ISSN: 2753-7919
British Journal of Healthcare and Medical Research; Vol. 13 No. 01 (2026) (204-216)
https://doi.org/10.14738/bjhr.1301.19722

Spatial Eigenvector Predictive Mapping Using a Second Order Eigenfunction Eigen-decomposition Autocorrelation Specification in a Regional Convolution Neural Network in a Geo-Intelligent Smartphone App to Identify Lead Contamination in Primary and Secondary Schools in Hillsborough County, USA

Janessa Monchery[1], Namit Choudhari[2], Anusha Parajuli[3], Aarya Satardekar[4], Caleb Jaramillo[5], Sasha Mosich[6], & Benjamin G. Jacob[7]

1. Samuel P. Bell III College of Public Health, University of South Florida, Tampa, FL, USA
2. School of Geosciences, College of Arts & Sciences, University of South Florida, Tampa, FL, USA
3. Department of Epidemiology, Samuel P. Bell III College of Public Health, University of South Florida, Tampa, FL, USA
4. Department of Health Policy and Systems Management, Samuel P. Bell III College of Public Health, University of South Florida, Tampa, FL, USA
5. Global Communicable Diseases, Samuel P. Bell III College of Public Health, University of South Florida, Tampa, FL, USA
6. Department of Epidemiology, Samuel P. Bell III College of Public Health, University of South Florida, Tampa, FL, USA
7. Department of Biostatistics and Data Science, Samuel P. Bell III College of Public Health, University of South Florida, Tampa, FL, USA

Abstract: Currently, there are no lead-contamination models to describe the locations of aggregation/non-aggregation sites (i.e., hot-and-cold spots) in Hillsborough County, Florida, USA. We generated a Global Moran's index (I) of spatial autocorrelation to identify hot-and-cold spots of lead contamination in Hillsborough County. The data used were based on lead sampling of water sources within the internal environments of elementary, middle, and high schools. A second-order eigenfunction eigen-decomposition algorithm embedded in a regional convolutional neural network (R-CNN) machine learning [ML] in a geo-spatial artificial intelligence [geo-AI] smartphone application yielded a model that identified Kingswood Elementary School as the most clustered geographic location with the highest lead contamination concentration levels. The spatial autocorrelation model also identified Essrig Elementary School as the least clustered geographic capture point. The Moran's I diagnostic summary plot revealed a final Z-score of 8.347 and a p-value of 0.000. Mapping hotspots of lead concentration using a second-order eigenfunction eigen-decomposition algorithm within an instantaneous R-CNN, ML, geo-AI, iOS pipeline can allow school administrative boards and policymakers to allocate resources to at-risk areas with lead contamination.

Keywords: Lead contamination, Hot spot, Eigen-decomposition Algorithm, R-CNN, ML, Geo-AI, iOS, Hillsborough County, Blood Lead Level

INTRODUCTION

Lead poisoning can be defined as having elevated levels of lead present in the blood [1]. There is no safe blood lead level for children [2]. Currently, the Centers for Disease Control and Prevention (CDC) has set a blood lead level [BLL] of 3.5 µg/dL as a higher-risk marker for adverse physiological effects [3,4]. Even though efforts have been made to eliminate lead in paint, gasoline, and other products, it is still present in household items such as cans, ceramic dishes, crayons, and brightly painted toys, particularly those imported from different countries. When children are exposed to lead, it can cause decreased intelligence, dysfunctional behavior, and difficulties with learning [2]. The effects of lead and the range of their severity worsen as the Intelligence Quotient (IQ), underperformance in school, and the development of Attention-Deficit/Hyperactivity Disorder (ADHD) in children [3]. Lead exposure is not readily detectable in children unless they are exposed to extremely high levels; the Centers for Disease Control and Prevention (CDC) recommends that professionals use blood lead tests for children at high risk of exposure [4]. The adverse effects of exposure are more of a concern in younger, developing children, who can absorb up to 4-5 times more ingested lead from a particular source than adults [2].

It is estimated that over 170 million Americans were exposed to high lead levels in their early childhood, with 31% of the population having childhood BLL of above 10 µg/dL, three times higher than the CDC's reference level for clinical concern [5]. Hillsborough County has a disproportionately high number of lead cases, ranking second-highest in Florida [6]. Currently, there are no studies examining childhood lead exposure and lead-contaminated water in Hillsborough County, specifically regarding schools as a potential source of exposure.

Spatial eigenvector mapping may provide a robust statistical framework for analyzing the distribution of lead contamination prevalence at fine geographic scales such as primary and secondary schools. By deriving eigenvectors from spatial weighting matrices—typically based on geographic contiguity or distance matrix —an epidemiologist or clinician could isolate latent spatial patterns that traditional approaches for lead contamination prevention mapping have not identified in the literature. These eigenvectors may map fine-scale and localized clustering of lead-contaminated schools, thereby allowing the delineation of true spatial processes (e.g., environmental, socioeconomic, or health-system-access factors) and other artifacts that may initiate or exacerbate mental and physical development problems, which may be driven by spatial autocorrelation at the zip code level.

