DOE SunShot Initiative Rooftop Solar Challenge
Sun Rise New England – Open for Business team

Analysis Report

ROOFTOP SOLAR PV ADOPTION PATTERNS 2004-2012: HOTSPOT AND DENSITY ANALYSIS

February 8, 2012

Presented by

Marcello Graziano

Analyst

Connecticut Center for Economic Analysis

University of Connecticut, Storrs, CT

And

William Waite

Manager of Research Projects

Connecticut Center for Economic Analysis

University of Connecticut, Storrs, CT

To

Connecticut Clean Energy Finance & Investment Authority

This material is based upon work supported by the U.S. Department of Energy award number DE-EE0005688.

ROOFTOP SOLAR PV ADOPTION PATTERNS 2004-2012: HOTSPOT AND DENSITY ANALYSIS

EXECUTIVE SUMMARY

This study uses the spatial weighted approach known as "Getis-Ord Statistics" at Block Group level to identify physical and socioeconomic conditions that limit the adoption of residential solar roof-top PV in Connecticut. Our findings show that higher per-thousands adoption rates of residential solar roof-top cluster in rural areas. The Z-scores in highly populated metro-areas increase moving outside the core of each city. Low adoption in inner-cities is not a novelty. Clustering of low economic conditions associated with poverty-related segregation and tenure affect the outcome. However, this study identifies housing density as one of the spatial best-match for explaining the lower rates. This finding may be linked to lack of clear incentives for rent properties and for owners living in multi-family buildings (e.g. prohibition of sub-metering in residential complexes). Further research is required to quantify the relationship between socioeconomic variables and the adoption rates, to understand the role of the housing type and to assess the role of legal barriers to adoption.

The present analysis represents the first step towards a more comprehensive research effort undertaken and will be incorporated as part of a PhD thesis by the authors. The thesis title is '*Adoption of Diffused Renewable Energy Technologies: Patterns and Drivers of Residential Photovoltaic Systems in Connecticut, 2005-2013*'.

ACKWOLEDGMENTS

The authors thank Prof Kenneth Gillingham, Center for Business and the Environment, Yale University, New Haven (CT) for his support and guidance in preparing the dataset. Additionally, the authors express their gratitude to Professor Fred V. Carstensen, Professor of Economics, Department of Finance, School of Business, University of Connecticut, Storrs (CT) for his support as director of the Connecticut Center for Economic Analysis (CCEA).

SPATIAL ANALYSIS: BACKGROUND AND METHODS

The use of spatial distribution analysis is well established in Geography as well as in other disciplines since the 1960s (Hagerstrand, 1967; Brown, 1981 and Cressie, 1993). More recently, Hotspot analysis in has been widely used by crime experts and scholars in Health Economics and Health Geography (Robinson, 2000; Brundson, 2007; Yannakoulis, 2011).

The patterns of adoption and adoption-per-capita of solar PV systems throughout the state of Connecticut will shed some light upon two different sets of explanatory variables:

1) The role played by the structure of housing units and their distribution; and
2) The role played by socio-economic variables.

The present study uses two different methodological approaches:

a) Kernel Density Analysis (KDA), which is an interpolation techniques which forecasts the spatial distribution of point-features over a specified surface using actual observations points; and
b) Getis-Ord Gi* Statistics (GOG), which uses Census Block Groups data to identify hotspots where higher (lower) values cluster non-randomly.[1]

KDA and GOG proceed on two different scales. KDA uses actual observation points to simulate what the distribution would be in those areas where no observation points occur. The final result is a surface where the density of observation points is shown, with higher values where the observation points cluster together. One advantage of KDA lies in the use of actual observed values rather than aggregated data. GOG proceeds differently. The data are aggregated at a specified scale: in the present study, we aggregated the data at Census Block. We selected this scale as it is the smallest scale at which socio-economic have an acceptable error term (U.S. Census, 2010). The algorithm then uses a weighted value to ascertain the concentration value of each observation, or the value the "PV roof-top systems per thousand residents" in the present study. The input data are listed in Appendix I.

The present analysis used ArcGIS 10.1 and the built-in modules for calculating, displaying and testing the results.

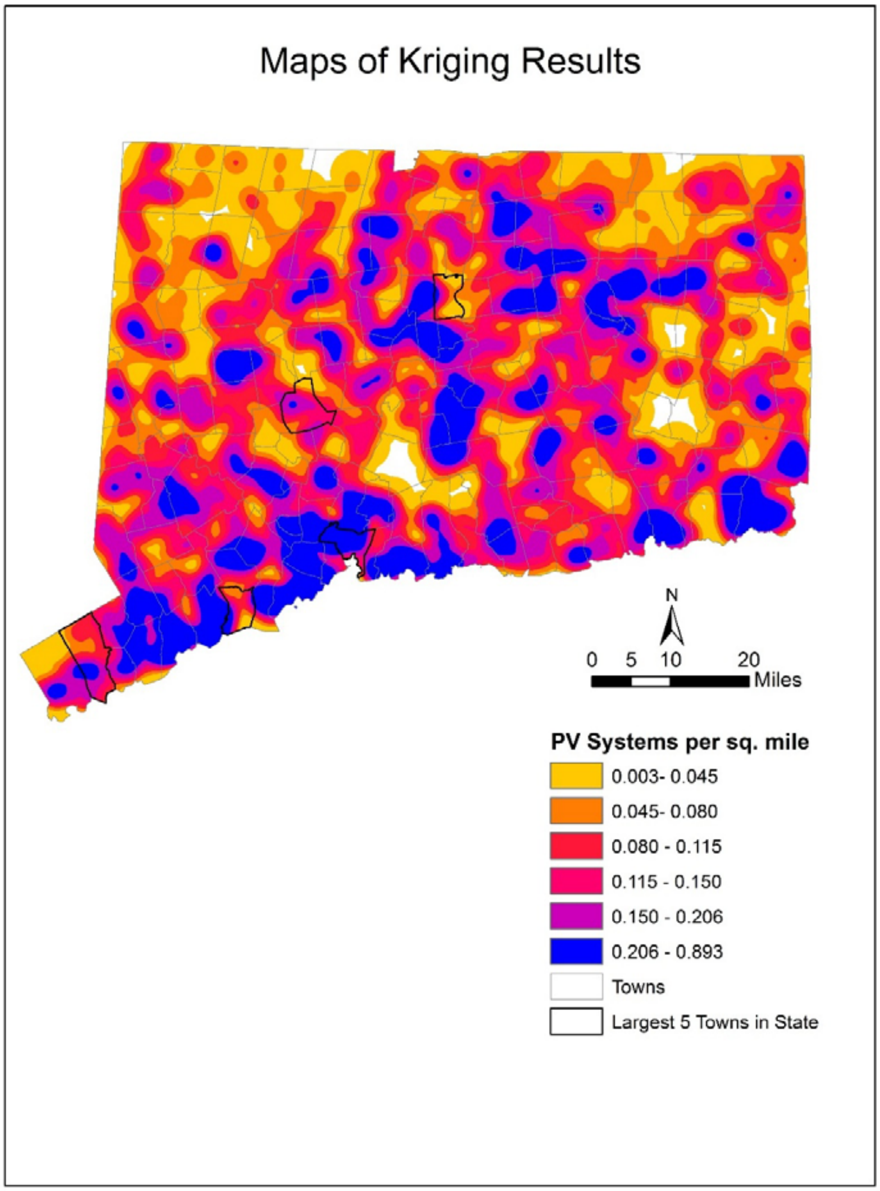
[1] See Appendix II for mathematical formulations.

