Landcover Data to Derive Related Ancillary Variables for Dasymetric Modeling

Stefan Leyk, Barbara P. Buttenfield, Nicholas N. Nagle and Alexander K. Stum

ABSTRACT: A growing need for demographic small area estimates beyond the scales of census units has led to increased interest in dasymetric modeling. Landcover-based residential areas are typically used as limiting ancillary variables to spatially refine demography to sub-tract level. This paper examines statistical associations between landcover-derived variables and census-based housing characteristics at fine spatial scale defined by areas within census block groups. To validate, parcel data are allocated to these areas to create fine-scale 'ground truth' data for the attribute of interest (number of different building types) which is highly correlated to census housing attributes. A number of related ancillary variables are tested such as target zone area and measures of 'urban-ness' which are called Inner Dimension Metrics (IDMs). IDMs indicate for each pixel within the developed landcover class the minimum number of neighbors of the same class in each direction which is then averaged for each target zone. For each landcover class Poisson Adjusted Generalized Linear Models (GLMs) are computed to estimate each of the housing attributes at the target zone scale. We demonstrate our methods for Boulder County, Colorado.

The results show highly significant relationships between the number of various building types and related ancillary variables across the different landcover classes. Maps of model residuals identify spatial clusters of over- and under-predictions of individual models. These preliminary results indicate that the related ancillary variables have predictive power for estimating census housing attributes at fine spatial scale. This research could improve dasymetric models for small area population estimates by establishing a statistical basis for selecting related variables. It may also eliminate the need for limiting variables, since all dasymetric ancillary information could be considered as related variables.

KEYWORDS: Dasymetric modeling, parcel data, landcover, related ancillary variables

Introduction

An urgent need exists to improve small area estimates of demographic characteristics beyond spatial scales provided by census units such as tracts or block groups. Methods to create finer resolution estimates include dasymetric mapping (Semenov-Tian-Shansky, 1928; Wright, 1936) - a special type of areal interpolation that makes use of other ancillary variables. Ancillary variables, which are assumed to have finer spatial resolutions than the source zones of the demographic data, allow the partitioning of space into homogeneous (target) zones minimizing variation within each zone and showing steepest changes in values at the boundaries (Mennis, 2009). The growing body of literature on dasymetric modeling describes the incorporation of various types of ancillary data such as landcover (Mennis, 2003), road density (Reibel and Bufalino, 2005), Landsat TM (Yuan et al., 1997), parcel data (Tapp, 2010) or address points (Zandbergen and Ignizio, 2011). Two types of ancillary variables are commonly considered. Limiting ancillary variables constrain the study area to regions that can be populated versus areas that have zero population. Related ancillary variables are intended to define more complex relationships with population data to constrain or amplify demographic estimates at finer spatial scale. How to identify such meaningful relationships objectively has been discussed for some time (e.g., Harvey, 2002; Li and Weng, 2005; Mennis and Hultgren, 2006; Maantay et al., 2007; Zandbergen and Ignizio, 2010) and remains a persistent challenge (Mennis, 2009). One main reason is the absence of 'ground truth' data to establish statistics-based relationships between ancillary and population data at the scale of the target zones. Due to confidentiality issues publicly available population (census) data are aggregated and thus dasymetric population refinements at finer scales are commonly difficult to verify.

Landcover data derived from remotely sensed imagery are the most common ancillary data type used in dasymetric mapping, frequently applied as a limiting variable to predetermine residential area. Attempts to use landcover as related variables usually suffer from above described problems and are often based on subjective assumptions (Eicher and Brewer, 2001; Mennis and Hultgren, 2006; Zandbergen and Ignizio, 2010). Thus one question is whether landcover can be used in a statistical framework to dasymetric modeling. This can only be done if statistical relationships between population data and related ancillary data can be established in an objective, validation-driven way. This question is highly relevant because landcover data are often released as national datasets. Moreover, establishing objective statistical associations would allow researchers to carry out the same analysis for any study area within the same country.

