism. Much of the concern centers on subjective variable selection and various methodologies in use, thus making indices themselves uncertain (Tate, 2012; Rufat et al., 2019). However, the unifying intent of indexing is to identify vulnerable populations within a community, as well as spatial concentrations of those populations (Muttarak et al., 2015). This underscores frequently overlooked issues of data completeness and appropriateness in the derivation of vulnerability measures. In pursuit of covering broad temporal and spatial scales, social vulnerability measures generally rely on census data or equivalent survey information (Mah et al., 2023). However, this data does not fully account for crucial demographic considerations, namely unauthorized immigrants, non-citizen residents, homeless populations, rural residents, etc. (Bazuin and Fraser, 2013; Meyer et al., 2023).

Among the myriad of groups that are unaccounted for, the unhoused are an especially important segment of the socially vulnerable, perhaps even more so when considering hazards. Homelessness is a challenging issue facing society. Unfortunately, statistics associated with the unhoused are often confusing or unreliable due to their mobility and invisibility. As an example, agencies like the U.S. Census Bureau do attempt to provide counts and estimates of group quarter populations with sub-categories referring to people having lived in emergency shelters, but it has been critiqued for undercounting, especially the unsheltered homeless (Meyer et al., 2023). Further, the reasons for becoming homeless are diverse, including chronic illness, housing affordability, addiction, domestic violence, etc., which makes differentiation and characterization challenging. Thus, available census indicators such as rent/income ratio, housing cost, or group quarter population can only be used to infer the potential for becoming homeless. Existing data generally fails to reflect this particular population, and more broadly, the heterogeneity of vulnerability within society.

The homeless population can be categorized into two groups: people facing housing insecurity and people experiencing homelessness (PEH). They can be further subdivided into sheltered and unsheltered. The unsheltered refers to individuals who live in places not meant for human habitation, such as along streets, in cars, or in temporary makeshift shelters (Montgomery et al., 2016; Vickery, 2018; County of Santa Barbara, 2023). In the case of natural disasters and severe weather, such as wildfire, flooding, storm, etc., the unsheltered have the greatest direct exposure to physical stressors (e.g., safety, fire, heat, smoke, debris, etc.), with significantly less opportunity to be warned, more difficulty evacuating, a higher probability of mortality, and more difficulty accessing healthcare resources (Montgomery et al., 2016; Vickery, 2018; Meyer et al., 2023; Vargo et al., 2023). There is growing recognition that the unsheltered are inhabiting national forests, grasslands, watersheds, and other public lands due to less public resistance, lower costs, and greater availability of resources in these areas (Cerveny and Baur, 2020). This exacerbates their spatial marginality, placing them at the periphery of both social and physical landscapes (Morrow, 1999; Gaillard et al., 2019). As a result, the unsheltered are not only more vulnerable to hazards but also may increase the risk of a hazard event occurring. This is especially true for wildfire hazards where individuals who inhabit fire-prone wildlands both elevate risk of the hazard occurring (unintentional fire ignition), and as a result elevate their own exposure to the hazard (wildfire, wildfire smoke, ash, heat, debris) (Palaiologou et al., 2019; Cerveny and Baur, 2020).

This study considers unsheltered PEH experiencing homelessness to evaluate the enduring marginalization of this group in society. Given the difficulties of measuring vulnerability and the importance of studying the unsheltered, three central research questions drive this research: (1) To what extent do current social vulnerability measures reflect regional unsheltered people experiencing homelessness?; (2) How does available information on unsheltered people experiencing homelessness agree or diverge?; and, (3) What methods can offer a more complete and accurate spatial representation of unsheltered people experiencing homelessness in the assessment of vulnerability? The next section offers background on social vulnerability frameworks in the context of disasters, focusing in particular on wildfire risk. This is followed by a review of spatial analytics that can be applied to examine the relationship between social vulnerability, the distribution of the unsheltered, and wildfire risk. A case study involving the unsheltered in coastal Santa Barbara County is detailed. The thesis ends with a discussion and concluding observations.

Background

Since the United Nations initiated the International Decade for Natural Disaster Reduction in the 1990s, the social vulnerability paradigm has become increasingly central to disaster management strategies (Thomas et al., 2013; Coppola, 2015). Analysis, planning, policy, and decision-making have focused on the identification, measurement, and representation of social vulnerability. Other fields have long developed similar metrics of vulnerability, often referred to as fairness or equity measures. For example, Sloggett and Joshi (1994) and Shouls et al. (1996) used social deprivation and neighborhood profiling to explain the differentiated health outcomes across communities. Murray and Davis (2001) constructed a public transportation need index based on the distribution of demographic information for a region. Harner et al. (2002) reviewed different coefficient measurements that explain the spatial distribution of toxic material exposure in the context of environmental justice. Cutter et al. (2003) established a vulnerability framework in the context of hazards based on factors from U.S. Census data to identify populations facing inequalities or special needs. These factors were then summarized into an index and mapped to highlight communities requiring targeted intervention. Cutter et al. (2024) highlight flooding while notable methodology extensions and refinements include that of Flanagan et al. (2011) and Rufat et al. (2019) examining hurricane impacts.

Wildfires have also been identified as a hazard contributing to social vulnerability. Early vulnerability assessments primarily focused on hazard likelihood, fire intensity, and weather conditions. Only in recent years have the social dimensions of wildfire impacts gained attention (see Thomas et al., 2022; Lambrou et al., 2023). Inspired by Cutter et al. (2003), studies like Gaither et al. (2011), Wigtil et al. (2016), and Palaiologou et al. (2019) have explored the co-occurrence of high social vulnerability and elevated wildfire potential. Pludow (2024) demonstrated social vulnerability change due to wildfire information variability and altered assessment focus.

A hallmark of wildfire social research is its focus on experiences. Studies have highlighted marginalization of vulnerable groups in wildfire planning and emergency management. Davies et al. (2018) examined the unequal pressure faced by race and ethnic minorities in areas with medium or low fire hazard potential when responding to wildfire. Méndez et al. (2020) interviewed undocumented immigrants to reveal their invisibility in wildfire mitigation planning. Wong et al. (2021) studied evacuation behavior through group surveys and found that the homeless receive less social trust and miss out on opportunities for mutual assistance.

Unlike other natural hazards, the ignition, spread and extinguishment of wildfires not only depends on weather conditions but also on parcel or household-level mitigation and resource prepositioning. Hence, efforts to undertake vulnerability mapping have aimed to improve spatial accuracy. Paveglio et al. (2018) integrated household demographic information and risk perception metrics through surveys. More recently, Murray et al. (2023) applied areal interpolation techniques to derive parcel-level social vulnerability metrics from multiple existing indices.

While a range of approaches have been explored for assessing social vulnerability in the context of hazards, a challenge remains for a more precise representation of social vulnerability in disaster management. In particular, consideration of the actual spatial distribution of a particularly vulnerable population is critical.

Data

The study area is Santa Barbara, California. It is bordered by the Pacific Ocean to the south and the Santa Ynez Mountains to the north, with a range of geographic characteristics that make it susceptible to most natural hazards, including earthquakes, flooding, drought, tsunamis, landslides, storm surge, extreme heat, and wildfire. This too is characteristic of much of southern California. Climate change has exacerbated and amplified these hazards, making them more frequent and more extreme. The Santa Barbara region is home to a vast vulnerable population, including senior citizens (16% of the population), renters (35% of the households), people without a high school degree (20% of the population) as well as the unhoused, making vulnerability assessment and monitoring a high priority (County of Santa Barbara, 2021a).

Social vulnerability is reported and characterized by the Centers for Disease Control and Prevention (CDC) (2020) using an SVI, and is examined in this study. The CDC calculates the SVI by the percentile rank method, i.e., order and rank tracts based on the sum of the individual percentile ranking for each input variable (Flanagan et al., 2018; Mah et al., 2023). As a result, the index ranges from 0 to 1, with higher values indicating a relatively more vulnerable area (see Figure 1). CDC produces two SVI estimates by Census tract, one a comparison of Census tracts in the corresponding state and the second based on all Census tracts across the nation (CDC, 2022). Other SVI estimates also exist, including at the Census block group (Bryan 2022), 1 × 1 km grids across the US (NASA SEDAC 2023), and individual parcels (Murray et al. 2023), and are based in part on CDC SVI. These SVI estimates are summarized in Table 1.