SPATIAL ANALYSIS: RESULTS

KDO

The figure below shows the result form the KDO analysis.

Figure 1. Map of KDA Results – Cumulative Adopters, 2004-2012



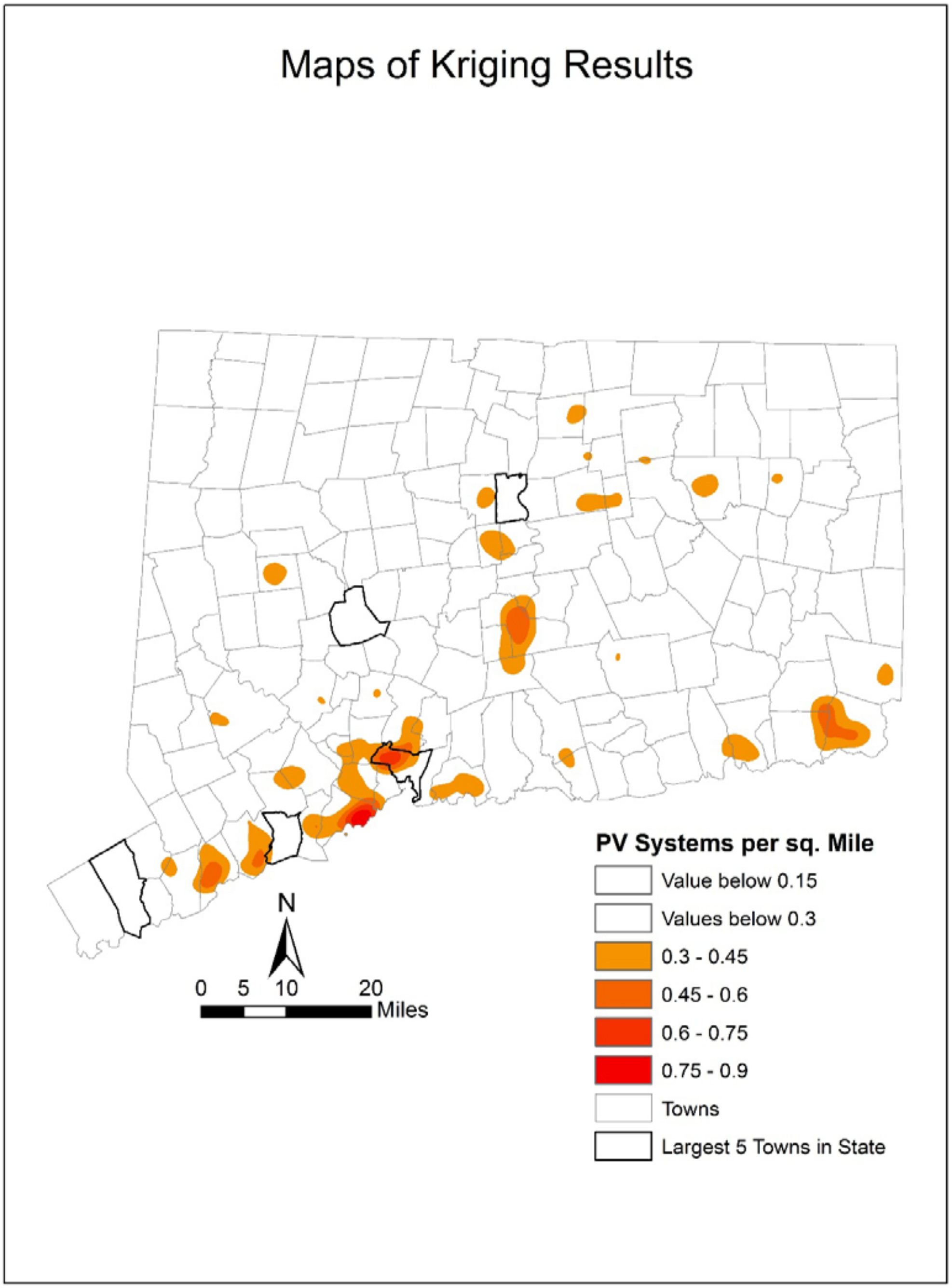
The results are displayed using a quantile classification: as such, each break point contains 16.67% of the total observations. Due to this repartition of data, higher values of density are grouped together in a large class, spreading from 0.2 to 0.8 PV systems per square mile. Looking at the upper-end of the values, only about 6.65% of total results have density values higher than 0.37. These results suggest very low density values in Connecticut. The few higher values are shown in the map below. In this map, break-points have been manually identified at 0.30 to highlight the differentiation among higher values.

KDA has several limitations:

1) It does not weight the results in terms of population, income, population density or any other socioeconomic variable. The results are not surprising in that the highest densities are in highly populated areas along the Connecticut River Valley and in the South-Western coast; and
2) The searching-patterns of the model embedded in Arc 10.0 tend to smooth the values. The surface obtained is commonly constituted by circles around concentration areas very similar to iso-lines.

Due to this limitation, the study adopts GOG in order to identify concentrations of solar panels adopters by population and discover spatial patterns across the state.

Figure 2. Map of KDA Results – Higher Values, 2004-2012



GOG – 2004-2012

The GOG analysis was initially conducted using the perspective of 2012, using the total number of PV systems adopted from 2004 to 2012.

