The use of in situ gamma radiation measurements as a method of determining radon potential in urban environments

Andrew S. Berens

Follow this and additional works at: https://scholarworks.gsu.edu/geosciences_theses

Recommended Citation
https://scholarworks.gsu.edu/geosciences_theses/89
THE USE OF IN SITU GAMMA RADIATION MEASUREMENTS AS A METHOD OF DETERMINING RADON POTENTIAL IN URBAN ENVIRONMENTS

by

ANDREW BERENS

Under the Direction of Jeremy E Diem, PhD

ABSTRACT

Radon is a radioactive gas that is the leading cause of lung cancer in non-smokers. While radon is natural and ubiquitous, higher concentrations greatly increase cancer risk. As such, understanding the spatial distribution of radon potential is key to planning and public health efforts. This project tests a method of determining radon potential using in situ measurements of gamma radiation. The in situ measurements were used to create a raster of gamma emissions in the study region using kriging. The resulting model showed that the operational scale of gamma radiation in the study region was 4.5 km. Indoor radon concentrations were then assigned gamma emission rates from the raster and the two were compared. While there was evidence of an association between higher gamma and high radon, the gamma readings were not quantitatively predictive. As such only categorical predictions of radon potential and risk could be made.

INDEX WORDS: Radon, Gamma, Kriging, Spatial interpolation, Statistical modeling
THE USE OF IN SITU GAMMA RADIATION MEASUREMENTS AS A METHOD OF PREDICTING RADON POTENTIAL IN URBAN ENVIRONMENTS

by

ANDREW BERENS

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in the College of Arts and Sciences
Georgia State University
2016
THE USE OF IN SITU GAMMA RADIATION MEASUREMENTS AS A METHOD OF PREDICTING RADON POTENTIAL IN URBAN ENVIRONMENTS

by

ANDREW BERENS

Committee Chair: Jeremy Diem

Committee: Dajun Dai
Lawrence Kiage

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
May 2016
DEDICATION

I would like to thank the hard work of Dr. Jeremy Diem, who put a great deal of effort and time into making this project be as good as it is. I would also like to thank my parents, Matt and Eileen, for all their love and support of me throughout the years. I also want to thank my siblings, Theresa, Sean, Bryan, and Laura, for always being on my team. I also thank my mother/father in-law, Trish and Gord, and brothers/sisters in-law, Leigh, Heather, Scott, Kasey, Kristin, Matthew, and Sarette, for being there for me. Finally I want to thank my wife, Anna. Without her constant love and support I never could have done all of this. This thesis is for all of them.
ACKNOWLEDGEMENTS

I would like to thank and acknowledge my committee chair, Dr. Diem, and my committee members, Dr. Dai and Dr. Kiage. The three of them helped me make this thesis worth reading. In addition to my committee members and my committee chair, I want to formally acknowledge the hard work of all the students who volunteered to assist me with sampling during this project. Those students include the following: Hilina Melaku Alemayehu, Chelsea Bush, Greg Ciccone, Kate Connell, Chris Dobson, Lori Farley, Upendra Giri, Saira Hamid, Sam Holbrook, April Holloway, Jeremy Hughes, Lee Johnson, Jeff McKeon, Taylor McNeil, Daniel Mills, Amanda O’Kelley, Mark Rudell, Candice Simon, Caitlyn Smith, Diana Snyder, and Morgan Wright.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................................................. v

LIST OF TABLES ........................................................................................................... viii

LIST OF FIGURES ......................................................................................................... ix

1  INTRODUCTION ............................................................................................................. 1

1.1  Study Region ............................................................................................................. 6

2  Data and Methods......................................................................................................... 9

2.1  Gamma emission rate sampling ............................................................................... 9

2.2  Gamma emission rate analysis ............................................................................... 14

2.3  Spatial interpolation modelling .............................................................................. 14

2.4  Radon sampling and analysis .................................................................................. 16

2.5  Radon potential mapping ......................................................................................... 19

3  RESULTS ..................................................................................................................... 20

3.1  Gamma emission rates ........................................................................................... 20

3.2  Gamma emission rate variation by rock and soil type .............................................. 24

3.3  Spatial interpolation modelling .............................................................................. 26

3.4  Radon sampling ....................................................................................................... 29

3.5  Radon/gamma comparison and analysis .................................................................. 31

3.6  Radon potential map ............................................................................................... 34

4  Discussion ................................................................................................................... 36
LIST OF TABLES

Table 1 – Descriptive statistics of gamma emission rates for each bedrock type............ 24
Table 2 – Descriptive statistics of gamma emission rates for each soil type................. 25
Table 3 – Comparison of errors in various kriging models ........................................ 27
Table 4 – 2x2 contingency table relating indoor radon and gamma emissions ............. 34
LIST OF FIGURES

Figure 1 – Map of Georgia highlighting the Atlanta MSA and DeKalb County .......... 7
Figure 2 – Map of urban land in DeKalb County .................................................. 8
Figure 3 – Ratemeter and scintillation probe used in gamma emissions sampling .......... 9
Figure 4 – Example of a gamma emissions sample spot ............................................ 10
Figure 5 – Example of a gamma emissions sample site ............................................. 11
Figure 6 – Map of geology in DeKalb County ......................................................... 12
Figure 7 – Map of soils in DeKalb County ............................................................... 13
Figure 8 – Idealized semivariogram ........................................................................ 15
Figure 9 – Map of the radon sample locations before data cleaning .......................... 17
Figure 10 - Map of cleaned indoor radon concentration sampling locations ............ 18
Figure 11 – Map of gamma emission rate sampling locations .................................... 21
Figure 12 – Histogram of sampled gamma emission rates ....................................... 22
Figure 13 – Map of sampled gamma emission rate values ....................................... 23
Figure 14 – Semivariogram model of gamma emission rate sample locations .......... 26
Figure 15 - Map of predicted gamma emission rates .................................................. 28
Figure 16 - Semivariogram model of indoor radon concentrations ............................ 29
Figure 17 – Map of indoor radon concentrations ...................................................... 30
Figure 18 – Histogram of indoor radon concentrations .......................................... 31
Figure 19 – Histogram of predicted gamma emission rates ...................................... 32
Figure 20 - Scatterplot of gamma emission rates versus indoor radon concentrations ... 33
Figure 21 – Map of radon potential ......................................................................... 35
Figure 22 – Coverage map based on the operational scale of gamma emissions ......... 39
Figure 23 – Coverage map based on the optimal search radius of gamma emissions...... 40

