LEADED LOUIE

Tufts UNIVERSITY

Modeling the Risk of Lead Paint Hazards in St. Louis City

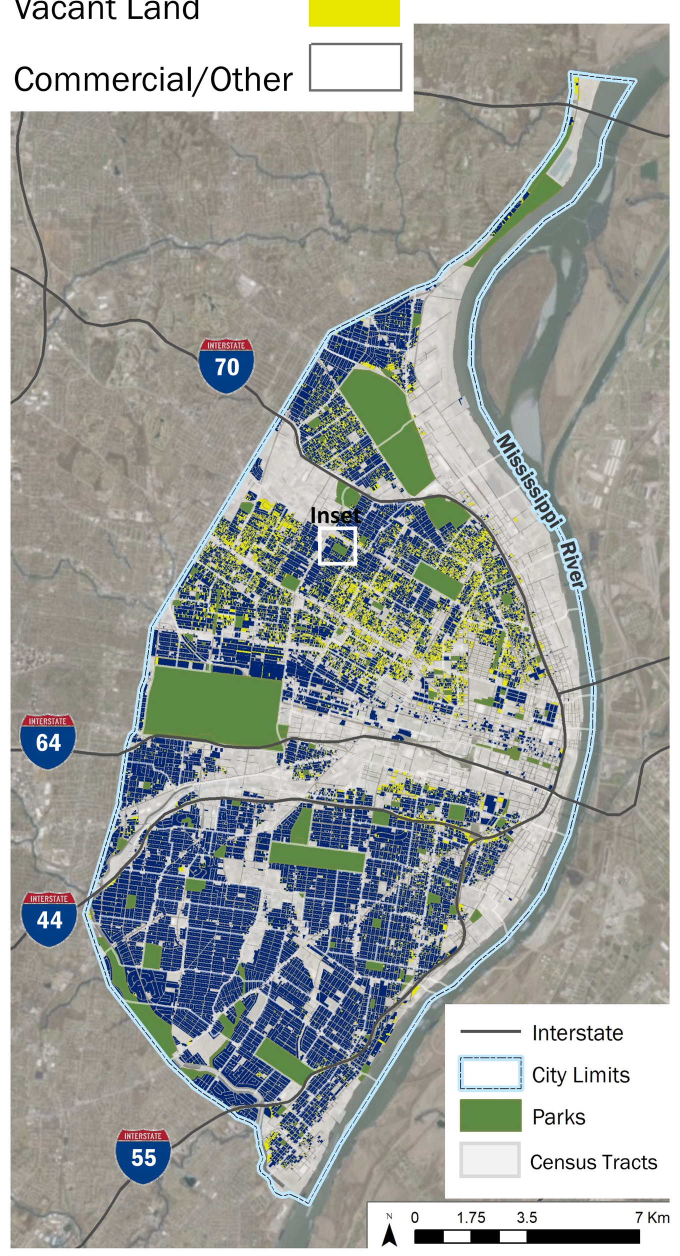
INTRODUCTION: The St. Louis Department of Health wants to conduct door-to-door screenings for childhood lead poisoning.. The entire city is classified as a high-risk area by the state government, and all children under the age of 6 must be screened annually under state guidelines. However, the Department lacks the resources to screen those 25,000+ children. In 2013, only 52% were screened, and 9.2% were found to have blood lead levels above the CDC's level of concern, 5µL/dL [2]. The Department wants to increase screening for the highest risk children. The most common cause of poisoning is the ingestion of dust or chips from lead-based paint at home [4]. The homes at greatest risk of lead paint hazards are those built before 1978 (the year lead paint was banned) in low-income areas, where chipping paint may be more likely [4]. Housing age is thought to be a stronger predictor of poisoning than income [4].

