

Rooftop Solar Photovoltaic Technical Potential in the United States: A Detailed Assessment

Pieter Gagnon, Robert Margolis,
Jennifer Melius, Caleb Phillips, and
Ryan Elmore
National Renewable Energy Laboratory

**NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC**

This report is available at no cost from the National Renewable Energy
Laboratory (NREL) at www.nrel.gov/publications.

Technical Report
NREL/TP-6A20-65298
January 2016

Contract No. DE-AC36-08GO28308



Rooftop Solar Photovoltaic Technical Potential in the United States: A Detailed Assessment

Pieter Gagnon, Robert Margolis,
Jennifer Melius, Caleb Phillips, and
Ryan Elmore
National Renewable Energy Laboratory

Prepared under Task No. SS13.1040

**NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC**

This report is available at no cost from the National Renewable Energy
Laboratory (NREL) at www.nrel.gov/publications.

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

Technical Report
NREL/TP-6A20-65298
January 2016

Contract No. DE-AC36-08GO28308

NOTICE

This report was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or any agency thereof.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Available electronically at SciTech Connect <http://www.osti.gov/scitech>

Available for a processing fee to U.S. Department of Energy and its contractors, in paper, from:

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
OSTI <http://www.osti.gov>
Phone: 865.576.8401
Fax: 865.576.5728
Email: reports@osti.gov

Available for sale to the public, in paper, from:

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Road
Alexandria, VA 22312
NTIS <http://www.ntis.gov>
Phone: 800.553.6847 or 703.605.6000
Fax: 703.605.6900
Email: orders@ntis.gov

Cover Photos by Dennis Schroeder: (left to right) NREL 26173, NREL 18302, NREL 19758, NREL 29642, NREL 19795.

NREL prints on paper that contains recycled content.

Acknowledgements

This work was funded by the Solar Energy Technologies Office of the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy under contract number DE-AC36-08GO28308. The authors would like to thank Mike Gleason, Carolyn Davidson, Sean Ong, Rebecca Hott, Steve Wade, Eric Boedecker, Paul Denholm, Donna Heimiller, Michael Bolen, Laura Vimmerstedt, and Paul Donohoo-Vallett for their reviews and input. We would like to thank Adnan Zahoor, Britney Sutcliffe, and Julian Abbott-Whitley for their assistance in processing data. Lastly, we would like to thank Jarett Zuboy for his diligence and attention to detail while editing this report.

List of Acronyms

AC	alternating current
ACS	American Community Survey
ANOVA	analysis of variance
CB ECS	Commercial Building Energy Consumption Survey
DC	direct current
DHS	U.S. Department of Homeland Security
EIA	U.S. Energy Information Administration
GIS	geographic information system(s)
GW	gigawatt
GWh	gigawatt-hour
kW	kilowatt
kWh	kilowatt-hour
Lidar	light detection and ranging
NCES	National Center for Education Statistics
NLCD	National Land Cover Database
NREL	National Renewable Energy Laboratory
PV	photovoltaic
RMSE	root-mean-square error
SAM	System Advisor Model
TMY3	Typical Meteorological Year 3
TWh	terawatt-hour
W	watt
ZIP code	Zoning Improvement Plan code

Executive Summary

This report quantifies the technical potential of photovoltaic (PV) systems deployed on rooftops in the continental United States, estimating how much energy could be generated by installing PV on all suitable roof area. The results do not exclude systems based on their economic performance, and thus they provide an upper bound on potential deployment rather than a prediction of actual deployment.

Although methods have been developed to estimate rooftop PV technical potential at the individual building level, previous estimates at the regional and national levels have lacked a rigorous foundation in geospatial data and statistical analysis. This report helps fill this gap by providing a detailed data-driven analysis of U.S. (national, state, and ZIP-code level) rooftop PV availability and technical electricity-generation potential. First, we use light detection and ranging (lidar) data, geographic information system (GIS) methods, and PV-generation modeling to calculate the suitability of rooftops for hosting PV in 128 cities nationwide—representing approximately 23% of U.S. buildings—and we provide PV-generation results for a subset of these cities. Second, we extend the insights from this analysis of areas covered by lidar data to the entire continental United States. We develop two statistical models—one for small buildings and one for medium and large buildings—that estimate the total amount of roof area suitable for hosting PV systems, and we simulate the productivity of PV modules on the roof area to arrive at the nationwide technical potential for PV.

Our analysis of the trends in the suitability of rooftops for hosting PV systems reveals important variations in this key driver of rooftop PV technical potential that have not been captured by previous approaches. Figure ES-1 shows the results—from our statistical modeling grounded in lidar data—for the percentage of small buildings that are suitable for PV in each ZIP code in the continental United States. In the figure we can identify regional trends in the suitability of small building rooftops, with high densities of suitable small buildings in California, Florida, and the West South Central census division. Such trends are also critical to estimating PV technical potential at finer resolution, and our report illustrates this with a high-resolution analysis of 11 representative cities.

Figure ES-2 shows the annual energy generation potential from rooftop PV as a percentage of each state's electricity sales in 2013. The estimates of energy generation are based on the rooftop suitability of small, medium, and large buildings as well as specific roof orientations, local solar resources, PV system performance assumptions, and building footprints.¹

¹ Because the medium and large building estimates are available only at the state level, the combined results are presented at that level.

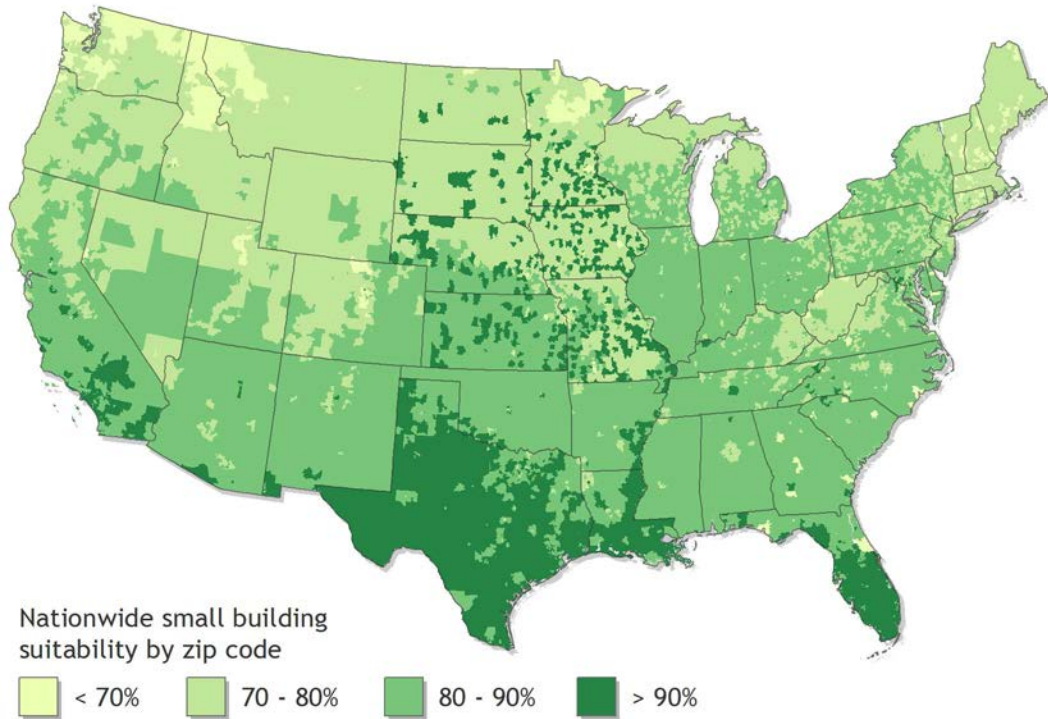


Figure ES-1. Percentage of small buildings suitable for PV in each ZIP code

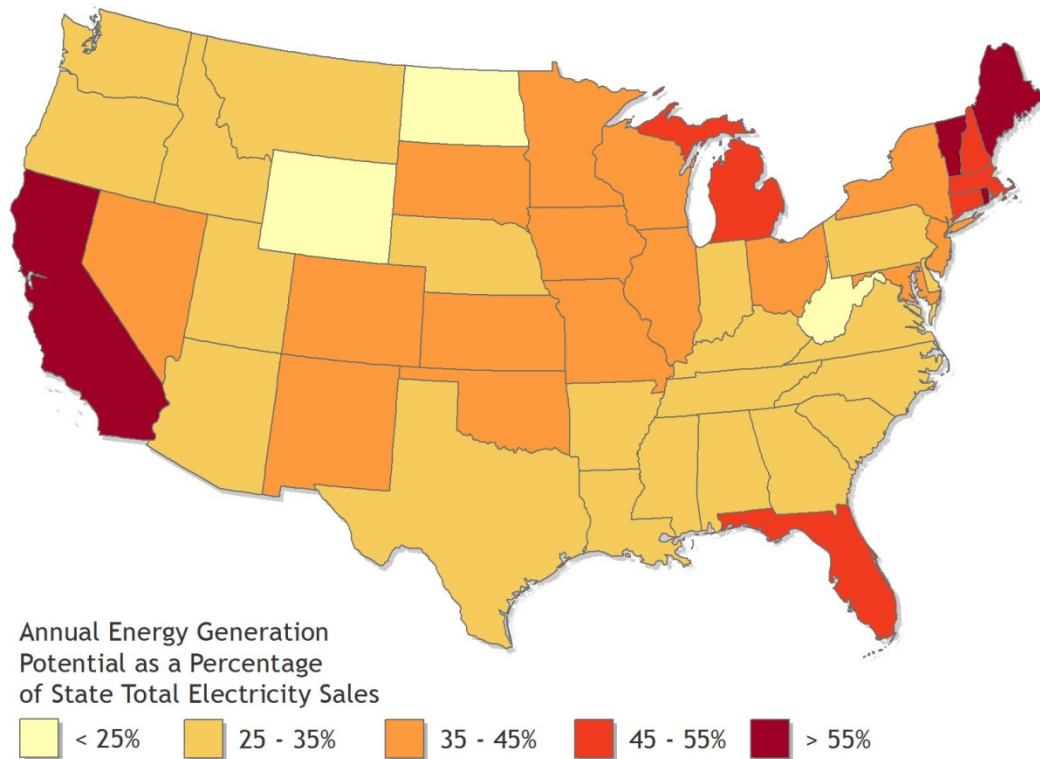


Figure ES-2. Potential rooftop PV annual generation from all buildings as a percentage of each state's total electricity sales in 2013

Figure ES-2 shows that California has the greatest potential to offset electricity use—its rooftop PV could generate 74% of the electricity sold by its utilities in 2013. A cluster of New England states could generate more than 45% because these states' low per-capita electricity consumption offsets their below-average solar resource. Washington, with the lowest population-weighted solar resource in the continental United States, could still generate 27%. Some states with below-average solar resource (such as Minnesota, Maine, New York, and South Dakota) have similar or even greater potential to offset total sales compared to states with higher-quality resource (such as Arizona and Texas).

The difference between Florida and other South Atlantic states illustrates the interplay between variables that affect technical potential. Florida can offset 47% of its total consumption despite having an average household consumption of 130% of the national average. This is largely explained by significantly below-average electricity consumption outside of the residential sector, which makes the state's total per-capita electricity sales slightly lower than the national average. In contrast, the other South Atlantic states range from a potential 23% to 35% of electricity offset owing to lower average rooftop suitability (see Figure ES-1²), slightly lower quality solar resource, and higher per-capita total electricity sales.

Table ES-1 shows our aggregate results.³ The total national technical potential of rooftop PV is 1,118 gigawatts (GW) of installed capacity and 1,432 terawatt-hours (TWh) of annual energy generation. This equates to 39% of total national electric-sector sales, and it is significantly greater than a previous National Renewable Energy Laboratory estimate of 664 GW of installed capacity and 800 TWh of annual energy generation (Denholm and Margolis 2008). The difference can be attributed to increases in module power density, improved estimation of building suitability, higher estimates of the total number of buildings, and improvements in PV performance simulation tools that previously tended to underestimate production.