Figure 24 – Radon sample location likely driven by building characteristics................. 43
1 INTRODUCTION

Indoor radon gas represents a serious public health risk. Behind smoking, radon is the leading cause of lung cancer, accounting for between 3 and 14 percent of all lung cancer deaths worldwide (Darby et al., 2001; World Health Organization, 2009; Noh et al., 2016; Oh et al., 2016; Sheen et al., 2016). This risk is so high in part because radon gas is one of the most common radioactive particles to which people are exposed (Kauppinen et al., 2000), accounting for as much as 50% of a person’s lifetime radiologic dose (Garcia-Talvera et al., 2007). Further, indoor concentrations of radon are often 10 times higher than average outdoor concentrations (UNSCEAR, 1994; Harley et al., 1988). This makes indoor radon especially troubling as increasing exposure, either through increased duration or increased magnitude, is directly correlated with heightened cancer risk (WHO, 2009; Planchard and Besse, 2015; Kang et al., 2016). Cancer risk due to radon is increased further by smoking, with the risk of cancer among smokers who are regularly exposed to the Environmental Protection Agency action level of radon (4 pCi/L) being almost 9 times higher than the risk to non-smokers exposed to similar levels (EPA, 2009). This problem is especially bad in the US, where radon accounts for approximately 21,000 deaths annually (epa.gov/radon).

Radon is a naturally occurring, ubiquitous gas. Various radon gases form through radioactive decay of unstable isotopes such as thorium-232 and uranium-238 (National Research Council, 1999; Peterson et al., 2007). The uranium-238 decay series specifically forms radon-222 via the alpha decay of radium-226 (Sakoda et al., 2011). This is important because radon-222 is the most common radon nuclei found in buildings (WHO, 2009) and thus the decay of uranium-238 and its daughters in soil and bedrock are a driver of indoor Radon potential.
One major control of indoor radon is geology. Because radon-222 is a byproduct of uranium-238 decay, the amount of uranium-238 contained in an area’s underlying bedrock will directly impact the amount of radon-222 released from the ground in that area; however, the amount of 238 is not uniform across all rock types, with some types (like granitic rocks and shales) having more uranium-238 on average (Muikku et al., 2007). Beyond just the amount of radioactive source material, bedrock characteristics such as permeability and porosity can affect the amount of radon-222 released into the overlying soil (or air in the case of rock outcroppings) (Bossew and Lettner, 2007). Fault activity can also affect radon-222 concentrations, with faults accumulating uranium-238 and providing pathways for radon to escape the ground (Pereira et al. 2010).

Another key control of radon in any home is that home’s underlying soil. Soil gas infiltration is the primary natural source of radon entering any home (WHO, 2009). As such, soils high in uranium-238 are generally expected to produce higher radon concentrations. Additionally, as with bedrock, the permeability of soil can affect the amount of radon-222 released from the ground (Bossew and Lettner, 2007). While soils high in uranium-238 are often the result of parent rock high in uranium-238, soils high in moisture or soils that have been transported from where they formed, such as floodplain soils, may result in very different radon emissions than would be expected based on their underlying geology (Grasty, 1997).

An important, non-natural control of indoor radon concentration is home construction and building materials. If a home is well built, lacking structural defects, said home is likely to have low radon-222 concentrations even if the emissions from the ground are very high (Vauptic et al., 2002). If there is a construction defect though, such as a foundation crack, radon will likely flow into the lower pressure of the home via the defect (Appleton, 2007). Additionally, climate
controls within the home will alter temperature and humidity, which can affect rates of radon-222 infiltration into the home (Akbari et al. 2013). Finally, building materials, especially concrete and wallboard, can contain uranium-238 and uranium-238 progeny such as Ra-226, can lead to radon-222 emission from floors and walls (Chen et al., 2010).

In response to the health hazard posed by radon, researchers have attempted to predict radon concentrations in various areas, called radon potential mapping, using indoor radon concentrations and geology. This is done by generalizing known radon concentrations to the underlying geologic unit and extrapolating this to areas without radon measurements based on that area’s bedrock (Cinelli et al., 2012). Areas with underlying geology consisting of granitic and gneissic rock, both generally high in uranium-238, show a trend toward higher indoor radon concentrations (Buttafuoco et al., 2007).

To further improve the accuracy of these geologic radon potential maps, other input variables, specifically gamma dose rate, have been also been investigated. Gamma radiation, measured in this study as the number of photons released/detected in a given area over a given time, is produced naturally as a result of the decay of potassium-40, thorium-232, and uranium-238 (Wilford, 2012). The latter two (thorium-232 and uranium-238) produce radon as well (National Research Council, 1999; Peterson et al., 2007), and uranium-238 concentrations in various sources is an important control of radon-222. The link between gamma radiation and radon-222 is improved since more gamma photons are released during the uranium-238 decay cycle than any other gamma producing cycle (world-nuclear.org). In fact, uranium-238, which again is the progenitor of radon-222, is so well linked to gamma radiation that gamma spectroscopy was used for uranium mining exploration (Wilford and Minty, 2007). In the field
this link has held true, with gamma emissions having a linear relationship to soil Ra-226 (Garcia-Talvera et al., 2013), which is in turn correlated to indoor radon-222 (Nason and Cohen, 1980).