MODEL: The model will attempt to estimate the relative risk of lead paint hazards in each residence in the city. The spatial unit of interest are tax parcels, which have one or more residences built on them. The spatial mechanism is the juxtaposition of one census tract-level risk factor, the prevalence of childhood lead poisoning, and two parcel-level risk factors, the age of the residence and the property value. For each parcel, risk factors are weighted to create a final risk score. Parcels with the highest quartile of risk score may be targeted for door-to-door screening. Since fine income data is difficult to come by, property value will serve as the proxy for housing quality. The model doesn't include a parameter for the area's population of children because the available data tends to be too spatially granular. Furthermore, the risk of missing an at-risk child for screening is deemed greater than surveying homes without a child under 6.

Residential Zoning

Code per tax parcel

Residences	
Vacant Land	
Commercial/Other	

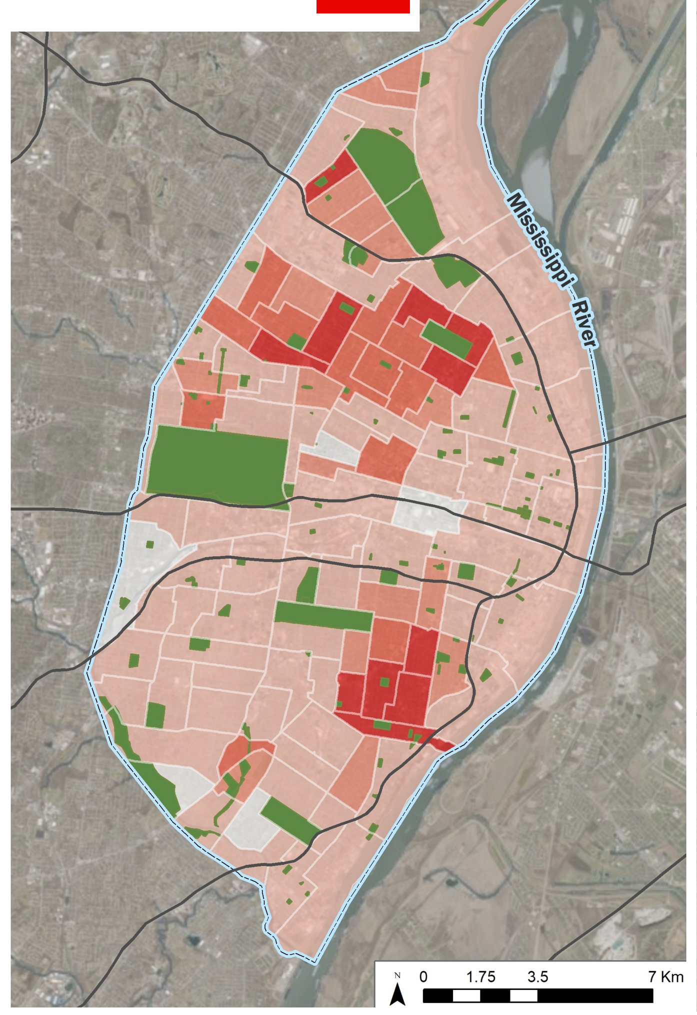


The municipal government provides a public dataset of tax parcel shape files, linked with zoning codes, vacancy status, the year of construction, building area, and assessed land value. This data was collected for use in city planning and tax assessment and updated in March, 2018. I place a high degree of trust in this dataset because of its use within local government. While the shapes of the parcels appear precise, visual checks show that parcels may not have consistent definition when it comes to multifamily housing. Using earth imagery, I see that some large buildings are considered a single parcel, while others are subdivided into smaller parcels.
Processing: Residential, non-vacant parcels were selected.

Quartile of Lead Poisoning Prevalence

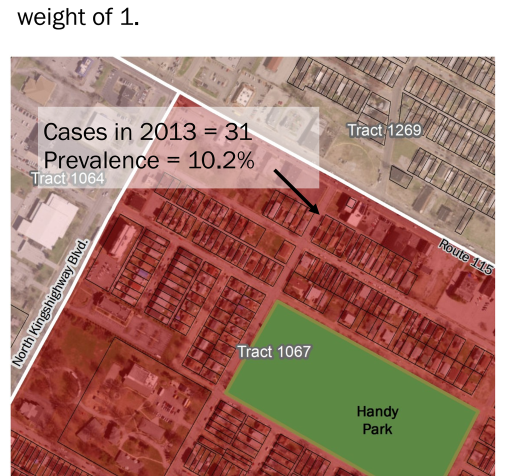
Cases per Child in Tract/Tract Area (hA)