This paper proposes an analytical procedure to establish statistical associations between parcel data building-related attributes, which are highly correlated to census housing attributes, and variables derived from the National Land Cover Database (NLCD), to build a dasymetric model at a target scale finer than the block group level of the U.S. census data. Parcel data will support formation of a fine-resolution validation dataset, and be tested here for Boulder County, CO. In recent years, as an increasing number of local governments release digital cadastral records, it has been argued that land parcel data will improve the precision and accuracy of dasymetric maps (Tapp, 2010). Parcel data is not available yet in some parts of the country however, and often available only at significant expense. Also, parcel data provide only a limited number of attributes that can be related to census data and processing them is labor- and computer-intensive. The aim of this paper is to examine whether parcel data with their landuse/building type information allow establishment of statistical associations to variables derived from the NLCD. Since parcel-based building type is assumed to be highly correlated to census housing variables such as 'Units in Structure' (Table H30 in the American Community Survey 2006-10, 5-year averages) the existence of such relationships would have considerable impact on small area estimation. For example dasymetric methods for estimating many different demographic attributes at the census tract level based on spatial allocation of household microdata (Leyk et al., in press) would greatly benefit from such advancement. It would allow use of these relationships to guide further allocation of households to fine-scale target zones inside census units and thus the computation of demographic summaries for all attributes found in public use microdata samples (PUMS) at the same fine scale.

Data and Pre-processing

Parcel data from Boulder county for the year 2008 (Figure 1) were obtained including geometry and attribute data. The data were filtered by the 'Land Use/ Building Type' attribute to generate a parcel dataset of residential land use. Parcels were identified as residential if there was residential use or multi-use with one being residential indicated in the building type attribute. The residential building type categories had to be reclassified in order to create data as comparable as possible to the 'Units in Structure' census attribute (see below). Since categories do not overlap perfectly and the perspective of the data is slightly different (parcel/property information vs. household/housing information) comparability is close but not perfect. The resultant aggregated 'Building Type' classes were 'One-Family Residences' (including detached and attached as well as farm residences), 'Duplex/Triplex', 'Multi-Unit Buildings with 4-8 Units', which included condominiums, 'Multi-Unit Buildings with 9 and more Units' and 'Manufactured Housing' including mobile homes.