Spatial autocorrelation is the correlation amongst observations that arises from their relative positions in geographic space [7]. Spatial autocorrelation reflects the fundamental idea that nearby geographic locations tend to be more similar (positive autocorrelation) or more dissimilar (negative autocorrelation) than would be expected under spatial randomness. Jacob et al. 2023 emphasized that non-zero autocorrelation is not noise but a meaningful structural property of spatial epidemiological data sets that must be accounted for in spatial statistical modeling [8].

In practical terms, a machine learning [ML] algorithmic, spatial eigenvector lead contamination model may provide a framework for treating spatial autocorrelation as a measurable pattern resulting from the configuration of units for optimizing predictive vulnerability modeling in primary and secondary schools, using online census tract data. An ML spatial eigenvector model is an approach that combines spatial autocorrelation modeling via eigenvectors generated using ML algorithms such as neural networks, which can integrate spatial adjacency using a regional convolutional neural network [R-CNN][9].

In this research, we employed ML algorithms for geographic prediction tasks in R-CNN to code county-and zip-code-level primary and secondary school census-tract outcomes to robustify hot-and-cold spot mapping of lead contamination vulnerability in Hillsborough County. We assumed that ignoring spatial autocorrelation patterns in a predictive, zip-code-based epidemiological lead contamination model may lead to biased estimates, inefficient models, and incorrect inference, as standard statistical techniques assume independence among the sampled sociodemographic and socioeconomic empirical georeferenced observations.

R-CNN is a convolutional neural network adapted for geographic regions instead of pixels. Traditional convolutional neural networks [CNN] work on images (structured grids). Counties/zip codes are irregularly shaped, so we converted them into a spatial adjacency matrix. Subsequently, the CNN understood the regional spatial patterns of the sampled, non-time-series-dependent zip code estimators for mapping regional lead contamination hotspots in Hillsborough County primary and secondary schools. We assumed that local spatial gradients would allow mapping lead contamination vulnerability at the zip code level, which in turn would allow differentiating socioeconomic gradients, demographic pockets, etc.

We initially incorporated multiple georeferenced zip codes, sampled sociodemographic and socioeconomic interpolative estimators, and a county-prognostic, eigen-decomposable variable into a hot-and-cold spatial autocorrelation model. Subsequently, we employed spatial eigenvector mapping in a CNN to extract eigenvectors that delineated varying scales of spatial autocorrelation. We operationalized the model estimators by employing eigenvector mapping to reveal zip-code-stratified, georeferenced, lead-contamination-stratified hot-and-cold spots in Hillsborough County. Our assumption was that spatial eigenvector lead contamination autocorrelation mapping would support nuanced, multilevel parameter estimation by accounting for spatial dependence across nested, hot-and-cold-spot geographic units, i.e., primary and secondary schools at the zip code level.

Counties can exhibit regional gradients in lead contamination, shaped by healthcare infrastructure, long-term care availability, or socioeconomic/sociodemographic conditions. At the same time, primary and secondary schools within Hillsborough County's zip codes may contain micro-contexts with distinct risk profiles. Incorporating spatial eigenvectors into a lead contamination vulnerability statistical model may enable the separation of these overlapping spatial influences, improving the accuracy of estimates and reducing biased inference from hot-and-cold spots, lead contamination, predictive, primary/secondary school zip code, and epidemiological risk models. Emerging evidence from a spatial eigenvector map may indicate the geographic location of lead contamination, down to the primary and secondary school neighborhood level, which may influence prevalence and incidence. Our research hypothesis was: Spatial eigenvector mapping can identify localized lead contamination, stratified into hot-and-cold spot clusters, and the sociodemographic and socioeconomic factors that contribute to them.

Our intention in this research was to implement a real-time social messaging platform inside a geo-spatial artificial intelligence [geo-AI] neural network iOS pipeline. ML, object-based eigen-algorithmic, generated lead-contamination-sampled county-level data was meshed into an R-CNN, which allowed combining georeferenced sociodemographic, socioeconomic, and zip code online census data into one unified network. Jacob et al. (2024) employed a real-time hot-and-cold spot detection network using an iOS R-CNN recognition pipeline [9]. The authors of Jacob et al. (2024) introduced a regional proposal network [RPN] that shared full-image convolutional autocorrelated features with a real-time detection geo-AI, ML, and interactive recovery capital application [app]. In this article, we introduced an RPN in a CNN deep convolutional network in a geo-intelligent smartphone app to aggregate and predict object bounds and objectiveness scores for implementing solutions to lead contamination in primary and secondary schools in Hillsborough County.

The research aims in this article were: 1. to employ a second-order eigenfunction eigen-decomposition spatial autocorrelation algorithm in an R-CNN to identify hot-and-cold spots of lead contamination in primary and secondary schools, and 2. to determine a social messaging lead contamination prevention program in Hillsborough County, Florida, USA, in a geo-intelligent smartphone application.