The GOG tool in ArcGIS 10.1 allows the user to define the search bandwidth (cut-off distance) of the model. This distance represents the maximum searching distance around each point beyond which the model will stop searching for neighbors and report a value of zero.[2] The bandwidth value used in the present analysis is 2.92 miles. This value represents the average distance between each ZIP code-5 centroid. This distance represents the median diagonal length of Connecticut ZIP-codes. Literature does not provide us a specific bandwidth beyond which peer-effects fade. However, in Bollinger and Gillingham (2012) find that this effect is present up to the ZIP-code. Because ZIP-codes differ in size across the U.S., we calculated this new value based on the size of Connecticut ZIP-codes. It should be noted that allowing the model to select its own cut-off distance overall results do not differ.

Figure 3 shows the results generated from the GOG analysis. The results displayed are those with significance level of 95% and, accordingly, z-values larger than +1.96 or smaller than -1.96. Negative values show those areas where is larger the concentration of Block groups with low adoption of PV. Positive (> 1.96) values show cluster of block groups with higher adoption per 1,000 people.

[2] The observations points in the present case are the geometric centroids of each Census Block Group.

Figure 3. Map of GOG Results – Adoptions per-1000 people, 2004-2012 Census Block Level

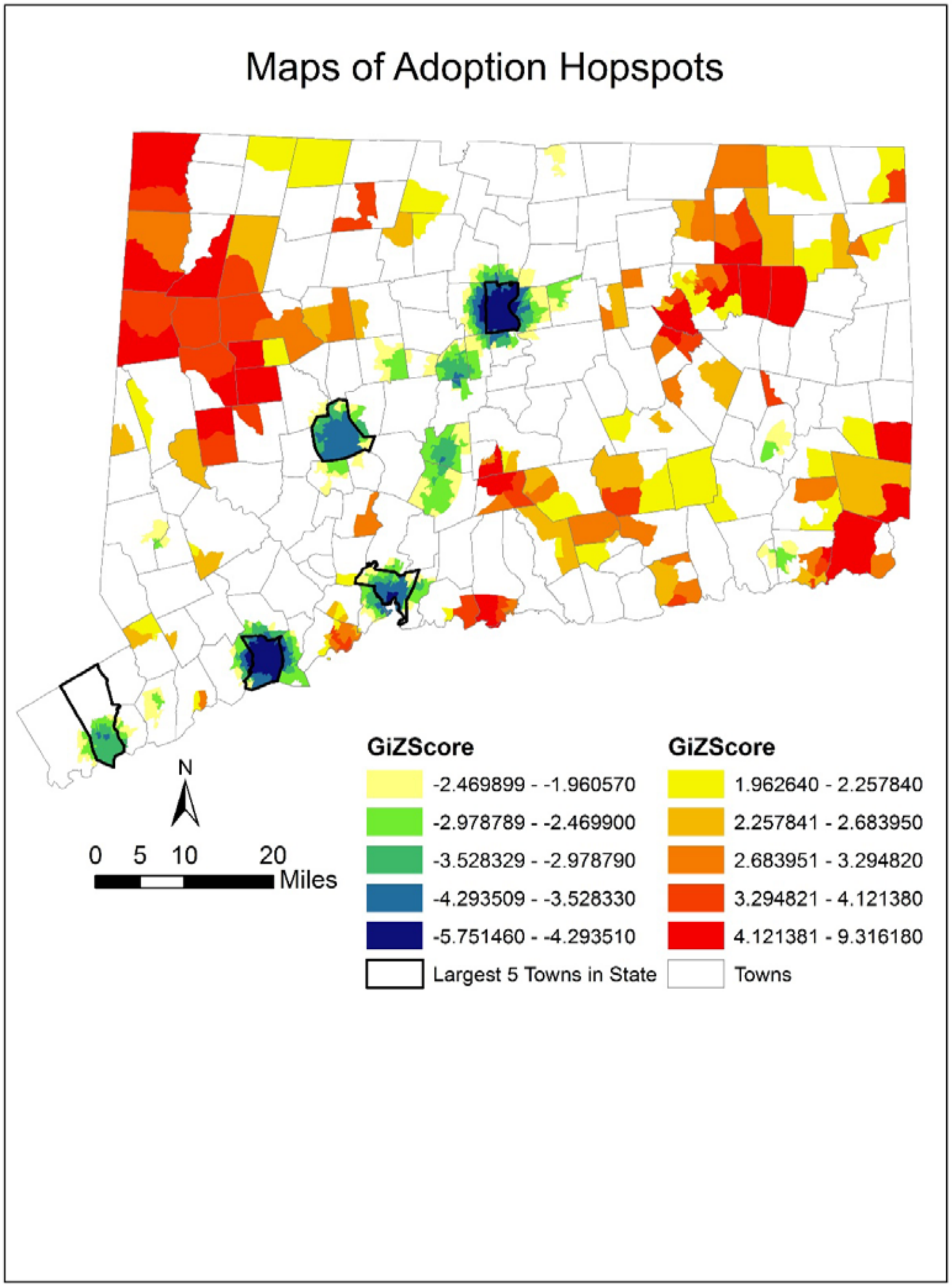
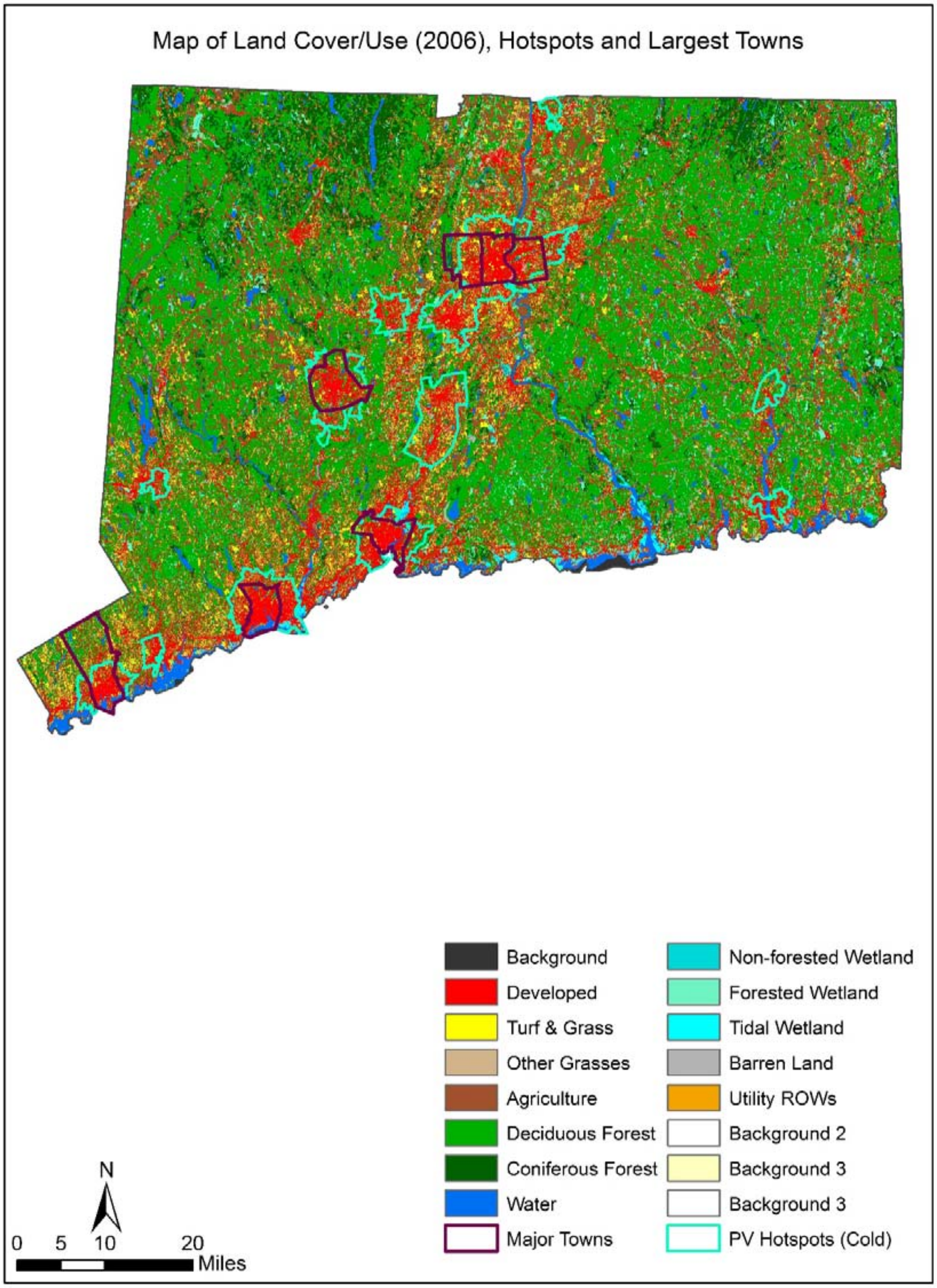