The association between gamma and radon has been used to help create more robust radon potential maps. This association, while indirect as explained above, is consistent, with gamma emissions correlating with indoor radon (Szegvary et al., 2007a). One study in Northern Ireland found that equivalent U concentration, which was derived from aerial gamma emission rate measurements, was the most important independent variable in predicting radon potential (Appleton et al., 2011a). Other studies have found that gamma dose rate accounts for as much as 58-60% of radon flux variability (Szegvary et al., 2007b; Griffiths et al., 2010).

Despite showing some ability to predict radon potential, methods of radon potential mapping focused on geology have serious shortcomings. Any study is only as accurate as the geologic data itself (Friedman and Groller, 2010). Often these studies only find correlations with some rocks (like granite, shales, and U enriched phosphate rocks) and radon concentrations (Buttafuoco et al., 2007), leaving variations of radon in other rock types unexplained. In some cases only a quarter of all variation in radon concentration can be explained by geology (Appleton and Miles, 2010), and this assumes geologic data are available and reliable, which simply is not true for some areas (Chen, 2009).

The shortcomings of geology focused radon potential mapping efforts would naturally lead one to prefer gamma radiation based efforts; however, these too have issues. While gamma emission rate functions as a proxy for uranium-238 concentration and can provide an immensely important independent variable when predicting radon-222 concentrations (Appleton et al., 2011a), the gamma data are collected primarily via aircraft to improve spatial coverage. The issue is that aerial gamma measurements have a 1 km plus spatial resolution (Appleton et al.,
2011b; Drolet et al., 2013). This poses problems because buildings and roads, which would naturally fall into those 1 km grid cells, can artificially increase or decrease gamma readings (Appleton et al., 2008). Further, aircraft must fly higher over cities, which introduces even more error because accuracy of gamma measurements decrease exponentially with distance from the ground (Appleton et al., 2008). Given the interference from the built environment, this data collection method is simply not accurate enough to be useful in a heavily urbanized environments, which is where the majority of people worldwide now live (United Nations, 2014).

Problems with both of the methods of radon potential mapping explained above prompted this study to ask the following research question: how effective are in situ gamma emission measurements at predicting radon potential in urbanized environments? This question led to two primary objectives:

(1) create a spatially complete database of forest-soil gamma emission for the entire study region, and

(2) examine the relationship between gamma emission rate and indoor radon concentration.

In answering the above question, this study attempts to determine if ground-truth measurements, which are intentionally taken in places with minimal influence from the built environment, can solve the problems associated with airborne gamma readings in a city. This would allow for the use of the gamma-based method of radon potential mapping to be used in environments where it would have previously been inaccurate.
1.1 Study Region

The study region of this project is DeKalb County, located near the center of the Atlanta metropolitan statistical area in northern Georgia (Figure 1). DeKalb, with 722,161 residents and 306,954 residential units in 2014 (census.gov), was selected for four reasons. First, DeKalb homes have been sampled for indoor radon enough to provide the data needed to achieve the second objective of this project. Second, DeKalb is heavily urbanized, with 13.5% of the county being classified as “urban” land by the USDA (Figure 2) and a population density of almost 2,700 residents per mi², while still having undisturbed, non-flood plain forest soils in some areas, allowing for completion of the first objective of this project. Third, DeKalb County has its own water system and very few private wells. This is important because radon can be found dissolved in water, especially water from drilled private wells (Vesterbacka et al., 2005). This additional radon exposure risk would have been very difficult to account for using the methods of this study; however, accounting for this exposure is not necessary in an area where these wells are uncommon, such as DeKalb. Fourth, DeKalb acts as a good case study for the type of area where knowing radon potential is important. Not all of the county is developed, despite being near Atlanta, a major, growing city. In the future, new development will lead to people living in previously undisturbed areas of the county. Knowing if those areas are at risk of radon exposure before development could help county officials make planning decisions to protect the population of newcomers prior to their moving in.
Figure 1 – Map of Georgia highlighting the Atlanta MSA and DeKalb County
Figure 2 – Map of urban land in DeKalb County (data source: USDA)
2 DATA AND METHODS

2.1 Gamma emission rate sampling

A geospatial database of gamma emissions in DeKalb County was created. Gamma emission readings were taken throughout the county using a Ludlum Measurements Inc. model 2221 scaler ratemeter attached to a scintillation probe. This device (Figure 3) measured the number of gamma photons that came in contact with the scintillation probe per minute. While this ratemeter measure the full spectrum of gamma radiation (i.e. from sources other than uranium-238 decay), readings can be used to relate uranium-238 content from one place to another because the proportion of gamma radiation produce by uranium-238 remains constant (Szegvary et al., 2007b). To ensure that the measurements of gamma emission were the result of authigenic radiation and not merely an artifact of some disturbance to the soil, all sample sites were forest areas with older growth trees outside of flood plains (e.g. Figure 4).

Figure 3 – Ratemeter and scintillation probe used in gamma emissions sampling
Sites were selected for sampling based on their suitability (i.e. whether they had undisturbed, forested soils) and their underlying geology. Residential areas were avoided to ensure investigators did not intentionally trespass on private property. Google Earth was used to determine if a site had both forested soils and was not on private property (Figure 5). Suitable sites found via Google Earth were than compared to DeKalb’s geology (Figure 6) to ensure sites did not all fall into a few geologic units. Once a site was selected, spots on top of each soil type (Figure 7) in the immediate vicinity of the site were selected. Three to five gamma emission readings would then be taken on top of each spot at the site (most sites had about three spots). The latitude, longitude, soil type, bedrock type, and result (in photons/min) of each reading
would be recorded at each spot. Bedrock was of particular importance, as uranium-238 is mostly in bedrock. Each of the three to five measurements at a spot were taken roughly a 5 to 10 meters apart to ensure that buried material was not distorting the reading and to ensure that small uranium-238 variations within soils and rock units could be captured. Each reading was one minute long. Once all readings for a specific spot were taken, the mean of those readings was assigned to that spot.