METHODS

The study area is Hillsborough County, Florida, home to 1.56 million residents, with land use including urban, rural, and farmland [10]. For this study, we are focusing on school-age children <5 years old and between 5 and 14 years old. We retrieved census data from the National Historical Geographic Information System (NHGIS) on the ages of county residents [11]. Then we categorized them in ArcGIS into two groups: ages below 5 years and ages between 5 and 14 years. Using ArcGIS, a bivariate choropleth map was generated to identify areas with populations of children under five and 5-14 years old (see Figure 1).

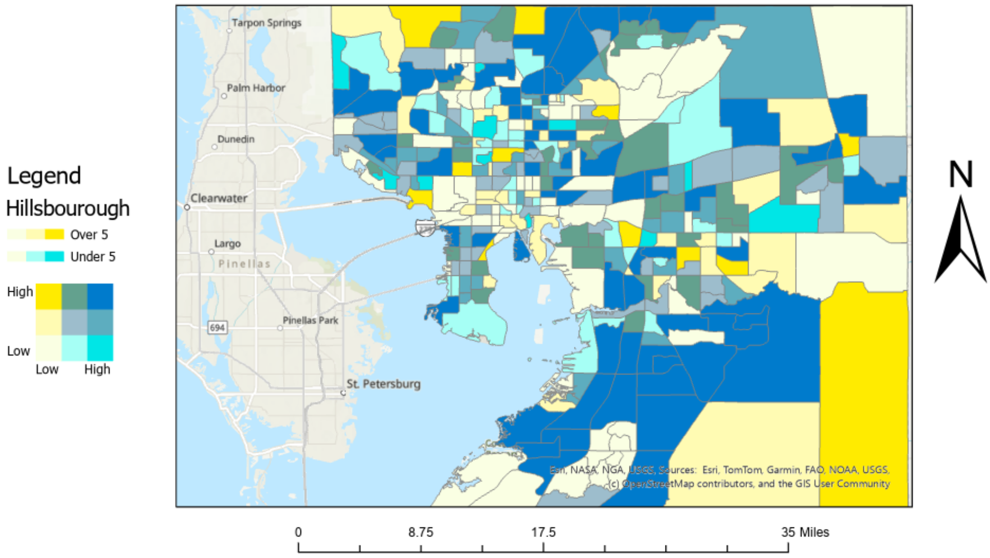

Figure 1: An ArcGIS bivariate choropleth map showing ages below 5 years and ages between 5-14 years in Hillsborough County

To generate a Moran's *I* second-order spatial autocorrelation, georeferenced primary and secondary schools lead water testing results were used. This data was published on the Hillsborough County Public Schools website between February 14th, 2025, and September 19th, 2025 [12]. We then evaluated the average lead amount per school by initially identifying sites with lead and then dividing the amount tabulated by the sum of the lead ppb (parts per billion) across testing samples. In so doing, we identified primary and secondary schools throughout the county.

We constructed a spatial eigenvector Moran's *I* lead contamination county zip code level hot-and-cold spot autocorrelation map in R. All sampled lead-related school covariates were then evaluated in a spatial error (SE) model. An autoregressive residual clustering model was employed, with a lead-related dependent variable Y, where Y was a function of nearby sampled Y lead values [i.e., an autoregressive response (AR) or spatial linear (SL) specification] in geographic space. The residuals of Y in the lead contamination hot-and-cold model were a function of nearby Y residuals [i.e., an AR or SE specification]. Euclidean distance between sampled sample sites was quantified in terms of an n-by-n geographic weights matrix, C, whose c_{ij} values were 1 if the sampled lead school locations *i* and *j* were deemed nearby and 0 otherwise. Adjusting this matrix by dividing each row entry by its row sum, with the row sums given by C1, converted this matrix to matrix W.

The n-by-1 vector x = [x1 ⋯ x n]T in the lead contamination model contained measurements of a quantitative variable for n spatial zip code units, which were embedded in an n-by-n spatial weighting matrix W. The formulation for the Moran's *I* of spatial autocorrelation used in this research was: $I(x) = \dfrac{n\Sigma_{(2)} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\Sigma_{(2)} w_{ij} \Sigma_{i=1}^{n}(x_i - \bar{x})^2}$

where $\sum_{(2)} \sum_{i=1}^{n} \sum_{j=1}^{n}$ with $i \neq j$.

The values w_{ij} in the lead contamination model were spatial weights stored in the symmetrical matrix W [i.e., ($w_{ij} = w_{ji}$)] that had a null diagonal ($w_{ii} = 0$). The matrix was initially generalized to an asymmetrical matrix W. Matrix W was generalizable by a non-symmetric matrix W^* by using $W = (W^* + W^{*T})/2$. Moran's *I* was rewritten using a matrix notation: $I(x) = \dfrac{n}{1^t W 1} \dfrac{x^T HHWHHx}{x^T HHx} = \dfrac{n}{1^T W 1} \dfrac{x^T HWHx}{x^T Hx}$ where H = (I - $11^T/n$) was an orthogonal projector verifying that $H = H^2$, (i.e., *H* was independent). Matrix W analyzed the lead sampled covariates. Matrix W is a stochastic matrix that expresses the georeferenced observed value y_i as a function of the average [7]. In this research, the sampled, sentinel site, zip code, and primary or secondary school location was delineated using *i*'s which were the neighboring lead counts in geographical space, allowing a single spatial autoregressive parameter, ρ, to have a maximum value of 1.