Although only 26% of the total rooftop area on small buildings (those with a footprint smaller than 5,000 ft²) is suitable for PV deployment, the sheer number of buildings in this class gives small buildings the greatest technical potential. Small building rooftops could accommodate 731 GW of PV capacity and generate 926 TWh/year of PV energy, which represents approximately 65% of rooftop PV's total technical potential. Medium and large buildings have a total installed capacity potential of 386 GW and energy generation potential of 506 TWh/year, which represents approximately 35% of the total technical potential of rooftop PV.

These results are sensitive to assumptions about module performance, which is expected to continue improving over time. For example, this analysis assumed a module efficiency of 16% to represent a mixture of various technology types. If a module efficiency of 20% were assumed instead, which corresponds to current premium systems, each of the technical potential estimates would increase by about 25% above the values stated in this report. Furthermore, our results are only estimates of the potential from existing suitable roof planes, and they do not consider the immense potential of ground-mounted PV. Actual generation from PV in urban areas could exceed these estimates by

² Figure ES-1 shows suitability results for only small buildings because more than 99% of medium and large buildings have at least one roof plane suitable for a PV system.

³ Because the relative magnitudes of the results are a strong function of the square footage used as a cutoff between building classes, these results should not be presented without that information.

installing systems on less suitable roof area, mounting PV on canopies over open spaces such as parking lots, or integrating PV into building facades.

Because our results are estimates of technical potential, they do not consider what would be required to use all the potential PV-generated energy. In practice, the integration of a significant quantity of rooftop PV into the national portfolio of generation capacity would require a flexible grid, supporting infrastructure, and a suite of enabling technologies.

Table ES-1. Estimated Suitable Area and Rooftop PV Technical Potential by Building Class

Building Class (Building Footprint)	Total Suitable Area (Billions of m²)	Installed Capacity Potential (GW)	Annual Generation Potential (TWh/year)	Annual Generation Potential (% of National Sales)
Small (< 5,000 ft ²)	4.92	731	926	25.0%
Medium (5,000–25,000 ft ²)	1.22	154	201	5.4%
Large (> 25,000 ft ²)	1.99	232	305	8.2%
All Buildings	8.13	1,118	1,432	38.6%

Many opportunities exist for expanding on the data, methods, and results from this report. Our base data set can be combined with numerous other data sets—such as insolation data, market data, and end-use electricity consumption data—to provide more precise and actionable information than has been available previously. To that end and to encourage municipalities, utility providers, solar energy researchers, and other PV stakeholders to use the data for their own purposes, we have posted supporting data and documentation of the methods we used to perform our analysis on the NSRDB Data Viewer website.⁴ The models developed here also can be applied generally in areas where lidar data may not be available, both for making estimates of rooftop areas for arbitrary geographies and in stochastic simulations where statistically representative buildings are needed.

⁴ See maps.nrel.gov/pv-rooftop-lidar.

Table of Contents

1	Introduction	1
2	Literature Review	3
3	Processing Lidar Data To Estimate Rooftop Area Suitable for PV Deployment	4
3.1	Input Data.....	5
3.2	Shading.....	7
3.3	Tilt.....	8
3.4	Azimuth.....	8
3.5	Application of Suitability Criteria.....	9
4	Analyzing the PV Potential of Lidar-Covered Cities	11
4.1	Classification of Suitable Roof Area by Orientation and Building Size	11
4.2	Trends in Rooftop Suitability for Different Building Sizes	12
4.3	Trends in Rooftop Tilt and Azimuth	12
4.4	Simulation of PV Productivity on Suitable Rooftop Area	12
4.5	Results for Select Cities: Small Buildings	14
4.6	Results for Select Cities: All Buildings.....	18
5	Modeling to Extend Small-Building PV Suitability and Technical Potential Estimates Nationwide	21
5.1	Modeled Estimates of National Small Building Rooftop Suitability	21
5.2	Estimated National PV Technical Potential from Small Buildings.....	22
5.3	Model Validation.....	27
6	Modeling to Extend Medium- and Large-Building PV Suitability and Potential Estimates Nationwide	30
6.1	Modeled Estimates of Total Developable Area from Medium and Large Building Rooftops Nationwide	30
6.2	Estimated National PV Technical Potential from Medium and Large Buildings	30
7	Estimated National PV Technical Potential for All Buildings	34
8	Conclusions and Future Research Directions	38
8.1	Data for Public Use	39
8.2	Future Work	39
	References	41
	Appendix A: Rooftop Tilt and Azimuth Model	43
	Appendix B: Small Building Suitability Model	48
	B.1 Predictive Regression Model.....	53
	B.2 Spatial Correlation	57
	B.3 Regression Kriging	59
	B.4 Prediction of the Number of Buildings.....	62
	Appendix C: Rooftop Plane Area Model	63
	Appendix D: Combined Suitable Area Model	69

List of Figures

Figure ES-1. Percentage of small buildings suitable for PV in each ZIP code.....	vi
Figure ES-2. Potential rooftop PV annual generation from all buildings as a percentage of each state’s total electricity sales in 2013.....	vi
Figure 1. Types of renewable energy potentials	1
Figure 2. Major steps for determining the suitability of roof area for PV	4
Figure 3. Lidar data coverage	6
Figure 4. Example of hourly shading and sunlight availability	7
Figure 5. Example of average daily hours of sunlight	8
Figure 6. Nine azimuth classifications.....	8
Figure 7. Rooftop azimuths included in final suitable planes data set.....	9
Figure 8. Tilt classes	11
Figure 9. Percent of small buildings with at least one plane suitable for PV by ZIP code in 11 select cities..	16
Figure 10. Small building average relative production for 11 select cities.....	17
Figure 11. Percentage of small buildings suitable for PV in each ZIP code in the continental United States.	22
Figure 12. Average rooftop PV production per small building at the ZIP-code level	23
Figure 13. U.S. solar resource.....	24
Figure 14. Small building average relative production at the state level	25
Figure 15. Histogram of error process for each azimuth and tilt category for small building area predictions	28
Figure 16. Annual rooftop PV generation potential for medium and large buildings	31
Figure 17. Total annual energy generation potential from rooftop PV for all building sizes	37
Figure 18. Potential rooftop PV annual generation from all buildings as a percentage of state total 2013 electricity sales.....	37
Figure A-1. Fraction of suitable rooftop planes in each tilt and azimuth category for small, medium, and large buildings.....	44
Figure A-2. Fraction of flat suitable rooftop planes for small, medium, and large buildings.....	44
Figure A-3. Fraction of small buildings with flat planes grouped by locale type.....	45
Figure A-4. Probability distribution functions for each combination of azimuth and tilt for small buildings.	45
Figure B-1. Distribution of suitable percentage among small buildings (within states).....	48
Figure B-2. Suitable percentage among small buildings (within states) and 95% confidence intervals	48
Figure B-3. Distribution of the percentage of small suitable buildings by ZIP code.....	49
Figure B-4. Percentage of small buildings containing a flat area with at least 10 m ² suitable for PV, by census division and aggregated locale description.....	50
Figure B-5. Census divisions	54
Figure B-6. Spatial autocorrelation of percentage of buildings that are small and suitable for PV, with data aggregated into ZIP codes, both before and after regression fitting	58
Figure B-7. Fitted semivariogram using an exponential semivariogram model (top) and residual error of regression Kriging model (bottom).....	60
Figure B-8. Performance of Regression Kriging model for Colorado.....	61
Figure C-1. Distributions of plane sizes and plane numbers for small buildings	63
Figure C-2. Exponential fits.....	64
Figure C-3. Observed size of suitable planes for buildings with 1–32 planes in lidar data	67
Figure C-4. Distribution of large building suitable areas (total usable rooftop area) in lidar data compared to model-calculated mean (green line) and assumed constant value in prior NREL studies (red line)	67
Figure C-5. Distribution of medium building suitable areas (total usable rooftop area) in lidar data compared to model calculated mean (green line) and assumed constant value from prior NREL studies (red line).....	68

List of Tables

Table ES-1. Estimated Suitable Area and Rooftop PV Technical Potential by Building Class	viii
Table 1. Assumptions for PV Performance Simulations	13
Table 2. Technical Potential of Rooftop PV from all Building Sizes within Boundaries of Cities Completely Covered by Lidar Data	19
Table 3. Estimated Rooftop PV Technical Potential for Small Buildings by State	26
Table 4. Relative Error for Each Tilt/Azimuth Combination ^a	28
Table 5. Estimated Rooftop PV Technical Potential of Medium and Large Buildings by State	32
Table 6. Total Estimated Technical Potential (All Buildings) for Rooftop PV by State	35
Table 7. Nationwide Estimated Suitable Area and Technical Potential by Building Class	38
Table A-1. National Average Tilt/Azimuth Distribution for Planes on Large Buildings	46
Table A-2. National Average Tilt/Azimuth Distribution for Planes on Medium Buildings	46
Table A-3. National Average Tilt/Azimuth Distribution for Planes on Small Buildings	46
Table A-4. Principle Component Loadings for Tilt/Azimuth	47
Table B-1. Suitable Percentage of Small Buildings by Locale/Division	51
Table B-2. Model Tuning and Performance, Final (Reduced) Model Marked with an Asterisk	55
Table B-3. Examples of Outliers in Unweighted Regression	56
Table B-4. ANOVA on Weighted Model	56
Table B-5. ANOVA on Reduced (Final) Model	56
Table B-6. Predictive Accuracy of Predicted Variables for National Potential Model	62
Table C-1. Building Planes: Small Buildings	63
Table C-2. Building Planes: Medium Buildings	65
Table C-3. Building Planes: Large Buildings	66

1 Introduction

Technical potential is a metric that quantifies the generation available from a particular technology in a given region; it considers resource availability and quality, technical system performance, and the physical availability of suitable area for development. An assessment of technical potential provides an established reference point for a renewable technology. For example, in 2012, an analysis by the National Renewable Energy Laboratory (NREL) calculated the technical potential of 10 technologies for the United States (Lopez et al. 2012).

Figure 1 illustrates the relationship between technical potential and three other types of potential often used to discuss energy technologies: resource potential, economic potential, and market potential. Resource potential, the largest of the three types, is the entire amount of energy in a particular form for the region under consideration. Technical potential, which is the topic of this report, estimates how much of that total resource could actually be captured, given physically available area and technology performance without considering economics. Economic potential is then the quantity of the possible generation that would result in a positive return on the investment of constructing the systems (Brown et al. 2015). Lastly, market potential estimates the quantity of energy expected to be generated from the deployment of a technology into the market, considering the impact of factors such as policies, competition with other technologies, and the rate of adoption on the actual deployment of a technology.

