

Title: Modeling and comparing central and room air conditioning ownership and cold-season in-home thermal comfort using the American Housing Survey

Carina J. Gronlund, PhD, MPH¹, Veronica J. Berrocal, PhD²

¹University of Michigan Institute for Social Research

²University of California, Irvine, Department of Statistics

Corresponding Author:

Carina Gronlund, PhD, MPH

Social Environment and Health Program

Survey Research Center

University of Michigan Institute for Social Research

426 Thompson St.

Ann Arbor, MI 48106

(734) 615-9215

gronlund@umich.edu

Running head: Air conditioning and cold discomfort model

Sources of Support:

Support for this research was provided by grants P30ES017885 (Berrocal and Gronlund), K99ES026198 (Gronlund) and R00ES026198 (Gronlund) from the National Institute of Environmental Health Sciences, National Institutes of Health and grant 1520803 from the National Science Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Science Foundation.

Keywords: climate change, vulnerability, air conditioning

ABSTRACT

Household-level information on central (cenAC) and room (rmAC) air conditioning and cold-weather thermal comfort are often missing from publicly available housing databases hindering research and action on climate adaptation and air pollution exposure reduction. We modeled these using information from the American Housing Survey for 2003-2013 and 140 U.S. core-based statistical areas employing variables that would be present in publicly available parcel records. We present random-intercept logistic regression models with either cenAC, rmAC or “home was uncomfortably cold for 24 hours or more” (tooCold) as outcome variables and housing value, rented vs. owned, age, and multi- vs. single-family, each interacted with cooling- or heating-degree days as predictors. The out-of-sample predicted probabilities for years 2015-2017 were compared to corresponding American Housing Survey values (0 or 1). Using a 0.5 probability threshold, the model had 63% specificity (true negative rate), and 91% sensitivity (true positive rate) for cenAC, while specificity and sensitivity for rmAC were 94% and 34%, respectively. Area-specific sensitivities and specificities varied widely. For tooCold, the overall sensitivity was effectively 0%. Future epidemiologic studies, heat vulnerability maps, and intervention screenings may reliably use these or similar AC models with parcel-level data to improve understanding of health risk and the spatial patterning of homes without AC.

INTRODUCTION

The primary function of housing is to protect us from “the elements,” and therefore housing quality is likely a critical mediator in the documented associations between weather, air quality, and health. In many key studies of the health effects of temperature and air pollution, large administrative databases of healthcare billing records or vital records have been linked to outdoor environmental exposure data with increasing spatial and temporal resolution (1-3). However, these studies usually lack information on housing quality or characteristics, or such information is available only at an ecologic scale (4, 5). Central air conditioning is likely a particularly important mediator of associations between ambient exposures and health, given its capacity to filter and cool air. This absence of housing information may lead to exposure misclassification and mis-estimation of the environmental health effects, given the imperfect, though moderate, correlations between outdoor and indoor exposures (6, 7). To understand how housing characteristics affect vulnerability to these ambient exposures, there has been increased interest in linking these data sets to tax assessor’s records. Available at the parcel level, these records include characteristics of the home relevant to establishing its value for taxation purposes. These characteristics usually include the year the home was built, occupation by a renter or owner (tenure), and the resulting assessed value. Sometimes, these records also include information about the number of stories, siding, basement type, and heating and cooling systems. However, the recorded characteristics vary widely between municipalities, and these more detailed characteristics, including central air conditioning, are often not available.

City health officials, particularly in temperate or cold climates, are also interested in the magnitude and variability of cold health effects. In many cities, in the present climate at least, cold-associated mortality and morbidity are greater than heat-associated mortality and morbidity (8-10). Though some evidence suggests that these burdens are borne disproportionately by homeless individuals, other lines of evidence suggest that individuals in homes with inadequate heating and cooling or even dangerous heating methods are also particularly vulnerable to increased morbidity or mortality during extreme heat and/or cold (11).

Therefore, having models to predict central air conditioning ownership and uncomfortably cold indoor temperature in residential homes could further our understanding of the mediating and modifying effects of these characteristics on health in the warm and cold seasons in epidemiologic studies. In turn,

this information could be used to screen households for allocation of weatherization resources, with potential prioritization of 1) homes that lack central air conditioning, 2) homes that are deemed too cold in the winter, 3) homes that use energy inefficiently, which can also be predicted from housing characteristics (12), and 4) households that can't afford to pay for utilities (13-18). These homes and households could also be prioritized in policies addressing health through improving housing conditions.

The American Housing Survey, started in 1973, originally represented two longitudinal surveys: a biennial national sample and a biennial metropolitan sample, consisting of a rotating subset of 67 U.S. core-based statistical areas (CBSAs). In 2015, the sampling strategy was altered, questions were revised, and an entirely new sample of participants was recruited (19). The surveys include a wide variety of questions about housing type and condition. Among the questions asked in the 2003-2013 surveys are "Does this housing unit have central air conditioning," "Does the housing unit have any room air conditioners," and "Last winter, for any reason, was your housing unit so cold for 24 hours or more that it was uncomfortable?". Using data from the 2003-2013 American Housing Survey, we first tested whether homes that lacked central air conditioning were the same homes that were reported as being uncomfortably cold for 24 hours or longer in the winter, to understand how strongly correlated these two temperature-related vulnerabilities were. Then, we constructed predictive models of central air conditioning ownership and cold indoor temperatures using commonly available parcel characteristics, hypothesizing that we could predict central AC ownership using year in time, year built, market value, and tenure in 140 U.S. CBSAs. We validated the out-of-sample predictions of our models using 2015-2017 American Housing Survey data. For central air conditioning ownership, we also tested the validity of the model using parcel data from Wayne County, Michigan, and, to illustrate a potential usage of our model predictions, we mapped the probability of air conditioning ownership at the parcel level in Wayne County, Michigan.