Figure 4. Map of GOG Results and Land Cover/Land Use



From Figures 3 and 4 we can infer that lower adoption-rates cluster together around urban areas, decaying then moving towards the suburbs. This radial decay is common among the five largest towns within the state: Bridgeport (144,229 residents), Hartford (124,755), New Haven (129,779), Stamford (122,643) and Waterbury (110,336; U.S. Census, 2010). These towns account for 17.59% of the state residents, and are shown in red in the map above. It should be noted that the radial decay around these towns goes beyond the towns' border, following what could be considered an extended urban area. For instance, the towns of Hartford, East Hartford and West Hartford are considered as economic cluster, despite having different mayors.[3] Figure 4 allows for an easier understanding of the how more developed (urban) areas overlap with cold adoption hotspots. This latter map shows seven contours for the largest towns, the additional two being East Hartford and West Hartford. These additional two towns will be used in the final part of the present analysis.

Higher values cluster together in rural areas in the north-western and eastern portions of the state. These Census Block Groups are usually considered as "rural". In particular, those Block Groups in within Litchfield County (North West) fall within the definition of "rural" given by five different federal agencies, including the U.S. Department of Agriculture. These results suggest that in those areas were housing (and population) density increase, the consequential multi-family and vertical development of buildings may pose an impairment to adoption. Further, suburban or rural population may be more inclined to adopt PV systems, possibly looking at an off-grid reliable source of electricity among other things. Finally, the demography of inner-cities may come in to play.

It may be suggested that state-wide incentive policies could acknowledge the structural differences and different challenges in these two environments, among others issues related to solar rights and sub-metering (Bronin, 2009 and 2012).

<u>GOG – ADOPTION AND OTHER VARIABLES</u>

The regression analyses previously performed by the research group were aimed at quantifying and showing relationships among the socioeconomic drivers of adoption of PV systems. The present part of this study seeks to explore how the results from the GOG analysis above relate to the socioeconomic variables characterizing the Census Block Groups of Connecticut in general, and those within the largest 5 cities in particular.

The density of solar PV systems in the largest urban areas in the state are particularly low. Local characteristics of the towns in terms of tenure, income, race and house density are thought to affect the decision process of adopters.

[3] http://www.bls.gov/eag/eag.ct_hartford_mn.htm, Accessed on 07/09/2012.

COMPARISON VARIABLES

The present analysis is conducted by comparing the results from adoption concentration and the values in each of the variables listed below.

- **Tenure**. The study uses the share of renters in each Census Block Group based on the American Community Survey (ACS), 5 years average 2006-2010. This variable describes the willingness of renters and landlords to invest in the property. Further, as suggested by Bird and Hernandez (2012) current incentives seldom favor landlords in investing on renewable energy systems.
- **Median Household Income**. The study uses the annual median household income in the previous 12 months. The data come from the ACS, 5 years average 2006-2010 Block Group level. This variable captures the availability of capital to the residents of the Block Groups.
- **Percentage of Self-defined Black in population.** The study uses the share of black population in each Block Group based on ACS, 5 years average 2006-2010. Race and income have long been related. As stated by Massey and Denton (1993), and tested empirically by Quillan (2012), *"[…] racial residential segregation and non-white group poverty rates combine interactively to produce spatially concentrated poverty*". The selection of African-American is based on the considerations by the authors cited above. Further, selecting one race only allows to avoid correlation problems[4].
- **Housing Density.** This variable has been derived using the number of housing units from the ACS, 5 years average 2006-2010 Block Group level and the land area of each Block Group. This variable is a proxy for capturing the housing composition of each block group. The areas with higher housing density will have more multi-family types housing units, thus working as a proxy for housing type and multi-family housing units.

[4] For a review of racial composition and income see **Wilson, W.J. (1987).** *The truly disadvantaged: the inner city, the underclass, and public policy*. Chicago, IL: University of Chicago Press.