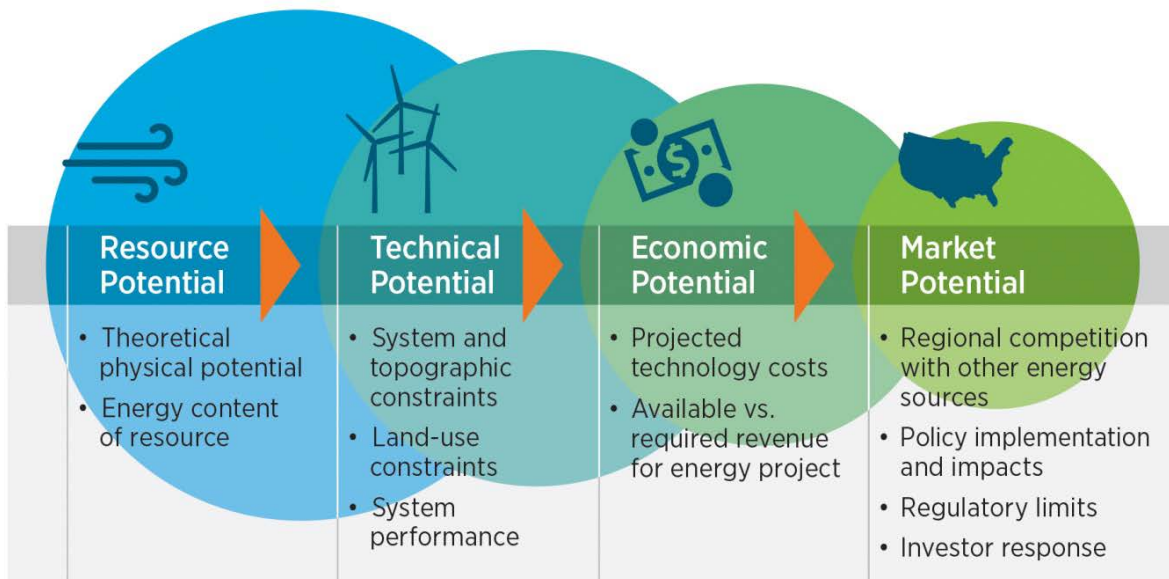


Figure 1. Types of renewable energy potentials

Source: Brown et al. 2015

This report focuses on quantifying the technical potential of photovoltaic (PV) systems deployed on existing suitable roof area in the United States.⁵ Rooftops provide a large expanse of untapped area for solar energy generation, and onsite distributed generation could potentially reduce the costs and losses associated with the transmission and distribution of electricity. Several existing tools and methods can be used to estimate the solar energy potential of a single home or building, and these tools are being actively developed to enable firms to market solar systems to individual building owners. Although there have been some detailed city- and state-level estimates (Bright and Burman 2010; Frantzis et al. 2007; Zhang et al. 2009), a limited number of national-level assessments of rooftop PV technical potential have been published to date (Chaudhari et al. 2004; Denholm and Margolis 2008; Paidipati et al. 2008). Further, most previous estimates have relied on rough engineering rules of thumb in estimating the fraction of rooftop area suitable for hosting PV systems.

To help fill these gaps, we provide a detailed data-driven analysis of U.S. (national, state, and ZIP-code level) rooftop PV suitability and technical electricity-generation potential. First, we use light detection and ranging (lidar) data, geographic information system (GIS) methods, and PV-generation modeling to calculate the PV suitability of rooftops for 128 cities nationwide—representing approximately 23% of U.S. buildings—and we provide PV-generation results for a subset of these cities. Next, we extend the insights from this analysis of areas covered by lidar data to the entire continental United States. We develop two statistical models—one for small buildings and one for medium and large buildings—and we populate them with geographic variables that correlate with rooftop PV suitability. We evaluate the accuracy of the models by comparing model estimates of rooftop suitability with analysis results in areas covered by the lidar data. Finally, we convert rooftop suitability into PV generation potential, sum the estimates for small, medium, and large buildings, and present nationwide results for rooftop PV technical potential. Our results could benefit a broad spectrum of stakeholders, including local municipalities, solar energy researchers, planners, utility companies, and policymakers.

⁵ The technical potential of rooftop PV is the amount of energy that could be generated by installing PV modules on all the roof area suitable for development. In contrast, resource potential would include all solar energy falling on rooftops, economic potential would include all energy generated if all economically viable sites were developed, and market potential would estimate energy generation considering factors such as maximum system sizes set by utilities.

2 Literature Review

A thorough literature review of more than 70 documents by NREL (Melius et al. 2013) identified the existing methods for estimating the fraction of rooftop area suitable for PV deployment. These include three main approaches: constant-value methods, manual selection, and GIS-based methods.

Constant-value methods for estimating rooftop suitability simply assume a certain percentage of building rooftop area is suitable for hosting PV. For example, a series of reports by Navigant and NREL (Chaudhari et al. 2004; Denholm and Margolis 2008; Frantzis et al. 2007; Paidipati et al. 2008) estimated that 22%–27% of residential rooftop area and 60%–65% of commercial rooftop area was suitable for PV. These values were applied to the total building stock to arrive at an estimated area available for PV deployment. This method is easy to use and provides results quickly, which makes it attractive to many researchers. However, it often has had little validation and does not consider specific nuances in the buildings, such as heating, ventilating, and air conditioning systems on large commercial buildings or differences in the tree canopy in neighborhoods with newer housing versus neighborhoods with older housing.

Alternatively, at the finest resolution, manual selection evaluates buildings individually to determine the total suitable area and potential energy generation. In this method, sources such as aerial photography, Google Earth, and NREL's PVWatts[®] Calculator provide visual clues to rooftop PV installation locations (Ordonez et al. 2010; Bright and Burman 2010; Zhang et al. 2009; Anderson et al. 2010). Manual selection provides the most precise estimate of the total rooftop area available for PV, but it is time consuming and cannot be replicated easily on a large scale.

The most robust methods for large-scale estimation of suitable rooftop area are based on GIS. These methods are more precise than constant-value methods, and they can handle much larger data sets than manual selection allows. Melius et al. (2013) provide examples of the many ways GIS-based methods have been used to derive rooftop suitability (e.g., Hofierka and Kanuk 2009; Compagnon 2004; Santos et al. 2011), and they develop and validate a suitability-estimation method based on the best practices found in the literature. We use this method for our analysis of lidar-covered cities, as shown in Sections 3 and 4.

3 Processing Lidar Data To Estimate Rooftop Area Suitable for PV Deployment

In this section, we present an overview of our approach to estimating the rooftop area that is suitable for PV. The major steps are illustrated in Figure 2. The inputs to the analysis are lidar and building footprint data sets. These data are processed to determine the shading, tilt, and azimuth of each rooftop at a horizontal resolution of 1 m². A set of criteria is then applied to determine what roof area is suitable for PV deployment. These results can then be aggregated to determine the total quantity of rooftop area suitable for PV systems at the building, ZIP code, utility service territory, state, or national levels of resolution. The following subsections provide details of each step. Complete documentation of every step can be found on the NSRDB Data Viewer website.⁶

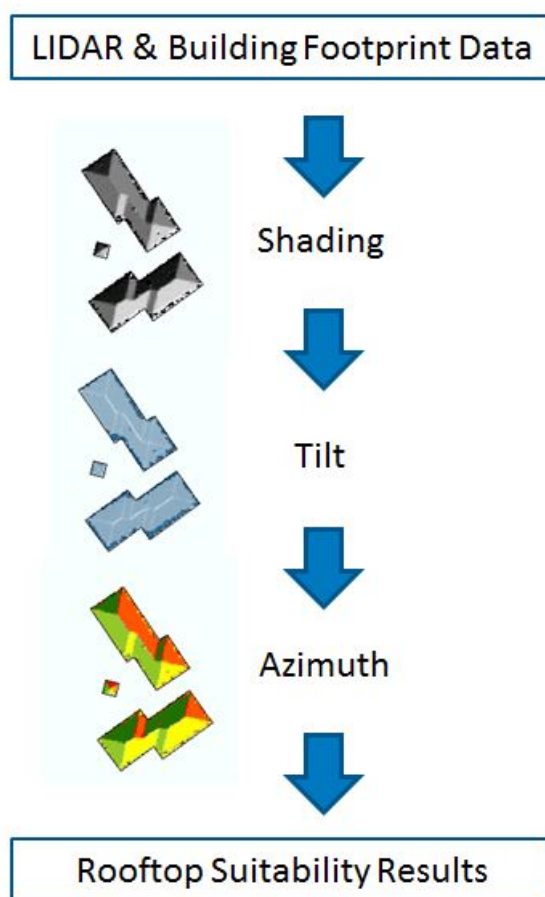


Figure 2. Major steps for determining the suitability of roof area for PV

⁶ The file containing the documentation can be accessed at maps.nrel.gov/pv-rooftop-lidar by selecting the “?” button to the right of the lidar Covered Areas raster listing and following the “Lidar Processing Methods” link at the bottom of the dialogue box that appears.

3.1 Input Data

The lidar data used in our analysis were obtained from the U.S. Department of Homeland Security (DHS) Homeland Security Infrastructure Program for 2006–2014. For each of the 128 cities in the data set, DHS provided (1) lidar data in raster format at 1-m by 1-m resolution and (2) a corresponding polygon shapefile of building footprints. The raster data are based on the reflective surface return (first return) of the lidar data, which correlates to the elevation of the first object detected and creates a digital surface model for each city.

The DHS data set also includes detailed data for about 26.9 million buildings and 7.7 billion m² of rooftop area, or about 23% of U.S. buildings (EIA 2009; EIA 2012). The area covered, shown in Figure 3, represents about 122 million people or 40% of the U.S. population.

Figure 3. Lidar data coverage

3.2 Shading

Our first step in processing the lidar data was to run a shading simulation on the digital surface model of each city.⁷ Figure 4 provides an example of output of the simulation, showing how the shadows move throughout a single day. Seasonal variation in shading was captured by running the simulation for four days: March 21, June 21, September 21, and December 21. The result was the number of hours of sunlight each square meter of roof area received on the simulated days.⁸ The hours of sunlight for the four days were averaged to determine an average number of hours of daily sunlight for each square meter, as shown in Figure 5. We used this metric to exclude roof area that is excessively shaded (see Section 3.5).

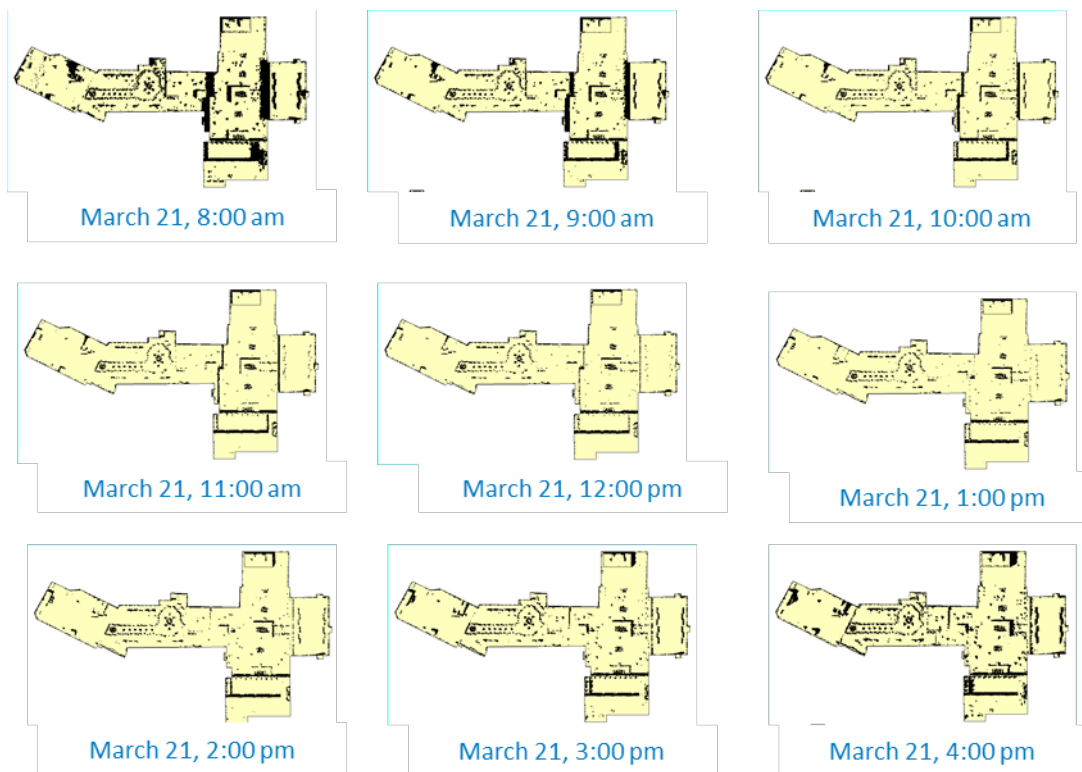


Figure 4. Example of hourly shading and sunlight availability

⁷ The standard ArcGIS Hillshade tool (available in the Spatial Analyst extension, ESRI 2014) was used for the shading simulation.