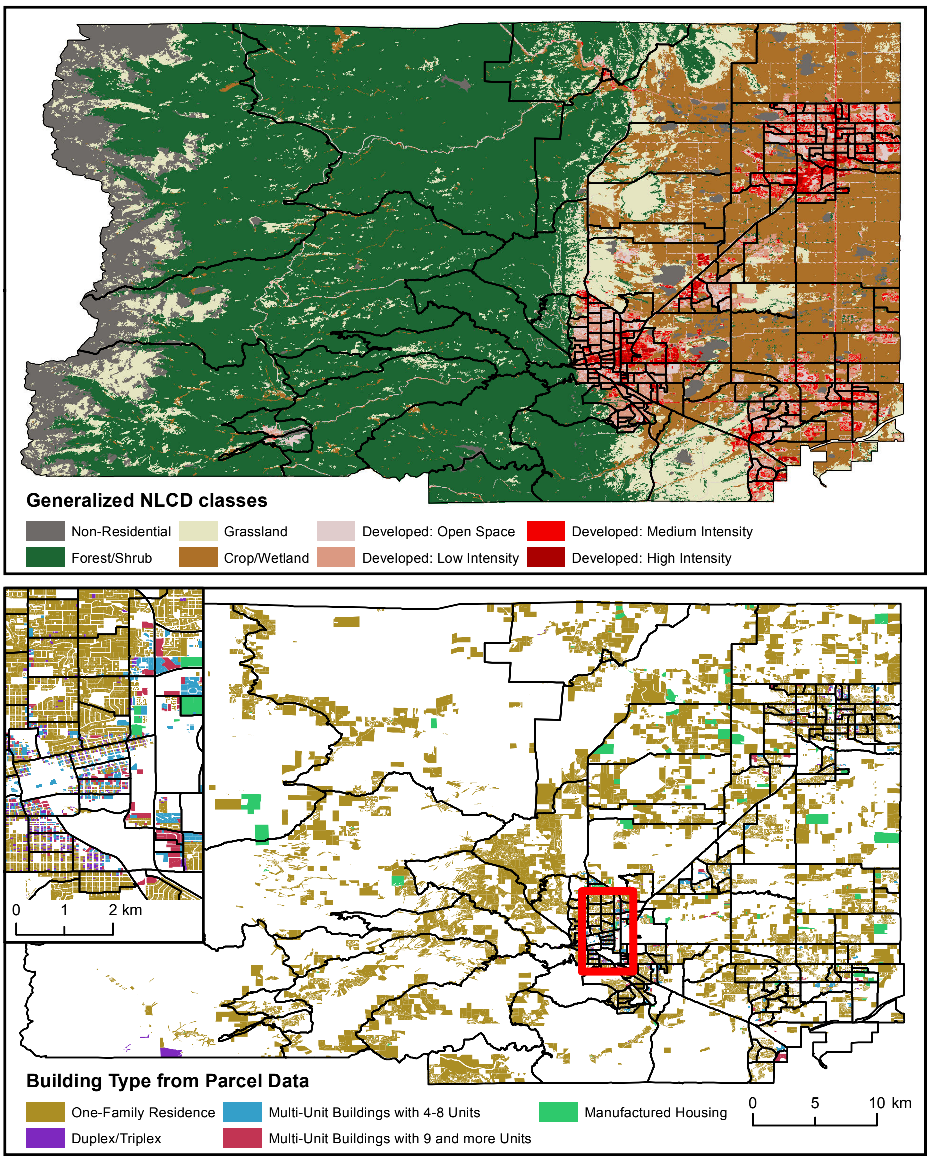


Figure 1: Study area: Boulder County, CO. Top: Reclassified NLCD data overlain by census block group boundaries. Bottom: Reclassified building types derived from parcel data.

The second source of data is the *2006 National Land Cover Database (NLCD)* produced by the U.S. Geological Survey. The NLCD is a 30 x 30 m raster grid of the entire United States, with each cell preclassified into a single landcover type based on remote sensing techniques. The NLCD classes deemed most similar in terms of expected population were aggregated (Figure 1). This resulted in one 'Non-Residential' class including water, barren and snow/ice, as well as three 'Vegetated' classes namely 'Forest/Shrub', 'Grassland' and 'Crop/Wetland'. These vegetated classes are kept separated because they are distinct classes and are found in different settings of landcover that can have different population and housing patterns. Finally the four 'Developed' landcover classes in NLCD ('Open Space', 'Low', 'Medium' and 'High' intensities) were left as single classes.

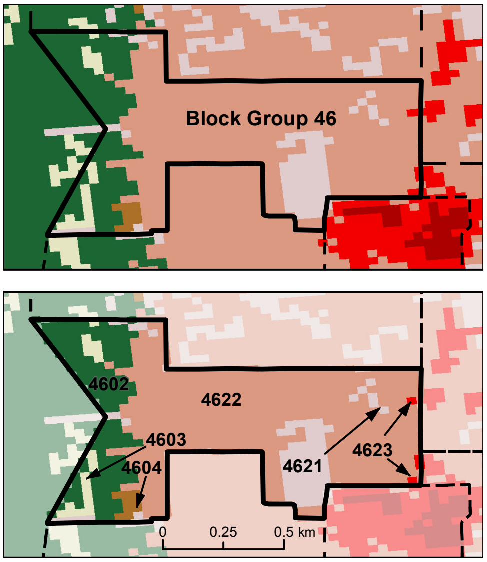
A third source of data is taken from *U.S. Census block group summaries* of the housing attribute 'Units in Structure' (Table H30) from the 2006-2010 5-year estimates of the American Community Survey (ACS). Reclassifying the categories of 'Units in Structure' resulted in the five classes 'One-Family Residences' (including detached and attached), 'Buildings with 2-4 Units', 'Buildings with 5-9 Units', 'Buildings with more than 9 Units' and 'Mobile Homes'. In order to test whether parcel data building types are indeed representative for 'Units in Structure' in ACS, correlations with block group summaries generated for parcel data building types will be calculated. The existence of such correlations is important to ensure that relationships between parcel-derived building types and landcover attributes are representative for the "Units in Structure"-landcover connection also at the scale of target units.

Method

The use of landcover data as a related dasymetric ancillary variable, and derivation of statistical relationships to census population at fine scales raises some interesting challenges. Landcover does not allow direct estimation of population attributes but appears to be indirectly linked to housing attributes as can be seen in definitions for different categories of developed land in NLCD. These definitions describe most common housing types. However, establishing statistical associations between census housing variables and landcover characteristics represents a difficult problem since public census population and housing data are aggregated to differing spatial units; spatially more precise data would be needed to examine whether associations exist at finer scales. Therefore the central question in this study is: Can building type information found in parcel data be used to identify statistical relationships between census housing attributes and variables derived from nationally available landcover datasets? Answering this question will help to better estimate the potential of landcover as a fine scale related dasymetric variable. Four different steps will be carried out in this study: construction of target zones; spatial allocation of parcel units to target zones; ancillary variable generation; and statistical modeling.

Constructing the Target Zones

Target zones in this paper were defined as areas of aggregated landcover classes found within block groups. As a consequence larger landcover patches, which span several block group units, were subdivided into groups of the same class each of which obtained



a unique identifier (Figure 2). Technically, the block group spatial boundaries were converted to raster data at NLCD spatial resolution. A simple map algebra operation created unique identifiers for each class of landcover found within block group areas. Since Boulder County contains 200 block groups, each landcover class could occur up to 200 times, also representing the maximum number of modeling units that can be used to fit a landcover-specific model to estimate the number of different building types.