No Cases	0
0 – 0.0005	1
0.0005 – 0.0008	2
0.0008 – 0.001	3
0.001 – 0.002	4



It stands to reason that homes in areas with a high prevalence of lead poisoning in 2013 pose a greater risk to residents in 2018, given that decontamination efforts are slow in St. Louis. In a 2013 report, the Childhood Lead Poisoning Prevention Program listed the annual prevalence of lead poisoning (blood lead levels > 5µL/dL) in children under 6 for each census tract, zip code, ward, and neighborhood. This report is meant for the public and other health agencies. Of the modifiable areal units, census tracts were found to be the least granular, i.e. have the lowest median area.

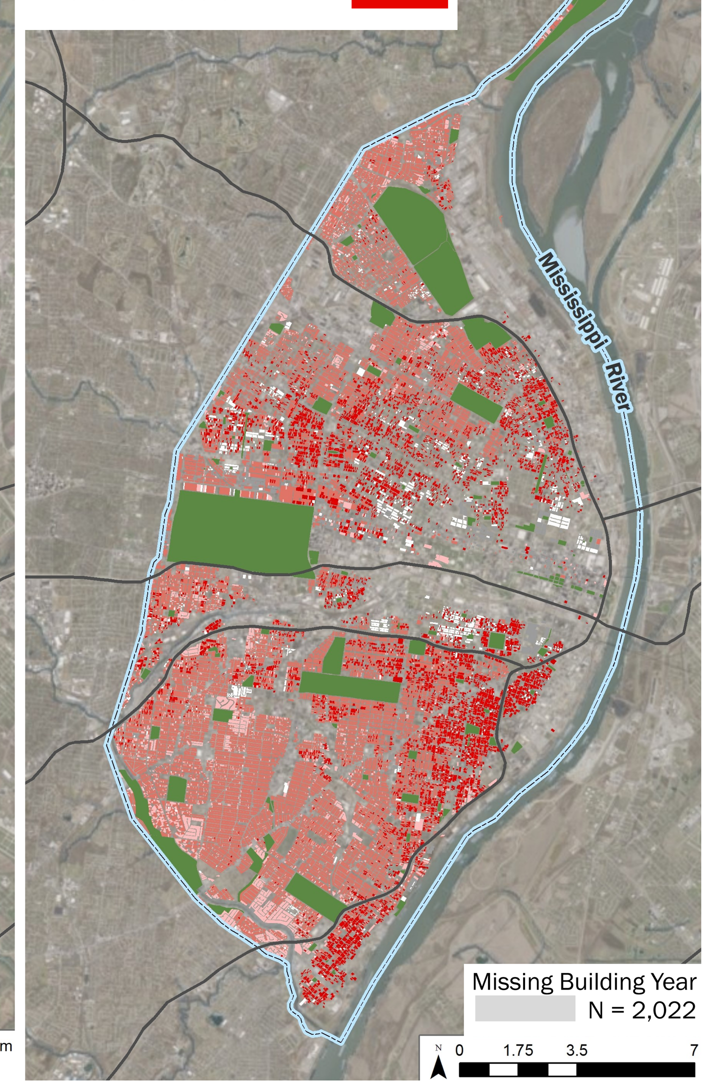
Processing: Parcels were spatially joined to tracts using their centroids. No parcels were found to be split between tracts. Prevalence was then divided into quartiles and scored with a weight of 1.