⁸ For each month, we determined a different threshold of illumination required to classify a cell as being in sunlight; March requires 60% illumination (values > 152), June requires 70% illumination (values > 178), September requires 60% illumination (values > 152), and December requires 50% illumination (values > 127).

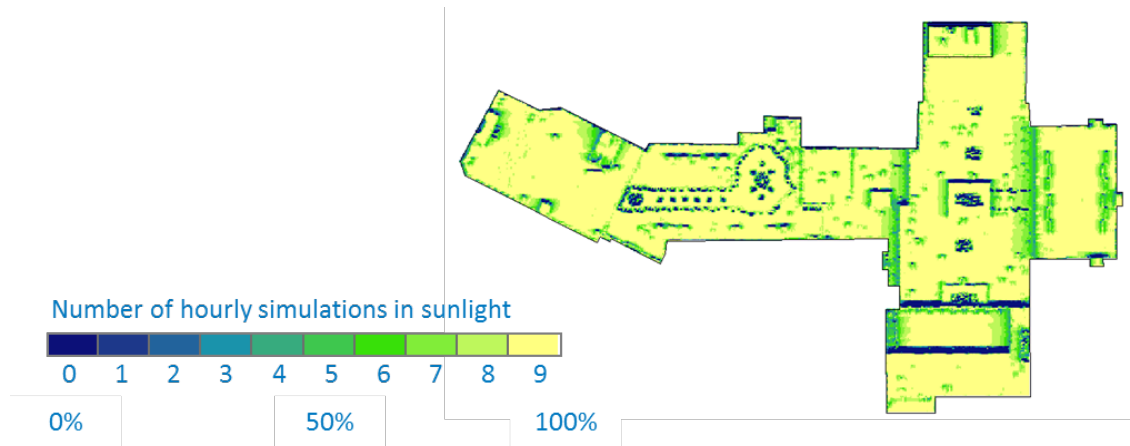


Figure 5. Example of average daily hours of sunlight

3.3 Tilt

The orientation of a roof plane is important for determining its suitability for PV and simulating the productivity of installed modules. We determined the first component of orientation—the tilt—for each square meter of roof area within our lidar data set. To be consistent with many roofers’ and PV installers’ definition of flat roofs, we defined all roof area with a tilt less than 9.5 degrees as “flat”. By comparing actual and predicted hours of sunlight for a subset of roof areas, we determined that our predictions underestimated the illumination of roofs with low angles of tilt. We determined, through iteration, that applying a multiplier of 1.5 to the estimated illumination of all flat roofs compensated for this bias.

3.4 Azimuth

We also determined the second component of roof plane orientation—the azimuth (aspect)—for each square meter of roof area. Each square meter was categorized into one of nine azimuth classes, shown in Figure 6, where tilted roof areas were assigned one of the eight cardinal and primary intercardinal directions; area with a tilt less than 9.5 degrees was classified as flat.

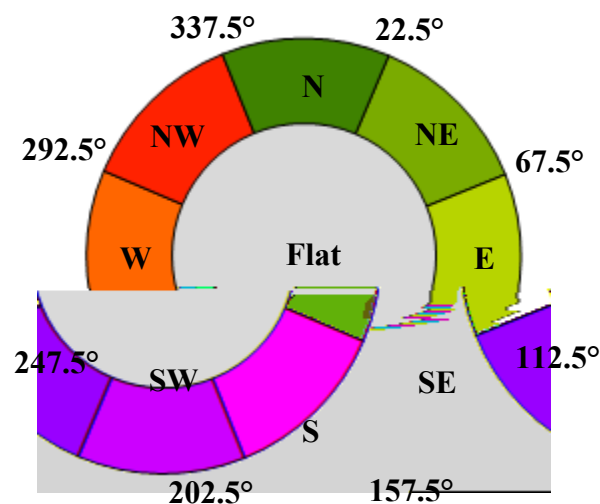


Figure 6. Nine azimuth classifications

The azimuth file was then run through a variety function, which returned the number of different values in the 3×3 neighborhood surrounding each square meter of roof area. Area bordered by more than three unique azimuths was excluded from the data set to remove areas of changing roof orientations and excessively noisy data.

We then used the azimuth values to identify roof planes by assuming contiguous areas of identical azimuth class were a unique plane, and we aggregated each of the individual square meters of roof area into polygons representing contiguous roof planes. For each of the individual roof planes, the ArcGIS Zonal Mean tool was applied to the tilt raster to determine the roof plane's mean tilt. The data set produced through this process consisted of a raster giving a single tilt value for each unique roof plane.

3.5 Application of Suitability Criteria

To determine the total roof area suitable for PV, we excluded any area that did not meet criteria for shading, tilt, azimuth, and a minimum amount of contiguous roof area. To determine what roof area met the shading criteria, a minimum allowable number of hours of sunlight was determined on a regional basis using the System Advisor Model (SAM), which was developed by NREL in collaboration with Sandia National Laboratories and the U.S. Department of Energy. For each city, SAM was used to calculate the number of hours a rooftop would need to be in sunlight to produce 80% of the energy produced by an unshaded system of the same orientation. Roof area that did not meet this shading criterion was excluded.

Roof planes were also excluded based on their orientation. All roof planes facing northwest through northeast (292.5–67.5 degrees; see Figure 7) were considered unsuitable for PV and excluded. All tilt values greater than 60 degrees were removed from the data set, based on the recommendation of PV installers.

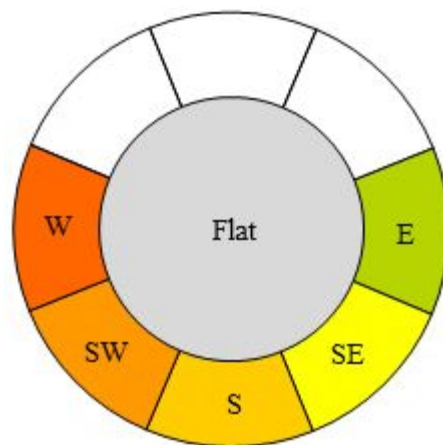


Figure 7. Rooftop azimuths included in final suitable planes data set

We also required a PV-suitable roof to have at least one contiguous plane with a projected horizontal footprint of 10 m² or greater that also meets the shading, tilt, and azimuth criteria. Doing so provides sufficient area to install a 1.5-kW system, assuming a 15%-efficient panel. We selected this minimum size threshold to represent a conservative lower-end estimate of viable PV system sizes, based on current PV performance and historical patterns in reported PV system sizing. Specifically, we reviewed reported system sizes for small PV systems (< 10 kW) through 2013 (Barbose et al. 2014) and determined that 96% of systems in this class were larger than 1.5 kW.

We calculated the area of each suitable roof plane, both as a projected area consistent with the building footprint and as a tilted area to determine the actual amount of developable area. Ultimately, we used the tilted-area values to calculate the installed PV capacity.

The final data set contains the suitable area of every roof plane in the 128 cities covered by lidar data. This data set can be aggregated to the level of a building, ZIP code, utility service territory, state, or any other region to develop summary statistics describing the suitability of the geographic region for rooftop PV.

4 Analyzing the PV Potential of Lidar-Covered Cities

This section describes our process for converting suitable roof area values to PV capacity and generation potential for the 128 cities covered by our lidar data, and it provides sample results from that data set. In Section 5 (small buildings) and Section 6 (medium and large buildings), we apply these insights and the same technical assumptions to produce national PV potential estimates using statistical models informed by the lidar data.

4.1 Classification of Suitable Roof Area by Orientation and Building Size

We categorized each roof plane into one of 21 orientation classes based on its tilt and azimuth. We defined a set of four tilt classes and a “flat” class, as shown in Figure 8. We modeled all roof planes within a given tilt class using a PV panel at the midpoint angle. For example, any roof plane with a tilt value between 47.4 and 60.0 degrees was placed in the same 53.7-degree tilt class. Because the northwest, north, and northeast azimuths were defined as unsuitable, we defined five non-flat azimuth classes (Figure 7).

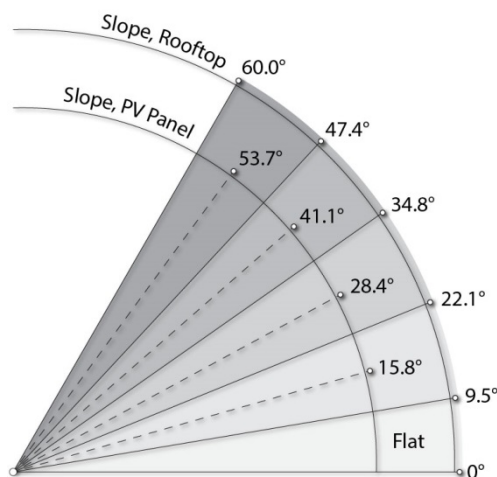


Figure 8. Tilt classes

To better understand the suitability and technical potential of buildings of various sizes, we subdivided all 26.9 million buildings into three classes according to the planar area of their footprints:

- Small: < 5,000 ft² (94% of buildings, 58% of rooftop area in our sample)
- Medium: 5,000–25,000 ft² (5% of buildings, 18% of rooftop area in our sample)
- Large: > 25,000 ft² (1% of buildings, 24% of rooftop area in our sample)

Finally, we aggregated the total suitable roof area in each class and building type to the ZIP-code level. This enabled analysis, predictions, and ultimately the presentation of results at fine spatial resolution.

4.2 Trends in Rooftop Suitability for Different Building Sizes

Based on the lidar data, small buildings—with their more diverse architectures (e.g. roof tilt) and more shadowing from trees and neighboring buildings—show substantially more variability in rooftop PV suitability than do medium and large buildings. Within the 128 cities covered by our lidar data, 83% of small buildings have a suitable PV installation location, but only 26% of the total rooftop area of small buildings is suitable for development.⁹ There is some variability among states, with central and southeastern states showing the greatest fraction of suitable rooftops. Appendix B provides additional detail on rooftop suitability for small buildings.

A very high fraction of medium and large buildings have at least 10 m² of roof area that is suitable for a PV system. Across the lidar data set, more than 99% of large and medium buildings have at least one qualifying roof plane. The percentage of total rooftop area that is suitable is also high relative to small buildings, with 49% suitable for medium buildings and 66% for large buildings. Across all building sizes, 32% of total rooftop area is suitable for PV deployment.

4.3 Trends in Rooftop Tilt and Azimuth

Throughout our entire lidar sample, flat planes are very common on large buildings (93% of planes on large building are flat) and medium buildings (74%) but less common on small buildings (26%). Most other suitable rooftop planes fall into the 28-degree tilt category, and steep rooftops (54 degrees) are an order of magnitude less common than the next category (41 degrees). Azimuths facing east, west, and south are most common, particularly among the 28-degree category of rooftops. These azimuths correspond to the alignment of buildings on a cardinal street grid. These observations appear to hold for small, medium, and large buildings alike.

Large cities have the most flat-roofed small buildings, with the fraction decreasing in more urbanized areas. Large cities also have a more homogeneous tilt/azimuth distribution than do small suburbs. Additional detailed calculations and statistics on roof plane orientation are provided in Appendix A.

4.4 Simulation of PV Productivity on Suitable Rooftop Area

Our next step was to simulate the productivity of PV modules covering the suitable roof area within the 21 different orientation bins for every ZIP code in the lidar data set. These PV performance simulations were executed using SAM (version 2015.1.30). SAM is a performance and economic model designed to facilitate decision making and analysis for renewable energy projects (Gilman and Dobos 2012). It uses hourly meteorological data, a PV performance model, and user-defined assumptions to simulate the technical performance of a solar installation.