Figure 5 – Example of a gamma emissions sample site (image credit: Google Earth)
Figure 6 – Map of geology in DeKalb County (data source: USGS)
Figure 7 – Map of soils in DeKalb County (data source: USDA)
2.2 Gamma emission rate analysis

General descriptive statistics of gamma emission rates were gathered. First the basic descriptive statistics (mean, range, standard deviation) of gamma emission rates were calculated in order to get a general sense of the data. Spatial autocorrelation was tested for using Moran’s I. The normality of the data was also checked. Mean and standard deviation of gamma emission rates were also calculated for the data after grouping by soil and bedrock type separately in order to ascertain if gamma emissions differed by rock and soil type.

In an effort to explore any possible connections between gamma emissions rates and soil/bedrock type, two one-way ANOVA test were run, one grouping gamma emissions by soil type and one grouping by bedrock type. A Tukey post-hoc analysis was then done for both to determine if any rock/soil type had consistently distinct mean gamma emission rates. Any soil/rock type with an n of only 1 was excluded as ANOVA requires a variance value to properly analyze a mean (with an n of 1 the variance is non-existent). The two-tailed critical value of each ANOVA test was based on \( \alpha = .05 \).

2.3 Spatial interpolation modelling

In addition to statistical analysis, the gamma emissions data were also used to create a predictive spatial interpolation model for gamma emission rates across DeKalb. This continuous surface allowed gamma emission rates to be extracted to each of the indoor radon concentration readings this project had access to, allowing the project to have a good sense of the gamma emission rate for any radon measurement without having to individually sample for gamma emissions at each radon sample location. The method used for the spatial interpolation model was kriging. Kriging was chosen because kriging models automatically account for errors due to some random variation and can calculate root mean squared errors (RMSE) as well as standard
errors. To ensure that the most accurate surface possible was created, four different kriging models (simple, empirical Bayesian, ordinary, and universal) were run separately and compared to one another using standardized RMSE (defined as \( \frac{\text{RMSE}}{\text{Standard Error}} \)), and the model with the standardized RMSE closest to 1 (the perfect value) was chosen. The resulting raster was then generalized to a scale of 1 km, in keeping with the scale used by the studies mentioned in the introduction.

In addition to creating a continuous, predictive surface, kriging modelling also produces a semivariogram, which represents the variance of paired points versus the distance between those points (Diem, 2003). Semivariograms often have a nugget, sill, and range (Figure 8). Break points within the semivariogram, which would include the range, can be understood to be the operational scale of a given phenomenon (Lam and Quattrochi, 1992; Diem, 2003). As such, the semivariogram can be used to determine if any region was sufficiently well sampled. Thus this study calculated the operational scale (defined here as the distance range of the semivariogram) and used half that distance as the optimal search radius to ensure full spatial coverage.
2.4 Radon sampling and analysis

Indoor radon concentration readings were collected. These readings were obtained through the ongoing NIH funded project. That project had two sources of radon data available. The first data source was 2,054 readings data provided by a private radon testing firm, Air Chek Inc, and the second was 200 readings collected by the NIH funded project’s own volunteer effort (Figure 9), which sampled for indoor radon in previously under sampled areas.

The resulting 2254 indoor radon concentration readings were then compiled together and cleaned. First, radon values of zero were removed as such a reading is more likely the result of error, either during sampling or entering into the database, than anything else. Radon is ubiquitous and is essentially always present in some small quantity. Next, readings with no coordinates were thrown out because they would not be able to be paired with a gamma emission rate. Finally, readings taken at the same location but at various times were removed, save the first reading. This was done to ensure that the reading used was the natural state of the house, and not the result of some mitigation. After cleaning 1358 points remained (Figure 10). All analysis of radon was done using these points.

Cleaned data was used to generate a semivariogram model of indoor radon. This was done to determine if any spatial autocorrelation present in the radon data was sufficient to help explain variations in said data. The nugget was also analyzed to determine if the sample scale was sufficiently fine to allow for interpolation.
Figure 9 – Map of the radon sample locations before data cleaning
Figure 10 - Map of cleaned indoor radon concentration sampling locations
Next, each reading was assigned a gamma emission rate value based on location. This was done by mapping the radon readings and extracting a gamma emission rate value to each point from the raster produced by the interpolation model mentioned above. Any points falling outside of the raster, and thus unable to be paired with a gamma value, were subsequently removed from the dataset.

General statistics were obtained for radon and relationships between radon and gamma were tested for. First basic descriptive statistics, including mean and standard deviation, of the indoor radon concentrations were calculated. Next, a Spearman correlation test was used to determine if there was a relationship between gamma and radon. Then the points were grouped by radon value into below and at or above EPA radon action level (4.0 pCi/L). Gamma emission rate means of the two groups were compared using a t-test. Finally, a chi-squared test was used to compare radon, grouped by EPA action level, and gamma, grouped by mean (i.e. above or below the mean). All test statistics had a two-tailed with critical values based on $\alpha = .05$.

### 2.5 Radon potential mapping

Based on the results of the gamma/radon comparison, a map of potential radon risk was created. This was done by stratifying the radon concentration values based on gamma emission rates into general categories of radon risk. Once categories were established, ANOVA testing comparing the mean radon concentrations of each category was done to ensure that the categories were in fact distinct from one another. This use of one-way ANOVA to ensure heterogeneity of radon values when they are grouped categorically by a parameter is in keeping with previous studies (Drolet et al., 2013; Drolet et al., 2014). To deal with issues of multiplicity if multiple ANOVA tests were needed to find heterogeneous categories, the two-tailed critical value was based on $\alpha = .01$. 
3 RESULTS

3.1 Gamma emission rates

After sampling, 400 valid locations were available for analysis throughout DeKalb without obvious spatial gaps (Figure 11). While there is some variation to the coverage, along I-85 for example there is more sampling immediately south of the interstate than there is immediately north, this variation is the result of land use. This project avoided trespassing on private lands, and some areas of DeKalb are exceedingly residential, making public lands difficult to find. However, even taking into account the sometimes varying density of sample spots, it is clear that no portion of the county lacks representation in the sampling scheme.