METHODS

The process for developing, validating, and applying models predicting central air conditioning (cenAC), room air conditioning (rmAC), and the home being too cold in the winter (tooCold) for a given parcel is outlined in Figure 1 and described in detail below.

Data

We downloaded publicly available American Housing Survey microdata for 2003-2017, for which the CBSA of the household is provided. Over 75% of respondents provided responses in only one survey year, so in instances where participation occurred in multiple years, only the most recent year of data was retained for that household. We retained variables for: year of survey, year the home was built, housing value, monthly rent, multi- vs. single-family, tenure (rented or owned by resident), central air conditioning ownership (cenAC), and "home was uncomfortably cold for 24 hours or more" (tooCold). To convert monthly rent values to a housing value, we assumed a 10% annual rent-to-value ratio (20).

Regional heating and cooling needs were estimated as average annual cooling-degree days (CDDs, cumulative number of degree days above 18 C per year) and heating-degree days (HDDs, cumulative number of degree days below 18 C per year), each averaged over a 30-year period. In R 3.5.0, CBSAs, from the *tigris* package, were superimposed using the *gIntersection* function (*rgeos* package) over state-specific climate divisions, 2-10 per state, using the definition from the Climate Prediction Center, and area-weighted averages of each climate division's 30-year (1981-2010) CDDs and HDDs (21) were extracted for each CBSA.

Statistical Analysis

Summary statistics were computed as follows. The weighted means or proportions of each housing characteristic were computed within each CBSA using the survey weights provided by the AHS. Then, means, interquartile ranges, minima, maxima, and 25th, 50th, and 75th percentile values were calculated across the 140 CBSAs. Separately, annual weighted means were computed using the national survey results only because different CBSAs are sampled in the metropolitan surveys on different years.

To predict whether a house does or does not have central air conditioning, we constructed the following logistic random intercept model using the lme4 package in R:

$$\begin{aligned} \text{logit}(P_{\text{cenAC}_{hm}}) = & \beta_0 + \beta_{1m}YRBLT_{hm} + \beta_{2m}RENTAL_{hm} + \beta_{3m}YEAR_{hm} + \\ & \beta_{4m}LNMARKVAL_{hm} + \beta_{5m}LNMARKVAL^2_{hm} + \beta_{6m}MULTIFAM_{hm} + \beta_{7m}MOBILEHOME_{hm} + \\ & \beta_{8m}(YRBLT_{hm} \times RENTAL_{hm}) + \beta_{9m}(YEAR_{hm} \times RENTAL_{hm}) + \beta_{10m}(LNMARKVAL_{hm} \times \\ & RENTAL_{hm}) + \beta_{11m}(LNMARKVAL^2 \times RENTAL_{hm}) + \gamma_m + \varepsilon_{hm} \end{aligned}$$

with β_{xm} defined as $\beta_{xm} = a_{1x} + a_{2x}CDD_m$, for $x = 1$ to 11. In the model above, we used the following notation: for household h in CBSA m , P_{cenAC} denotes the probability of cenAC, $YRBLT$ indicates the year the home was built, $RENTAL$ is an indicator for whether the home was rented or owned by the resident, $YEAR$ indicates the year of the survey, $LNMARKVAL$ denotes the natural log of the market value, $MULTIFAM$ is an indicator for whether the home was multi-family or single-family, $MOBILEHOME$ is an indicator for whether the home was a mobile home or not, and γ_m and ε_{hm} represent random effects for CBSA and households, respectively, each following normal distributions. In addition, CDD indicates cooling-degree days, and, assuming a modifying effect for this term, we interacted it with each of the household-level terms. The rmAC and tooCold models were similar, except heating degree-days were used in place of cooling degree-days. To examine for residual spatial dependence in P_{cenAC} , rmAC and tooCold not accounted for by the covariates we constructed and inspected semivariograms of the Pearson residuals using R's gstat package.

To validate the predictive power of the model, we used a hold-out validation procedure (22), comparing the out-of-sample predicted probabilities obtained using the logistic regression model fit using 2003-2013 AHS data with the AHS values for 2015-2017 data. The 2015 sampling year provided a natural cut-point for dividing the data into training and test data sets because of the substantial revisions to the survey in 2015 described above and the inclusion in 2015-2017 of only 35 of the 140 core-based-statistical areas available in the 2003-2013 sample. We calculated accuracy, sensitivity (true positive rate), and specificity (true negative rate), overall and by city using R's pROC package.

To compare cenAC, rmAC, and tooCold, we cross-tabulated each of the AC variables with tooCold in our entire set of the 2003-2017 AHS and calculated prevalence ratios and Wald confidence intervals using the epiR package in R.

To demonstrate the utility of these models, we created maps of the predicted probabilities in the City of Detroit by linking the cenAC and anyAC (cenAC or rmAC) with residential parcel data (23). Such records are commonly publicly available in the U.S. These records, based on tax assessor's records, do not indicate in Detroit whether a home has air conditioning but do include year built, rental (based on whether the owner's address matches the property address), assessed value, and whether the unit was multifamily. There are no mobile home parks within Detroit city limits, so we assumed that none of the parcels contained a mobile home.

Code Availability: the R code is available from Dr. Gronlund on request.

RESULTS

Across 140 CBSAs from 2003-2013, an average of 68% of homes in a CBSA had central AC, with central AC prevalence as low as 5% (Utica-Rome, NY) and as high as 99% (El Paso, TX). Housing units with room AC ranged from 0-72% by CBSA with a mean of 25%, while homes reported as having been uncomfortably cold for 24 hours or more in the winter in a CBSA ranged from 0-24%, with a mean of 8.7%. None of the homes in the sample had been built prior to 1910, and average year built ranged from 1937 to 1983. Mean housing value ranged widely, from \$48,000 to \$344,000, with a median of \$107,000 that was similar to the housing values geometric mean of \$114,000. Climatic conditions ranged widely as well, with annual degree-days ranging between 150-3,800 CDDs and 300-9,200 HDDs (Figure 2, Table 1). In addition to wide spatial variability in the selected housing characteristics, temporal trends were evident for some of the housing characteristics. The proportion of homes with central air conditioning rose from 0.50 in 2003 to 0.64 in 2013. Likewise, mean year built and mean housing value also rose over time (Table 2).