Next, a model specification was employed to describe the autoregressive variance uncertainty estimates in the lead contamination model. A spatial filter (SF) model specification was also used to describe both Gaussian and Poisson random lead sampled covariates. The resulting model specification took on the following form: $\mathbf{Y} = \mu(1 - \rho)\mathbf{1} + \rho\mathbf{W}\mathbf{Y} + \varepsilon,$ (2.1) where μ was the scalar conditional mean of Y, and ε was an *n*-by-1 error vector whose elements were statistically independent and identically distributed (i.i.d.) normally random variates. The spatial covariance matrix for equation (2.1) employed the sampled lead covariates, which were written as E [(Y - μ l)' (Y - μ l)] = Σ = $[(I - \rho W')(I - \rho W)]^{-1}\sigma^2$, where E (●) denotes the calculus of expectations. In the lead contamination model, I was the n-by-n identity matrix denoting the matrix transpose operation, and σ^2 was the error variance.

However, when a mixture of positive and negative eigen-spatial autocorrelation is present in a model, a more explicit representation of both effects leads to a more accurate interpretation of empirical results [7]. Alternatively, the excluded values may be set to zero, but if so, the mean and variance must be adjusted [8]. In this research, two different spatial autoregressive lead contamination stratified parameters appeared in the spatial covariance matrix model specification, which, in our model, were transformed as: $\Sigma = [(\mathbf{I} - <\rho>_{\text{diag}} \mathbf{W}')(\mathbf{I} - <\rho>_{\text{diag}} \mathbf{W})]^{-1}\sigma^2$, (2.2).

In the lead contamination, georeferenced, hot-and-cold-spot-county-zip code model, the diagonal matrix of autoregressive parameters, $<\rho>_{\text{diag}}$, contained two sampled parameters: ρ_+ for covariates displaying positive spatial dependency, and ρ for those covariates displaying negative spatial dependency. For example, by letting $\sigma^2 = 1$ and employing a 2-by-2 regular square tessellation, for the vector $\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$, enabled positing a positive relationship between the sampled lead contamination covariates, y_1 and y_2, a negative relationship between covariates, y_3 and y_4, and no relationship between covariates y_1 and y_3 and between y_2 and y_4. This covariance specification yielded:

$$\mathbf{Y} = \mu(\mathbf{I} - \rho_+ < \mathbf{I}_+ >_{\text{diag}} - \rho_- < \mathbf{I}_- >_{\text{diag}})\mathbf{1}, +(\rho_+ < \mathbf{I}_+ >_{\text{diag}} + \rho_- < \mathbf{I}_- >_{\text{diag}})\mathbf{W}\mathbf{Y} + \varepsilon$$

(2.3) where I_+ was a binary 0-1 indicator variable which denoted those lead contamination covariates displaying positive spatial dependency, and I_- was a binary 0-1 indicator variable denoting those sampled covariates displaying negative spatial dependency, using $I_+ + I_- = 1$. If either $\rho_+ = 0$ (and hence $I_+ = 0$ and $I_- = I$) or $\rho_- = 0$ (and hence $I_- = 0$ and $I_+ = I$), then equation (2.3) reduces to equation (2.1) [7,8].

A Jacobian estimation was implemented by utilizing the differenced indicator lead contamination stratified explanatory variables ($I_+ - \gamma I_-$), estimating ρ_+ and γ with ML techniques, and setting $\hat{\rho}_- = -\hat{\gamma}\hat{\rho}_+$. The Jacobian generalizes the gradient of a scalar-valued function of multiple variables, which itself generalizes the derivative of a scalar-valued function of a scalar [7]. A more complex specification was then posited by generalizing these binary indicator variables. We used $F: R^n \rightarrow R^m$ as a function from Euclidean n-space to Euclidean m-space, which was generated in the lead contamination hot-and-cold spot model using the Euclidean distance measurements between sampled covariates. Such a function was given by m covariate (i.e., component functions), $y_1(x_1, xn)$, $y_m(x_1, xn)$. The partial derivatives of all these functions were organized in an m-by-n matrix, where the Jacobian matrix J of F was delineated as follows: $J = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$. This matrix was denoted by $J_F(x_1,..., x_n)$ and $\frac{\partial(y_1,...,y_m)}{\partial(x_1,...,x_n)}$ where the i th row (i = 1,..., m) of this matrix was the gradient of the i^{th} component function y_i :(∇y_i). In this analysis, p was a sampled georeferenceable lead contamination stratified covariate in R^n, and F (i.e., sampled socio-demographic or socio-economic count variable) was differentiable at p; its derivative was given by $J_F(p)$. The model described by $J_F(p)$ was the best linear approximation of F near the capture point p (primary or secondary school) in the sense that: $F(x) = F(p) + J_F(p)(x - p) + o(\|x - p\|)$ (2.4).