Descriptive statistics of the gamma emissions that were sampled shed light on the phenomenon itself. With values ranging from 2,798 to 25,575 photons/min, the 400 sample spots had a mean gamma emission rate of 10,586 photons/min (95% CI: 10,230 to 10,942 photons/min, df = 399) (all gamma emission rates are rounded to the nearest whole number after analysis). The distribution of the gamma emission values is roughly normal despite a spike of below mean value between 7,000 and 8,000 photons/min (Figure 12). That uptick in below mean value coupled with a few very high (over 20,000 photons/min) values result in a histogram with a slight positive skew, but the deviation from normal is not enough to merit transforming the data and may simply be the result of some clustering of sample sites. Additionally, since no parametric statistical tests were run using these values, there is not an overwhelming need for the data to be perfectly normal.
Figure 11 – Map of gamma emission rate sampling locations
The distribution of gamma emission values across the county shows clear spatial autocorrelation, especially in a few areas (Figure 13). First there is an obvious grouping of low values in the far southwestern part of DeKalb. Additionally, there is a clear group of high values along the eastern border and in the southeast of the county. Finally there is an interesting intersection of a band of high values cutting across from the middle of the western border up to the northeast. This area of relatively high values is bordered on its north by a band of low values distributed in the same southwest to northeast trend. All these grouping are in keeping with gamma emissions’ strong spatial autocorrelation, with a Moran’s I of .49 (Z = 16.4, p <.0001). A Moran’s I that much greater than zero confirms that gamma emissions do vary spatially.

Figure 12 – Histogram of sampled gamma emission rates
Figure 13 – Map of sampled gamma emission rate values
3.2 Gamma emission rate variation by rock and soil type

While variation in mean gamma emission rates for various bedrock types varies somewhat (Table 1) there is limited value in knowing those variations. Specifically, while the ANOVA results suggest that there were some differences in gamma emissions by rock type ($F = 12.29, \ df = \frac{12}{386}, \ p < .05$), only one of the rock types has a consistently distinct mean from the rest. Excluding schist/gneiss, granitic gneiss has a significantly higher mean gamma emission rate than any other rock type. However similar distinctions cannot be made for the other rock types, making the predictive value of bedrock limited. This is further reinforced by the low $R^2$ value of the ANOVA, with only 27.6% of gamma variation being explained by rock type.

Table 1 – Descriptive statistics of gamma emission rates for each bedrock type

<table>
<thead>
<tr>
<th>Bedrock type</th>
<th>N</th>
<th>Mean (photons/min)</th>
<th>95% Confidence Interval</th>
<th>Coefficient of Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biotite Gneiss</td>
<td>59</td>
<td>8813</td>
<td>8253 - 9374</td>
<td>.25</td>
</tr>
<tr>
<td>Biotite Gneiss/Felsic Gneiss</td>
<td>7</td>
<td>9598</td>
<td>8299 - 10898</td>
<td>.18</td>
</tr>
<tr>
<td>Biotite Gneiss/Mica Schist</td>
<td>72</td>
<td>11573</td>
<td>10979 - 12167</td>
<td>.22</td>
</tr>
<tr>
<td>Granite</td>
<td>20</td>
<td>11426</td>
<td>10128 - 12724</td>
<td>.26</td>
</tr>
<tr>
<td>Granitic Gneiss*</td>
<td>39</td>
<td>14800</td>
<td>13452 - 16148</td>
<td>.29</td>
</tr>
<tr>
<td>Granitic Gneiss/Amphibolite</td>
<td>32</td>
<td>8637</td>
<td>7440 - 9835</td>
<td>.40</td>
</tr>
<tr>
<td>Mica Schist</td>
<td>33</td>
<td>11023</td>
<td>10043 - 12003</td>
<td>.26</td>
</tr>
<tr>
<td>Mica Schist/Gneiss</td>
<td>107</td>
<td>10406</td>
<td>9761 - 11050</td>
<td>.33</td>
</tr>
<tr>
<td>Schist</td>
<td>7</td>
<td>6138</td>
<td>5701 - 6575</td>
<td>.10</td>
</tr>
<tr>
<td>Schist/Gneiss</td>
<td>1</td>
<td>18154</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Quartzite</td>
<td>3</td>
<td>5969</td>
<td>4976 - 6961</td>
<td>.15</td>
</tr>
<tr>
<td>Quartzite/Mica Schist</td>
<td>6</td>
<td>9812</td>
<td>8783 - 10842</td>
<td>.13</td>
</tr>
<tr>
<td>Ultramafic</td>
<td>13</td>
<td>7979</td>
<td>5179 - 10779</td>
<td>.64</td>
</tr>
</tbody>
</table>

*This rock type has a mean that is statistically distinct from all the others based on ANOVA and a Tukey post-hoc comparison.
Variation among gamma emissions by soil type also provided minimal predictive value (Table 2). Unlike with bedrock, where at least one rock type was clearly distinct, no soil type was distinct from all the others in a significant way according to the ANOVA, though there were still difference (F = 5.11, df = \frac{15}{383}, p < .05). With few exceptions most of the soils showed no real distinction, indicating that soil does not appreciably impact gamma emission rates. This too is reinforced by a low $R^2$, with only 16.7% of gamma variations being explained.