In models of central AC, room AC and uncomfortably cold indoor winter temperatures, the semivariograms of the residuals from each of the three models did not reveal any residual spatial dependence. All of the characteristics included in the central AC and room AC models were statistically significant predictors of AC status, either alone or in interactions. When the continuous variables representing housing characteristics were scaled to an IQR increase, the housing characteristics most strongly associated with central AC ownership were year built, CDDs, and housing value, all of which predicted increased likelihood of central AC. In rental housing units, central AC was less likely, although this was attenuated in homes of higher value and in homes in CBSAs with more CDDs, i.e., warmer CBSAs. Year built, CDDs and housing value were also strong predictors of room AC. In contrast to central AC, these characteristics predicted a decreased likelihood of room AC. Being in a multi-family housing unit in a warmer climate was also associated with decreased likelihood of room AC. Indeed, central AC and room AC were strongly inversely associated, with a 0.04 (95% CI: 0.03, 0.04) odds of lacking room AC among individuals lacking central AC vs. those not lacking central AC.

For uncomfortably cold indoor winter temperatures, being in a rental unit, a home in a cooler climate (more HDDs), and a mobile home were the strongest predictors. Newer homes were associated with a lower likelihood of uncomfortably cold indoor winter temperatures. In contrast to the AC models, many of the interaction terms between the characteristics were not significant. For both the uncomfortably cold and the AC models, the likelihood of uncomfortably cold indoor winter temperatures or AC increased in later years of the survey indicating increasing trends in all three housing characteristics, holding other characteristics constant. For central AC in particular, we observed a 0.23 increase in the log-odds (i.e., a 1.26 times higher odds) of central AC for each 5-year increase in time (Table 3).

In comparing predicted central AC values to true central AC values in the 2015-2017 AHS hold-out data, the central AC model performed reasonably well (Table 4). If we use a probability threshold of 0.5 to declare a home as having central AC if the predicted probability is above it and not having central AC if the predicted probability is below it, the model has a 63% specificity. That means that the model correctly identified homes without central AC 63% of the time. On the other hand, the model has a 91% sensitivity, that is, it correctly identifies homes with central AC 91% of the time. Finally, the overall area under the receiver operating characteristic (ROC) curve, e.g. AUC, is 89%. By choosing a higher threshold of 0.9 at which to consider a home to have central AC, the model achieves a higher specificity (95%) at the cost of sensitivity, which drops to 59%. The thresholds at which sensitivity and specificity were maximized are close to 0.5 for both the central AC and the room AC models. At this threshold, the room AC model has a higher specificity of 94% but a lower sensitivity of 34%, and an overall area under the ROC curve of 80%. ROC curves for these two models suggested good model utility for many of the CBSAs

(Figure 3). However, the room AC models failed in Seattle-Tacoma-Bellevue, WA and San Francisco-Oakland-Hayward, CA where the areas under the ROC curve are close to 0.5.

We were able to apply this model to City of Detroit residential parcel data, for which information about central or room AC is entirely absent, but other model variables are present. The predicted probability of central AC varies widely across the city (Figure 4A), with mean probabilities of central AC less than 0.4 in 14 of the 349 census tracts (Figure 5A). For any AC, low probabilities are more dispersed throughout the city than are the low probabilities of central AC (Figure 4B). If we consider census tracts as our administrative area of interest, four of the 349 census tracts have mean any AC probabilities less than 0.7 (Figure 5B).

DISCUSSION

Using American Housing Survey housing data, we successfully created predictive models of central and room AC that yielded predictions of central and room AC status in the hold-out sample better than if one was randomly assigning central and room AC status to parcels. For applications such as epidemiologic studies, where correctly identifying both AC and lack of AC are equally valuable, accuracy should be maximized. For these applications, a threshold of 0.5 for separating homes with AC from those without is appropriate. At this threshold, homes without central AC were correctly identified in approximately 63% of the parcels without central AC (specificity) and homes with central AC were correctly identified in 91% of the parcels with central AC (sensitivity). In contrast, for room AC, at a threshold maximizing accuracy, specificity was high (94%), but sensitivity was low. For applications where specificity may be more important, i.e., in screening parcels for lack of AC, a model threshold of 0.9 resulted in a specificity of 95% for central AC and a specificity of 100% for room AC, that is, perfect identification of homes without room AC among the homes that were indeed without room AC. However, at this threshold of 0.9, sensitivity for room AC was only 1%.

Characteristics predicting central AC were often predictive of room AC but with opposite effect. For example, housing value and CDDs predicted a higher odds of central AC ownership but a lower odds of room AC ownership while the opposite was true for rental and mobile home units. Central AC and room AC ownership were inversely related, and it's likely that households that lack central AC purchase room AC units to compensate for a lack of central AC. However, our room AC variable did not include information on the number, use or effectiveness of the room AC units, and room AC has been shown to be less effective than central AC at lowering indoor temperatures in prior research (24). Therefore, central AC may be a more useful predictor of heat vulnerability than room AC.