The spatial structuring in the lead contamination model was constructed using a linear combination of a subset of the eigenvectors of a modified geographic weights matrix. We employed (I - 11'/n) C (I - 11'/n) in the numerator of the Moran's Coefficient (MC). Spatial autocorrelation can be indexed with an MC, a product moment correlation coefficient [8]. A subset of eigenvectors was then selected with a stepwise regression procedure. Because in the county lead contamination model (I - 11'/n) C (I - 11'/n) = E Λ E', occurred when E was an n-by-n matrix of eigenvectors and Λ was an n-by-n diagonal matrix of the corresponding eigenvalues, the resulting model specification was provided by: $\mathbf{Y} = \mu\mathbf{1} + \mathbf{E}_k\beta + \varepsilon$, (2.5) where μ the scalar mean of Y and Ek was an n-by-k matrix containing the subset of k <<n eigenvectors. This was selected in R using a stepwise regression technique where ß was a k-by-1 vector of regression coefficients.

RESULTS

A number of the eigenvectors were extracted from (I - 11'/n) C (I - 11'/n), which were associated with geographic patterns of the lead contamination covariates sampled in the Hillsborough County study site, indicating a negligible degree of eigen-spatial autocorrelation. Consequently, only k of the n eigenvectors were of interest for generating a candidate set for a stepwise regression procedure. A candidate eigenvector represents a level of spatial autocorrelation that can account for the redundant information in orthogonal map patterns [7]. The preceding eigenvector properties in the lead contamination spatial autocorrelation model resulted in $\hat{\mu} = \bar{y}$ and $\hat{\beta} = \mathbf{E}_k'\mathbf{Y}$ for equation (3.1). Expressing equation (3.2) in terms of the preceding 2-by-2 example yielded

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \mu \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 0.5 & -0.69048 & 0.15240 \\ -0.5 & 0.15240 & 0.69048 \\ -0.5 & -0.15240 & -0.69048 \\ 0.5 & 0.69048 & -0.15240 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix}, \text{ and}$$

$$\hat{\mu} = \frac{y_1+y_2+y_3+y_4}{4} \text{ and } \hat{\beta} = \begin{pmatrix} 0.5 & -0.69048 & 0.15240 \\ -0.5 & 0.15240 & 0.69048 \\ -0.5 & -0.15240 & -0.69048 \\ 0.5 & 0.69048 & -0.15240 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}. \quad (3.3)$$

Of note is that the 2-by-2 square tessellation rendered a repeated eigenvalue.

To identify lead-stratified spatial clusters of lead contamination in the Hillsborough County primary and secondary schools, we georeferenced stratified Thiessen polygons and generated a surface partitioning in R. To construct geographic neighbor matrices, we used MC in the eigen-spatial autocorrelation analysis. Entries in the matrix were 1 if two sampled zip code hot or cold spot geolocations shared a common Thiessen polygon boundary and 0 otherwise. Next, the linkage structure for each surface was edited to remove unlikely geographic neighbors, thereby identifying sampled lead contamination stratified covariates that share a common Thiessen polygon boundary. Attention was restricted to those map patterns associated with at least a minimum level of spatial autocorrelation, which, for our model implementation purposes, was defined by $|MC_j/MC_{max}| > 0.25$, where MC_j denoted the j th value and MC_{max} the maximum value of MC. This threshold value allowed two candidate sets of eigenvectors to be considered for substantial positive and substantial negative spatial autocorrelation, respectively. These statistics indicated that the detected negative spatial autocorrelation, may be statistically significant from a randomization perspective.

Of note is that the ratio of the PRESS (i.e., predicted error sum of squares) statistic to the sum of squared errors from the MC scatterplot trend line was 1.21 which was well within two standard deviations of the average standard prediction error value (roughly 1.13) for a sampled hot spot in the county study site. The maximum value of I was obtained by all the variation of z, as explained by the eigenvector u_1, which corresponded to the highest eigenvalue λ_1 in the eigen-spatial autocorrelation error matrix.

In the lead contamination model forecast, $cor^2(u_i, z) = 1$ (and $cor^2(u_i, z) = 0$ for $i \neq 1$), and the maximum value of I, was deductible for Equation (3.7), which was equivalent to $I_{max} = \lambda_1(n/1^TW1)$. The minimum value of I in the error matrix was obtained by showing that all the variation of z was explained by the eigenvector u_{n-c} corresponding to the lowest eigenvalue λ_{n-r} generated in the lead contamination stratified model. This minimum value was equal to $I_{min} = \lambda_{n-r}(n/1^TW1)$. If the sampled predictor variable was not spatialized, the part of the variance explained by each eigenvector was equal, on average, to $cor^2(u_i, z) = 1/n\text{-}1$. Because the sampled lead specified variables in z were randomly permuted, it was assumed that we would derive this result. In the county lead contamination model, the set of $n!$ random permutations, revealed that $\mathrm{E}_R(I) = \frac{n}{1^TW1(n-1)}\sum_{i=1}^{n}\lambda_i = \frac{n}{1^TW1(n-1)}trace(\Omega)$. It was easily demonstrated that $trace(\Omega) = -\frac{1^TW1}{n}$ and it followed that $\mathrm{E}_R(I) = -\frac{1}{n-1}$ in the Hillsborough County school lead contamination hot-and-cold spot autocorrelation model.