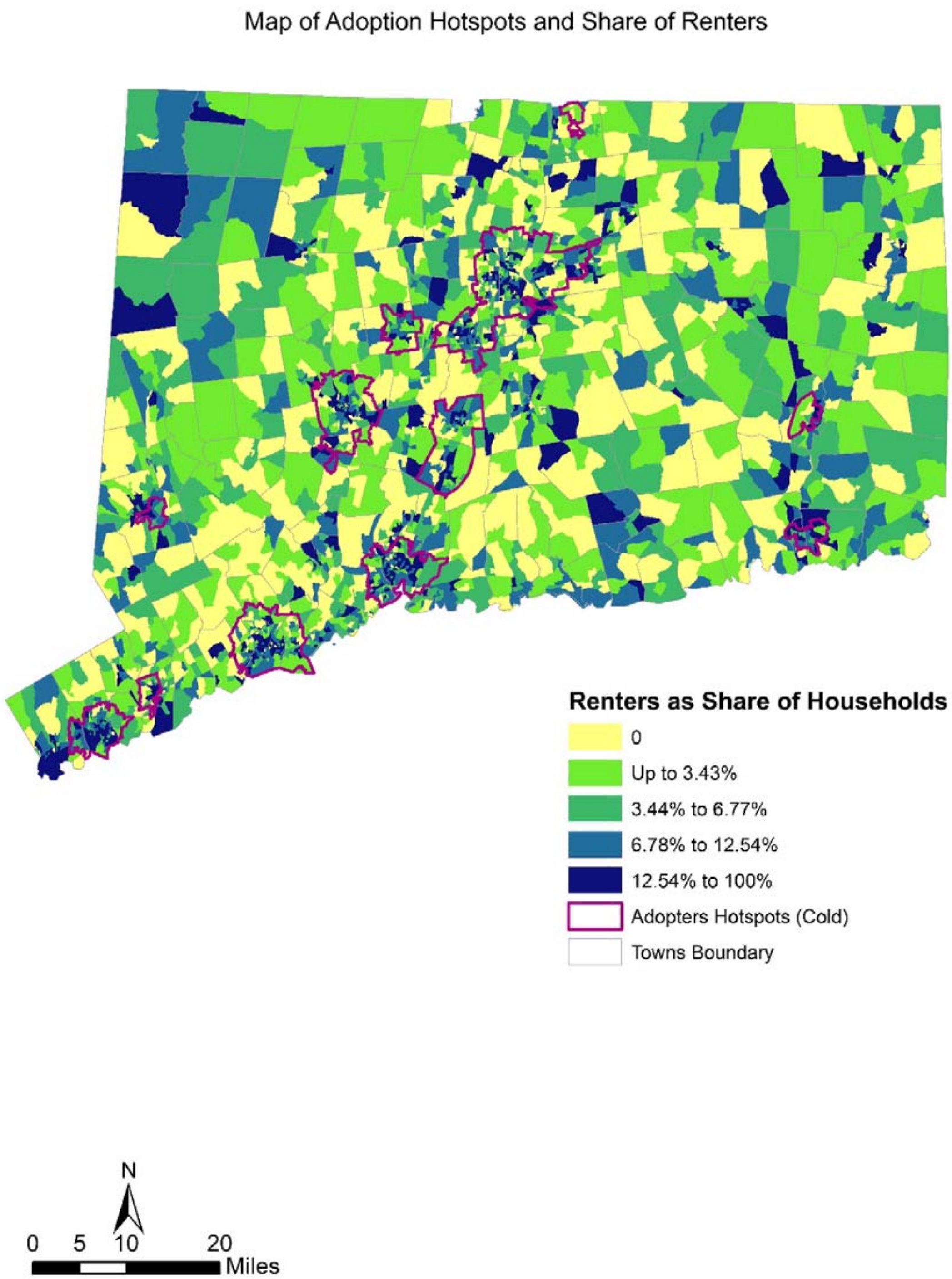
COMPARISON VARIABLES – ADOPTION AND TENURE

Tenure may affect the decision of adoption in multiple ways. Although the vast majority of tenants pay for electricity (88%; Bird and Hernandez, 2012) adopting any renewable energy source may be problematic both for the tenant and the landlord. On the tenant's side, limitations in the contract, the average tenure length and the loss of the capital once the tenure is over may lower the adoption of PV systems. On the other side, landlords may not find it profitable to install PV systems, as the premium price they may charge on the property would not be high enough. Further, problems may arise in multi-family building due to lack of clear solar rights and sub-metering (Bronin, 2009 and 2012).

Figure 5 shows the share of renters and the contours of the cold adoption hotspots. The map suggests that those areas with concentration of lower adoption rates also coincide with areas with higher share of renters. Further, higher adoption rates concentrate in areas where renters are fewer. Looking within the cold hotspots of adoption, the pattern followed by adoption are radial towards the outside of the inner-city areas. In other words, lower values of rent properties are found in the suburbs, that is, in those areas which in Figure 3 showed higher negative z-values.