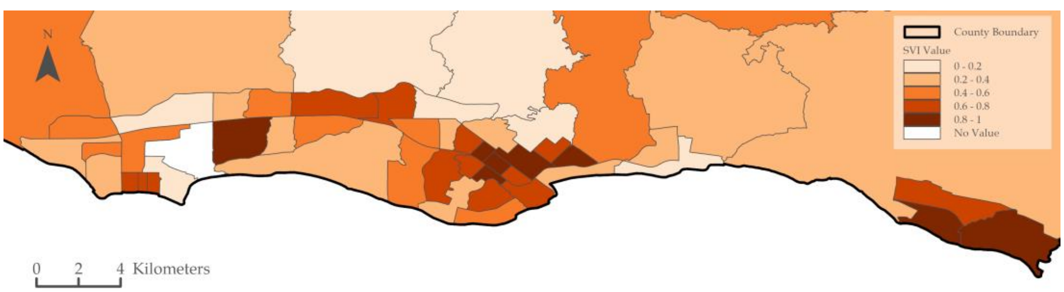


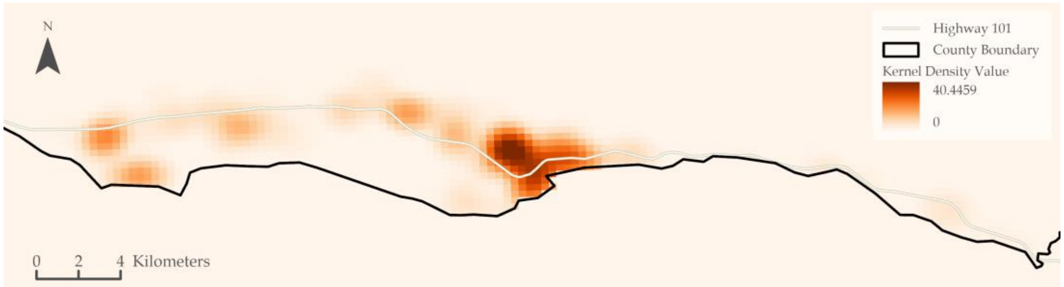
Figure 1. CDC SVI at census tract level

Table 1. Geographic SVI estimates

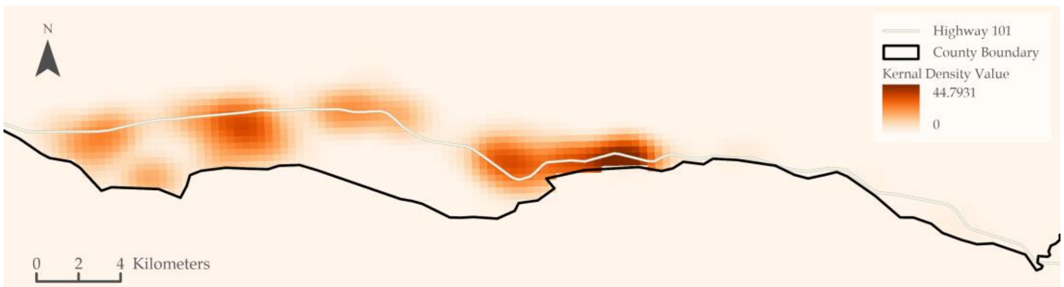
Resolution	Geographical Extent	Source	Sample Size
Census tract	Santa Barbara County	CDC (2022)	105
Census block group	Santa Barbara County	Bryan (2022)	315
1x1 km grid	Santa Barbara County	NASA SEDAC (2023)	5,721
Parcel	South Coast Santa Barbara County	Murray et al. (2023)	62,419

The County of Santa Barbara reported 1,887 PEH (people experiencing homelessness) in 2023, with 64% unsheltered and 93 households with minor children. In 2024 the county experienced a 12% jump in the total number of PEH, and more than half of people reported that they were experiencing homelessness for the first time (County of Santa Barbara, 2024). These statistics come from the Point-In-Time Count (PITC), a federally required enumeration and unduplicated survey of each unsheltered PEH in the entire administrative area at least every two years (HUD, 2023). The 2023 PITC in the County of Santa Barbara was conducted by volunteers and staff on January 25 between 5 and 9 am. A summary of the spatial distribution of PEH based on the PITC is shown in Figure 2a.

Another PEH information source is the Encampment Response Protocol (ERP). This is part of a county-funded initiative, Encampment Resolution Strategy, that validates and monitors reported encampments to target and offer outreach services (County of Santa Barbara, 2021b). Reported encampments come from community members, the unsheltered PEH themselves, and outreach service providers; it is a volunteered geographic information (VGI) database. The accumulated information is longitudinal, beginning in August 2021 through October 2023. A summary of the spatial distribution of PEH based on ERP is shown in Figure 2b. Worth noting is that the report-and-respond approach reflected in the ERP is generally geographically limited to human interaction and activity intensity (like all VGI), differing from the PITC. Further, both information sources only show the number of unsheltered PEH encampments, including belongings, tent-like structures, and vehicles, not the actual size of the encampment population. There also exists the possibility of one encampment being recorded in both information sources repeatedly. Thus, the PEH locational sites may not be completely independent of each other; instead, they represent a location that has been selected by unsheltered PEH at some point as their encampment site. Nevertheless, the PITC and the ERP represent two different information sources for the unhoused, offering insight into this social issue and its geographic heterogeneity.



(a) 2023 Point-In-Time Count



(b) 2021-2023 Encampment Response Protocol

Figure 2. Kernel density map of unsheltered PEH encampments

Method

This study uses an exploratory spatial data analysis (ESDA) framework that applies multiple methods, including geographic information systems (GIS), statistics, and geostatistics, to better understand the social dimensions of PEH. The framework is summarized in Figure 3.

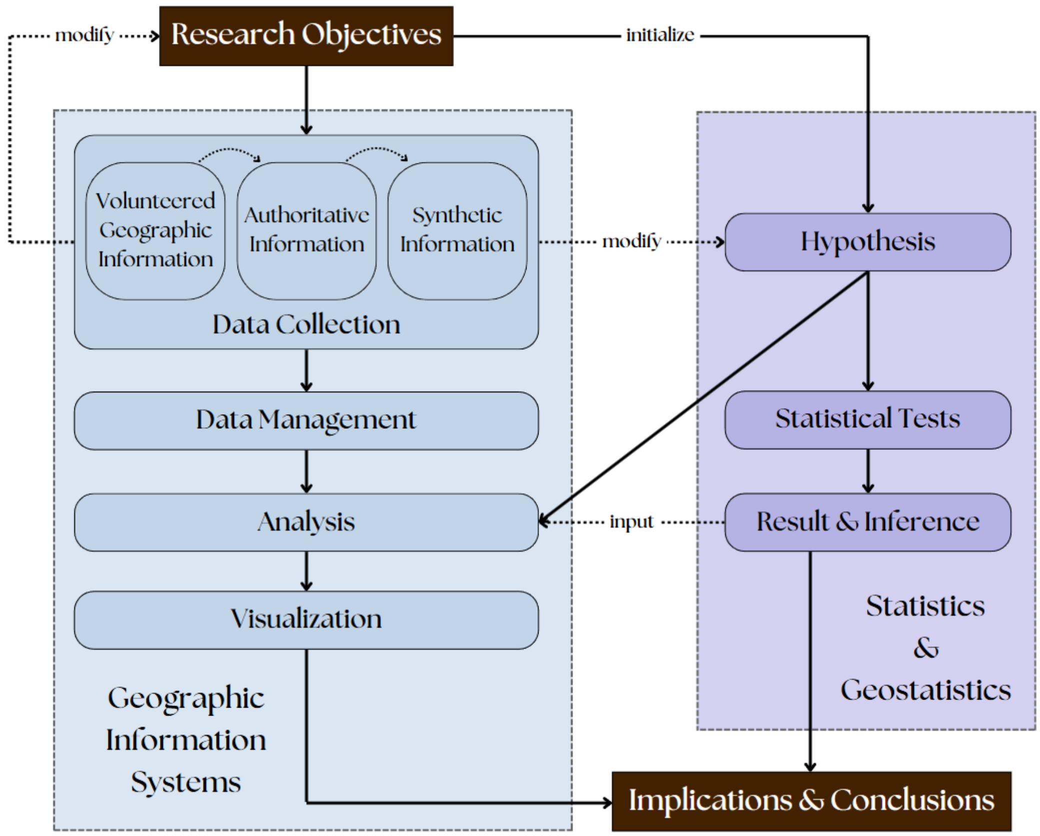


Figure 3. Research framework for understanding the geographic nature of PEH social vulnerability