Category of Construction Year

Year per residential parcel

After 1978	0
1950 – 1978	10
1900 – 1950	20
Before 1900	30



Housing age is the strongest predictor of childhood lead poisoning in multiple studies [5]. In 1978, lead-based paint was outlawed, meaning that later construction is at no risk for lead paint hazards However, its use had been gradually declining since 1950. Before 1950, the older the home, the higher the likelihood of lead paint contamination [5]. Housing age is part of the municipal parcel dataset, and is used for property tax purposes. It is missing for 0.01% of parcels.
Processing: Housing age was divided into four categories, with attributed risk increasing by a factor of 10. The weight is arbitrary; in reality, risk may increase linearly with age.
Statistics: The housing stock in St. Louis is very old; 83% of residential parcels were build before 1950. There's no association between normalized value and age (p = 0.09).

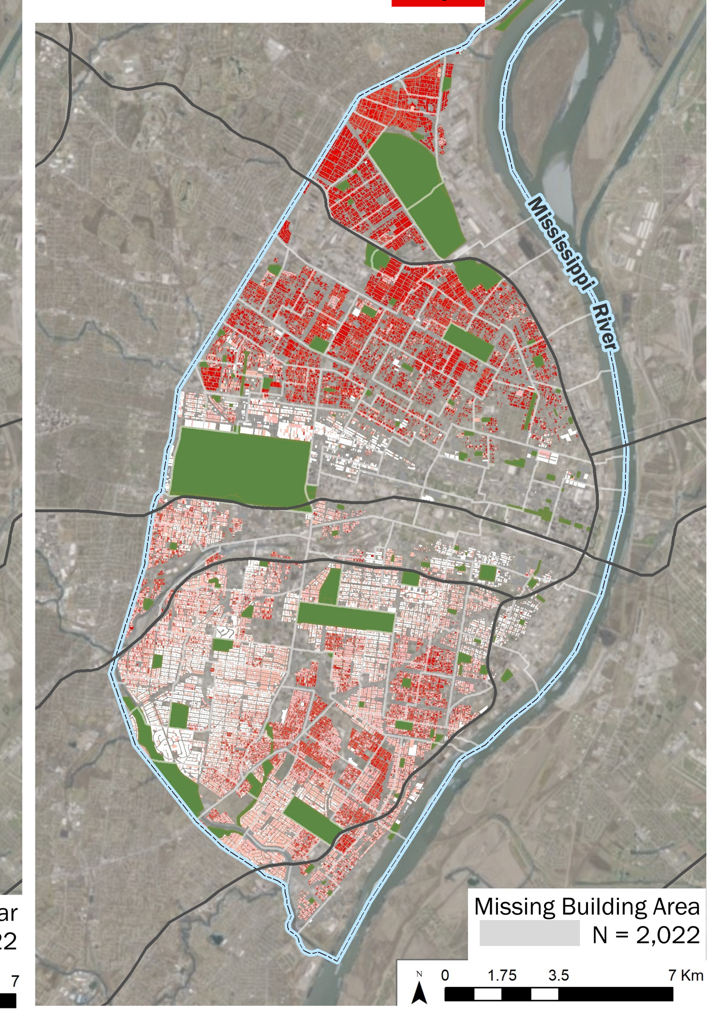
Property Value Quartile

Age Category	# Parcels	Q1	Q2	Q3	Q4	Q5
	After 1978	201	154	404	889	1713
	1950 – 1978	869	1126	1933	3360	3291
	1900–1950	13144	13525	12355	11336	9871
	Before 1900	3263	2672	2782	1892	2601

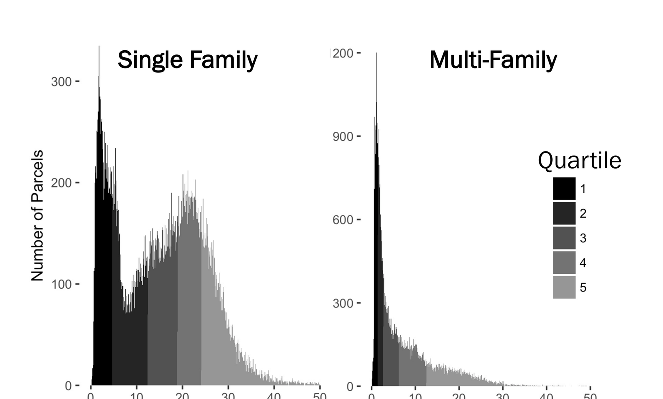
Quartile of Normalized Property Value

Value ($) / Building Footprint Area (m^2)