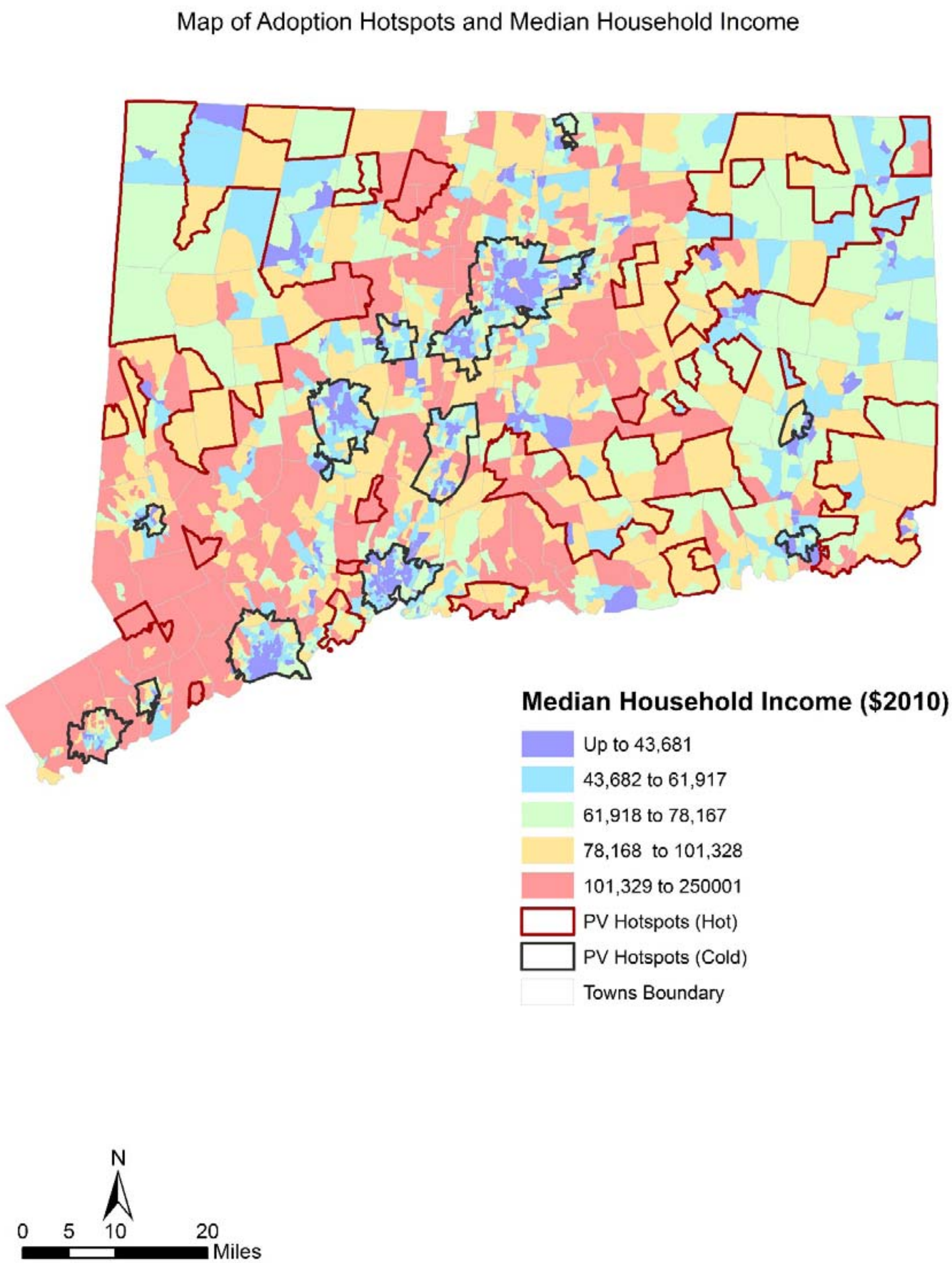
These results suggest that inner-cities tenure may contribute to the lower adoption rates found in those Block Groups.

Figure 5. Map of Hotspots Comparison: Adoption vs. Share of Renters



COMPARISON VARIABLES – ADOPTION AND INCOME

Figure 6. Map of Hotspots Comparison: Adoption vs. Income



The map in the Figure6 is slightly different from the map styles used in in the previous and following sections. Differently from renters and race (see below), it is of interest to look at income hotspots and PV systems hotspots altogether. Once again, with few exceptions, higher Z-scores in adoption rates do not display high concentration rates of wealthier Block Groups. The majority of hot hotspots of adoptions encompass areas Block Groups of the state with income within the upper three quintiles. However, the vast majority falls within the third and fourth quintile, rather than in the highest one. These results suggest that, once reached a certain income level, the decision on adoption depends on other factors, including, tenure and housing type.

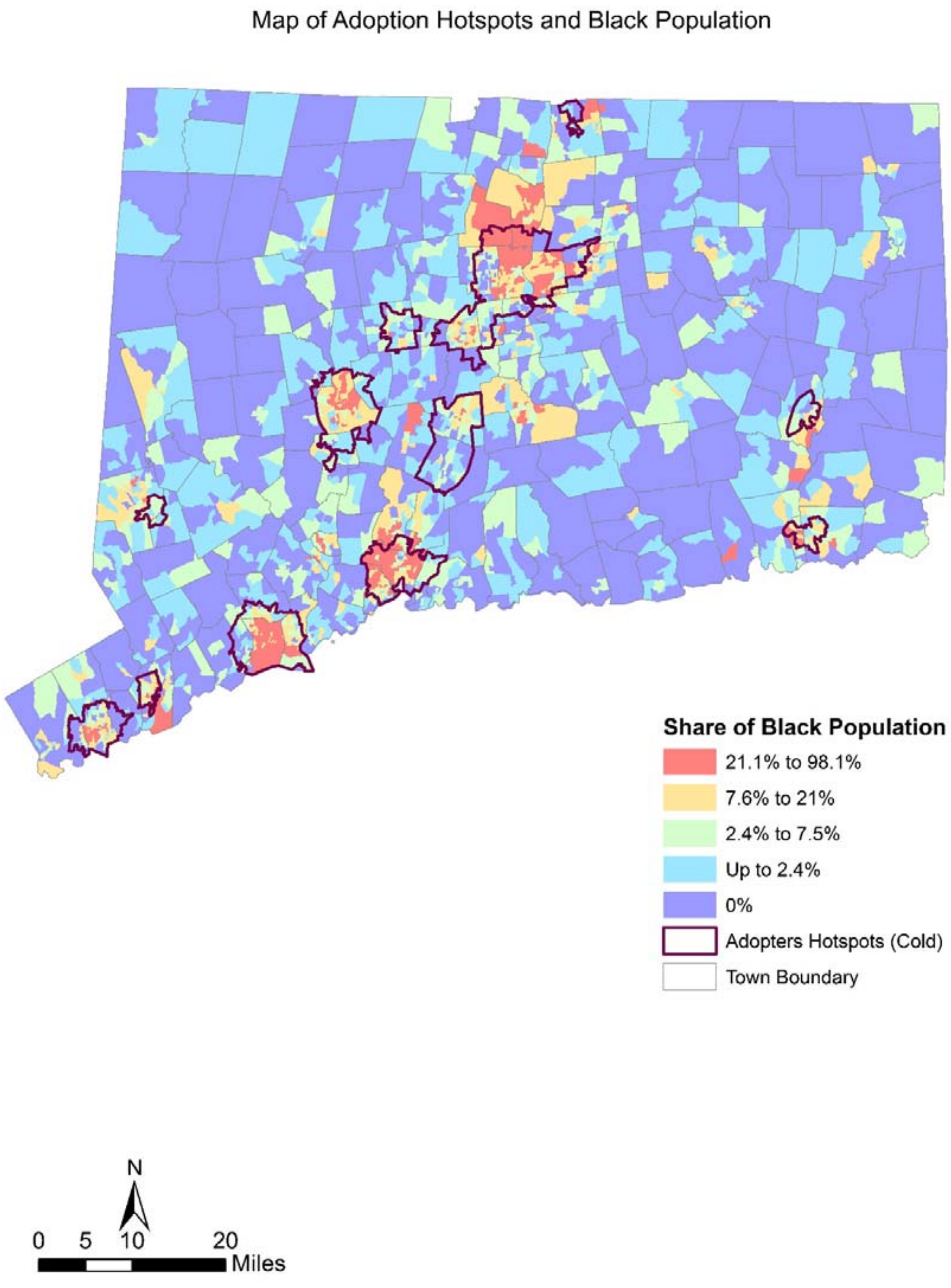
The results for the cold hotpots are quite easier to infer. Each of the cold hotspots polygons has at its core low income Block Groups. Some Higher income values are recorded towards the outside of the polygons, in reflecting the higher Z-scores recorded in the hotspots analysis.