<table>
<thead>
<tr>
<th>Soil series</th>
<th>N</th>
<th>Mean</th>
<th>95% Confidence Interval</th>
<th>Coefficient of Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Altavista</td>
<td>4</td>
<td>10169</td>
<td>4901 - 15437</td>
<td>.53</td>
</tr>
<tr>
<td>Appling</td>
<td>12</td>
<td>11120</td>
<td>10151 - 12089</td>
<td>.15</td>
</tr>
<tr>
<td>Ashlar</td>
<td>16</td>
<td>12896</td>
<td>11039 - 14752</td>
<td>.29</td>
</tr>
<tr>
<td>Ashlar-Wedowee</td>
<td>28</td>
<td>12934</td>
<td>11145 - 14722</td>
<td>.37</td>
</tr>
<tr>
<td>Cecil</td>
<td>51</td>
<td>10906</td>
<td>10090 - 11723</td>
<td>.27</td>
</tr>
<tr>
<td>Chestatee</td>
<td>5</td>
<td>13750</td>
<td>9896 - 17632</td>
<td>.32</td>
</tr>
<tr>
<td>Gwinnett</td>
<td>31</td>
<td>9023</td>
<td>7441 - 10606</td>
<td>.50</td>
</tr>
<tr>
<td>Hiawassee</td>
<td>1</td>
<td>5660</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Iredell</td>
<td>1</td>
<td>2990</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Madison</td>
<td>52</td>
<td>10850</td>
<td>10103 - 11597</td>
<td>.25</td>
</tr>
<tr>
<td>Musella</td>
<td>2</td>
<td>12083</td>
<td>10168 - 13998</td>
<td>.11</td>
</tr>
<tr>
<td>Pacolet</td>
<td>162</td>
<td>10036</td>
<td>9545 - 10527</td>
<td>.32</td>
</tr>
<tr>
<td>Sweetapple-Grover</td>
<td>8</td>
<td>12043</td>
<td>10699 - 13387</td>
<td>.16</td>
</tr>
<tr>
<td>Udorthent</td>
<td>4</td>
<td>15368</td>
<td>11052 - 19683</td>
<td>.29</td>
</tr>
<tr>
<td>Wedowee</td>
<td>13</td>
<td>11187</td>
<td>9017 - 13358</td>
<td>.36</td>
</tr>
<tr>
<td>Wilkes</td>
<td>9</td>
<td>5327</td>
<td>4376 - 6278</td>
<td>.27</td>
</tr>
</tbody>
</table>
3.3 Spatial interpolation modelling

The semivariogram model indicates that gamma emissions are strongly spatially dependent. Aside from allowing the production of a continuous surface for gamma emissions, this also confirms that enough sampling has been done. The operational scale (defined as the range of the semivariogram in this study, as explained in the section 2.3) of gamma in DeKalb is roughly 4.5 km (Figure 14). An operational scale of 4.5 km further indicates that the optimal search radius in this project’s study region would be about 2.25 km.

Based on comparisons of errors among multiple kriging models, an empirical Bayesian kriging model was determined to be the most accurate. The standardized RMSE (defined in section 2.3) of the empirical Bayesian kriging model, .977 (a value of 1 is the optimal standardized RMSE), indicates it is the best model to predict gamma emission rates based on the available sample data (Table 3).
Table 3 – Comparison of errors in various kriging models

<table>
<thead>
<tr>
<th>Model type</th>
<th>Root mean squared error</th>
<th>Standardized RMS error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple kriging</td>
<td>2260.04</td>
<td>0.902</td>
</tr>
<tr>
<td>Empirical Bayesian kriging</td>
<td>2241.41</td>
<td>0.977</td>
</tr>
<tr>
<td>Ordinary kriging</td>
<td>2251.78</td>
<td>0.921</td>
</tr>
<tr>
<td>Universal kriging</td>
<td>2323.23</td>
<td>1.115</td>
</tr>
</tbody>
</table>

The kriging surface resulting from the selected model provides an accurate assessment of variation of gamma emissions in DeKalb. Similar to Figure 9, the results of the empirical Bayesian kriging model show high values in the eastern/southeast border of DeKalb and very low values in the southwest and north/northeast parts of the county (Figure 15). Notably the central/south central part of the county shows a relatively high degree of variability though it should be noted that this region is not as well sampled as some of the other parts of the county (Figure 11) and sits on top of a fairly uniform portion of DeKalb geology consisting mostly of granitic gneiss (some with amphibolite, some without).
Figure 15 - Map of predicted gamma emission rates created using an empirical Bayesian kriging model optimized to the sampled gamma data with grid cells resized to 1 km square
3.4 Radon sampling

The distribution of radon values is not as spatially autocorrelated as that of gamma emissions. With a Moran’s I of .09 ($Z = 10.8$, $p < .0001$), radon is autocorrelated, but not a particularly strong way. This lack of autocorrelation can be seen in the semivariogram model of radon (Figure 16).

While there is minimal spatial autocorrelation there were clearly some areas with consistently low radon (Figure 17). Most notably the far southwestern part of the county contained none of the 176 actionable homes were located. Additionally, the central/south central portion of the county appears to be a pocket of relative low for radon.

Figure 16 - Semivariogram model of indoor radon concentrations
Radon in DeKalb is on average below the EPA action level, though the data is not normal. The 1,352 readings had a mean concentration of 2.31 pCi/L (95% CI: 2.17 to 2.45 pCi/L, \(df = 1351\)). Based on the histogram it is clear that the radon readings are not normally distributed (Figure 18). This is because, while radon levels cannot fall below 0.0 pCi/L, there is no theoretical upper limit to how high indoor radon concentrations can be.

![Histogram of indoor radon concentrations]

**Figure 18 – Histogram of indoor radon concentrations**

### 3.5 Radon/gamma comparison and analysis

The descriptive statistics of predicted gamma emission rates extracted to the indoor radon reading locations show a similar, albeit muted version of the gamma emission phenomenon relative to the sampled rates. The 1,348 values (four radon reading were taken outside the area of coverage of the kriging model and thus did not have gamma emission rates extracted to them) had a mean of 10,293 photons/min (95% CI: 10,178 to 10,408 photons/min, \(df = 1,347\)) with value ranging from 5,178 to 22,208 photons/min. It is worth noting that the predicted values of gamma emissions had a less extreme range than the sampled values. Further, the predicted values
were much closer to the shape of the normal distribution (Figure 19), potentially because of the significantly higher N or the fact that the radon sample sites were less clustered than gamma sites.