Figure 2: Creation of target zones for dasymetric modeling: landcover patches within block groups.

Spatial Allocation of Parcel Units to Target Zones

In order to generate reliable spatial estimates of the numbers of buildings of different types at the scale of the target zones (landcover patch within a census block group; Figure 2), parcel units were allocated to the various target zones. This allocation problem is not trivial and impeded by several factors. First, the landcover data and thus the target zones are in raster format with 30m resolution while parcel boundaries are in vector format. Second, the ratio between the relative sizes of a parcel unit and of a target zone can invert between urban and rural areas. In urban areas parcels can be very small (even smaller than a landcover raster cell) but parcels in rural areas can be very large and frequently include areas of different landcover classes within the same block group unit. Third, parcel units can intersect with block group boundaries such that landcover patches found within one parcel unit could be contained in different block groups.

A spatial allocation was carried out as follows. First, the target zone areas in raster format were vectorized and intersected with the parcel units. The resultant vector dataset consisted of more than 181,000 features many of which were created as small fractions of parcels which crossed target zone boundaries. All polygons were retained to maintain the pycnophylactic property (Tobler, 1979). The pycnophylactic or mass-preserving property requires that while population is re-aggregated from source to target zones, the total number of entities must remain unaltered.

Second, each parcel unit was distributed across the target zones it spatially intersected with according to proportions of these target zones within the parcel unit as well as conditional on the existence of certain landcover types. The following logical assignment procedure was applied:

For each parcel:

- Identify how many and what type of target zones can be found inside the parcel;
- Calculate the proportion occupied by each of these target zones;
- If only one target zone was found, allocate the whole parcel unit to this target zone;

- Else, if one or more target zones of any type of 'Developed Land' are detected the parcel is allocated to these target zones only; each developed land target zone obtains a fraction of the parcel that corresponds to its proportion on the total area of developed land within the parcel; all other landcover classes are excluded;
- Else, if there are only target zones of type vegetated land (no developed) the parcel unit is proportionally allocated to each of the vegetated classes (relative to their area); if non-residential land is identified this area is excluded and proportions recalculated to the total area vegetated land.

Summing up fractions of parcels allocated to target zones results in a statistically precise solution which preserves the pycnophylactic property and allows for derivation of meaningful summary statistics for small areas. Allocation results were created for total counts of buildings (i.e., parcels) but also separately for each building type category, by adding up fractional values only if the corresponding building type has been found as attribute. This allocation also allowed to create sub-block group maps of the number of parcels (i.e., buildings), as well as the number of each individual building type. Maps of building and building type densities could be created at the same scale. The resultant allocation represents both a dasymetric map of a housing attribute ('Building Type' which is related to census-derived 'Units in Structure') and at the same time reflects some kind of ground truth data. As these ground truth data are created at a spatial scale defined by landcover data they can be used to examine if statistical associations exist between these housing attributes and characteristics of landcover-derived target zones.

Creating Related Ancillary Variables for Dasymetric Modeling

To ensure that related variables are generic (i.e., they can be defined the same way for any other study area) they have to be directly derived from landcover data and extracted from within target zones. The first related variable used for statistical modeling was the area of the target zones since it can be assumed that the number of buildings as well as the number of particular building types are related to the extent of the spatial unit.

It can also be assumed that the degree of 'urban-ness' or simply 'how far inside of a developed region a target zone is located' could influence the number of buildings and in particular of different types of buildings. For example, it could be expected that the further inside a developed region a target zone is located the more urban the building types will become, conditional on the underlying landcover class of the target zone itself. A measure was derived, which is called the 'inner dimension metric' (*IDM*). *IDM* quantifies for each developed land pixel in the NLCD the minimum number of neighbors which also belong to one of the developed classes, in all directions. The higher the metric becomes the further inside a developed region a considered pixel is located (Figure 3).

For each target zone the mean *IDM* value was extracted as a zonal statistic. Three variants of this mean *IDM* variable were derived: IDM_{all} which includes all developed classes ('Open Space', 'Low', 'Medium' and 'High Intensity'), IDM_{low} which includes only 'Open Space' and 'Low Intensity' developed land, and IDM_{high} which includes 'Medium' and 'High Intensity' developed land.