The framework in Figure 3 is driven by overarching research questions, with various analytics and methods integrated to support evaluation, assessment, and understanding. A central element of the Figure 3 framework is GIS, used to collect and manage different types of spatial data as well as support analysis through visualization and summary. The data collection component of Figure 3 illustrates the use of EPR (Encampment Response Protocol) data as volunteered geographic information (VGI), PITC (Point-In-Time Count conducted by the federal government) as authoritative information, and the Census-based SVI (Social Vulnerability Index) as synthetic information. As described previously, although each information has been processed at different levels, all rely on GIS for spatial positioning.

Different data characteristics shape research objectives into two primary components- relationship and spatial pattern-and leaves a big challenge for analysis which is to transform information across multiple scales. One common analytical method in Figure 3 is offered by GIS, often used in the curation of SVI as well, is spatial interpolation with a generalized form:

$$z_X = \sum_{i=1}^{n} w_i z_i \tag{1}$$

where z_X is the estimated value at location X based on the know values of surrounding points i, z_i, with assigned weight w_i. Areal weighted interpolation can proportionally re-distribute data from larger to smaller geographic areas, while diffusion interpolation, using a heat equation as the weight function, models the gradual spread of point attributes across the surface. Mapping, in itself, serves as an intuitive approach for understanding spatial patterns and relationships. Kernel density mapping is another analytical method in Figure 3, offering valuable insights into the point distribution, and takes the form of:

$$\hat{f}(x) = N_p / A_p \tag{2}$$

where a location x is assigned a density measure based on the number of points or observations in the associated window p around x, denoted as N_p, divided by its area, A_p (Silverman, 1986; Murray, 2020; Sweeney and Arabadjis, 2022).

As reflected in Figures 1 and 2, visual inspection to assess pattern similarity makes visualization in the Figure 3 framework critical. However, it is difficult to interpret statistical significance and validation. Statistics and geostatistics in Figure 3 then are important and can be applied to review the strength and direction of relationships across multiple spatial patterns, analysis such as correlation, regression, and statistical testing are chosen based on the scale, distribution, and format of the information.

Given spatial unit $i = \{1, 2, \dots, n\}$, with SVI measure x_i and number of unsheltered PEH encampments y_i, the first relationship examined in this study is whether a linear relationship exists between x_i and y_i. As there are considerable regions having $y_i = 0$; three different analyses are introduced to reduce the potential bias caused by ties. Spearman's tie-corrected rank correlation (Kendall and Gibbons, 1990) might be useful:

$$\rho = \frac{\frac{n^3-n}{6} - \sum_{i=1}^{n}(\hat{x}_i - \hat{y}_i)^2 - \sum t_x - \sum t_y}{\sqrt{\left(\frac{n^3-n}{6} - 2\sum t_x\right)\left(\frac{n^3-n}{6} - 2\sum t_y\right)}} \tag{3}$$

where $\hat{x}_i$ and $\hat{y}_i$ are the tied ranking for each observed value x_i and y_i, respectively. The correction factor for ties in X, t_x, and in Y, t_y is $t^3 - t_j$, where t_j represents the length of tie j. Another option is Goodman and Kruskal's cross-classification correlation (1979):

$$\gamma = \frac{\Pi_s - \Pi_d}{\Pi_s + \Pi_d} \tag{4}$$

where Π_s and Π_d represent the probability of observing the same or different orders in a cross-classification table in which both x_i and y_i are assigned five ordinal rankings based on quintiles. The final test to handle the excess zeros from y_i is the zero-inflated negative binomial regression (Gelman et al., 2021), specified as:

$$y_i = \begin{cases} 0, & \text{if } S_i = 0 \\ \sim\text{negative-binomial}\left(e^{x_i\beta}, \varphi\right), & \text{if } S_i = 1 \end{cases} \tag{5}$$

where S_i is the indicator of whether unsheltered PEH will appear in area i and $P(S_i = 1) = logit^{-1}(x_i\gamma)$.

The analysis then moves on to explore the relationship between PEH point patterns A and B. Now, y_i^A and y_i^B represent the aggregated counts of PEH encampments in patterns A and B. However, the aggregation does not adhere to the boundaries used in the SVIs; instead, it utilizes finer and more regular tessellations through a method known as quadrat analysis, which examines observations in terms of their deviation from spatial expectations. A comparison of multiple point distributions is possible through a chi-square (χ^2) goodness-of-fit test or a Monte Carlo χ^2 test:

$$\chi^2 = \frac{\sum(O-E)^2}{E} = \frac{\sum\left(y_i^A - y_i^B\right)^2}{y_i^B} \tag{6}$$

While quadrat analysis indicates whether the point patterns are the same overall, it does not identify specific locations where similarities or differences exist. Therefore, a distance-based analysis, such as Ripley's K function, is necessary to more comprehensively address the research objectives. For multivariate spatial pattern analysis, the K function can be expressed as:

$$K_{AB}(r) = \lambda_B^{-1} E_B(r) \tag{7}$$

where $E_B(r)$ is the expected number of type B points within distance r of a randomly chosen type A points and λ_B is the density of type B point per unit area (Dixon, 2002). A transformation of K function is $L(r) = \sqrt{K(r)}$ where $L(r) = r$ represents a homogeneous Poisson point pattern, or so-called complete spatial randomness (Dixon, 2002; Sweeney and Arabadjis, 2022). The L function can be estimated to understand the spatial tendency for each PEH point pattern. Then, to capture the difference, D function, $D(r) = L_{BB}(r) - L_{AA}(r)$, is calculated and compared with the $\widehat{D}(r)$ under random labeling of points from a superposition set combining point pattern A and B with Monte Carlo simulation (n = 1000) (Diggle and Chetwynd, 1991; Sweeney and Arabadjis, 2022). Further, instead of a randomly chosen type A point, $L_{i,AB}(r)$ and $L_{i,AA}(r)$ can be estimated for each point i from type A, and some correlation coefficients such as Pearson's r as:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \tag{8}$$

can be used to understand the local location variation of points around A_i (Komladzei, 2021). Similarly, $L_{i,BA}(r)$ and $L_{i,BB}(r)$ can be calculated for B_i. In this approach, the spatial pattern, rather than point attributes, serves as the input for statistical analysis.

Collectively, the framework illustrated in Figure 3 attempts to show how research questions, spatial data, and supporting methods are utilized in an exploratory fashion to develop insights, formalize hypotheses, carry out statistical testing, and reach informed management and policy decisions.

Results

Social vulnerability in Santa Barbara can be examined in a number of ways. However, it is very much dependent on the context in which it is being considered. In this thesis, the primary interest is the degree to which people may be vulnerable to wildfire, but any hazard or context could be of interest.

The CDC SVI for Census tracts in Santa Barbara is depicted in Figure 1 suggests one perspective on social vulnerability. Another is shown in Figure 2 using the kernel density of unsheltered PEH encampments. Visual inspection suggests that the spatial patterns of unsheltered PEH are not the same. The 2023 PITC (Figure 2a) shows higher density in the central downtown area, whereas the 2021-2023 ERP (Figure 2b) has a more dispersed density with greater intensity in south-eastern downtown. A more direct comparison between SVI and PEH can be viewed in Figure 4, where a bivariate map depicts the PITC counts and SVI estimates at the Census block group level. There is substantial disagreement as there are numerous areas with low SVI scores but high PEH counts, and vice versa.