In contrast to the models for central and room AC, the model for uncomfortably cold indoor winter temperatures did not provide good validation summary statistics in the hold-out sample, and we would not recommend using the model based on 2003-2013 housing characteristics to attempt to predict uncomfortably cold indoor winter temperatures in subsequent years. The uncomfortably cold model may not have validated well because of stark changes over time in the relationships between the independent variables and self-reported perceptions of uncomfortably cold indoor winter temperatures. Given that the construct of primary interest is housing with chronic under-heating problems, data sets using actual indoor temperatures and information on weatherization and utility usage rather than self-reported discomfort may be better suited to predicting under-heating in the winter.

Despite better-than-random performance of the central and room AC models, neither central nor room AC are perfectly predicted on the basis of the age of the home, market value or tenure. These models are suitable to be used in studies estimating the effects of AC ownership when other options for assigning AC exposure are not available, but the respective central and room AC uncertainties should be

taken into account. When additional characteristics that may predict AC ownership beyond those used in our models are available in parcel records, we recommend that users construct and validate new predictive models with these additional characteristics. Other statistical techniques, including multiple imputation, may be more appropriate depending on the quantity and missingness of the housing data available to the analyst. Another limitation is that although AC ownership is predicted by these models, AC use is not. Lacking the financial resources to operate an air conditioner is likely a substantial barrier to home cooling (13-18, 25). Cognitive impairment, having medical conditions aggravated by AC use, concerns about energy efficiency, and additional barriers to resources related to racial discrimination and neighborhood of residence (13, 18) may also limit home cooling. Further research and action is needed to quantify socio-economic drivers of air conditioning usage and reduce racial disparities in access to cool spaces.

Additionally, we did not have information on indoor “uncomfortably warm temperatures” from the American Housing Survey. The U.S. Energy Information Administration’s Residential Energy Consumption Survey (RECS) does include responses on summer indoor temperatures when someone is and is not home as well as air conditioning usage (26), although information on the respondent’s comfort level is not available and sample sizes are smaller (the total number of responding households nationally was 5,686 in the 2015 RECS (27)). In future research, air conditioning usage and indoor temperatures could be modeled using household occupant and building characteristics from RECS or other surveys, as new years of data become available. Such research could advance our ability to predict indoor summer heat exposure for epidemiologic research and heat vulnerability screening.

To demonstrate the utility of the AC models, we applied the models to City of Detroit parcel data, from which information on AC ownership is entirely lacking. Room AC appeared fairly evenly dispersed across the city, but central AC ownership by census tract was more heterogeneous. Several of the census tracts with low central AC prevalence are in Southwest Detroit in an area of the city known for high levels of ambient air pollution given the presence of an oil refinery, steel mills, a wastewater treatment facility, a coal burning power plant, and truck traffic across the nearby international U.S.-Canadian border (28). That these homes also lack the air filtration offered by AC units (7) may increase personal exposure to these air pollutants. Furthermore, these areas of the county experience high outdoor and land-surface temperatures due in part to high levels of heat-retaining surfaces (29-31), thereby also increasing personal exposure to high summer temperatures. Efforts to reduce heat and air pollution exposures, such as green roofs (32) or weatherization interventions (33), may be particularly impactful in these areas of the city. In general, both heat and air pollution exposures disproportionately affect communities of color and the economically disadvantaged (34, 35), further motivating research and action addressing the independent and cumulative effects of these exposures to reduce disparities in heat- and air pollution-related health effects (13, 36).

In examining the AHS sample as a whole, we did see that uncomfortably cold indoor winter temperature was associated with central AC and, to a lesser extent, any AC, confirming that homes that are uncomfortably cold may also lack AC and may require weatherization measures for both summer and winter temperatures (Table 5). In cities such as Detroit, where weatherization measures currently focus on keeping indoor temperatures warmer in winter, additional measures, such as education on keeping homes cool in the summer or AC installation, may also be appropriate to help these same residents adapt to a warming climate.

In conclusion, central AC and room AC were reasonably predicted by housing value, year built, climate, and tenure. Future U.S. epidemiologic studies and heat vulnerability maps may reliably use the central or room AC models, or similar models, with readily available parcel-level data to predict AC ownership

when AC ownership information is not available from other sources. However, users should validate or amend this model or use alternative modeling strategies when additional housing information is available. Outside of the U.S., where AC prevalence is often lower (37), we recommend the development of region-specific models by adding air conditioning ownership and/or usage questions to national or regional surveys, particularly as air conditioning ownership increases (37). These efforts will improve understanding of mortality and morbidity risk and the spatial patterning of homes without AC to prioritize weatherization or utility assistance and heat-wave preparedness measures.

ACKNOWLEDGEMENTS

We thank our “Heatwaves, Housing, and Health” community-academic partnership, which includes Marie O’Neill, Larissa Larsen, and Tony Reames of the University of Michigan, Guy Williams and Wibke Heymach of Detroiters Working for Environmental Justice, Justin Schott of EcoWorks, Sara Clark of Southwest Detroit Environmental Vision, Michelle Lee and Rebecca Nikodem of Jefferson East, Inc., and Zachary Rowe of Friends of Parkside for providing the research hypotheses and feedback on preliminary results. Support for this research was provided by grants P30ES017885 (Berrocal and Gronlund), K99ES026198 (Gronlund) and R00ES026198 (Gronlund) from the National Institute of Environmental Health Sciences, National Institutes of Health and grant 1520803 from the National Science Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the National Science Foundation.

CONFLICTS OF INTEREST

The authors declare no competing interests.