The lower-income profile of inner cities has been long studied (Wilson, 1986 and 1997). Inner-cities, with higher density population density and concentration of lower income residents seem to impose a socio-spatial restriction on the adoption of PV systems.

COMPARISON VARIABLES – ADOPTION AND MINORITIES

Looking at the racial composition of local tenants allows for taking in to account segregation effects and in understanding if there is any connection among adoption, race and income. Figure 8 shows the results of the comparison between the hotspots of PV systems adoption and African Americans as share of population. Other minorities could have been taken in to account. However, African-Americans are currently the one showing the lowest level of median income. As such, understanding the spatial location of where this minority is located is fundamental to understand income and adoption patterns.

Figure 8. Map of Hotspots Comparison: Adoption vs. "Black" Residents[5]

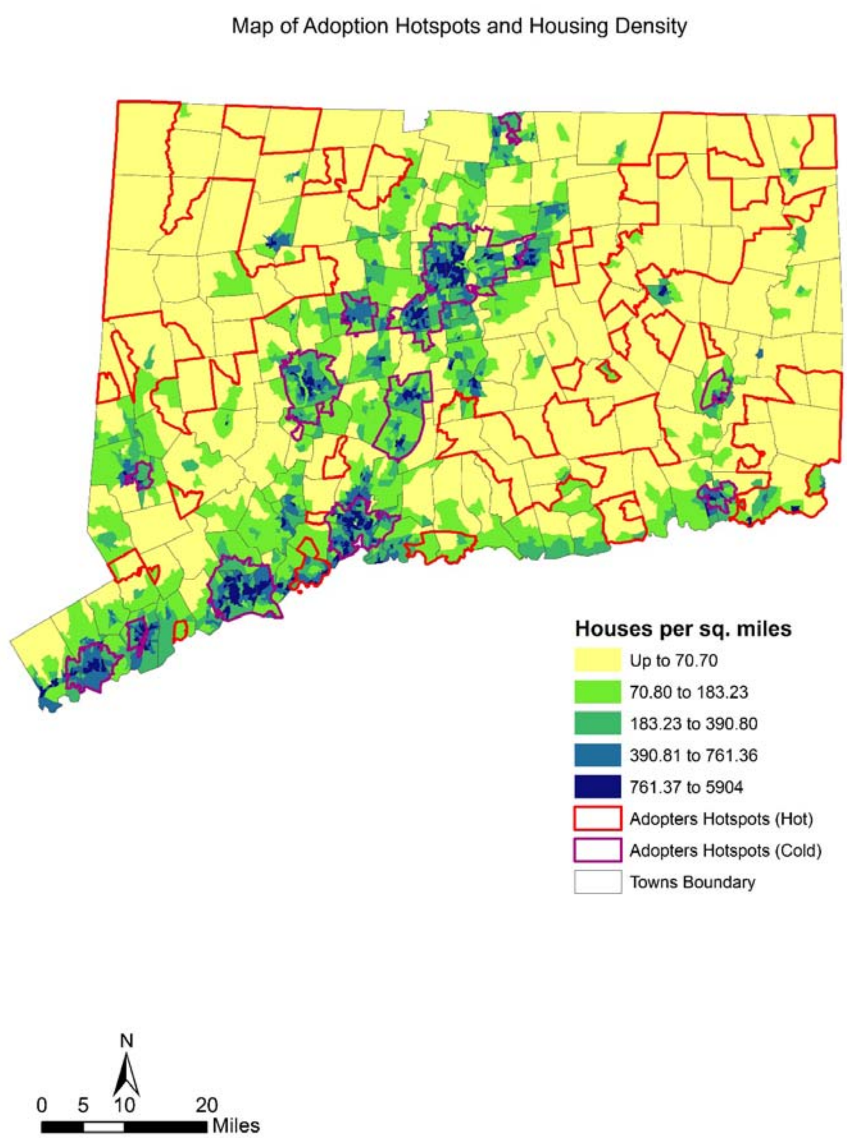


[5] The term "Black" is specific descriptor used by the U.S. Census Bureau in its demographic surveys. In this report we use "Black" and "African American" interchangeable in reference to the same demographic group.

The results in Figure 7 show quite clearly the overlap between hot low adoption rates of PV systems hotspots and high concentration of African Americans. Further, the same results can be compared with those on income and tenure type. It is quite clear that each of the low variables overlap in the major urban areas such as New Haven, Hartford, Bridgeport and Waterbury. The results do not suggest that (self-reported) African Americans are less likely to adopt PV systems. Rather, it shows how low adoption rates, income, race and tenure are strictly spatially related.

COMPARISON VARIABLES – ADOPTION AND HOUSING DENSITY

Figure 8. Map of Hotspots Comparison: Adoption vs. Housing Density



The density of housing units works as a proxy for housing characteristics. The results shown in Figure 8 are quite straightforward. Every set of polygons cold polygons matches areas with very high housing density. Vice versa, almost all the hot hotspots of adoption perfectly overlap areas with low housing density. There, within each set of polygons, exceptions. However, considering the increasing radial pattern of hotspots around larger metro areas it is easy to infer that these higher (lower) values reflect higher (lower) Z-scores.

COMPARISON VARIABLES – CONCLUSIONS

The comparison approach used in the previous sections is based on quantitative and qualitative tools. Despite not directly quantifying the correlation of values among the variables, the comparison approach allows the analyst to seek and follow spatial patterns based on actual observations in the area of study. Several considerations can be made following these results.