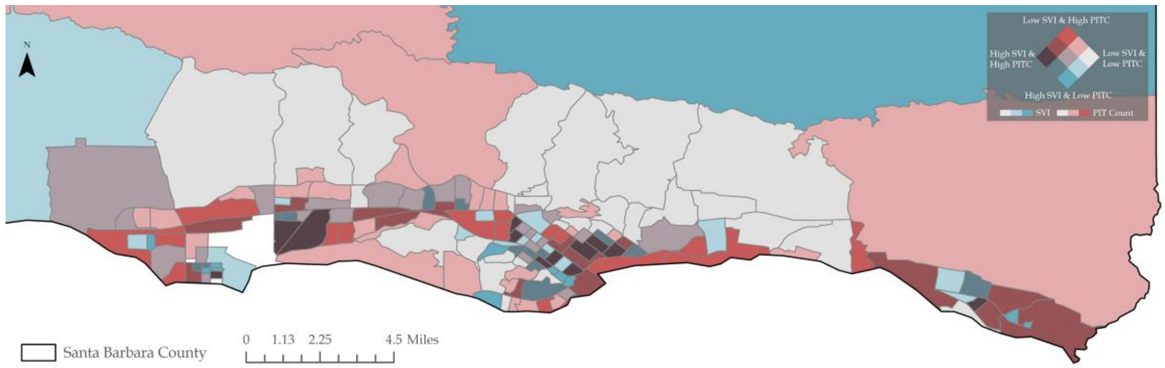


Figure 4. Bivariate map of SVI and PITC PEH Counts

In particular, the analysis tests whether a positive linear relationship exists between existing SVI metrics and encampment counts, which was posited based on the expected effectiveness of those measures in representing truly vulnerable regions. For Census tract, block group or 1x1 km grid spatial representations, both SVI and PEH counts within each areal unit in Santa Barbara County are examined. At the parcel level, SVI is compared with PEH counts within 100, 300, and 500 meters of the parcel boundary. Counts from both PITC and ERP are used given their different methodologies involved. As the SVI estimates use data from the same source or are derived from each other, the correlation among SVI metrics is not meaningful in this analysis. A summary of the statistical analysis is given in Tables 2 and 3. ρ and γ range from -1 (perfect negative correlation) to 1 (perfect positive correlation). In terms of the ZINB estimates, the sign and value represent the direction and magnitude of association when using SVI to fit the negative binomial distribution of PEH Counts. Although all t test results for correlation analysis are significant, confirming the existence of relationship between SVI and PEH counts, the magnitude of the coefficient still suggests a little to no positive linear correlation. In terms of ZINB regression performance, the likelihood ratio test suggests that the inclusion of SVI as the predictor improves the performance of a reduced intercept only regression model. While the majority of coefficients are positive, a quite surprising result emerges between SVI and PEH counts within 100 meters of the parcel, where negative associations appear for the ZINB estimates.

Table 2. Statistical assessment of SVI and PITC Counts

	CT_{US}[1]	CT_{CA}[2]	CBG[3]	1km	100m[4]	300m[4]	500m[4]
Spearman's correlation							
ρ	0.325	0.316	0.245	0.044	0.207	0.384	0.466
p value[5]	0.001	0.001	0.000	0.001	0.000	0.000	0.000
Goodman & Krustal's correlation							
γ	0.306	0.309	0.249	0.220	0.425	0.441	0.463
p value[5]	0.001	0.001	0.000	0.000	0.000	0.000	0.000
Zero-inflated negative binomial regression							
β	0.971	0.981	0.626	0.532	-0.294	1.489	1.983
p value[6]	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note. [1] SVI ranked based on all census tracts in the U.S., [2] SVI ranked based on all census tracts in California, [3] Census Block Group, [4] distance from the parcel, [5] t test with null hypothesis that coefficient is zero, [6] likelihood ratio test with reduced intercept

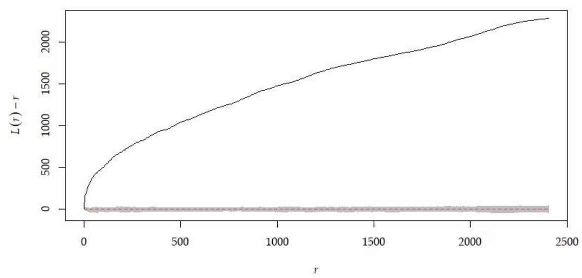
Table 3. Statistical assessment of SVI and ERP Counts

	CT_{US}[1]	CT_{CA}[2]	CBG[3]	1km	100m[4]	300m[4]	500m[4]
Spearman's correlation							
ρ	0.279	0.267	0.291	0.032	0.189	0.354	0.449
p value[5]	0.004	0.006	0.000	0.015	0.000	0.000	0.000
Goodman & Krustal's correlation							
γ	0.234	0.271	0.315	0.173	0.460	0.449	0.478
p value[5]	0.016	0.005	0.000	0.000	0.000	0.000	0.000
Zero-inflated negative binomial regression							
β	0.302	0.423	0.189	0.172	-0.021	1.192	1.699
p value[6]	0.000	0.000	0.000	0.000	0.000	0.000	0.000

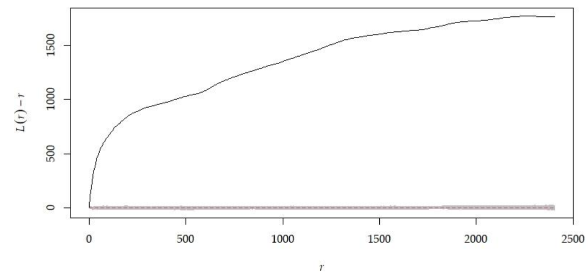
Note. [1] SVI based on all census tracts in the U.S., [2] SVI based on all census tracts in California, [3] Census Block Group, [4] distance from the parcel, [5] t test with null hypothesis that coefficient is zero, [6] likelihood ratio test with reduced intercept

There is an observed trend across various geographic scales for different methods and PEH count sources. Specifically, for PITC, SVI estimates at the Census tract level have a relatively stronger correlation with PEH conditions whereas for ERP the Census block group level SVI is slightly higher. The 1x1 km grid subset consistently has the weakest correlation. The divergence between Census tract level SVI estimates is minimal for PITC but becomes pronounced for ERP. At the parcel level, the trend demonstrates an increased correlation when the buffer from the parcel increases. It appears that comparison between SVI and PEH counts within 500 meters from parcels has the greatest correlation, but this is weak at best from a statistical perspective.

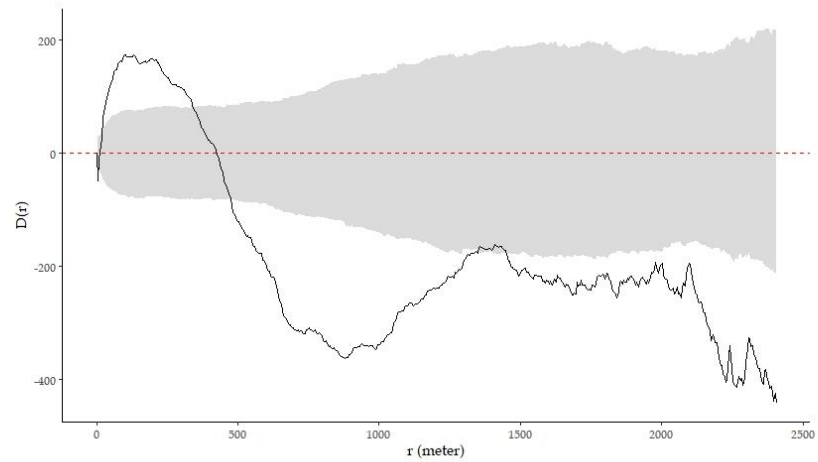
The second research objective is to see if there are differences between the spatial distribution of two PEH encampment point patterns. In the case of point pattern analysis, the pattern from PITC (n=664) is treated as "expected" and the one from ERP (n=1217) serve as "observed". The quadrat analysis finds that all χ^2 goodness-of-fit tests reject the null hypothesis that the observed distribution of frequency in ERP is the same as the expected PITC distribution (see Table A1). In addition to tests based on χ^2 distribution table, to control the unbalanced observation numbers from original point patterns, the null hypothesis is modified to see if the difference between two patterns under their actual labeling is the same as under random labeling. In this case, 664 points are sampled from ERP to generate empirical critical value of χ^2 distribution with Monte Carlo simulation (n=1000). The χ^2 distribution of random labeling of points with Monte Carlo simulation (n=1000) is then used to generate the p-value. Results from this procedure show inconsistency across different quadrat shapes and sizes (see Table A2), suggesting instability. However, the majority of test results reject the null hypothesis.