The solar resource and meteorological data used for this analysis are from the Typical Meteorological Year 3 (TMY3) data set of the National Solar Radiation Database (Wilcox and

⁹ Because of obstructions, the tilt of a small fraction of roof area within the lidar data set was unknown. Statements about the total percentage of suitable roof area therefore assume the obstructed rooftops follow the same distribution of tilt as the rest of the stock.

Marion 2008). The TMY3 data set includes hourly representative profiles for 1,001 stations throughout the United States. For a given simulation, we used the TMY3 station profile closest to the boundary of the ZIP code under consideration.¹⁰ Because the TMY3 stations are frequently located in or near major cities, the average distance from a ZIP code to a station for the lidar data set was 9 kilometers.

In addition to the solar resource varying geographically, the technical performance of PV systems can vary depending on the equipment used and design choices of the installer. For this analysis, a set of technical assumptions was made to represent the average performance of PV systems as they are being installed in 2015, shown in Table 1. We used these values in SAM, in conjunction with the TMY3 solar resource and meteorological profiles, to determine the electrical output of PV systems.¹¹

Table 1. Assumptions for PV Performance Simulations

PV System Characteristics	Value for Flat Roofs	Value for Tilted Roofs
Tilt	15 degrees	Midpoint of tilt class (Figure 8)
Ratio of module area to roof area	0.70	0.98
Azimuth	180 degrees (south facing)	Midpoint of azimuth class (Figure 7)
Module power density		160 W/m ²
Total system losses		14.08%
Inverter efficiency		96%
DC-to-AC ratio ^a		1.2

^a A system’s direct current to alternating current (DC-to-AC) ratio is the ratio of the nameplate capacity of the PV modules to the AC-rated capacity of the inverters. For example, a system with a DC-to-AC ratio of 1.2 would have 8.33 kW of inverters installed for every 10 kW of nameplate PV capacity.

The power density value used in this analysis corresponds to a module with approximately 16% efficiency. This value is the median module efficiency from approximately 48,000 systems installed during 2014 (Barbose and Darghouth 2015). This value was selected to represent an installed mixture of monocrystalline-silicon, multicrystalline-silicon, and thin-film modules, as opposed to universal installation of premium systems.

The losses from soiling, shading, snow, wiring, and other sources are captured in the total system losses parameter, which was chosen to remain at the SAM default value for this analysis. The inverter efficiency value also remained at the SAM default level. These levels have been selected to be representative of typical systems. A DC-to-AC ratio of 1.2 was selected based on existing literature on the optimal sizing of inverters to minimize the cost of PV-generated electricity (Mondol et al. 2009).

¹⁰ When simulations were performed at state-level resolution—as is described in Section 6 for medium and large building predictions—the PV performance values used were a population-weighted average of the PV performance values for each TMY3 station within the state.

¹¹ Documentation of the mathematical models used by SAM can be found internally within the program, under the “Help” section. For more information, see sam.nrel.gov.

For flat roofs, the ratio of module area to roof area was assumed to be 0.7 to reflect the row spacing necessary to incur only approximately 2.5% losses from self-shading for south-facing modules at 15-degree tilt. For tilted roofs, the value was assumed to be 0.98 to reflect 1.27 cm of spacing between each module for racking clamps.¹²

Using the above assumptions, we ran simulations in SAM to estimate the installed capacity and annual energy generation per square meter of PV systems in the 21 orientations described in Sections 3 and 4. These productivity values were then multiplied by the total suitable roof area, accounting for tilt, within each orientation bin. The results were summed across the 21 orientation bins to arrive at total production values for a given ZIP code. Appendix C provides details about the distribution of rooftop plane sizes in the data.

4.5 Results for Select Cities: Small Buildings

To provide additional detail on the various patterns of PV potential, we mapped results for 11 geographically diverse cities. We chose these cities based on (1) good coverage of the ZIP codes within each city's boundaries and (2) how the cities illustrate the data variation geographically. Figure 9 shows the percentage of small buildings that have suitable roof planes at the ZIP-code level. Only small building suitability is mapped, because over 99% of medium and large buildings have at least one roof plane suitable for PV deployment. Figure 9 and Figure 10 both show the nominal city boundaries, as defined by the U.S. Census Bureau 2013 TIGER/Line Shapefiles.

Figure 9 shows only a weak trend of high building density driving down the suitability of small buildings. Most of the highly developed downtown ZIP codes in the 11 cities have suitability similar to the suitability in other ZIP codes within the city boundaries, although some suburban ZIP codes outside city boundaries do show higher levels of suitability.

Figure 10 shows the *average relative production* of small buildings, which is defined here as the annual electricity generation potential for an average small building as a percentage of the average household annual electricity consumption in that city's state (EIA 2009). Because the national building stock is estimated to contain 78 million single-family households but only 3.2 million commercial buildings with a footprint less than 5,000 ft², Figure 10 can be interpreted as approximately comparing the potential electricity production of an average single-family household in a given ZIP code with the state average household electricity consumption.¹³ This metric should not be confused with the ability of small building PV to offset a state's total electricity sales, which is given in Section 5. Furthermore, because this metric includes buildings unsuitable for PV and presents an average for each ZIP code, it should not be interpreted as

¹² Representative spacing between modules for racking clamps was obtained from a SnapNrack Series 100 UL installation manual, a SunFix Plus Installation Guide, and an IronRidge Roof Mountain System Design Guide. These racking systems are meant to illustrate existing products; mentioning them does not constitute an endorsement.

¹³ Because the consumption value is a state average, it is constant across all ZIP codes for a given city and therefore does not capture household-level variation in consumption that would be driven by socioeconomic status, building size, or other household-specific factors. Therefore, the average relative production value mapped in Figure 10 should only be interpreted as a simple estimation of the potential ability for a group of households in a given ZIP code to offset its consumption.

predicting the productivity of individual buildings, which would vary significantly within each ZIP code.

Figure 10 shows strong regional variation in the average relative production of small buildings. The average productivity of households within a ZIP code is driven by average suitability, household footprint, and solar resource. The average relative production is then also a function of the state average household consumption. For example, high-quality solar resource and low state average household energy consumption lead to a high average relative production for small buildings in Los Angeles. However, the outcome of the interaction among these four drivers is not always obvious. For example, although Colorado has low state average household consumption (7.4 MWh/year or 65% of the national average), low suitability, moderate solar resource, and moderate household sizes lead to generally low average relative production for small buildings within the city boundaries of Denver. In contrast, despite Florida's relatively high state average household consumption of 14.8 MWh/year (130% of the national average) and low state average square footage per housing unit (85% of the national average), Miami's high suitability and good solar resource result in generally high average relative production for small buildings. This demonstrates that one or even two metrics are not sufficient for predicting the ability of aggregations of households to offset their consumption.

Although it is generally understood that a household with adequate roof area can generate greater than 100% of its annual energy consumption with PV, the variation in rooftop suitability and building characteristics makes it less obvious whether that trend holds for groups of buildings. This analysis suggests that in many parts of the United States ZIP-code-sized aggregations of households can collectively generate enough electricity to offset their expected annual consumption. However, notable exceptions include Atlanta and Portland, which have relatively few ZIP codes in which annual energy generation would match expected consumption as estimated by state average household consumption.

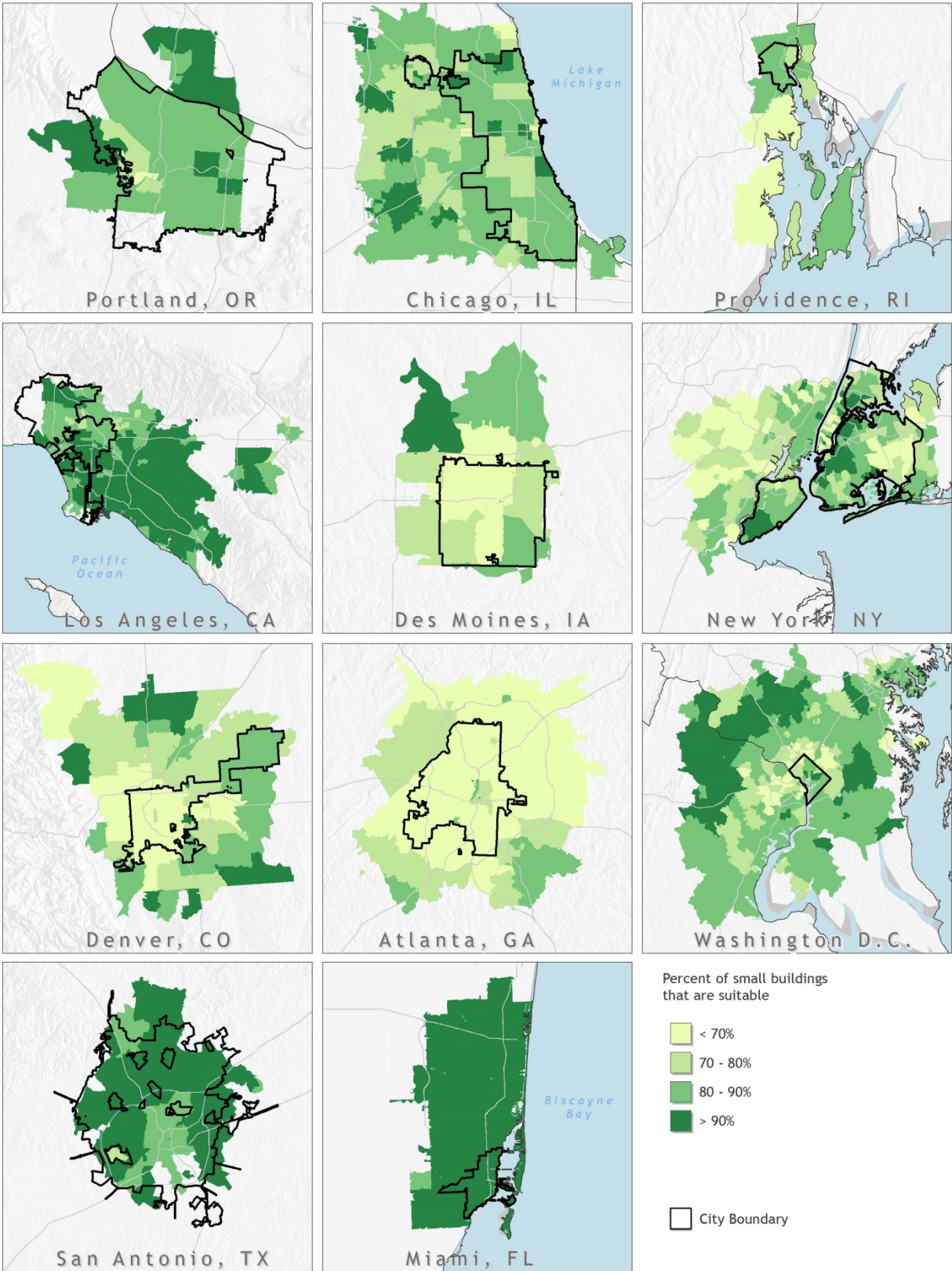


Figure 9. Percent of small buildings with at least one plane suitable for PV by ZIP code in 11 select cities