Highest Value	0
Upper-middle Value	1
Middle Value	2
Lower-middle Value	3
Lowest Value	4



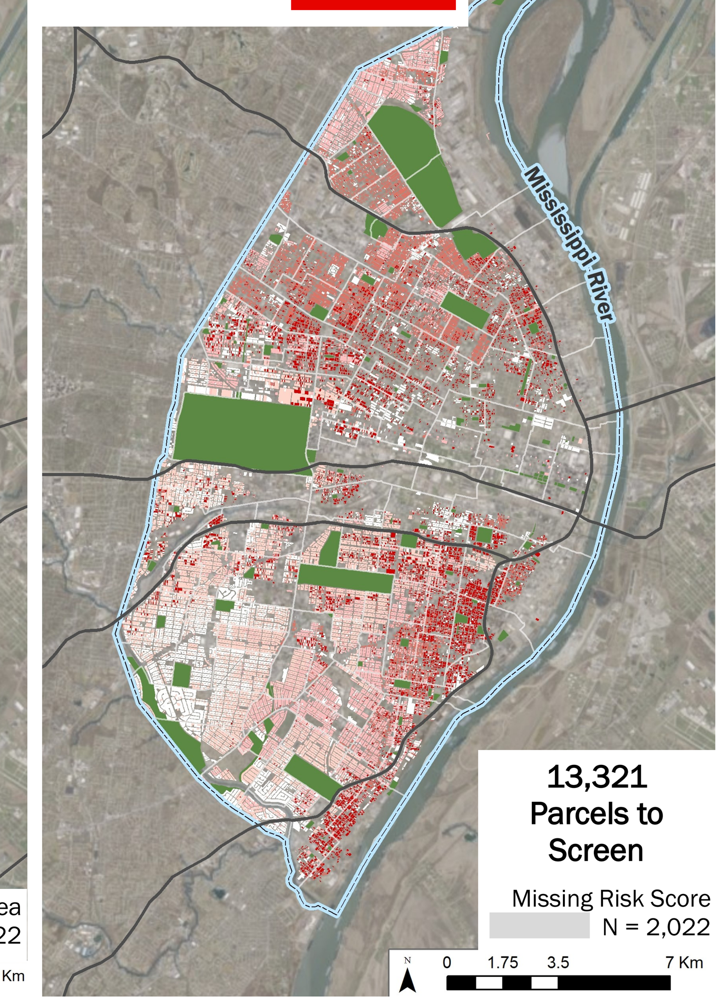
Socioeconomic status predicts lead poisoning risk in many studies, via ecological and dietary mechanisms [5]. In lieu of household income, I used low property value as a proxy for poor housing quality, especially chipping lead-based paint [5]. Parcels are linked with property value, which is assessed for property tax purposes. The year of assessment is not given, which could be a source of bias. Though all parcels have an assessed value, 0.01% are missing building areas.
Processing: Value is normalized by the building footprint area. Then, single-family zoned residences are divided into quartiles separately from multi-family residences, since their distributions follow strikingly different patterns. Quartiles have a weight of 1.