(a) L-function for PITC



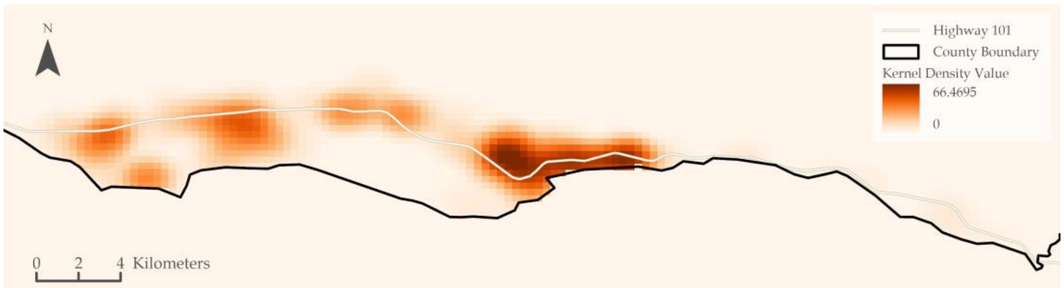
(b) L-function for ERP



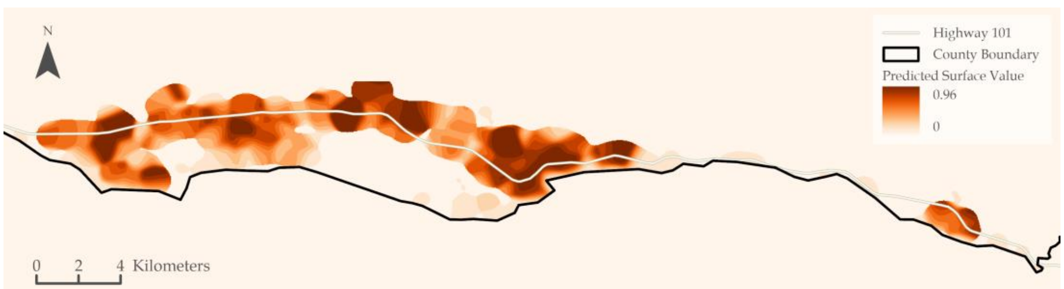
(c) D-function

Figure 5. Distance-based point pattern analysis

To figure out where exactly those clustering differences happened, distance-based analysis is carried out. These estimations are summarized in Figure 5, where L functions in Figures 5a and 5b show that both PITC and ERP patterns are clustered compared to the complete spatial randomness. Under the same first-order effect for both populations, the second-order effect can be reviewed by the result of D function. In Figure 5c, the gray envelope is the 97.5% confidence interval of the D-function curve range under random labeling of PITC and ERP, and the black line represents the correct labeling. This shows that the points from ERP are significantly more clustered than the points from PITC at the range around 250 meters, and consequently, more dispersed after 1000 meters. Meanwhile, if the analysis scope zooms to around 400 meters, there exists a good chance that two point patterns can be observed similarly.



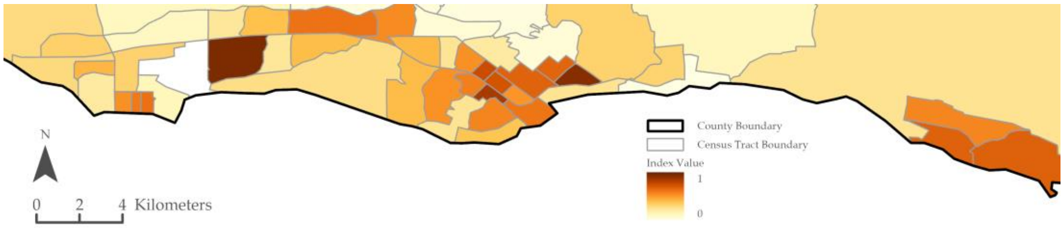
(a) Kernel smoothing of all points



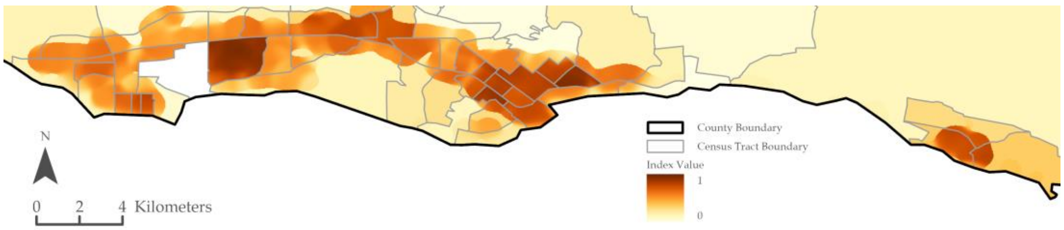
(b) Diffusion interpolation based on similarity

Figure 6. Spatial pattern of unsheltered PEH

Finally, the analysis explores ways to create a spatial representation of PEH patterns. One approach is shown in Figure 6 that considered kernel density and diffusion interpolation. Kernel density is shown in Figure 6a. However, such an approach assumes the independence of points across information sources, which does not match with how information is collected, as discussed previously. Hence, the idea here is to see if a point can be verified by both PEH counts, thereby making it a more credible observation. The interpolated surface in Figure 6b shows similar hot spots or clustering patterns with Figure 6a, while capturing more details in gradient changes.



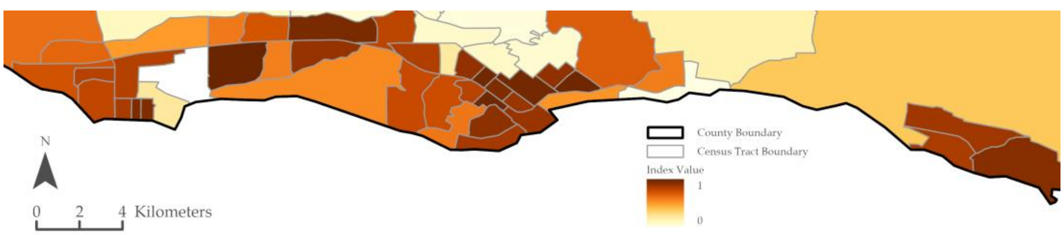
(a) Original SVI



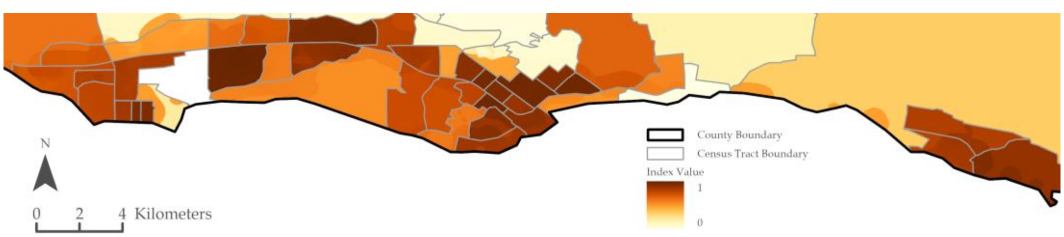
(b) Weighted SVI

Figure 7. Incorporating PEH condition as a weighted index

Note. $Weighted\ SVI = \alpha * Original\ SVI + (1 - \alpha) * PEH\ Condition$, $\alpha = 0.5$ in (b).