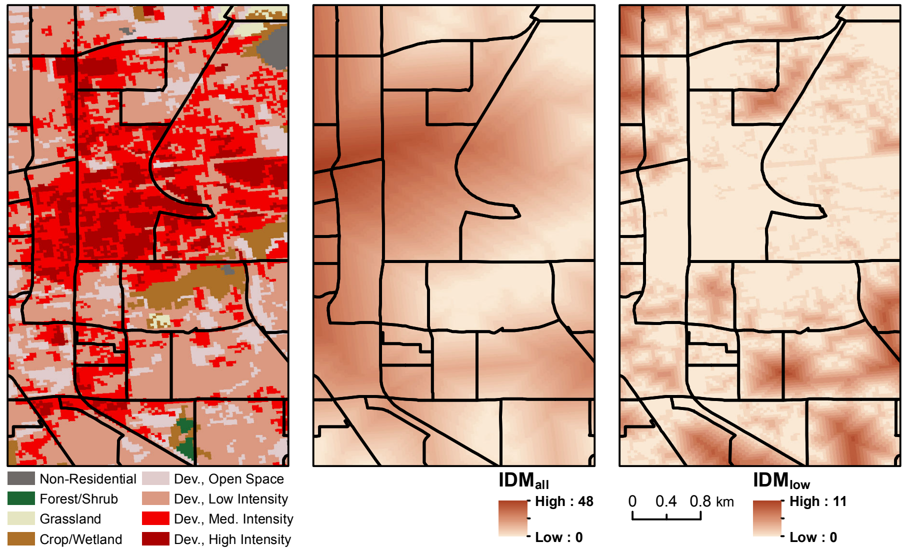


Figure 3: Creating *IDM* surfaces; reclassified landcover (left), IDM_{all} which includes all developed NLCD classes (middle) and IDM_{low} which includes only 'Open Space' and 'Low Intensity' developed land (right).

Statistical Modeling to Examine Related Ancillary Variables

The three *IDM* metrics were input to develop different landcover-specific models, since it has to be expected that relationships between outcome variables and related variables would be different within different landcover classes. For example, a model for 'Low Intensity' developed land would be computed based on all target zones that were created by regions of 'Low Intensity' developed land found in block groups. Since Boulder County contains 200 block groups, models can be fit on 200 entities at a maximum if every block group contained at least one pixel of the corresponding landcover class. In addition, individual models were computed for each of the six outcome variables (i.e., number of buildings and number of different building types) because within the same landcover class each outcome variable will have a different relationship to the individual related variables. This model setup resulted in the computation of 48 models in total. It also allowed for each landcover class the estimation of each of the outcome variables of interest and thus a very detailed examination for the existence of statistical relationships that would be useful for a dasymetric model based solely on landcover-derived variables.

Since the statistical distributions of the number of buildings and numbers of different building types follow a Poisson distribution, Poisson Adjusted Generalized Linear Models (McCullagh and Nelder, 1989) were fit. The coefficient values, their p-values as well as the Akaike Information Criterion (AIC) establish the strength of observed associations between outcome and related variables. Maps of model residuals were created to examine their spatial patterns across the study area, to visualize where the model performs most accurately and where it over- or under-estimates.

Results and Discussion

All correlations between the reclassified census attributes ('Units in Structure') and the parcel-derived 'Building Types' were significant ($p<0.01$) on the block group level.

Correlation coefficients ranged from 0.57 to 0.95 for different categories. Some discrepancies between the sources will be addressed below, however these correlations provide evidence that parcel data 'Building Type' estimates can be used as reliable, spatially more precise, ground truth data for the 'Units in Structure' attribute in the ACS.

In Figure 4 dasymetric maps are shown for two building types illustrating the spatial distribution of the outcome variables as well as the calculated densities at the spatial scale of the target zones. Such maps are useful in representing direct results when parcel data are used in combination with landcover data to create related ancillary variables which spatially refine the distribution of the housing attribute 'Units in Structure'. Second, these maps show how related ancillary variables directly derived from landcover data can be used in the absence of parcel data if the statistical relationships described below prove to be stable.