REFERENCES

1. Kloog I, Nordio F, Lepeule J, Padoan A, Lee M, Auffray A, et al. Modelling spatio-temporally resolved air temperature across the complex geo-climate area of France using satellite-derived land surface temperature data. *International Journal of Climatology*. 2017;37(1):296-304.
2. Wang Y, Shi LH, Lee M, Liu PF, Di Q, Zanobetti A, et al. Long-term Exposure to PM_{2.5} and Mortality Among Older Adults in the Southeastern US. *Epidemiology*. 2017;28(2):207-14.
3. Abel DW, Holloway T, Harkey M, Meier P, Ahl D, Limaye VS, et al. Air-quality-related health impacts from climate change and from adaptation of cooling demand for buildings in the eastern United States: An interdisciplinary modeling study. *PLoS Med*. 2018;15(7):27.
4. Bell ML, Ebisu K, Peng R, Dominici F. Adverse Health Effects of Particulate Air Pollution: Modification by Air Conditioning. *Epidemiology*. 2009;20(5):582-686.
5. Ostro B, Rauch S, Green R, Malig B, Basu R. The effects of temperature and use of air conditioning on hospitalizations. *Am J Epidemiol*. 2010;172(9):1053-61.
6. Quinn A, Tamerius JD, Perzanowski M, Jacobson JS, Goldstein I, Acosta L, et al. Predicting indoor heat exposure risk during extreme heat events. *Sci Total Environ*. 2014;490:686-93.
7. Allen RW, Adar SD, Avol E, Cohen M, Curl CL, Larson T, et al. Modeling the residential infiltration of outdoor PM_{2.5} in the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Environ Health Perspect*. 2012;120(6):824-30.
8. Gasparri A, Guo Y, Hashizume M, Lavigne E, Zanobetti A, Schwartz J, et al. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *Lancet*. 2015;386(9991):369-75.
9. Noe RS, Jin JO, Wolkin AF. Exposure to natural cold and heat: hypothermia and hyperthermia Medicare claims, United States, 2004-2005. *Am J Public Health*. 2012;102(4):e11-8.

10. Schwartz JD, Lee M, Kinney PL, Yang S, Mills D, Sarofim MC, et al. Projections of temperature-attributable premature deaths in 209 U.S. cities using a cluster-based Poisson approach. *Environ Health*. 2015;14(1):85.
11. Gronlund CJ, Sullivan KP, Kefelegn Y, Cameron L, O'Neill MS. Climate change and temperature extremes: a review of heat-and cold-related morbidity and mortality concerns of municipalities. *Maturitas*. 2018;114:54-9.
12. Bednar DJ, Reames TG, Keoleian GA. The intersection of energy and justice: Modeling the spatial, racial/ethnic and socioeconomic patterns of urban residential heating consumption and efficiency in Detroit, Michigan. *Energy and Buildings*. 2017;143:25-34.
13. Gronlund CJ. Racial and socioeconomic disparities in heat-related health effects and their mechanisms: a review. *Curr Epidemiol Rep*. 2014;1(3):165-73.
14. Sampson N, Gronlund CJ, Buxton MA, Catalano L, White-Newsome JL, Conlon KC, et al. Staying cool in a changing climate: Reaching vulnerable populations during heat events. *Global Environ Chang*. 2013;23:475-84.
15. Sheridan S. A survey of public perception and response to heat warnings across four North American cities: an evaluation of municipal effectiveness. *Int J Biometeorol*. 2007;52(1):3-15.
16. Nitschke M, Hansen A, Bi P, Pisaniello D, Newbury J, Kitson A, et al. Risk Factors, Health Effects and Behaviour in Older People during Extreme Heat: A Survey in South Australia. *Int J Environ Res Public Health*. 2013;10(12):6721-33.
17. Banwell C, Dixon J, Bambrick H, Edwards F, Kjellstrom T. Socio-cultural reflections on heat in Australia with implications for health and climate change adaptation. *Glob Health Action*. 2012;5.
18. Madrigano J, Lane K, Petrovic N, Ahmed M, Blum M, Matte T. Awareness, Risk Perception, and Protective Behaviors for Extreme Heat and Climate Change in New York City. *International Journal of Environmental Research and Public Health*. 2018;15(7):11.
19. American Housing Survey. AHS Metropolitan Area Histories 2013 [Available from: <https://www.census.gov/content/dam/Census/programs-surveys/ahs/tech-documentation/AHS%20Metropolitan%20Area%20Histories.pdf>].
20. Aten B. Rental equivalence estimates of national and regional housing expenditures 2017 [Available from: <https://www.bea.gov/system/files/papers/WP2017-5.pdf>].
21. National Weather Service, Climate Prediction Center. Degree Day Statistics 2019 [Available from: https://www.cpc.ncep.noaa.gov/products/analysis_monitoring/cdus/degree_days/].
22. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning*. 2nd ed. New York: Springer; 2013.
23. CoreLogic Parcel Level Real Estate Data: Tax Records 2015 [Available from: <https://www.lib.umich.edu/database/corelogic-parcel-level-real-estate-data>].
24. Quinn A, Kinney P, Shaman J. Predictors of summertime heat index levels in New York City apartments. *Indoor Air*. 2017;27(4):840-51.
25. Ito K, Lane K, Olson C. Equitable Access to Air Conditioning: A City Health Department's Perspective on Preventing Heat-related Deaths. *Epidemiology*. 2018;29(6):749-52.
26. U.S. EIA. U.S. Energy Information Administration (EIA) 2015 Residential Energy Consumption Survey: Housing Characteristics Tables 2018 [Available from: <https://www.eia.gov/consumption/residential/data/2015/>].
27. U.S. EIA. U.S. Energy Information Administration (EIA) 2015 Residential Energy Consumption Survey: Methodology 2018 [Available from: <https://www.eia.gov/consumption/residential/data/2015/>].
28. Environmental Protection Agency (EPA). EJSCREEN: Environmental Justice Screening and Mapping Tool 2019 [Available from: <http://epa.gov/ejscreen>].