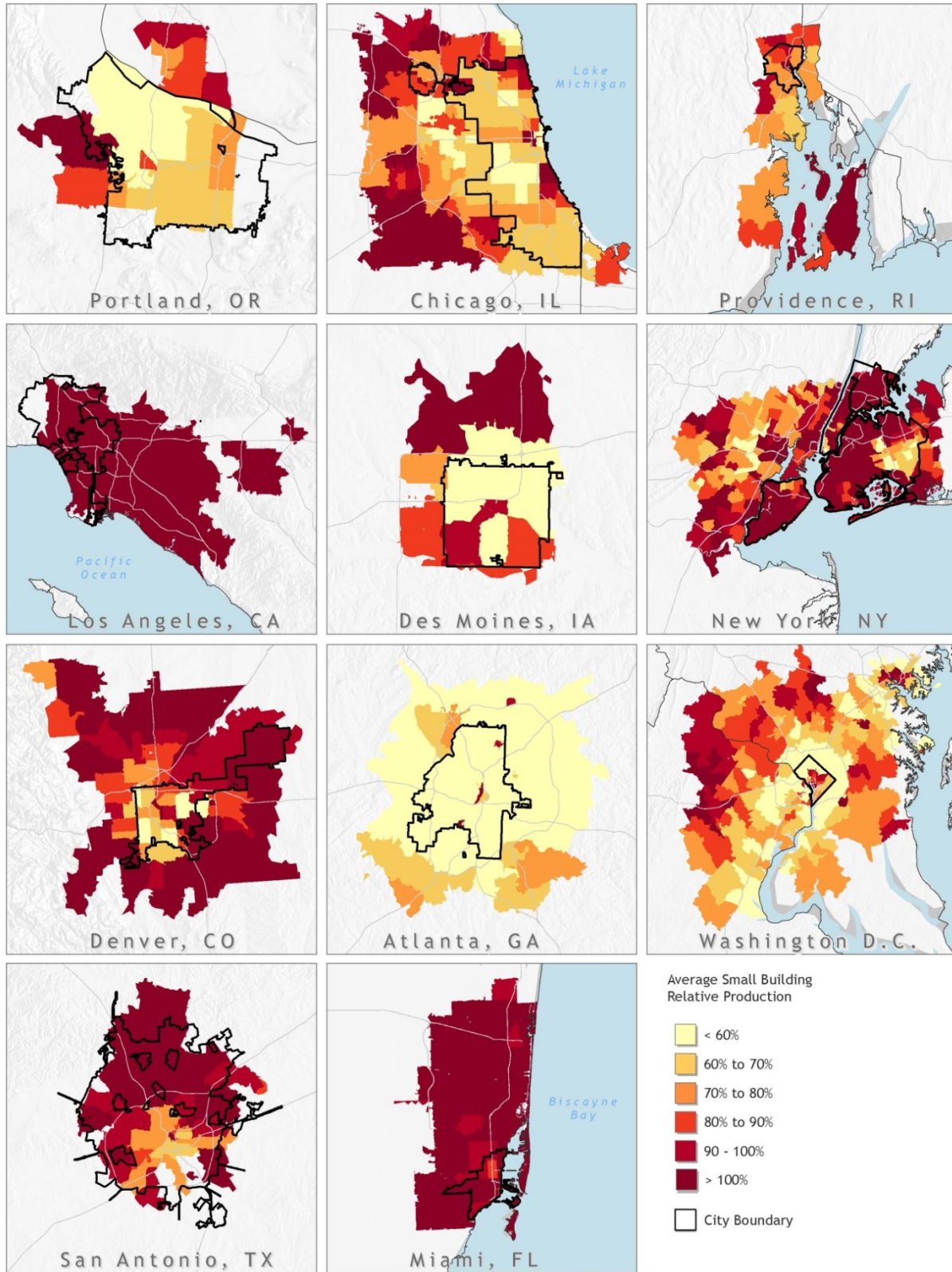


Figure 10. Small building average relative production for 11 select cities
 (average small building PV production / state average household consumption)

4.6 Results for Select Cities: All Buildings

To summarize the technical potential of the lidar regions at an accessible resolution level, the ZIP code results were aggregated for 47 cities whose ZIP codes all have at least some coverage by the DHS lidar data and have ZIP code boundaries that approximately align with city boundaries (U.S. Census Bureau 2013 TIGER/Line Shapefiles). In contrast to the previous section, which only explored data for small buildings, the results here aggregate the productivity of all building sizes. Table 2 gives the estimated total installed capacity and annual energy generation potential for the 47 cities. Many cities have lidar data that extend beyond official city boundaries. Zip codes outside the city boundaries were not included in calculations of the total capacity and energy estimates.

To enable a simple estimation of the ability of these cities to offset their electricity consumption with PV, each state's total electric-industry sales were distributed to its cities by population weight. For example, Florida has 222 TWh of annual sales, and 1.5% of Florida's population lives within the boundaries of Tampa; therefore, the estimated consumption of Tampa is 3.33 TWh. This approximation will overestimate the potential for PV to meet a city's actual consumption for cities that consume more per capita than the state average, and it will underestimate the potential for cities that consume less.

Owing to their size and building density, the cities with the largest potential installed capacity are Los Angeles and New York, with 9.0 GW and 8.6 GW, respectively—illustrating that, even in dense urban areas, shading from buildings does not prevent appreciable installation of PV. Even with large potential capacities in these dense cities, however, PV cannot meet the same percentage of city electricity demand as can be met in some smaller cities. For example, Syracuse and New York City have similar solar resources, but Syracuse can generate 57% of its associated consumption with rooftop PV, whereas New York City can generate only 18%. The total percentage of roof area suitable for PV is similar in the two cities (48% in Syracuse and 46% in New York City), suggesting the difference is driven by low roof area per capita in New York City.

Mission Viejo has relatively high per capita production, driven in part by a low proportion of multi-unit households (which constitute only 15% of total housing units, as compared to an average of 30% throughout the rest of California), resulting in a high quantity of residential roof area per resident. When combined with a relatively low average state per capita consumption and high quality solar resource, the city can generate 88% of its estimated consumption using rooftop PV.

The values in Table 2 for installed capacity and annual generation are the technical potential of existing rooftops that meet our criteria for suitability. The potential to install PV capacity within any of these cities could go beyond these estimates by deploying PV on less suitable roof area, mounting PV on canopies over open spaces such as parking lots, or integrating PV into building facades. Furthermore, continued increases in PV module performance and innovations in racking could increase the technical potential in these cities over time.

Table 2. Technical Potential of Rooftop PV from all Building Sizes within Boundaries of Cities Completely Covered by Lidar Data

City	Installed Capacity Potential (GW)	Annual Generation Potential (GWh/year)	Ability of PV to meet Estimated Consumption
Mission Viejo, CA	0.4	587	88%
Concord, NH	0.2	194	72%
Sacramento, CA	1.5	2,293	71%
Buffalo, NY	1.2	1,399	68%
Columbus, GA	1.1	1,465	62%
Los Angeles, CA	9.0	13,782	60%
Tulsa, OK	2.6	3,590	59%
Tampa, FL	1.4	1,952	59%
Syracuse, NY	0.6	657	57%
Amarillo, TX	0.7	1,084	54%
Charlotte, NC	2.6	3,466	54%
Colorado Springs, CO	1.2	1,862	53%
Denver, CO	2.3	3,271	52%
Carson City, NV	0.2	386	51%
San Antonio, TX	6.2	8,665	51%
San Francisco, CA	1.8	2,684	50%
Little Rock, AR	0.8	1,099	47%
Miami, FL	1.4	1,959	46%
Birmingham, AL	0.9	1,187	46%
St. Louis, MO	1.5	1,922	45%
Cleveland, OH	1.7	1,881	44%
Toledo, OH	1.4	1,666	43%
Providence, RI	0.5	604	42%
Worcester, MA	0.5	643	42%
Atlanta, GA	1.7	2,129	41%
New Orleans, LA	2.1	2,425	39%
Hartford, CT	0.4	404	38%
Baltimore, MD	2.0	2,549	38%
Bridgeport, CT	0.4	435	38%
Detroit, MI	2.6	2,910	38%
Portland, OR	2.6	2,811	38%
Milwaukee, WI	2.1	2,597	38%
Boise, ID	0.5	760	38%

City	Installed Capacity Potential (GW)	Annual Generation Potential (GWh/year)	Ability of PV to meet Estimated Consumption
Des Moines, IA	0.8	1,026	36%
Cincinnati, OH	1.0	1,176	35%
Norfolk, VA	0.8	1,047	35%
Wichita, KS	1.1	1,537	35%
Newark, NJ	0.6	764	33%
Philadelphia, PA	4.3	5,289	30%
Springfield, MA	0.3	370	29%
Chicago, IL	6.9	8,297	29%
St. Paul, MN	0.8	903	27%
Pittsburgh, PA	0.9	907	27%
Minneapolis, MN	1.0	1,246	26%
Charleston, SC	0.3	407	25%
New York, NY	8.6	10,742	18%
Washington, DC	1.3	1,660	16%

5 Modeling to Extend Small-Building PV Suitability and Technical Potential Estimates Nationwide

In the previous section, we presented results for regions where we have lidar coverage. The remaining sections show how we leveraged these data to build predictive models and estimate PV's technical potential for larger regions where lidar coverage is incomplete or non-existent. The overall model, described in its complete form in Appendix D, is composed of three major component models:

- Azimuth-tilt (Appendix A)
- Rooftop suitability (Appendix B)
- Rooftop plane area (Appendix C)

These models predict average and upper and lower bound estimates for suitable rooftop area. Upper and lower estimates include the combined model uncertainty at the 95% confidence level. This section exclusively presents the results for small buildings, predicted at the ZIP-code level for the entire nation. The observation that, within the regions covered by our lidar data, 58% of all roof area came from small buildings, suggests that small buildings contribute significantly to the total national potential of rooftop PV. Medium and large buildings are discussed in Section 6.

Previous estimations of U.S. rooftop PV technical potential focused on estimating results at a regional or national level (Denholm and Margolis 2008). However, here we find significant variation in the percentage of small buildings that are suitable across the various census divisions and locale types. Our predictive model leverages this variability—along with additional attributes such as ground cover classification, population density, and spatial characteristics—to tune our predictions at fine resolution.

5.1 Modeled Estimates of National Small Building Rooftop Suitability

We used our statistical model and building-count data from the 2011 U.S. Census American Community Survey (ACS) to calculate the percentage of small buildings that are suitable for PV by ZIP code throughout the continental United States. The resulting map is presented in Figure 11. Where actual lidar data exist (as shown in Figure 3), the map shows actual data.¹⁴

Developable area for rooftop PV is, by its nature, highly correlated geographically with population. Most potential for PV energy generation is condensed in the relatively small fraction of the country's land space that is developed. National maps such as the one shown in Figure 11, therefore, overemphasize the weight of rural regions if used to approximate visually the impact of a given statistic on the magnitude of total national rooftop PV potential. Nonetheless, such maps can be useful for observing broad geographic trends. For example, the low suitability of northern Minnesota has little impact on the state's total technical potential, but it does illustrate the effect of heavy forestation on rooftop suitability.

¹⁴ For the small fraction of ZIP codes with no predictor information, the model assumed the predictor of the ZIP code's nearest neighbor. This assumption had little impact on the results, because the ZIP codes with missing predictor information covered less than 0.01% of the country's population.

Figure 11 confirms the existence of meaningful regional trends in small building suitability, as suggested previously in our detailed city maps (Figure 9). The highest densities of high-suitability ZIP codes are in southern California, Florida, Louisiana, and Texas. The percentage of small buildings that are suitable tends to be higher in regions without significant tree canopy coverage; for example, the relatively unforested southeast portion of Washington has a higher percentage of suitability than the northeastern region of the state.