1) Inner-city areas in Connecticut display concentrations of low adoption rates per capita. This fact is shown and quantified through the use of the GOG statistics. The five largest towns in the state have radial, increasing Z-score value which overlap almost perfectly with the land cover of the metro areas.
2) Level of minorities, tenure type and housing concentration in the inner-cities explain quite well low concentrations of adoption rates.
3) It appears more problematic to relate higher adoption rates with lower (or higher) concentration values of the above mentioned variables, with income being among the most difficult to interpret. However, this is not true for the GOG performed using adoption rate of PV systems.
4) Higher Z-scores tend to cluster in rural areas with income levels not necessarily belonging to the highest quintile. However, the Block Groups in these areas usually belongs to the third or fourth quintile. Wealthier regions in the South-West part of the state show mixed results, suggesting that other variables other than capital influence the decision process of adoption. Thus, policies aimed at lowering the cost of capital or capital requirements in general for solar PV may prove less effective in wealthier areas where these other factors intervene.
5) The results on the housing density match almost perfectly in the lower-end of the Z-score (lower densities) with those on adoption, thus indicating a relation between building type, occupants' characteristics and adoption rate.

The results stated above indicate that several steps may be considered in the design of policies aimed at spreading the adoption of solar PV roof-top systems. In particular, sub-metering, split incentives and design of solar-rights policies may be considered in order to remove the barriers potential adopters may face, even when capital is available.

FUTURE RESEARCH

The present study represents the first step of a PhD dissertation currently under development. Future research will include:

1. The use of Ripley's K function and fishnetting to analyze the weighted point density of PV systems diffusion. This methodology will allow to better understand the pattern of the diffusion process (Cressie, 1993);

2. The introduction of econometric models to understand what spatial and socio-economic factors influence the adoption of PV systems at state level and at town level. The models used will be Zero-Inflated Poisson regression and fixed-effect panel model;

3. The development of spatial variables to account for spatial peer-effects in the diffusion process of PV system (Brown, 1981); and

4. The development of spatial variables to account for differences in building density and area geography at Block Group level in the models listed at point 2. These variables will control for the influence of the area geography on the diffusion process, particularly in the absence of geography-specific incentives and unclear rights, as suggested by Bronin (2009 and 2012).

LITERATURE CITED

Bollinger, B. and Gillingham, K. (2012). *Peer Effects in the Diffusion of Solar Photovoltaic Panels*. Marketing Science, 31(6):900-912.

Bird, S. and Hernandez, D. (2012). *Policy for the split incentive: Increasing energy efficiency for low-income renters*. Energy Policy, 48:506-514.

Bronin, S.C. (2009). *Solar Rights*. Boston University Law Review, 89:1917.

Bronin, S.C. (2012). Building-related renewable energy and the case of 360 street. Vanderbilt Law Review, 65(6):1875-1934.

Brown, L.A. (1981). *Innovation Diffusion*. New York, NY, USA: Methuen.

Brundson, C. (2007). *Visualizing space and time in crime patterns: A comparison of methods*. Computers, Environment and Urban Systems, 31(1):52-75.

Cressie, N. (1993). *Statistics for Spatial Data*. New York, NY, USA: J.Wiley.

Ord, J.K. and Getis, A. (1995). *Local Spatial Autocorrelation Statistics.* Geographical Analysis. 27(4): 287-306.

Hagestrand, T. (1967). *Innovation diffusion as a spatial process [Innovationsförloppet ur korologisk synpunkt]*. Postscript and translation by Allan Pred; Translated with the assistance of Greta Haag. Chicago: University of Chicago Press.

Johnston (2003). *R.R. Geography & Geographers. Anglo-American Human Geography since 1945*. London: Arnold.

Massey, D.S. and Denton, N. (1993). *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard University Press.

Quillan, L. (2012). *Segregation and Poverty Concentration: The Role of Three Segregations*. American Sociological Review. 77:354.

Robinson, T.P. (2000). *Spatial statistics and geographical information systems in epidemiology and public health*. Advances in Parasitology, 47:81-128.

Silverman, B. W (1986). *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall.

United States Census Bureau (U.S. Census, 2010). *Decennial Census 2010 - Table P.1 Total Population*. U.S. Census.

United States Economic Research Service (U.S. ERS, 2012). *Three rural definitions based on Census places. U.S. Department of Agriculture*. At http://www.ers.usda.gov/datafiles/Rural_Definitions/StateLevel_Maps/CT.pdf, Accessed on 08/09/2012.

Yiannakoulias, N. (2011). *Spatial aberration vs. geographical substance: Representing place in public health surveillance*. Health & Place, 17(6):1242-1248.

Wilson, W.J. (1987). *The truly disadvantaged: the inner city, the underclass, and public policy*. Chicago, IL: University of Chicago Press.

Wilson, J.W. (1996). *When work disappears: the world of the new urban poor*. New York, NY: Knopf.

APPENDIX I: INPUT DATASETS OF ADOPTION GOG

DATA	SOURCE	LINK (if applicable)
Block Group Boundaries with Population	University of Connecticut Mapping and Geographic Center (MAGIC)	http://magic.lib.uconn.edu/, 10/10/2012
State and Town Boundaries	University of Connecticut Mapping and Geographic Center (MAGIC).	Area Boundaries and clipping features
Land Cover/Use Raster	Center for Land Use Education – University of Connecticut	http://clear.uconn.edu/projects/landscape/download2.htm#connecticut, 01/10/2013
Median Household Income, Race composition data, tenure type data and Housing Density	U.S. Census ACS, 5 years average, Block Group Level, 2006-2010	http://magic.lib.uconn.edu/, 08/14/2012
Solar Adopters Data	Elaboration on CEFIA Database, 2004-2012.	N/A

APPENDIX II: GOG and KDA EQUATIONS

GOG

$$G_i(d) = \frac{\sum_j w_{ij}(d)x_j}{\sum_j x_j}$$

"In this equation, the x_j are the weighted values of the points in the study area. w_{ij} is a binary, symmetric weights matrix with ones for all points *j* within distance *d* of point *i* and zeros otherwise.

There are two variants of the local G statistic. The G_i statistic excludes the value at *i* from the summation and is used for spread or diffusion studies, while the G_i* includes the value at *i* in the summation and is most often used for studies of clustering".[6]

GOG as implemented in ArcGIS 10.0 is based on the work of Getis and Ord (1995).

KDA

"Conceptually, a smoothly curved surface is fitted over each point. The surface value is highest at the location of the point and diminishes with increasing distance from the point, reaching zero at the **Search radius** distance from the point. Only a circular neighborhood is possible. The volume under the surface equals the **Population field** value for the point, or 1 if NONE is specified. The density at each output

[6] From http://www.biomedware.com/files/documentation/clusterseer/Getis-Ord_Local_G_Test/Getis-Ord_LocalG__Statistic.htm, Accessed on 08/09/2012.

raster cell is calculated by adding the values of all the kernel surfaces where they overlay the raster cell center. The kernel function is based on the quadratic kernel function described in Silverman (1986; equation 4.5)".