29. Oswald EM, Rood RB, Zhang K, Gronlund CJ, O'Neill MS, White-Newsome JL, et al. An investigation into the spatial variability of near-surface air temperatures in the Detroit, MI metropolitan region. *J Appl Meteorol Climatol*. 2012;51(7):1290-304.
30. White-Newsome JL, Brines SJ, Brown DG, Dvonch JT, Gronlund CJ, Zhang K, et al. Validating satellite-derived land surface temperature with in situ measurements: a public health perspective. *Environ Health Perspect*. 2013;121(8):925-31.
31. Zhang K, Oswald EM, Brown DG, Brines SJ, Gronlund CJ, White-Newsome JL, et al. Geostatistical exploration of spatial variation of summertime temperatures in the Detroit metropolitan region. *Environ Res*. 2011;111(8):1046-53.
32. Santamouris M. Cooling the cities – A review of reflective and green roof mitigation technologies to fight heat island and improve comfort in urban environments. *Solar Energy*. 2014;103:682-703.
33. Tonn B, Rose E, Hawkins B. Evaluation of the US department of energy's weatherization assistance program: Impact results. *Energy Policy*. 2018;118:279-90.
34. Mikati I, Benson AF, Luben TJ, Sacks JD, Richmond-Bryant J. Disparities in Distribution of Particulate Matter Emission Sources by Race and Poverty Status. *American Journal of Public Health*. 2018;108(4):480-5.
35. Koman PD, Romo F, Swinton P, Mentz GB, de Majo RF, Sampson NR, et al. MI-Environment: Geospatial patterns and inequality of relative heat stress vulnerability in Michigan. *Health & place*. 2019;60:102228.
36. Hajat A, Hsia C, O'Neill MS. Socioeconomic Disparities and Air Pollution Exposure: a Global Review. *Curr Envir Health Rpt*. 2015;2(4):440-50.
37. International Energy Agency (IEA). The Future of Cooling Paris2018 [Available from: <https://www.iea.org/reports/the-future-of-cooling>].

TABLES

Table 1. Summary statistics of climate and housing characteristics across 140 U.S. core-based statistical areas (CBSAs), 2003-2013.

Characteristic	Mean	Minimum	25 th Percentile	Median	75 th Percentile	Maximum
Number of households sampled ¹	1,306	19	52	105	2,698	8,042
Proportion of households with central air conditioning (cenAC) ¹	0.67	0.046	0.50	0.73	0.89	0.99
Proportion of households with room air conditioning (rmAC) ¹	0.25	0.00	0.13	0.20	0.34	0.72
Proportion of households uncomfortably cold for 24 hours or more (tooCold) ¹	0.085	0.00	0.059	0.085	0.11	0.24
Annual cooling degree-days (CDDs) ²	1,300	150	560	910	1,900	3,800
Annual heating degree-days (HDDs) ²	4,500	300	2,700	4,600	6,400	9,200
Mean year built ¹	1963	1937	1957	1964	1971	1983
Mean year of survey ¹	2010	2005	2010	2011	2011	2011
Proportion of rental homes ¹	0.38	0.21	0.32	0.37	0.45	0.61
Proportion of multi-family homes ¹	0.29	0.04	0.22	0.27	0.35	0.64
Proportion of mobile homes ¹	0.026	0.000	0.000	0.018	0.039	0.187
Geometric mean housing value (dollars) ^{1,3}	112,000	48,000	85,800	107,000	140,000	344,000

American Housing Survey

National Weather Service Climate Prediction Center

³Dollar values were not discounted across years.

Table 2. Means or proportions of selected housing characteristics by survey year, American Housing Survey National Survey, 2003-2013.

	2003	2005	2007	2009	2011	2013
Proportion of households with central air conditioning (cenAC)	0.50	0.54	0.60	0.59	0.65	0.64
Proportion of households with room air conditioning (rmAC)	0.32	0.30	0.28	0.30	0.28	0.28
Proportion of households uncomfortably cold for 24 hours or more (tooCold)	0.10	0.08	0.10	0.10	0.10	0.09
Mean year built	1957	1957	1959	1959	1964	1960
Proportion of rental homes	0.40	0.49	0.54	0.49	0.36	0.41
Proportion of multi-family homes	0.35	0.42	0.46	0.43	0.30	0.33
Proportion of mobile homes	0.06	0.02	0.02	0.01	0.02	0.01
Geometric mean housing value (dollars)	117,000	110,000	112,000	125,000	143,000	140,000

Table 3. Change in the log-odds of central air conditioning ownership (cenAC), room air conditioner ownership (rmAC) and “home was uncomfortably cold for 24 hours or more” (tooCold) for presence vs. absence or interquartile-range (IQR) increase¹ in each characteristic in 140 U.S. core-based statistical areas, 2003-2013.

	cenAC		rmAC		tooCold	
	Coefficient	Standard Error	Coefficient	Standard Error	Coefficient	Standard Error
Intercept	1.51	0.11***	-1.60	0.07***	-2.50	0.04***
Year built	1.12	0.01***	-0.87	0.01***	-0.23	0.01***
Rental	-0.52	0.02***	0.31	0.02***	0.36	0.02***
DDs	1.16	0.12***	-0.42	0.07***	0.35	0.06***
Year of survey	0.23	0.01***	0.03	0.01***	0.20	0.01***
ln(Housing value)	0.62	0.01***	-0.42	0.01***	-0.14	0.01***
ln(Housing value) ²	0.09	0.00***	-0.07	0.00***	0.00	0.00
Multi-family	0.03	0.02	-0.08	0.02***	-0.20	0.02***
Mobilehome	-0.56	0.04***	0.74	0.03***	0.41	0.04***
Year built x rental	0.04	0.02*	0.26	0.01***	-0.09	0.02***
Year built x DDs	0.27	0.01***	-0.16	0.01***	0.15	0.02***
Rental x DDs	0.16	0.03***	-0.23	0.02***	-0.06	0.04
Year of survey x rental	-0.20	0.02***	0.07	0.02***	-0.20	0.02***
Year of survey x DDs	-0.16	0.01***	0.04	0.01***	0.09	0.02***
ln(Housing value) x rental	0.36	0.02***	-0.14	0.02***	0.04	0.02
ln(Housing value) x DDs	0.24	0.01***	-0.15	0.01***	0.11	0.02***
ln(Housing value) ² x rental	0.12	0.01***	-0.06	0.01***	-0.02	0.01*
ln(Housing value) ² x DDs	0.01	0.00***	-0.01	0.00**	0.02	0.01***
Multi-family x DDs	0.33	0.02***	-0.43	0.02***	0.16	0.04***
Year built x rental x DDs	-0.20	0.02***	-0.18	0.02***	-0.02	0.03
Year of survey x rental x DDs	0.19	0.02***	-0.06	0.02***	-0.09	0.04**
ln(Housing value) x rental x DDs	0.06	0.03*	-0.13	0.03***	-0.06	0.04
ln(Housing value) ² x rental x DDs	0.03	0.01***	-0.02	0.01*	-0.01	0.01