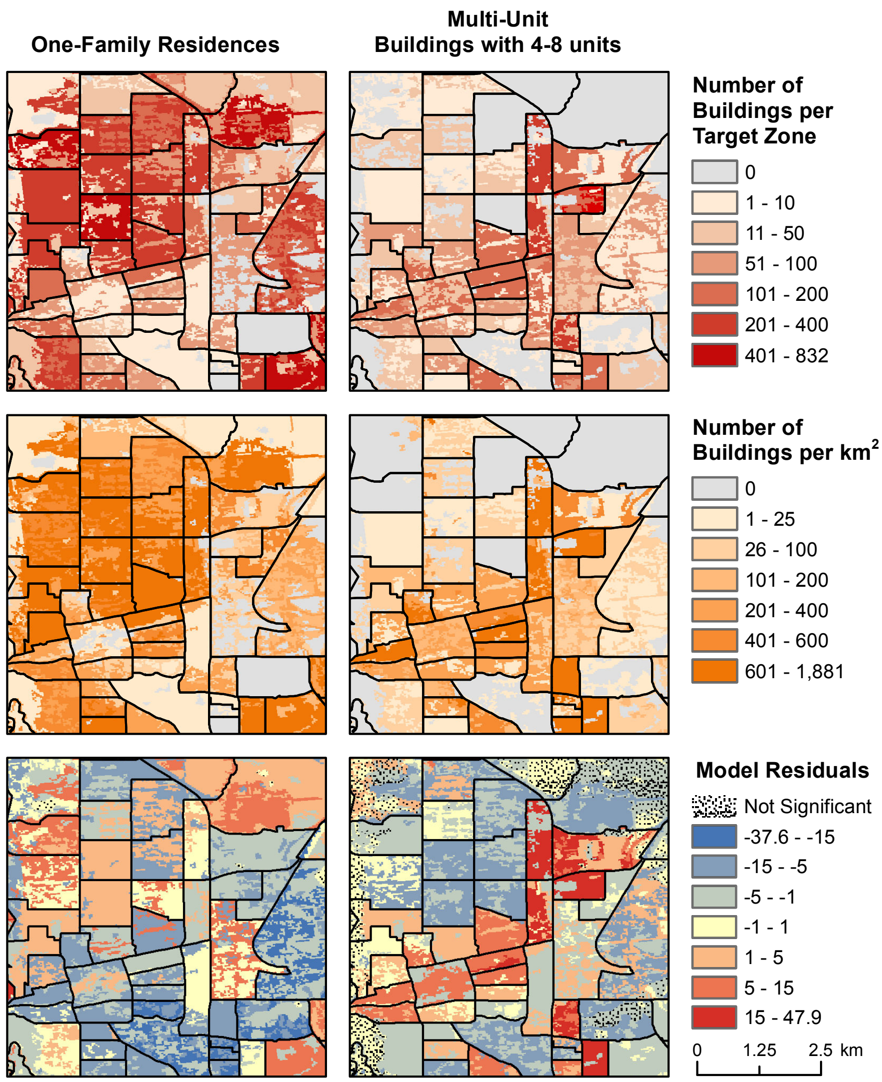


Figure 4: Top: Dasymetric maps for number of 'One-Family Residences' (left) and 'Multi-Unit Buildings with 4-8 Units' (right); Middle: The same attributes presented as densities; Bottom: Spatial distributions of residuals from models for the two building types.

Table 1 reports model fits for the different types of target zones and landcover classes. If a model could not be fit because there were no or very few target zones that overlapped with any residential parcels of the considered building type, it was assumed that no association exists between ancillary variables derived from these target zones and the corresponding outcome variable.

As for the *‘Non-residential’* and *‘Vegetated’* classes it should be noted that only one related variable (the area of the corresponding target zone) could be used for modeling. Only two ‘Non-residential’ target zones had values greater than zero, meaning there is no association between the number of any building type and the area of these non-residential landcover patches. For ‘Forest/Shrub’, ‘Grassland’ and the ‘Cropland/Wetland’ classes highly significant but small coefficients can be observed for the related ancillary variable when modeling ‘One-Family Residences’ as well as the total number of buildings, which are very similar in these cases. Very few target zones showed counts greater than zero for the other building types suggesting that there is no statistical relationship.

Table 1: Model coefficients of variables tested significant ($p<0.01$) sorted by landcover type.

Landcover	***Building types***	***Area***	***IDM_{all}***	***IDM_{low}***	***IDM_{high}***
Forest/ Shrub	1-family	0.0167	x	x	x
	All	0.0167	x	x	x
Grassland	1-family	0.1111	x	x	x
	All	0.1111	x	x	x
Cropland/ Wetland	1-family	0.0633	x	x	x
	All	0.0611	x	x	x
Developed Open Space	1-family	0.9667	-0.0357	0.2094	x
	Dup/Tri			0.2655	x
	Multi (4-8)	-0.2333	-0.0395	-0.1193	x
	Mobile	1.2333			x
	All	0.9111	-0.0357	0.1924	x
Developed low intensity	1-family	1.0956	0.0033	0.1266	x
	Dup/Tri	-0.7444	0.0395	0.1096	x
	Multi (4-8)		0.0199		x
	Multi (9+)	-2.1222	0.0578		x
	All	0.9700	0.0078	0.1039	x
Developed medium intensity	1-family	2.2111	-0.0243	x	-0.946
	Dup/Tri	0.9900	0.0458	x	0.4001
	Multi (4-8)		0.0098	x	0.743
	Multi (9+)		0.0511	x	0.587
	All	1.5889	-0.0084	x	-0.150
Developed high intensity	1-family	-0.6111	-0.0398	x	
	Multi (4-8)	0.9422	0.022	x	0.711
	All			x	0.478