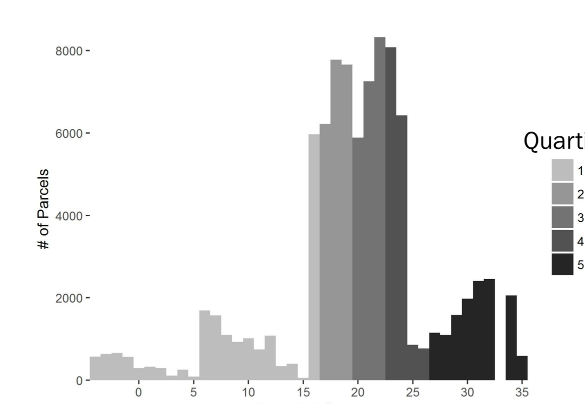
Quintile of Final Risk Scores

Score per residential parcel

Lowest Priority	-4 – 16
	17 – 19
	20 – 22
	23 – 26
TARGET GROUP	27 – 35



The final risk score attempts to estimate which residences have the greatest likelihood of lead paint hazards, which poses a risk to young children. Parcels missing construction year and building area weren't scored. Once the cut-off for what score constitutes a screening-worthy residence, the addresses linked to each parcel can be easily distributed to volunteers for a door-to-door screening campaign.
Processing: The weighted scores from prevalence quantile, construction year category, and property value quantile are totaled to get a final risk score. Here, the cut-off for targeting a parcel is the a score in the 5th quintile. This is entirely arbitrary.



Risk Score Quintile	# Parcels	Avg. Year	Avg. Norm Value	Avg. Prev
1	18,724	1958	20.53	5.2%
2	21,647	1924	17.15	4.1%
3	21,455	1920	7.04	9.7%
4	16,131	1915	3.11	14.4%
5	13,321	1892	7.19	10.1%

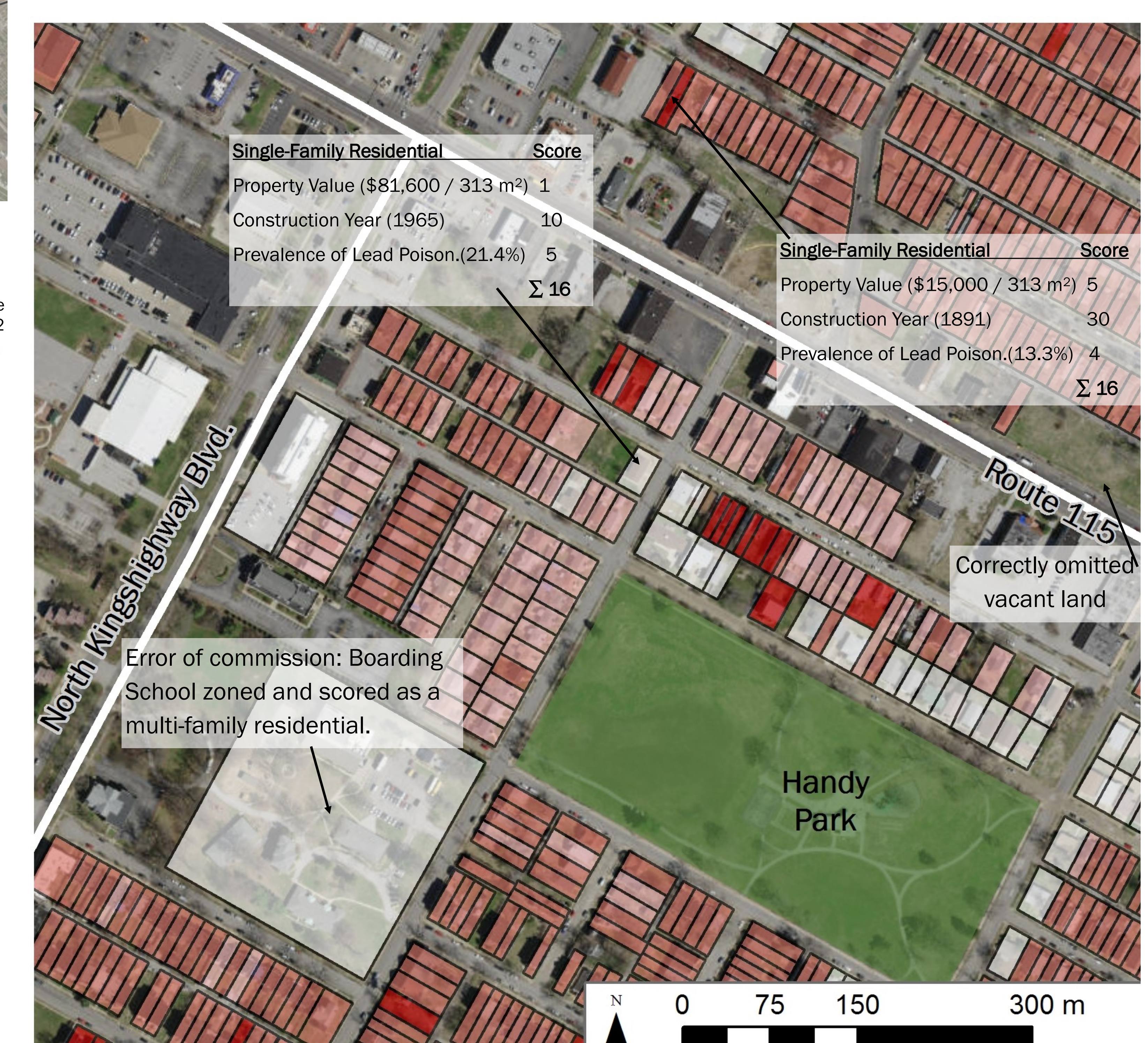
EVALUATION: This model is an attempt to distribute scarce resources in an unbiased manner, based on known risk factors for lead paint hazards. However, the model is highly arbitrary in its own ways. Other GIS studies that created similar models used individual-level data to validate them. St. Louis doesn't have that luxury, as the state doesn't distribute individual-level data.