(a) Re-standardized SVI without PEH condition



(b) Re-standardized SVI with PEH condition

Figure 8. Recalculating CDC SVI with PEH condition as a variable

Given the PEH clustering pattern in Figure 6b, some integration can be made to promote the original SVI to a more inclusive assessment. One straightforward approach is to create a weighted index based on the current SVI. Comparing Figure 7a and Figure 7b, it is evident that the additional weight of the PEH condition does not alter the classification of the most vulnerable area. But in moderately vulnerable communities, the focus areas can be accentuated. For instance, the three census tracts in the southeastern corner of the study region have areas around their boundary intersection notably vulnerable. Alternatively, the PEH condition can be included as one of the variables in use. Following the CDC SVI procedure, Figure 8a and Figure 8b show the recalculated and standardized SVI value with and without PEH conditions for grids in the study region respectively. Similarly, the identification of the most vulnerable areas remains consistent while gradual differences can be observed for other areas.

Discussion

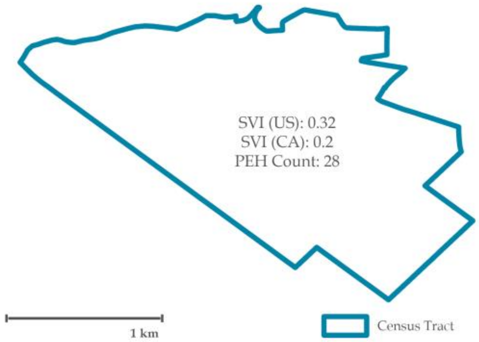
This thesis utilized an exploratory spatial data analysis framework to consider homelessness in Santa Barbara County concerning broadly used interpretations of social vulnerability using SVI. Two different information sources of PEH are assessed and compared based on their spatial distribution to several existing CDC SVI measures at different scales. The results highlighted the under-representation of unsheltered PEH using SVI estimates along with inherited impacts due to scale and boundary selection. In addition, a persistent spatial pattern of unsheltered PEH across multiple information sources. Finally, geostatistical methods are applied to estimate PEH distribution, and showed how such measures could be incorporated in existing SVI calculations.

As consideration of social equity continues to grow in hazard mitigation and response, a better understanding of social conditions and social vulnerability is of paramount importance.

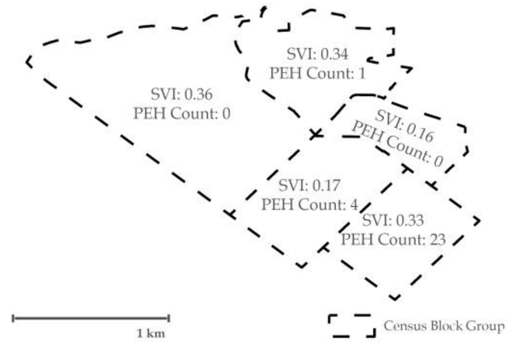
The CDC SVI is one of the many latent indices that aggregate a pool of population variables into one measure across a region. Though scholars critique and attempt to improve such an indexing approach, research generally focuses on how to aggregate and what variables to use (Tate, 2012, 2013; Rufat et al., 2019), yet ignores population and geographic representation. SVI and unsheltered PEH were considered as a case study to investigate missing components.

The weak correlation shown in Table 2 and Table 3 is not unexpected, given that unsheltered encampment condition is never a direct indicator in any SVI metrics. PEH has been largely ignored in global disaster management practice. Wisner (1998) identified various challenges faced by this group in disaster preparation, including lack of access to community plans due to self-isolation, social amnesia that prevents them from being recognized and considered, and persistent stereotypes that exaggerate social ostracism. The state of invisibility from policymakers of this population is the outcome of systematic inequalities, reflecting their loss of power and control. Homelessness is epitomized in modern society as a social problem, an economic dilemma, or a human rights issue, making it a distinct phenomenon of study, which makes it disconnected from other fields. Our findings demonstrate that the indirect socio-demographic factors are not enough to represent the unsheltered PEH, and further suggest that in the context of the social vulnerability framework, homelessness should be viewed as an independent characteristic, alongside the elderly, the disabled, etc. as part of the demographic make-up of a community.

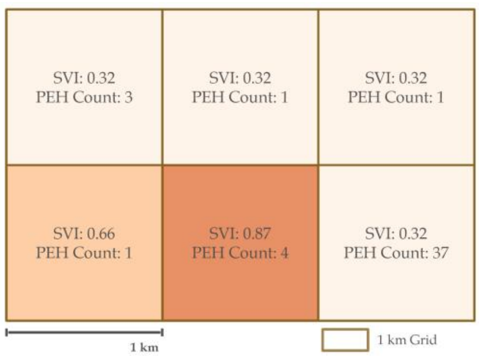
However, this is not the only reason contributing to the discrepancy. A closer look at how the assessment is made would reveal that SVI estimates essentially blur internal heterogeneity within the region. Figure 9 portrays a small region around one census tract, where unsheltered PEH are noticeably concentrated in the lower right corner of the region.



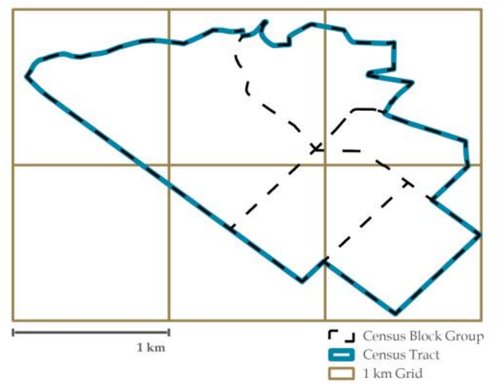
(a) Census Tract



(b) Census Block Group



(c) 1km Grid



(d) Overview

Figure 9. SVI and PEH Counts across geographic scale in the same region

First, it is clear that using census tract level SVI would directly obfuscate regional variations (Figure 9a). Second, as depicted by Figure 9c, the 1 km dataset is essentially a rasterization and areal interpolation of the original CDC SVI across the U.S., so cells within one census tract would have the same value (NASA SEDAC, 2023). This explains the sharp weakening in correlation for census tract and 1 km in Tables 2 and 3. Since the 1 km grid considers the variation of unsheltered PEH, it reflects the homogeneity of the SVI estimate. Third, as shown in Figure 9b, the census block group SVI is a relatively better representation. Tables 2 and 3 suggest such a hypothesis holds true for ERP, but not for PITC, which relates to the representation of PEH. Figure 4, Figure 9, Table 2, and Table 3 illustrate the importance of scale in vulnerability assessment, presenting issues and limitations associated with existing approaches.

It then raises the question of whether some boundaries (census geography, population grid, etc.) frequently relied upon are suitable for population vulnerability analysis. This problem is thornier when it comes to homelessness research. Unsheltered PEH are much less likely than other groups to choose an encampment site because that portion of land is attached to a particular census tract. The selection can be quite discrete for personal preference. Hence, the traditional way of aggregating points into the predefined boundary and making the summed count as an indicator seems to be inadequate. Instead, the study attempts to see if a site is an enduring spot for homeless encampment so that people commonly return or select surrounding spots with similar characteristics.

This expectation redirects the study back to its data collection and management phase to inspect available information on homeless encampments. Although PITC has been heavily criticized for its representativity of regional unsheltered PEH using a snippet a single night in January and potentially biased statistical methodology, it is the only mandated national data available, so it remains a ubiquitous presence in homelessness studies (Smith and Castañeda-Tinoco, 2019). On the flip side, studies that actually review PITC are limited because it depends on whether or not there is another dataset available for the region. In recent years, with the development of GIS and its popularization in everyday use, community engagement and collective efforts to provide geographic information have been largely promoted. However, the credibility of volunteered geographic information (VGI) generates concerns about its source, quality, and authenticity (Flanagin and Metzger, 2008). As discussed, the ERP information used in this study is one type of VGI. While the ground truth of encampment has been verified by the governmental agency, the data documentation and potential duplication can cause bias in the data quality (e.g., one PEH can be reported by multiple observers at different times.) Overall, the distinct geographic coverage, temporal resolution, and sample size across information sources amplify challenges in inferring the actual PEH activity.