Among the models for *‘Open Space’* landcover interesting patterns could be detected. All three variables – *Area*, IDM_{all} and IDM_{low} – were tested significant ($p<0.01$) when estimating the total number of buildings as well as the number of ‘One-Family Residences’ with very similar coefficient values. The negative coefficient for IDM_{all} suggests an inverse relationship while the positive coefficient for IDM_{low} indicates that higher IDM_{low} values are associated with higher numbers of 'One-Family Buildings' (note that IDM_{low} is derived from 'Open Space' and 'Low Intensity' developed classes only). When estimating the number of 'Duplex/Triplex' building types only IDM_{low} was tested significant with a positive coefficient. This indicates that for target zones of 'Open Space' landcover type higher numbers of this building type can be expected further inside of larger patches of lower intensity developed land. When estimating 'Multi-Unit Buildings with 4-8 Units' both *IDM* variables are tested significant ($p<0.01$) with negative coefficients. Finally, when estimating the number of mobile homes, target zone area was tested significant with a positive coefficient.

As for the models for the '*Low Intensity Developed'* landcover class, significant ($p<0.01$) variables could be found for most building type outcomes. All three related variables, target zone area, IDM_{all} and IDM_{low}, were tested significant when modeling total number of buildings, 'One-Family Residences' as well as 'Duplexes/Triplex'. IDM_{all} was tested significant when modeling 'Multi-Unit Buildings with 4-8 Units' as well as 'Multi-Unit Buildings with 9 and more Units'. All these *IDM* coefficients were positive suggesting that in models across target zones of 'Low Intensity Developed' landcover IDM_{all} and/or IDM_{low} are positively related to the number of all building types except for mobile homes which were rarely found within this landcover class.

Among the models fit for the '*Medium Intensity Developed'* landcover class negative relationships were observed between the *IDM* variables (IDM_{all} and IDM_{high}) and both the total number of buildings and the number of 'One-Family Residences'. When modeling any of the multi-unit building types the significant regression coefficients of *IDM* variables were positive. These observations suggest that the further inside developed land a target zone of this landcover type is located the more multi-unit buildings and the fewer 'One-Family Residences' can be expected. Finally, the models for the '*High Intensity Developed'* landcover class show negative associations between 'One-Family Residences' and both *IDM* variables (IDM_{all} and IDM_{high}) as well as positive associations between 'Multi-Unit Buildings with 4-8 Units' and *IDM*. These trends are similar to the models for the 'Medium Intensity Developed' landcover class.

The two maps at the bottom of Figure 4 show the spatial distribution of model residuals computed for the two outcome variables 'One-Family Residences' and 'Multi-Unit Buildings with 4-8 Units' (note that for target zones of different landcover type the residuals come from different models). These maps demonstrate where a model tends to under- and overpredict. The residuals computed from the model for 'One-Family Residences' shows a distinct pattern of clusters of over- and under-predictions and indicates spatial non-stationarity in the error terms. The model for 'Multi-Unit Buildings with 4-8 Units' shows a somewhat similar pattern. This outcome variable is less correlated with its corresponding census category on the block group level and thus the relationships are less reliable.

Conclusions

This study describes an attempt to shed light on the problem of establishing statistically based relationships among related ancillary variables and demographic attributes in dasymetric modeling as one persistent challenge for demographic small area estimation. The ability to use nationally available landcover data to derive related ancillary variables at fine scales would empower existing approaches of dasymetric modeling considerably. As shown in this paper, parcel data represent an appropriate data source to examine whether statistical relationships exist between building type characteristics highly correlated with census housing attributes and variables directly derived from the NLCD at fine spatial scales. The model results indicate that there are several highly significant relationships identified for Boulder county that would allow dasymetric models to be computed using NLCD-derived related ancillary variables only. Thus there would be no need for incorporating any limiting variables. This could be an important finding if these relationships are valid for other study areas as well. Also, this study demonstrated that parcel data and NLCD in combination (where available) provide a very valuable set of input data to create fine-scale dasymetric maps.