*P-value < 0.05; **P-value < 0.01, ***P-value < 0.001

Note that the odds ratios for presence vs. absence or an interquartile-range increase can be found by taking exp(coefficient).

DDs = mean annual cooling degree-days (CDDs, in air conditioning models) or mean annual heating degree-days (HDDs, in tooCold model).

The value of the random effect for each core-based statistical area is provided in Supplemental Material Table 1 as well as instructions for computing parcel-level probabilities.

The standard deviations of the random effects were 1.18, 0.69, and 0.34 for the cenAC, rmAC, and tooCold models, respectively.

¹The IQR increases for year built, CDDs, HDDs, year of survey, and ln(Housing value) were calculated on the raw data rather than the CBSA means (as in Table 1) and therefore differ from those in Table 1. They are 30 years, 1000 CDDs, 3000 HDDs, 5 years, and 1 log-dollar, respectively.

Table 4. Model sensitivity and specificity for logistic regression models of central air conditioning (cenAC), room air conditioning (rmAC) and “home was uncomfortably cold for 24 hours or more” (tooCold). Sensitivity and specificity are obtained by comparing predicted probabilities for the 2015-2017 period to the 2015-2017 American Housing Survey data.

	Model Threshold ¹ = 0.1	Model Threshold ¹ = 0.5	Model Threshold ¹ = 0.9
cenAC			
Specificity ²		63%	95%
Sensitivity ³		91%	59%
Accuracy ⁴		84%	68%
rmAC			
Specificity ²		94%	100%
Sensitivity ³		34%	1%
Accuracy ⁴		82%	80%
tooCold			
Specificity ²	57%	100%	
Sensitivity ³	62%	0%	
Accuracy ⁴	57%	93%	

¹Model threshold refers to the modeled probability above which a housing unit is considered to have AC or be tooCold.

²True negative rate, i.e., the rate at which homes without AC were predicted to lack AC or homes that were not tooCold were predicted to be not tooCold.

³True positive rate, i.e., the rate at which homes with AC were predicted to have AC or homes that were tooCold were predicted to be tooCold.

⁴Percent of the values predicted correctly by the model.

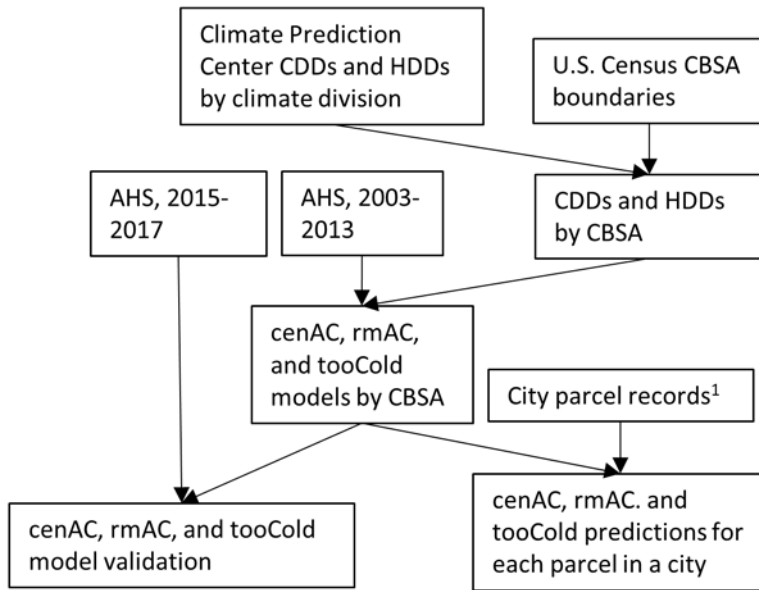
See Supplemental Material Table 2 for city-specific cenAC results, including results for a model threshold of 0.95.

Table 5. Cross tabulations of central air conditioning presence (cenAC) or any air conditioning (central or room, anyAC) and “home was uncomfortably cold for 24 hours or more” (tooCold) and the odds of tooCold among those without vs. with AC (or, identically, the odds of no cenAC among those tooCold vs. not tooCold) from the 2003-2017 American Housing Survey.

	tooCold	not tooCold	Odds Ratio (95% Confidence Interval)
no cenAC	4.2%	27.7%	2.13 (2.07, 2.18)
cenAC	4.5%	63.7%	
no anyAC	1.4%	10.1%	1.57 (1.52, 1.63)
anyAC	7.2%	81.3%	

FIGURES

Figure 1. Process for developing and applying models predicting central air conditioning (cenAC), room air conditioning (rmAC), and uncomfortably cold indoor winter temperatures (tooCold) for a given parcel.



AHS = American Housing Survey, CDDs = cooling degree-days, HDDs = heating degree-days, CBSA = core-based statistical area

¹City parcel records from Detroit, Michigan, USA were used.