After constructing a second-order spatial autocorrelation lead contamination predictive map, we generated an inverse-distance-weighted interpolation (IDW). This model revealed that nearby, stratified socioeconomic and georeferenced capture points had more influence on hot-and-cold spots than more distant ones. IDW weights use Euclidean distances from hot-and-cold spot covariates to quantify power parameters for controlling the rate of decay [7]. In this research, we generated an IDW model in R.

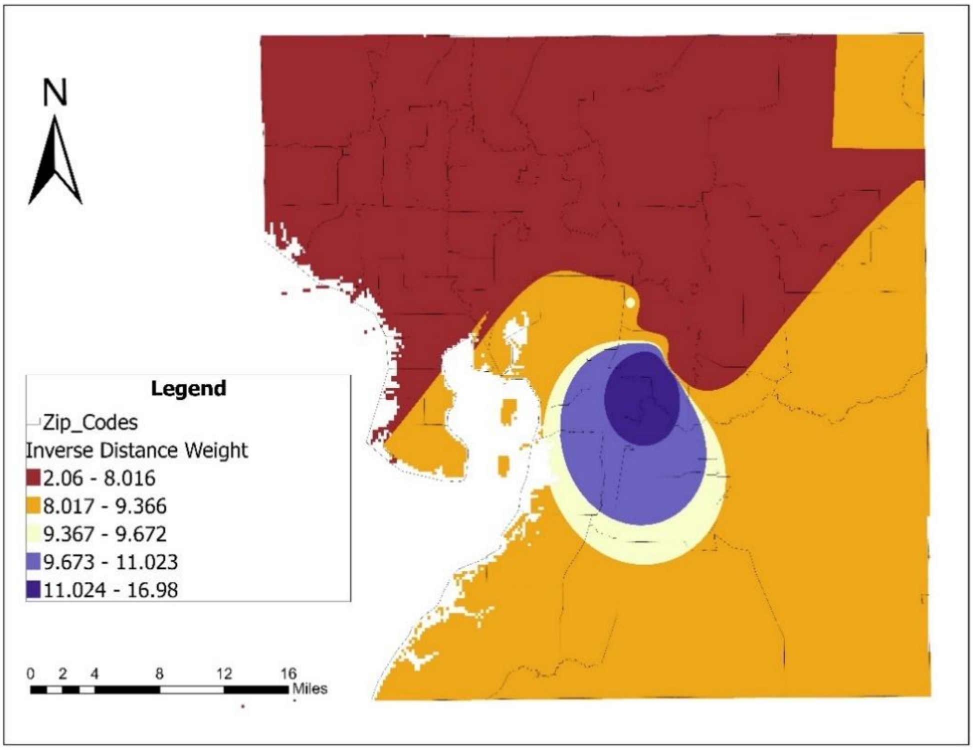

Figure 2: Inverse Distance Weight Map

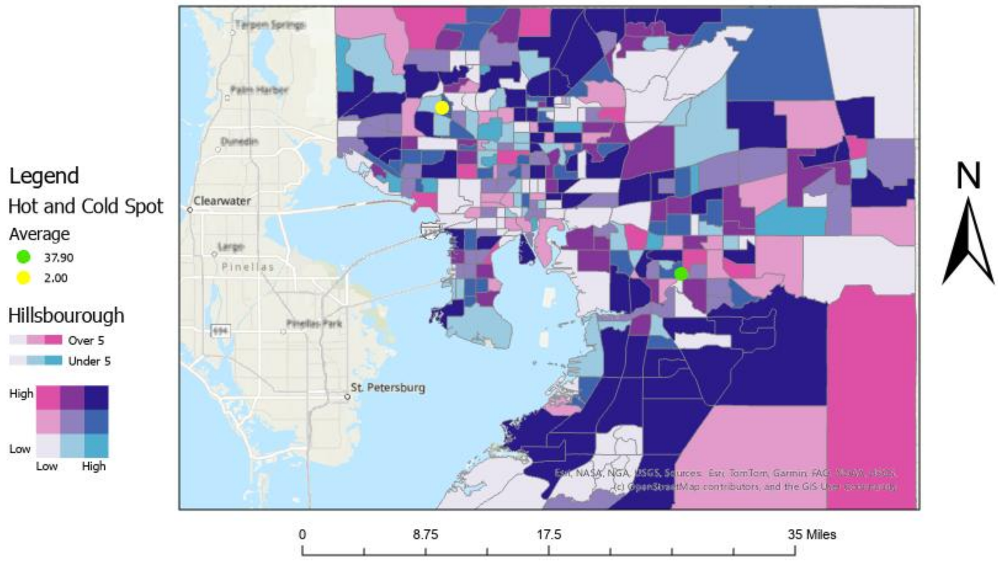

Figure 3: Hot-and-Cold spot map of Hillsborough County, Florida

DISCUSSION

Kingswood Elementary was identified as the primary hotspot for lead exposure in Hillsborough County. With a ppb level of 37.90, Kingswood's lead level is over 2.5 times the EPA action limit of 15 ppb [a measure of the effectiveness of corrosion control treatment in water systems] [13,14]. This school is located in an area where there is an equal dispersion of children under 5 years old and 5-14 years old, revealing that there are many vulnerable children who could face developmental and cognitive impairments because of this exposure. Elevated BLL is seen to be positively associated with increased attention dysfunction, aggression, and delinquent behavior, with prior research showing that children with elevated BLL have higher rates of dropping out of school and reading disabilities [15].