The model identifies an excessive number of parcels for screening (N = 13,321). The cut-off value for targeting should be raised, or the risk factors re-evaluated. Greater differentiation between parcels could be achieved by weighting the risk factors in their continuous forms, or increasing by the number of categories. The continuous variables for year, prevalence, and property value are as statistically significant in predicting the final risk score as their categorical counterparts ($p < 2 \times 10^{-6}$ for all β). Categorization makes the model simpler to explain, and doesn't appear to bias the final risk score.

SOURCES OF BIAS: The aggregation of prevalence into census tracts was used for this model. If the available aggregations by neighborhoods, wards, or zip codes had been used instead, the model would have given different results. This is a consequence of the modifiable areal unit problem. In future models, the overlap between these irregularly-shaped areal units can be leveraged to create a less granular estimate of where case distribution. Prevalence suffers from epidemiological biases as well. In areas with low healthcare access, only symptomatic children may be screened [2]. Healthcare providers' perception of risk also affects the likelihood that children will be screened annually [2].

Property value is a very rough proxy for housing quality. Value is normalized by the building footprint area, which isn't equal the total area for multistoried buildings. This may bias buildings' value in denser and lower income parts areas, where multifamily residences are more common. A more precise property value would come from sales data, but the municipal dataset for this was incomplete.

Though residential lead paint hazards are the most common source of lead, children could be exposed outside the home. Any property where children under 6 spend more than 10 hours per week, like that of a caretaker or family member, poses a risk [2] . Other environmental sources of lead such as heavy traffic and industry, which tend to cluster in low-income areas, aren't directly accounted for [2]. Therefore, the final risk scores likely underestimate the differences between areas.



GIS Data Sources

[1] Zoning, Construction Year, Property Value:
Parcels 1997 – 2013; January 2015; City of St. Louis Real Estate Records, Planning & Urban Design Agency
[2] Lead Poisoning Prevalence:
Childhood Lead Poisoning in the City of Saint Louis; 2013; City of St. Louis Department of Health.
[3] Base Map
World Imagery. Esri, HERE, Garmin, OpenStreetMap Contributers, and the GIS user community.

Literature Sources

[4] Miranda, M.L., D.C. Dolinoy, and M.A. Overstreet, Mapping for prevention: GIS models for directing childhood lead poisoning prevention programs. Environmental Health Perspectives, 2002. 110(): p. 947-953.
[5] Jacobs, D.E., et al., *The prevalence of lead-based paint hazards in U.S. housing.* Environmental Health Perspectives, 2002. **110**(10): p. A599-A606.

Daniele Todorov
Tufts Friedman School of Nutrition Science and Policy
NUTR 231: Fundamentals of GIS
May 16th, 2018

Eastern Missouri State Plane Projection (NAD 1983) was used for all maps. Maps were constructed in ArcMap. Statistics and graphs were constructed in R.