Figure 19 – Histogram of predicted gamma emission rates

The results the f correlation test show a significant, though very weak, correlation between gamma emissions and radon. Specifically, Spearman’s rho was .059 (N = 1348, p < .05). While this correlation was significant, the R² makes clear that this correlation is mostly a result of a high N (Figure 20). As such this correlation does not provide any valuable predictive ability beyond indicating some positive association.
Figure 20 - Scatterplot of predicted gamma emission rates versus indoor radon concentrations

Despite weak correlation, there is some relationship between gamma emission rates and indoor radon concentrations. The t-test found that mean predicted gamma emission rates of dwellings with indoor radon concentrations below 4.0 pCi/L (mean = 10244, s = 2161, n = 1172) were significantly lower than mean predicted gamma emission rates in dwellings with indoor radon concentrations of 4.0 pCi/L (the EPA action level) or more (mean = 10619 photons/min, s = 2045, n = 176) (t = -2.158, df = 1346, p < .05). This would mean that generally one would expect to find higher gamma values at homes with actionable radon readings, even if only marginally.

This association between actionable radon concentrations and higher gamma emissions is strengthened more by the results of the chi-squared test. The chi-squared test indicates that
actionable radon concentrations are related to above mean gamma emission rates ($\chi^2 = 4.728$, df = 1, p < .05) (Table 4).

<table>
<thead>
<tr>
<th>Indoor Radon concentration</th>
<th>Gamma Emission Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Above mean</td>
</tr>
<tr>
<td>At or above EPA action level</td>
<td>106 (93)</td>
</tr>
<tr>
<td>Below action level</td>
<td>603 (614)</td>
</tr>
</tbody>
</table>

Expected values, which are in parentheses, have been rounded to the nearest whole number value

3.6 Radon potential map

The high variability among the radon readings meant that the county could only be stratified into 3 categories of risk based on gamma emissions, with basically all of the county either “at risk” or “highly at risk” (Figure 21). The break points between the categories were 7,500 and 13,000 photons/min. The ANOVA indicated that each of the categories had a distinct average indoor radon concentrations ($F = 5.45$, df = $\frac{2}{1345}$, p < .01), with a Tukey post-hoc comparison indicating that all 3 are distinct (minimally at risk: mean radon = 2.0 pCi/L, 7.8% of readings at/above 4.0 pCi/L; at risk: mean radon = 2.3 pCi/L, 13.3% of readings at/above 4.0 pCi/L; highly at risk: mean radon = 3.0 pCi/L, 18.5% of readings at/above 4.0 pCi/L).
Figure 21 – Map of radon potential
4 DISCUSSION

4.1 Gamma emissions

Gamma emissions varied by rock type in an unremarkable way. Only granitic gneiss had statistically distinct, specifically higher, gamma emissions, which is in line with the literature, as granitic rocks generally have more gamma source material (Muikku et al., 2007). Additionally, the amount of variation in gamma emissions explained by bedrock type, 27.6% based on the ANOVA in section 3.2, is very similar to the amount of variation of radon explained by rock type, about 25% (Appleton and Miles, 2010). Past studies have also found that only a few, uranium-238 rich rock types may be predictive of radon flux (Buttafuoco et al., 2007), which would seem to hold true for gamma emissions based on this study.

Ultramafic rock produced unexpected null results, possibly as a result of inaccuracies of the geologic map used. Ultramafic rocks are generally understood to be low in uranium-238, but in the ANOVA in section 3.2 showed the gamma emissions rate of ultramafic rock was not distinctly lower than all, or even most of the other rock types. In fact, based on mean gamma emissions, ultramafic is actually the third lowest. While this might be a result of some geologic feature that has concentrated uranium-238 in this unit, it may also be a function of coarse resolution in the geologic map used. At a resolution of 1:500,000, the geologic map used for this project likely had some inaccuracies, especially in defining the edges of various geologic units. The accuracy of geologic maps is a known problem in radon potential mapping (Friedman and Groller, 2010), and it follows that this problem would extend to gamma emissions, as gamma and radon share natural source materials. For the ultramafic unit of southwest DeKalb, this may have meant that some, higher uranium-238 rock types adjacent to the ultramafic unit were
erroneously classified as ultramafic. This would help explain why ultramafic has the highest coefficient of variation at .64.

Variations in gamma emissions across soil series were even less remarkable than variations by geologic unit. The predictive power of soil was not only lower, with only 16.7% of variation explained based on the ANOVA in section 3.2, but there was also no consistently distinct gamma emission rate for any soil type. The closest to being distinct, Wilkes (which had significantly lower gamma emissions than all but one of the other soil series) might shed light on why soil explains any variation. Wilkes is formed from mafic rocks, which are generally lower in gamma’s radioactive source materials. This means that the low gamma emission rates found over Wilkes soils may simply be a function of the parent rock Wilkes sits on top of. In that way, the difference in explained variation between rock type and soil series may simply be a function of soil series being an imperfect proxy for bedrock type. This makes sense as the areas that depart most from expected radon based on geology areas with soils high in moisture or transported from other locations (i.e. areas with soils that are not indicative of underlying geology) (Grasty, 1997).

4.2 Spatial interpolation

While the primary focus of the spatial interpolation modeling effort was to create a continuous surface of gamma emissions, the results of the semivariogram model show that the first objective of this project was achieved and that the 400 sample points were sufficient to spatially cover DeKalb. Two parts of the semivariogram would indicate that this project sampled enough locations. The first is the range, which as a break in the semivariogram provides the operational scale for gamma emissions (Lam and Quattrochi, 1992; Diem, 2003). The range provides the absolute furthest distance sample points can be from one another. Based on Figure 14 the range of gamma in DeKalb is 4.5 km. That means that, based on the operational scale, all
of DeKalb is sufficiently covered (Figure 22). Half the operational scale, also known as the optimal search radius, provides the ideal maximum distance between sample points to ensure coverage accurately represents a given phenomenon. For this study the optimal search radius was 2.25 km. Comparing that distance to the gamma sample points it is clear that, with one exception in south central DeKalb, the county is sufficiently well sampled (Figure 23).