Figure 2. Average annual cooling-degree days (CDDs, cumulative number of degree days above 18 C per year), for the 1981-2010 climatological period, in each of the 140 available core-based statistical areas (CBSAs), American Housing Survey, 2003-2013.

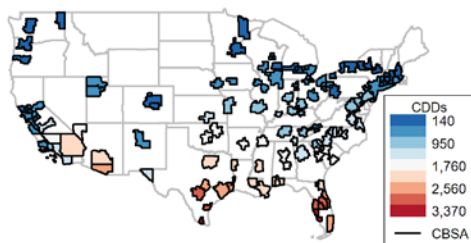


Figure 3. Receiver operating curves (ROCs) for each of the 35 core-based statistical areas (CBSAs) with data available during the training period of 2003-2013 and the out-of-sample prediction period 2015-2017. The ROC refers to predicted probabilities of central air conditioning (A), room air conditioning (B), and home was uncomfortably cold (C). CDDs = cooling-degree days and HDDs = heating-degree days.

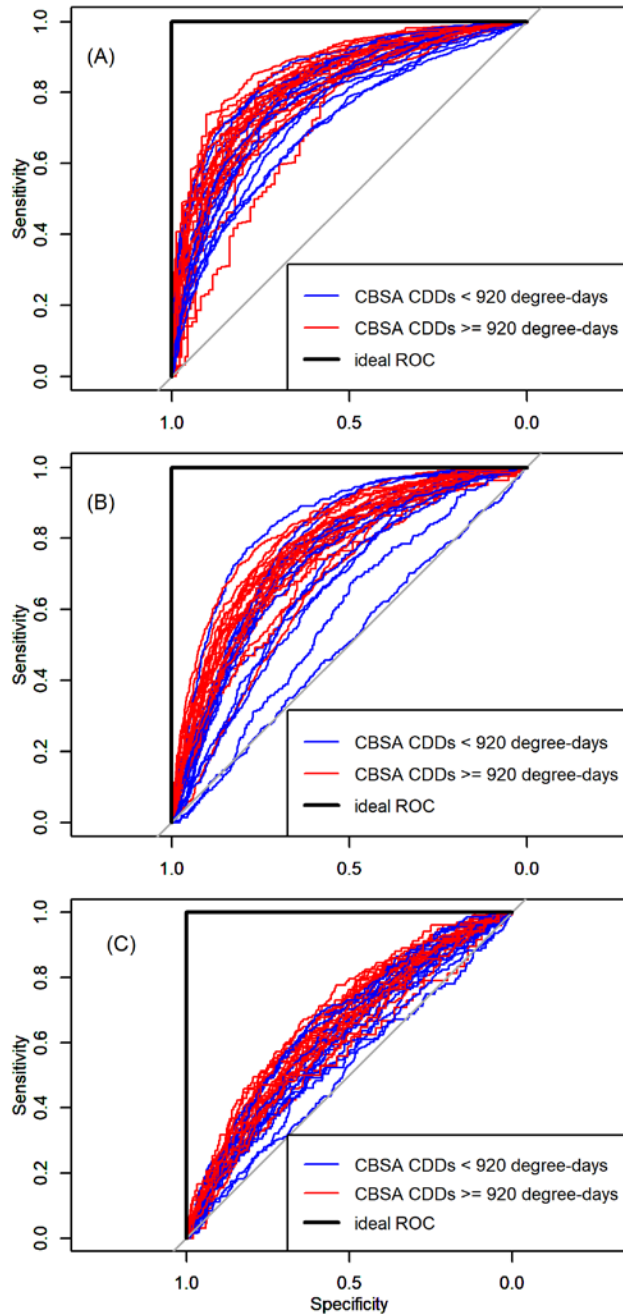


Figure 4. Model-predicted probability (Prob) of central air conditioning (AC) (A) or any (central or room) AC (B) by parcel (N = 390,668), Detroit, MI, 2016.

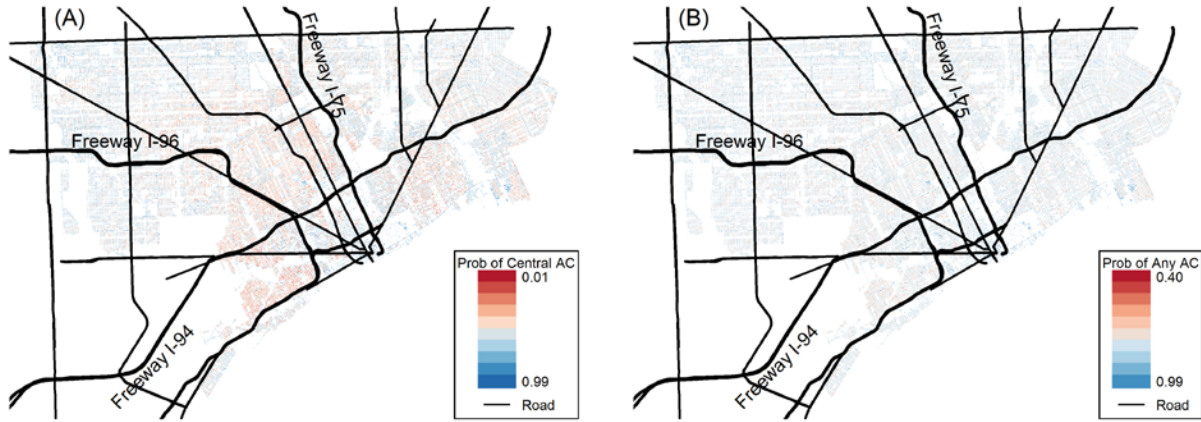


Figure 5. Model-predicted probability (Prob) of central air conditioning (AC) (A) or any (central or room) AC (B) by tract, Detroit, MI, 2016.