Point pattern analysis is introduced to examine the spatial trends of point clustering while avoiding introducing extra bias from boundary selection. The K function and its multiple extensions allow us to get a glimpse of the differences between two unsheltered PEH behavioral patterns (Sweeney and Arabadjis, 2022). In this study, for the south coastal region of Santa Barbara County, the D function gives us a fairly straightforward answer that the clustering pattern of PITC and EPR are similar around 400 meters, while EPR is relatively more compact than PITC at closer distance intervals. With point coordinates only, we are still agnostic about the exact underlying causation, but some assumptions can be made to understand the difference. As mentioned before, the origins of ERP information differ, thus data reporting tendency can lead to a more clustered pattern in ERP. For example, the compact alignment of points under tree canopy along the creek and highway is more likely to have been recorded by staff during outreach activities. In the case of community members, much of this is tied to a willingness to report, i.e., residents might be more proactive in reporting seemingly threatening or unsafe encampments in their own neighborhoods or activity zones. Although PITC also has some differing interviewing emphasis due to volunteer attention to different unsheltered PEH subpopulations (e.g., with vehicle, substance abuse disorder, youth, etc.), PITC has made sure to report on every encampment observed during the period (Smith and Castañeda-Tinoco, 2019). A comparison between Figure 2 and Figure 6 shows that both datasets catch some of the persistent hot spots of unsheltered PEH encampments. While it is still debatable, this analysis validates the baseline quality of both information sources.

Finally, the ESDA framework enables the transformation of point-level information into a surface representation that can serve as an indicator of unsheltered PEH conditions. Spatial interpolation is one method to better amplify the gradient changes of PEH encampment hotspots over space, compared with density mapping. As shown in Figures 7 and 8, noticeable changes occur after incorporating this factor, highlighting the future research needs of multiscale information synthesis and inclusive social vulnerability portrayal. The identification of critical areas requiring targeted interventions is essential to understand the complex dynamics of social vulnerability and enhance community resilience in terms of disaster risks.

Conclusion

Current social vulnerability measures have inadequate correlation with observed unsheltered PEH, suggesting the need to re-examine SVI in practice. The divergent results from statistical tests indicate that coarse resolution tends to eliminate the heterogeneity across a region, thus diminishing the effectiveness of SVI. Moreover, point pattern analysis finds PEH encampments are more compact in local records than annual federal counts. It is evident that PEH encampment selection is not random and requires additional focus on spatial variation. Consequently, this paper explores an alternative perspective of interpreting spatial representation of PEH. Instead of using each point location, an interpolated surface based on prevalence of PEH encampment selection. This integration process of point and polygon in the study also signals the scale and boundary in current SVI estimates might not be suitable for studying specific population dynamics.

In social vulnerability assessment, population representation is a missing dimension. However, if SVI fails to encompass truly vulnerable populations, it violates the need and intent of analysis. Moreover, such approaches are likely to be counterproductive for resource allocation and public policy. This paper underscores the necessity of understanding the social dynamics of at-risk populations when wildfire hazard risk is increasing. More empirical evidence on risk intersections can better support robust data-driven community planning. This study conducts the rigorous spatial analysis of a hard-to-count population with a supportive framework. Such a framework not only advances homelessness research by identifying clustering patterns but also enables more inclusive disaster preparedness by overcoming barriers to include this population in risk mitigation efforts. Additionally, the framework shows ways of selecting appropriate analytical methods based on the understanding of information characteristics, including collection methods, data distribution, uncertainty, and potential bias. Its application can extend to other social groups facing similar challenges.

References

Bazuin, J. T., & Fraser, J. C. (2013). How the ACS gets it wrong: The story of the American Community Survey and a small, inner city neighborhood. *Applied Geography*, *45*, 292–302.

Bryan, M. (2022). US Social Vulnerability by Census Block Groups [dataset]. https://doi.org/10.7910/DVN/ARBHPK

Centers for Disease Control and Prevention (CDC). (2022). CDC/ATSDR Social Vulnerability Index 2020 Database [dataset]. https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_download.html

Cerveny, L. K., & Baur, J. W. R. (2020). Homelessness and nonrecreational camping on national forests and grasslands in the United States: Law enforcement perspectives and regional trends. *Journal of Forestry*, *118*(2), 139–153.

Coppola, D. P. (2015). *Introduction to international disaster management* (Third edition). Elsevier/Butterworth-Hein.

County of Santa Barbara. (2021a). *Climate change vulnerability assessment*.

County of Santa Barbara. (2021b). *County of Santa Barbara encampment response protocol*.

County of Santa Barbara. (2023). *Encampment resolution strategy: progress report year 1 (September 1, 2021-August 31, 2022)*.

County of Santa Barbara. (2023). *Santa Barbara County Point In Time Count 2023.*

County of Santa Barbara. (2024). *Santa Barbara County Point In Time Count 2024*.

Cutter, S. L. (2024). The origin and diffusion of the social vulnerability index (SoVI). *International Journal of Disaster Risk Reduction*, *109*, 104576.

Cutter, S. L., Boruff, B. J., & Shirley, W. L. (2003). Social vulnerability to environmental hazards. *Social Science Quarterly*, *84*(2), 242–261.

Davies, I. P., Haugo, R. D., Robertson, J. C., & Levin, P. S. (2018). The unequal vulnerability of communities of color to wildfire. *PLOS ONE*, *13*(11), e0205825.

Diggle, P. J., & Chetwynd, A. G. (1991). Second-order analysis of spatial clustering for inhomogeneous populations. *Biometrics*, *47*(3), 1155.

Dixon, P. M. (2002). Ripley's K function. *Encyclopedia of Environmetrics*, *3*, 1796–1803.

Flanagan, B. E., Gregory, E. W., Hallisey, E. J., Heitgerd, J. L., & Lewis, B. (2011). A social vulnerability index for disaster management. *Journal of Homeland Security and Emergency Management*, *8*(1).

Flanagan, B. E., Hallisey, E. J., Adams, E., & Lavery, A. (2018). Measuring community vulnerability to natural and anthropogenic hazards: The Centers for Disease Control and Prevention's social vulnerability index. *Journal of Environmental Health*, *80*(10), 34–36.

Flanagin, A. J., & Metzger, M. J. (2008). The credibility of volunteered geographic information. *GeoJournal*, *72*(3–4), 137–148.

Gaillard, J. C., Walters, V., Rickerby, M., & Shi, Y. (2019). Persistent precarity and the disaster of everyday life: homeless people's experiences of natural and other hazards. *International Journal of Disaster Risk Science*, *10*(3), 332–342.

Gaither, C. J., Poudyal, N. C., Goodrick, S., Bowker, J. M., Malone, S., & Gan, J. (2011). Wildland fire risk and social vulnerability in the Southeastern United States: An exploratory spatial data analysis approach. *Forest Policy and Economics*, *13*(1), 24–36.

Gelman, A., Hill, J., & Vehtari, A. (2021). *Regression and other stories*. Cambridge University Press.

Goodman, L. A., & Kruskal, W. H. (1979). *Measures of association for cross classifications*. Springer.

Harner, J., Warner, K., Pierce, J., & Huber, T. (2002). Urban Environmental Justice Indices. *The Professional Geographer*, *54*(3), 318–331.

Kendall, M. G., & Gibbons, J. D. (1990). *Rank correlation methods* (5th ed.). E. Arnold.

Komladzei, S. C. (2021). Co-localization analysis of bivariate spatial point pattern. *University of New Orleans Theses and Dissertations*.

Lambrou, N., Kolden, C., Loukaitou-Sideris, A., Anjum, E., & Acey, C. (2023). Social drivers of vulnerability to wildfire disasters: A review of the literature. *Landscape and Urban Planning*, *237*, 104797.