The second part of the semivariogram in Figure 14 that indicates enough sampling of gamma emissions was done was the nugget. The nugget, explained in Figure 8, can be understood to show some systemic error or a variation in the phenomenon occurring at distances well below the sampling interval (Burrough and McDonnell, 1998). The lack of a nugget in Figure 14 would indicate that most of the variation in gamma emissions in DeKalb has been captured by the sampling interval. This likely means that, while additional sample sites may make the predictions of the interpolation model more robust, further sampling is unlikely to change the outcome. Taken together, the results of the semivariogram and interpolation model generally indicate that this project succeeded in completing its first objective.
Figure 22 – Coverage map of gamma sampling based on the operational scale of gamma emissions
Figure 23 – Coverage map of gamma sampling based on the optimal search radius of gamma emissions
4.3 Radon and gamma

There is a clear, positive relationship between gamma emissions and indoor radon concentrations, even if that relationship is weak. The results of the correlation, t-test, and chi-squared test all point to the same conclusion: higher gamma emissions are generally associated with higher indoor radon concentrations. This follows the literature that establishes the same positive relationship (Jackson, 1992; Szegvary et al., 2007a; Szegvary et al., 2007b). This is conceptually sensible as uranium-238 is an important geologic driver of both radon and gamma emissions (Garcia-Talvera et al., 2007; Peterson et al., 2007; Sakoda et al., 2011; Wilford 2012).

While there is clearly a relationship between gamma and indoor radon, said relationship is too weak to be quantitatively predictive. The extremely low $R^2$ associated with the significant correlation between gamma and indoor radon in this study is the first indication that this is true. The poor $R^2$ is largely caused by confounders, which result in variation in indoor radon reading that does not vary spatially or in a way that gamma could predict. The large nugget in Figure 16 and low Moran’s I for indoor radon both indicate that a large portion of radon variability cannot be explained by spatial variation. These variations are likely caused by non-natural drivers such as construction quality and building materials (Vauptic et al., 2002; Appleton, 2007; Chen et al., 2010). In fact, building characteristics may influence indoor radon more than natural controls of radon (Borgoni et al., 2014). All told the positive relationship between gamma and indoor radon found in this project, even if weak, clearly satisfies the second objective laid out in the introduction.

4.4 Radon risk

Despite not being able to quantitatively predict indoor radon using gamma emissions, a broader assessment of radon potential or general risk is possible. Such an assessment occurs in
Figure 21. Defining most of the county as either “at risk” or “highly at risk” is actually in line with the EPA’s assessment of DeKalb, which it labels a level 1 radon risk (level 1 is the highest level of risk) (epa.gov/radon). Additionally the ANOVA confirms that these categories are not arbitrary, with only 7.8% of tested dwellings in the “minimally at risk” region being at or above the EPA action level for radon versus 18.5% of tested dwelling in the “highly at risk” region. The “at risk” region had a similar percentage of dwellings at or above the action level as the dataset as a whole (13.3% and 13.1% respectively).

This broad, categorical distinction of radon potential or risk is an appropriate way to approach this kind of radon potential mapping. Outdoor gamma emissions have been found to be most useful as a qualitative predictor of indoor radon (Quindos et al., 2008). This is largely because gamma emissions are linked outdoor controls of radon which are then linked to indoor radon (Nason and Cohen, 1980; Szegvary et al., 2007a; Garcia-Talvera et al., 2013). This indirect link between external gamma emissions and indoor radon is confounded by indoor, non-natural controls of radon, making any prediction of radon potential necessarily general. However, the cheapness and ease of this method make the cautious use of these more general predictions valuable (Voltaggio et al., 2006; Smethurst et al., 2008).

Using the radon risk map produced by this study might help locate sites where confounders play a role in increasing radon above what natural sources would produce. Specifically, 13 of the 176 dwelling with radon at or above the EPA action level fell into the “minimally at risk” zone (Figure 24). All 13 of these homes are in the northern part of the county and seven of them are all in one area, Tucker. This cluster of dwellings likely have high radon as result of confounders such as building characteristics. Alternatively there may be some geologic features unique to this area that were not fully accounted for.
Figure 24 – Radon sample location likely driven by building characteristics
5 CONCLUSIONS

Gamma emissions are effective at predicting radon potential, if only in a categorical way. This data should be used cautiously, but the methods used here can be a key first step to further research. While gamma emissions cannot be used to give an exact value for indoor radon, the ease and low cost of collecting gamma data can help establish what areas are at increased risk or merit additional scrutiny when it comes to radon. This project would indicate that the use of in situ measurements can allow investigators to use gamma emissions as a variable in radon potential mapping, even in a city where the traditional aircraft collected data would not be accurate. Combining this data with other variables may further improve predictive power.

A key result of this project is the operational scale of gamma emissions. Future research focused on gamma radiation, especially in the southeast United States could use the 4.5 km operational scale found in this study to inform sampling protocols. This operational scale will also help determine if an area has been sufficiently sampled before attempting to interpolate gamma radiation data. It should be noted that, under different geologic conditions, the operational scale of gamma may vary. The 4.5 km value could be a useful starting point either way however, especially in the Piedmont geologic region.

An important conclusion of this project is that the use of indoor radon measurements in determining geogenic radon potential might be problematic. While a sufficient number of variables and controls may yield a more predictive model, the possibility of cofounders causing spurious correlations is always high with indoor radon measurements. Future research should focus not only on including multiple variable to explain radon variation, but also on using radon data that has been as scientifically controlled as possible. Ideally, if using indoor radon readings, these reading should come from homes that are as similar as possible. If possible, future research
should attempt to create a database of radon measurements taken in a controlled way directly from the ground to determine if gamma is a more quantitative predictor of geogenic radon, even if gamma only categorically predicts indoor radon.
REFERENCES