Essrig Elementary was identified as a cold spot for lead exposure with an average ppb of 2.00. Even though this school has lead in its water, the levels are below the EPA-mandated action limit. Cold-spot aggregation locations of lead contamination in primary and secondary schools require annual monitoring and surveillance to ensure they have not transitioned into hot spots. Regrettably, due to changing sociodemographic and socioeconomic conditions at the zip code level in Hillsborough County, this may result in varying levels of lead contamination in schools.

As previously mentioned, lead can have detrimental effects on children's school performance. This is evident in students' test scores at Essrig and Kingswood Elementary. When evaluated consistently within all categories, Essrig Elementary outperformed Kingswood Elementary students in English, Math, and Science [16]. Except for math, the majority of students at Kingswood are not performing at grade level.

To evaluate the socioeconomic status of children at the two schools, we used free- or reduced-lunch eligibility as a metric. At Essrig Elementary School, 52% of students are receiving free or reduced lunches [17]. In contrast, at Kingswood Elementary School, an average of 75% of students received free or reduced lunches, indicating that the school may be populated by students from lower socioeconomic backgrounds [18]. Currently, the percentage of schoolchildren receiving free or reduced-price lunches in Florida is 53.8% [19]. Therefore, Essrig's student body is economically on par with the average socioeconomic status of Florida.

In contrast, in terms of the state's percentage, Kingswood schoolchildren are less economically advantaged. Based on our research findings, providing families with the resources to identify their child's BLL correctly is pivotal to addressing the state's growing youth population. An example of such would be Florida's Lead Poisoning Prevention Program, which conducts blood lead testing and promotes screening for high-risk populations, an intervention that would be ideal to address this issue [20].

Introducing novel RPNs that share convolutional layers with an R-CNN, state-of-the-art ML, geo-AI algorithmic object detection, and real-time networks, infused into an interactive iOS app geo-intelligent dashboard, can aid in implementing lead contamination prevention strategies (e.g., calculating hot spots using socioeconomic and sociodemographic stratified zip codes) in Hillsborough County. Here, we employed a smartphone app's website navigation software and GPS functionality to delineate lead contamination stratified hot-and-cold spots. By sharing convolutions in real time at test time, the marginal cost of computing these proposals would be small (e.g., 10ms per retrieval of a georeferenced lead contamination capture point, sentinel site, or geolocation). A neural network architecture can retrieve lead-contaminated, covariate convolutional feature attribute data capture points throughout the County's primary and secondary schools by using R-CNN in an interactive app.

Convolutional features of lead contamination, primary and secondary georeferenced schools, constructed in an RPN within an interactive mobile app, would allow for adding a convolutional layer that simultaneously regresses region bounds and objectness scores at each sampled lead contamination capture point, employing a stratifiable zip code grid. Hence, the RPN in the app would be a fully convolutional network. As such sentinel site information data of lead contamination stratified capture points could be trained end-to-end specifically for the task of generating robust data, which can provide real-time information to family members, and other support specialists of lead in primary and secondary schools in Hillsborough County in so doing an epidemiologist or other research collaborator could establish a data capture point of hot-and-cold lead contamination stratified socioeconomic and sociodemographic sentinel sites. A suggested iOS geo-intelligent app could include: 1. Create a QR code that can be scanned at lead contamination meeting halls, 2. Click on the "Start your lead contamination prevention lecture"– available to all pathways hosted on the app, 3. Click on and watch till the end for any of the guided instructions on the app dashboard 4. Click on instructional videos developed by lead contamination specialists in the county 5. Click on and watch/listen to the lead contamination prevention in primary and secondary schools on Hillsborough County podcasts, 6. Submission of support message on the app's social media platform, 7. Completion of a lead contamination assessment, 8. Completion/update of a lead contamination prevention management plan.

Python-coded interactive apps are a revolutionary technology that may change the face of programming. This allows users to write code in Python, one of the most popular and readily accessible programming languages currently. With its simple user interface and powerful features, Python can be infused into an interactive, geo-intelligent smartphone app designed to help lead contaminated primary and secondary schools in Hillsborough County on their journey to recovery. Python coding can enable the construction of real-time, graphical time series models for determining changes in blood levels in tested school children based on academic performance. The app introduces users to important coding concepts, such as explanatory data variables related to lead contamination, loops, and functions for gamification, participant personal data logging, advanced security encryption, and more. The lead contamination smartphone app can transcribe professional and personal interviews with specialists in real time. This material can then be analyzed thematically by coding and implementing various lead contamination prevention tactics (e.g., scoring dashboards and token-based incentive programs) for primary and secondary schools in Hillsborough County.