Mah, J. C., Penwarden, J. L., Pott, H., Theou, O., & Andrew, M. K. (2023). Social vulnerability indices: A scoping review. *BMC Public Health*, *23*(1), 1253.

Méndez, M., Flores-Haro, G., & Zucker, L. (2020). The (in)visible victims of disaster: Understanding the vulnerability of undocumented Latino/a and indigenous immigrants. *Geoforum*, *116*, 50–62.

Meyer, B. D., Wyse, A., & Corinth, K. (2023). The size and Census coverage of the U.S. homeless population. *Journal of Urban Economics*, *136*, 103559.

Montgomery, A. E., Szymkowiak, D., Marcus, J., Howard, P., & Culhane, D. P. (2016). Homelessness, unsheltered status, and risk factors for mortality: Findings from the 100 000 homes campaign. *Public Health Reports*, *131*(6), 765–772.

Murray, A. T. (2020). Spatial clustering, detection and analysis of. In Kobayashi, A. (Ed.), *International encyclopedia of human geography (second edition)*, pp. 367–373. Elsevier.

Murray, A. T., & Davis, R. (2001). Equity in Regional Service Provision. *Journal of Regional Science*, *41*(4), 557–600.

Murray, A. T., Baik, J., Echeverri Figueroa, V., Rini, D., Moritz, M. A., Roberts, D. A., Sweeney, S. H., Carvalho, L. M. V., & Jones, C. (2023). Developing effective wildfire risk mitigation plans for the wildland urban interface. *International Journal of Applied Earth Observation and Geoinformation*, *124*, 103531.

Muttarak, R., Lutz, W., & Jiang, L. (2015). What can demographers contribute to the study of vulnerability? *Vienna Yearbook of Population Research*, 1–13.

NASA Socioeconomic Data and Applications Center (NASA SEDAC). (2023). U.S. Social Vulnerability Index Grids, revision 01 [dataset]. https://doi.org/10.7927/2ev1-g103

O'Keefe, P., Westgate, K., & Wisner, B. (1976). Taking the naturalness out of natural disasters. *Nature*, *260*(5552), 566–567.

Palaiologou, P., Ager, A. A., Nielsen-Pincus, M., Evers, C. R., & Day, M. A. (2019). Social vulnerability to large wildfires in the western USA. *Landscape and Urban Planning*, *189*, 99–116.

Paveglio, T. B., Edgeley, C. M., & Stasiewicz, A. M. (2018). Assessing influences on social vulnerability to wildfire using surveys, spatial data and wildfire simulations. *Journal of Environmental Management*, *213*, 425–439.

Pludow, B.A. (2024). Considerations of Space for Prioritizing Wildfire Risk Reductions in the Wildland-Urban Interface. PhD dissertation. University of California at Santa Barbara.

Rufat, S., Tate, E., Emrich, C. T., & Antolini, F. (2019). How valid are social vulnerability models? *Annals of the American Association of Geographers*, *109*(4), 1131–1153.

Shouls, S., Congdon, P., & Curtis, S. (1996). Modelling inequality in reported long term illness in the UK: Combining individual and area characteristics. *Journal of Epidemiology & Community Health*, *50*(3), 366–376.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall.

Sloggett, A., & Joshi, H. (1994). Higher mortality in deprived areas: Community or personal disadvantage? *BMJ*, *309*(6967), 1470–1474.

Smith, C., & Castañeda-Tinoco, E. (2019). Improving homeless point-in-time counts: Uncovering the marginally housed. *Social Currents*, *6*(2), 91–104

Spielman, S. E., Tuccillo, J., Folch, D. C., Schweikert, A., Davies, R., Wood, N., & Tate, E. (2020). Evaluating social vulnerability indicators: Criteria and their application to the Social Vulnerability Index. *Natural Hazards*, *100*(1), 417–436.

Sweeney, S., & Arabadjis, S. (2022). Spatial point patterns. In *Handbook of spatial analysis in the social sciences* (pp. 262–276). Edward Elgar Publishing.

Tate, E. (2012). Social vulnerability indices: A comparative assessment using uncertainty and sensitivity analysis. *Natural Hazards*, *63*(2), 325–347.

Tate, E. (2013). Uncertainty analysis for a social vulnerability index. *Annals of the Association of American Geographers*, *103*(3), 526–543.

Thomas, A. S., Escobedo, F. J., Sloggy, M. R., & Sánchez, J. J. (2022). A burning issue: Reviewing the socio-demographic and environmental justice aspects of the wildfire literature. *PLOS ONE*, *17*(7), e0271019.

Thomas, D. S. K., Phillips, B. D., Lovekamp, W. E., & Fothergill, A. (2013). *Social Vulnerability to Disasters, Second Edition*.

U.S. Department of Housing and Urban Development (HUD). (2023). *The 2023 Annual Homeless Assessment Report (AHAR) to Congress. part 1: Point-in-time estimates of homelessness*.

Vickery, J. (2018). Using an intersectional approach to advance understanding of homeless persons' vulnerability to disaster. *Environmental Sociology, 4*(1), 136–147.

Wigtil, G., Hammer, R. B., Kline, J. D., Mockrin, M. H., Stewart, S. I., Roper, D., & Radeloff, V. C. (2016). Places where wildfire potential and social vulnerability coincide in the coterminous United States. *International Journal of Wildland Fire*, *25*(8), 896.

Wisner, B. (1998). Marginality and vulnerability. *Applied Geography*, *18*(1), 25–33.

Wong, S. D., Walker, J. L., & Shaheen, S. A. (2021). Trust and compassion in willingness to share mobility and sheltering resources in evacuations: A case study of the 2017 and 2018 California Wildfires. *International Journal of Disaster Risk Reduction*, *52*, 101900.

Appendix A

Table A1. Regular chi-square test result from quadrat analysis

Regular Tiling	Area (m^2)	χ^2	df	p-value
Square	10000	182.320734	7	0.00000*
Square	40000	80.226449	7	0.00000*
Square	90000	104.999134	8	0.00000*
Square	160000	66.365830	8	0.00000*
Square	250000	30.871416	8	0.00015*
Square	360000	37.708542	8	0.00001*
Square	490000	36.561500	8	0.00001*
Hexagon	10000	300.707260	7	0.00000*
Hexagon	40000	72.148281	7	0.00000*
Hexagon	90000	65.355289	8	0.00000*
Hexagon	160000	61.879256	8	0.00000*
Hexagon	250000	46.269518	8	0.00000*
Hexagon	360000	36.970821	8	0.00001*
Hexagon	490000	49.687410	8	0.00000*
Triangle	10000	175.688010	7	0.00000*
Triangle	40000	98.906485	7	0.00000*
Triangle	90000	100.484967	8	0.00000*
Triangle	160000	46.477683	8	0.00000*
Triangle	250000	42.892951	8	0.00000*
Triangle	360000	55.271408	8	0.00000*
Triangle	490000	26.470460	8	0.00087*

Table A2. Modified chi-square test with Monte Carlo simulation

Regular Tiling	Area (m^2)	Critical Value	p-value
Square	10000	54.852315	0.003*
Square	40000	71.988320	0.007*
Square	90000	29.425879	0.180
Square	160000	35.114233	0.178
Square	250000	95.856366	0.000*
Square	360000	64.949773	0.020*
Square	490000	51.651787	0.056
Hexagon	10000	44.761774	0.007*
Hexagon	40000	39.662908	0.079
Hexagon	90000	47.250742	0.036*
Hexagon	160000	50.361296	0.086
Hexagon	250000	27.139858	0.239
Hexagon	360000	52.846479	0.033*
Hexagon	490000	67.417787	0.014*
Triangle	10000	35.854472	0.026*
Triangle	40000	42.602282	0.043*
Triangle	90000	52.687021	0.083
Triangle	160000	52.489574	0.030*
Triangle	250000	34.215126	0.222
Triangle	360000	50.008915	0.104
Triangle	490000	77.754849	0.008*

Note. Critical value is calculated by 95th percentile of the empirical χ^2 distribution; p-value $= \frac{\text{\# of simulated result exceeding critical value}}{\text{\# of Monte Carlo simulation}}$