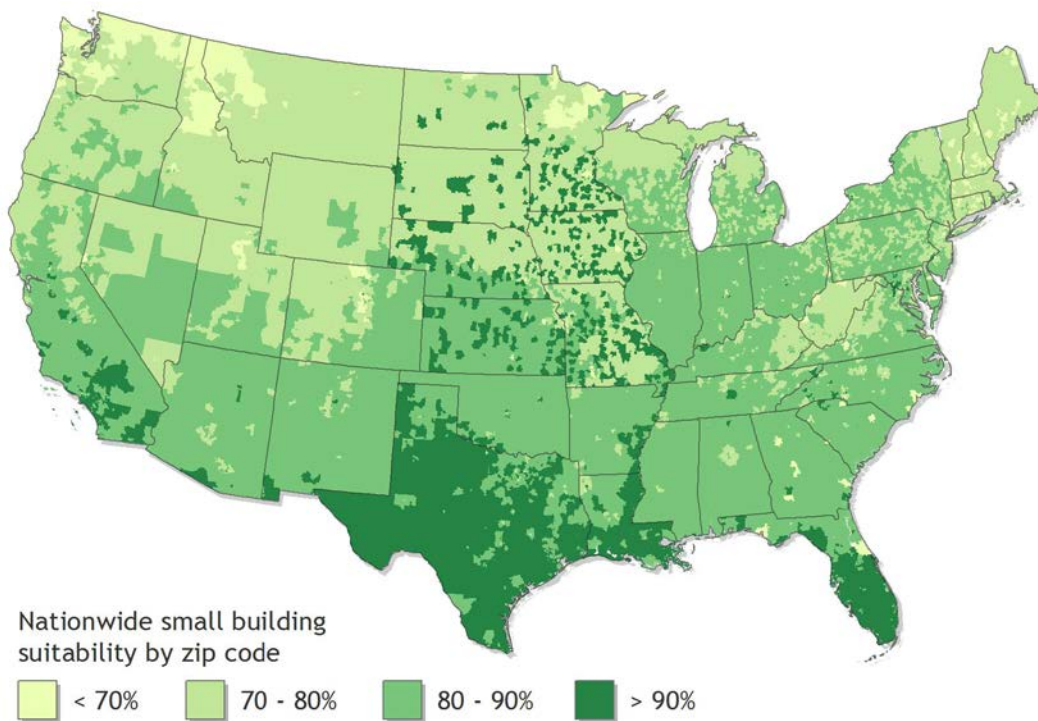


Figure 11. Percentage of small buildings suitable for PV in each ZIP code in the continental United States

The results shown in Figure 11 are used as a key input in calculating rooftop PV technical potential. Furthermore, the figure demonstrates the importance of using a high spatial resolution for small buildings when estimating rooftop PV technical potential. Capturing the considerable variation in rooftop suitability is critical for understanding how rooftop PV potential varies on local, regional, and national levels. Previous estimates of rooftop PV technical potential typically used a single national value or a few regional values to estimate suitability nationwide. Figure 11 represents a significant step forward in understanding the suitability of small building rooftops for PV deployment at fine geographic resolution throughout the United States.

5.2 Estimated National PV Technical Potential from Small Buildings

We used the complete model described in Appendix D—which combines statistical models for rooftop suitability, plane size, tilt, and azimuth—to estimate the total amount of suitable rooftop area in each of the 21 orientation classes for small buildings, on a ZIP-code level across the country. The same technical assumptions and simulation approach described in Section 4 for the lidar-covered cities were used to estimate the installed-capacity and energy-generation potential of modules installed on the predicted small building rooftop area nationwide.

Figure 12 shows the average estimated rooftop PV production per small building at the ZIP-code level.¹⁵ For comparison, Figure 13 shows the simulated energy generation from generic hypothetical PV panels tilted at latitude, illustrating the varying intensity of the U.S. solar resource. Broadly speaking, average small building production strongly correlates with the solar resource; however, there exists significant local variation driven by average household footprint and suitability. For example, the simulated average production in Florida is 12,100 kWh/year per small building (130% of the national average), owing to an above-average solar resource, but it ranges from 5,300 kWh/year to 30,100 kWh/year on a ZIP-code level because of variation in suitability and building footprint.

Differences in suitability can drive differences in total productivity between regions with similar solar resource. For example, lower suitability in the South Atlantic states (see Figure 11) leads to lower average small building productivity than on the Florida peninsula, despite a solar resource of similar quality.

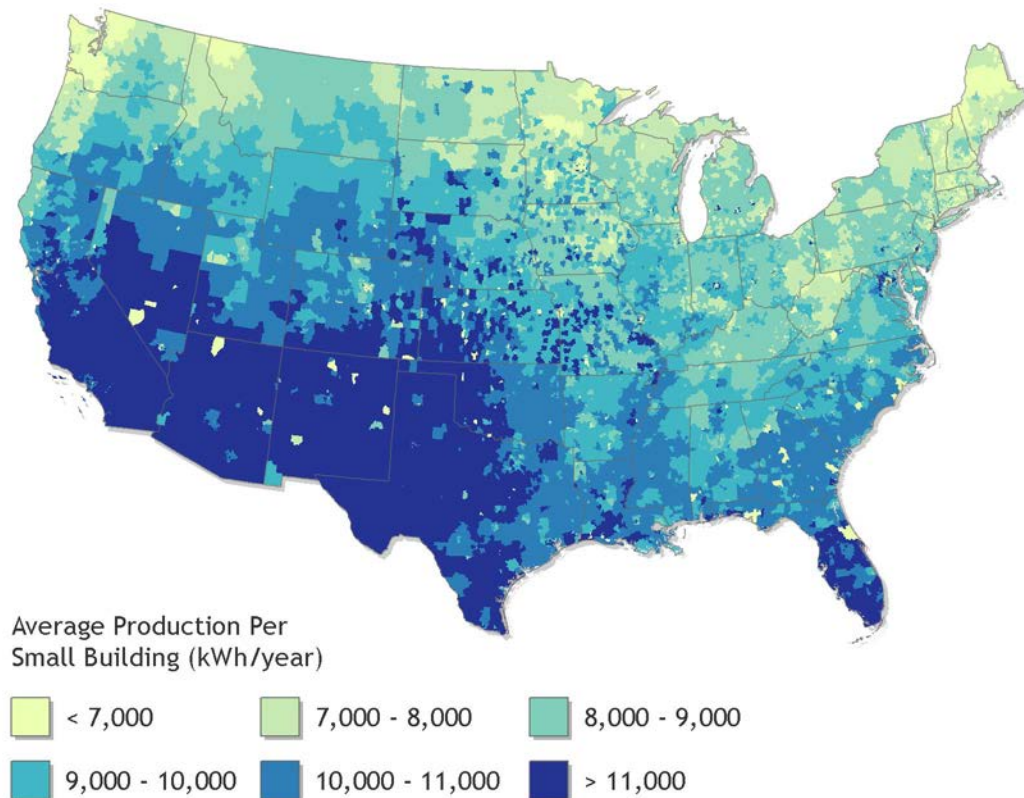


Figure 12. Average rooftop PV production per small building at the ZIP-code level

¹⁵ This average includes non-suitable small buildings. For example, if half of the buildings in a given ZIP code had an annual PV generation of 10,000 kWh/year, and the other half were completely unsuitable for PV deployment, the average small building PV generation potential of that ZIP code would be 5,000 kWh/year.

Figure 13. U.S. solar resource

Figure 14 shows the average relative production of small buildings at the state level, which is considered here to be the annual rooftop PV generation of an average small building as a percentage of each state's average annual household consumption. This metric was considered previously for the 11 cities shown in Figure 10.

These results show that a relatively poor solar resource does not preclude the residential sector from offsetting a significant percentage of its consumption. An average small building across all of New England's states except Rhode Island could generate greater than 90% of the electricity consumed by an average household in the region. This is driven by the low average household consumption of 8,011 kWh/year in the region (70% of the national average), which is due in part to high use of natural gas and oil for heating as well as relatively low summer cooling requirements.

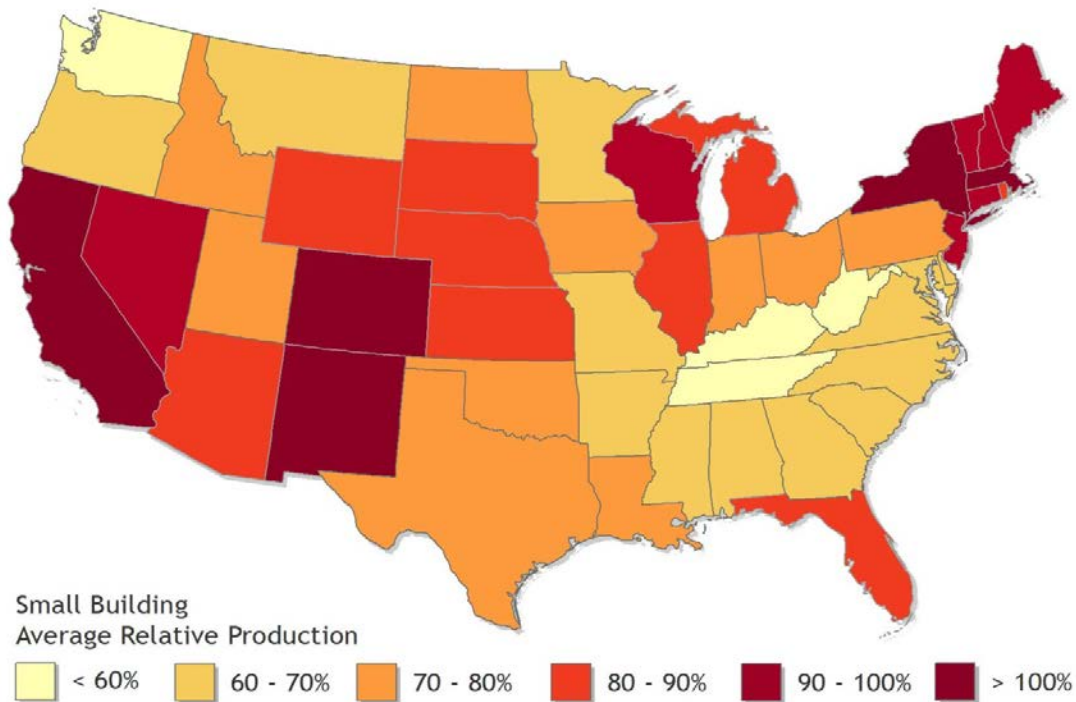


Figure 14. Small building average relative production at the state level

(average small building PV production / state average household consumption)

Table 3 gives each state’s small-building installed capacity potential, annual energy generation potential, and total area suitable for PV deployment. Also given is the potential electricity generation of small buildings as a percentage of each state’s total electricity sales from the electric industry in 2013.¹⁶ For example, Colorado has an estimated potential to offset 27% of its total electricity sales with rooftop PV generation from small buildings. This differs from the average relative production of small buildings that was shown in Figure 14, and it instead presents their ability to offset their state’s total sales of electricity.

Mirroring the trend in small-building average relative production in Figure 14, Table 3 shows that California, New Mexico, Florida, Arizona, and several New England states can offset the highest percentage of their states’ total electricity sales. In contrast, states such as North Dakota and Wyoming, which have moderate average relative production from their small buildings, have some of the lowest potential to offset total state sales, presumably owing to an above-average proportion of their electricity consumption happening outside the residential sector. Nineteen states can generate 25% or more of their annual electric sales with rooftop PV, and 40 states can generate 20% or more.

¹⁶ The quantity of electricity sold was obtained from the U.S. Energy Information Administration’s (EIA’s) “Retail Sales of Electricity (Megawatthours) by State by Sector by Provider” data file (EIA 2014). The quantities are calculated from utility companies’ responses to Form EIA-861 and estimates derived from Form EIA-861S. The quantities, therefore, closely approximate total demand, because they represent all generation except the small quantity of electricity generated on site by consumer-owned systems. Data from 2013 were used because 2013 was the most recent year with data available.