In future research, we would like to scale up the hot-and-cold-spot model across Florida. In doing so, we could develop a statewide model for lead contamination assessment and prevention, enabling the spatial targeting of schools across Florida's county ZIP codes. This would allow school administrative boards and policymakers to further allocate resources [blood lead level testing in schools] to regions of concern.

REFERENCES

[1] Mayo Foundation for Medical Education and Research. (2025). Lead poisoning. Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/lead-poisoning/symptoms-causes/syc-20354717

[2] Heidari, S., Mostafaei, S., Razazian, N., Rajati, M., Saeedi, A., & Rajati, F. (2022). The effect of lead exposure on IQ test scores in children under 12 years: A systematic review and meta-

analysis of case-control studies. Systematic Reviews, 11(1). https://doi.org/10.1186/s13643-022-01963-y

[3] Donzelli, G., Carducci, A., Llopis-Gonzalez, A., Verani, M., Llopis-Morales, A., Cioni, L., & Morales-Suárez-Varela, M. (2019, January 29). The association between lead and attention-deficit/hyperactivity disorder: A systematic review. International Journal of Environmental Research and Public Health. https://pmc.ncbi.nlm.nih.gov/articles/PMC6388268/

[4] Kamat, A. (2024). Screening for lead poisoning | Pediatric Care Online. Screening for Lead Poisoning. https://publications.aap.org/pediatriccare/article/doi/10.1542/aap.ppcqr.396030/62/Screening-for-Lead-Poisoning

[5] McFarland, M. J., Hauer, M. E., & Reuben, A. (2022). Half of US population exposed to adverse lead levels in early childhood. Proceedings of the National Academy of Sciences, 119(11). https://doi.org/10.1073/pnas.2118631119

[6] Centers for Disease Control and Prevention. (2024). Florida. Centers for Disease Control and Prevention. https://www.cdc.gov/lead-prevention/success-stories-by-state/florida.html

[7] Griffith DA (2003) Spatial autocorrelation and spatial filtering: Gaining understanding through theory and scientific visualization. Springer, Berlin, Germany.

[8] Jacob BG, Izureta R, Bell J, Parikh J, Aceng JR, et al. (2023) Approximating non-asymptoticalness, skew heteroscedascity and geo-spatiotemporal multicollinearity in posterior probabilities in Bayesian eigenvector eigen-geospace for optimizing hierarchical diffusion-

[9] Jacob, B., McDonald, H., & Bohn, J. (2024). Closing the gap on addiction recovery engagement with an AI-infused convolutional neural network technology application—a design vision. American Journal of Neural Networks and Applications, 10(1), 1–14. https://doi.org/10.11648/j.ajnna.20241001.11

[10] Hillsborough County - Data Commons. Home - Data Commons. (n.d.). https://datacommons.org/place/geoId/12057?q=hillsbourough%2Bcounty%2Bflorida

[11] Schroeder, J., Van Riper, D., Manson, S., Knowles, K., Kugler, T., Roberts, F., & Ruggles, S. (2025). IPUMS National Historical Geographic Information System (Version 20.0) [Data set]. IPUMS. https://doi.org/10.18128/D050.V20.0

[12] Hillsborough County Public Schools. Current Water Test Results | Hillsborough County Public Schools. (2025). https://www.hillsboroughschools.org/o/hcps/page/current-water-test-results

[13] Voluntary lead testing program for schools and childcare facilities. Florida Department of Environmental Protection. (2024). https://floridadep.gov/water/certification-restoration/content/voluntary-lead-testing-program-schools-and-childcare

[14] Environmental Protection Agency. (2025). What do you mean when you say the action level has been exceeded for a drinking water system?. EPA. https://www.epa.gov/lead/what-do-you-mean-when-you-say-action-level-has-been-exceeded-drinking-water-system

[15] Lead exposure in children. Home. (2025). https://www.aap.org/en/patient-care/lead-exposure/lead-exposure-in-children/?srsltid=AfmBOope0fzLTFCVvvD6PqIwuCKGBdbr_oMt9BeZ8z0HF5r7AG19r7ut

[16] Florida Assessment of Student Thinking and B.E.S.T. Assessments. Florida Department of Education. (2025). https://www.fldoe.org/accountability/assessments/k-12-student-assessment/results/2025.stml

[17] Search for Public Schools - ESSRIG ELEMENTARY SCHOOL (120087002507). National Center for Education Statistics (NCES). (2024a). https://nces.ed.gov/ccd/schoolsearch/school_detail.asp?Search=1&Zip=33694&Miles=5&ID=120087002507

[18] Search for Public Schools - KINGSWOOD ELEMENTARY SCHOOL (120087000966). National Center for Education Statistics (NCES). (2024). https://nces.ed.gov/ccd/schoolsearch/school_detail.asp?Search=1&ID=120087000966

[19] Digest of Education Statistics. Number and percentage of public school students eligible for free or reduced-price lunch, by state: Selected school years, 2000-01 through 2022-23. (2024). https://nces.ed.gov/programs/digest/d23/tables/dt23_204.10.asp

[20] About the Lead Poisoning Prevention Program. Florida Department of Health. (2024). https://www.floridahealth.gov/environmental-health/lead-poisoning/index.html