There are some limitations that will need attention in the near future. First, the correlations between block group level parcel and census attributes indicate that the partitions were sub-optimal and need reconsideration. This observation might be explained to some degree by the subjectivity that can be expected when respondents are answering the census questions on ‘Units in Structure’. Second, parcels are considered here as a spatially more precise data source but have also limitations in detecting residential buildings. If parcels in rural areas have larger spatial extents this aspect becomes important and needs to be addressed. Future steps will include roads and imperviousness surfaces also released as national datasets to further improve the described analysis by extending this research to other study areas.

Acknowledgments

This research is funded by the National Science Foundation: “Collaborative Research: Putting People in Their Place: Constructing a Geography for Census Microdata”, project BCS-0961598 awarded to University of Colorado - Boulder and University of Tennessee - Knoxville. Funding by NSF is gratefully acknowledged.

References

Eicher, C. L. and Brewer, C. A. (2001) Dasymetric Mapping and Areal Interpolation: Implementation and Evaluation. *Cartography and Geographic Information Science* 28, 2, pp. 125-138.

Harvey, J. T. (2002) Estimating census district populations from satellite imagery: Some approaches and limitations. *International Journal of Remote Sensing* 23, pp. 2071–2095.

Leyk, S. Buttenfield, B. P. and Nagle, N. N. (in press) Modeling Ambiguity in Census Microdata Allocations to Improve Demographic Small Area Estimates. *Transactions in Geographic Information Science.*

Li, G. and Weng, Q. (2005) Using Landsat ETM+ Imagery to Measure Population Density in Indianapolis, Indiana, USA. *Photogrammetric Engineering and Remote Sensing* 71, pp. 947–958.

Maantay, J. A. Maroko, A. R. and Herrmann, C. (2007) Mapping Population Distribution in the Urban Environment: The Cadastral-Based Expert Dasymetric System (CEDS)." *Cartography and Geographic Information Science* 34, 2, pp. 77-102.

Mennis, J. (2006). Socioeconomic-Vegetation Relationships in Urban, Residential Land: The Case of Denver, Colorado." *Photogrammetric Engineering and Remote Sensing* 72, 8, pp. 911-921.

Mennis, J. (2009) Dasymetric Mapping for Estimating Population in Small Areas." *Geography Compass* 3, 2, pp. 727-745.

Mennis, J. and Hultgren T. (2006) Intelligent Dasymetric Mapping and its Application to Areal Interpolation. *Cartography and Geographic Information Science* 33, 3, pp. 179-194.

Mennis, J. (2003) Generating Surface Models of Population Using Dasymetric Mapping. *The Professional Geographer* 55, pp. 31–42.

Reibel, M. and Agrawal, A. (2007) Areal Interpolation of Population Counts Using Pre-Classified Land Cover Data. *Population Research and Policy Review* 26, 5, pp. 619-633.

Semenov-Tian-Shansky, B. (1928) Russia: Territory and Population: A Perspective on the 1926 Census. *Geographical Review* 18, 4, pp. 616-640.

Tapp, A. F. (2010) Areal Interpolation and Dasymetric Mapping Methods Using Local Ancillary Data Sources. *Cartography and Geographic Information Science* 37, 3, pp. 215-228.

Tobler, W. (1979) Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association* 74, pp. 519-530.

Wright, J. K. (1936) A Method of Mapping Densities of Population: With Cape Cod as an Example. *Geographical Review* 26, 1, pp. 103-110.

Yuan, Y. Smith, R. M. and Limp, W. F. (1997) Remodeling census population with spatial information from LandSat TM imagery. *Computers, Environment and Urban Systems* 21, pp. 245–258.

Zandbergen, P. A. and Ignizio, D. A. (2010) Comparison of Dasymetric Mapping Techniques for Small-Area Population Estimates. *Cartography and Geographic Information Science* 37, 3, pp. 199-214.

Stefan Leyk, Assistant Professor, Department of Geography, University of Colorado at Boulder, Boulder, CO 80309. Email <stefan.leyk@colorado.edu>

Barbara P. Buttenfield, Professor, Department of Geography, University of Colorado at Boulder, Boulder, CO 80309. Email <buttenfield@colorado.edu>

Nicholas N. Nagle, Assistant Professor, Department of Geography, University of Tennessee, Knoxville, TN 37996. Email <nnagle@utk.edu>

Alexander K. Stum, PhD Student, Department of Geography, University of Colorado at Boulder, Boulder, CO 80309. Email <alexander.stum@colorado.edu>