Table 3. Estimated Rooftop PV Technical Potential for Small Buildings by State

State	Annual Generation Potential (% of sales)	Installed Capacity Potential (GW)	Annual Generation Potential (TWh/year)	Total Roof Area Suitable for PV Deployment (millions of m²)
California	43.6%	76.8	114.0	525.5
Vermont	40.3%	2.0	2.3	13.6
Maine	40.0%	4.2	4.7	28.0
New Mexico	32.8%	4.6	7.6	32.6
New Hampshire	32.4%	3.2	3.6	21.5
Arizona	31.6%	15.0	23.9	103.1
Rhode Island	31.2%	2.1	2.4	14.0
Michigan	30.6%	28.3	31.5	189.2
Florida	30.3%	50.3	67.3	343.4
South Dakota	29.5%	2.9	3.6	19.2
Missouri	28.1%	18.8	23.5	126.3
Connecticut	27.6%	7.2	8.2	48.2
Wisconsin	27.6%	16.3	19.0	109.0
Oklahoma	27.4%	12.2	16.4	80.9
Kansas	27.3%	8.2	10.9	55.3
Colorado	27.1%	10.0	14.5	67.4
Montana	26.7%	3.0	3.7	20.1
Massachusetts	25.5%	12.3	14.1	82.8
Utah	25.1%	5.4	7.7	36.2
New Jersey	24.9%	15.6	18.6	104.6
Iowa	24.9%	9.8	11.6	65.8
New York	24.8%	31.3	36.7	210.0
Illinois	23.6%	28.4	33.5	192.4
North Carolina	23.5%	23.9	30.6	160.1
Oregon	23.4%	9.7	11.2	65.3
Pennsylvania	23.2%	29.6	33.9	198.8
Ohio	23.1%	31.0	34.7	206.5
Minnesota	23.0%	13.9	15.8	92.7
Arkansas	22.5%	8.3	10.5	55.7
Idaho	22.2%	4.0	5.4	26.7
Delaware	22.1%	2.0	2.5	13.7
Texas	22.0%	62.7	83.2	424.6
Tennessee	22.0%	17.0	21.3	114.6

State	Annual Generation Potential (% of sales)	Installed Capacity Potential (GW)	Annual Generation Potential (TWh/year)	Total Roof Area Suitable for PV Deployment (millions of m ²)
Nevada	21.6%	4.8	7.6	32.9
Georgia	21.6%	22.4	28.1	149.6
Maryland	21.4%	10.9	13.3	72.1
Virginia	20.6%	18.3	22.8	121.6
Mississippi	20.5%	7.8	10.0	51.7
Indiana	20.3%	18.3	21.4	122.1
Nebraska	19.9%	4.8	6.1	32.1
Alabama	19.1%	13.2	16.8	88.1
South Carolina	18.9%	11.4	14.9	76.2
Louisiana	18.3%	12.5	15.7	83.5
North Dakota	18.3%	2.4	2.9	16.2
Washington	18.2%	15.7	16.9	105.6
West Virginia	17.3%	4.8	5.4	32.3
Kentucky	16.0%	11.6	13.6	77.6
Wyoming	12.5%	1.5	2.1	10.2
Washington, DC	4.2%	0.4	0.5	2.8
Continental U.S. Total	25.0%	731.1	926.4	4,922.3

5.3 Model Validation

To characterize the accuracy of the small building predictions, we carried out a validation experiment for ZIP codes that have lidar data. We trained the complete model on the 3,312 ZIP codes with 90% or greater lidar coverage, predicted the results from these same ZIP codes using the fitted model, and analyzed the difference between the predicted and actual values. This validation experiment is a best-case performance method, because it is predicting the same data on which it is trained and cannot be used to estimate the combined error in areas where ground-truth data are not available. This model validation is also limited to small buildings, because key input data for large and medium buildings are only available at a census division scale; thus, large and medium buildings do not have sufficiently fine-resolution input data to enable a comparison to ZIP-code-level aggregated lidar data.

Figure 15 shows the error in predictions of suitable area for each azimuth and tilt combination. The black bars in the histogram correspond with differences between the mean prediction and the actual observed value. The smaller red bars are size of error outside of the predicted range (lower and upper predictions) when the actual observed value is above our upper prediction or below our lower prediction. The 54-degree tilt category has been removed, because it is sparsely represented in the data. The bars are skewed to the left, because the predicted range is a conservative estimate and generally under-predicts the true value.

Figure 15. Histogram of error process for each azimuth and tilt category for small building area predictions

Black bars are the mean prediction error. Red (lighter) bars are the error outside of the estimated range (lower/upper).

The average error per ZIP code in each tilt and azimuth category is 317 m² and is approximately normally distributed, with some skew to the left. The average error per ZIP code across all tilt and azimuth combinations is 6,603 m². Larger aggregations come closer to the true values owing to a balance in under- and over-predictions (as seen visually by the symmetry in Figure 15). The total error for all 3,312 ZIP codes is 21.9 km², which is a relative error of 2.55%. Table 4 gives relative error percentage for each tilt/azimuth combination.

Table 4. Relative Error for Each Tilt/Azimuth Combination^a

Region	15°	28°	41°	54°
East	27.3%	1.6%	-1.7%	-149.7%
Southeast	18.4%	-6.3%	-11.7%	-93.1%
South	25.2%	1.8%	-4.8%	-62.4%
Southwest	19.3%	-6.2%	-12.4%	-101.2%
West	28.9%	-0.1%	-1.6%	-131.4%

^a The error for flat roofs is 4%.

From the table, it is clear that some tilt/azimuth combinations come much closer to the true values than others. The flat area case and several other cases have less than 5% relative error, including tilt of 28 degrees and azimuth of east, south, and west—and tilt of 41 degrees and azimuth of south, east, and west. The errors for the 54-degree tilt category appear inflated, because there are so few rooftops with that tilt in each ZIP code. The overall error of the range is 0.47% (4.0 km²), with an average error per ZIP code of 1,215 m² and per azimuth/tilt category of 58 m². Based on these validation results, future studies that use ZIP-code-level estimates can use the upper and lower bounds to provide a range of estimates of suitable area.

6 Modeling to Extend Medium- and Large-Building PV Suitability and Potential Estimates Nationwide

In contrast to the previous section, which presented the technical potential of PV on small buildings, this section presents the estimated technical potential of PV on medium and large buildings. Medium buildings have a footprint between 5,000 ft² and 25,000 ft², and large buildings have a footprint larger than 25,000 ft². Estimates of suitable rooftop area for these size classes were obtained using the same model framework as with small buildings, leveraging what was observed in the lidar data set to make predictions in areas not covered by lidar data. We then use the same set of technical assumptions and simulation method as before to predict the nationwide technical potential of PV on medium and large buildings.

6.1 Modeled Estimates of Total Developable Area from Medium and Large Building Rooftops Nationwide

In contrast to the small building rooftop area estimates, the number of medium and large buildings per ZIP code is not available as a model input. As an input for these medium and large building predictions, we used Commercial Building Energy Consumption Survey (CBECS) building count data, which are available on a census division scale. Because greater than 99% of medium and large buildings within the lidar data had at least one suitable plane, we assumed that all medium and large buildings were suitable. Estimates of the total suitable area on each building were then derived using the same modeling process used for small buildings and described in full in Appendix D, with constituent models described in Appendices A–C.

To allow state-level predictions, we distributed the estimated rooftop area in a census division to its states by population weight.¹⁷ This method of distribution assumes that medium and large building counts are linearly related to population at a sufficiently large enough geographic scale. The actual number of buildings in states with above-average building counts per capita would therefore be underestimated, and states with below-average building counts per capita would be overestimated.¹⁸

6.2 Estimated National PV Technical Potential from Medium and Large Buildings

Figure 16 shows the total annual rooftop PV generation potential for medium and large buildings geographically. The total installed capacity, annual energy generation, suitable roof area, and annual energy generation as a percentage of state electricity sales are shown in Table 5. Across all states, small buildings have greater potential than the combined potential of medium and large buildings. On the national level, medium and large buildings have the potential to generate

¹⁷ To leverage what was known about existing medium and large building counts from the lidar data, known building counts within a census division were subtracted from CBECS estimates of building count. The rooftop area from the remaining buildings was then distributed among states, weighted by the population of the state that was not covered by the lidar data. Because approximately half of the country's estimated medium and large buildings were covered by lidar data, this significantly improved the final state-level estimates.

¹⁸ Because it is not possible to thoroughly evaluate the error introduced with this weighting, we caution the reader to consider state-level estimates as a visualization only.

506 TWh/year of electricity (14% of total national electricity sales), which is about half of the small building potential of 926 TWh/year (25% of total national sales).

Differences in potential between building size classes can largely be explained by the total suitable roof area and the utilization of the available space. Despite the higher percentage of medium and large building roof area suitable for PV deployment,¹⁹ the far smaller numbers of these buildings resulted in a lower total developable area (3.2 billion m²) compared to small buildings (4.9 billion m²). Additionally, the generally tilted roof area on small buildings could be used more efficiently than the predominantly flat roofs of large buildings, because flat roofs require greater spacing between modules to prevent excessive losses from shading. As innovative racking and module-packing techniques are developed for flat roofs, the technical potential from medium and large buildings could increase.

Figure 16. Annual rooftop PV generation potential for medium and large buildings

The relative magnitudes of the state-level results for medium and large buildings (Table 5) largely reflect the small building results (Table 3), with a few differences. Washington, D.C. can offset a much greater percentage of its total sales from medium and large buildings, because it differs from all the lower 48 states by having more developable area on medium and large buildings than on small buildings. Relative to other states, Arizona, Idaho, New Mexico, and Utah have less potential generation from these building classes than from small buildings, driven primarily by fewer large buildings per capita in that region compared with the national average.

¹⁹ Within the lidar data, the roof area suitable for development as a fraction of total roof area was 66% for large buildings, 49% for medium buildings, and 26% for small buildings.

Table 5. Estimated Rooftop PV Technical Potential of Medium and Large Buildings by State

State	Annual Generation Potential (% of sales)	Installed Capacity Potential (GW)	Annual Generation Potential (TWh/year)	Total Roof Area Suitable for PV Deployment (millions of m²)
California	30.6%	52.2	80.0	435.3
Rhode Island	25.4%	1.6	2.0	13.7
Connecticut	22.1%	5.6	6.6	46.3
Massachusetts	21.5%	10.1	11.9	82.4
New Hampshire	21.0%	2.0	2.3	16.5
Maine	20.1%	2.1	2.4	16.7
Vermont	19.7%	1.0	1.1	7.9
Nevada	18.0%	3.9	6.4	33.8
Maryland	17.3%	8.5	10.7	70.4
Colorado	16.9%	6.1	9.0	51.9
Oklahoma	16.7%	7.2	10.0	59.5
Florida	16.2%	25.9	35.9	213.6
New Jersey	15.5%	9.3	11.6	79.0
Minnesota	15.4%	9.1	10.6	75.7
Michigan	15.3%	13.7	15.8	114.1
Missouri	14.5%	9.4	12.1	78.1
Kansas	14.3%	4.2	5.7	34.7
Nebraska	14.2%	3.3	4.4	27.7
Illinois	13.4%	15.7	19.0	131.6
Texas	12.7%	35.1	48.1	289.9
New York	12.6%	15.3	18.6	129.7
Wisconsin	12.6%	7.3	8.7	60.2
Ohio	12.2%	15.8	18.3	131.3
Georgia	12.2%	12.2	15.9	101.9
Virginia	11.8%	10.2	13.1	83.9
Louisiana	11.6%	7.6	9.9	62.6
North Carolina	11.4%	11.1	14.7	92.3
Pennsylvania	11.3%	14.0	16.5	116.8
Washington DC	10.9%	1.0	1.2	8.1
Oregon	10.8%	4.4	5.1	35.6
Arkansas	10.8%	3.9	5.0	31.9
Alabama	10.7%	7.2	9.4	58.6

State	Annual Generation Potential (% of sales)	Installed Capacity Potential (GW)	Annual Generation Potential (TWh/year)	Total Roof Area Suitable for PV Deployment (millions of m²)
Iowa	10.7%	4.1	5.0	33.5
Mississippi	10.6%	3.9	5.2	31.8
New Mexico	10.6%	1.5	2.4	12.6
Tennessee	9.9%	7.4	9.6	60.4
South Dakota	9.3%	0.9	1.1	7.1
Kentucky	9.2%	6.4	7.8	52.9
Indiana	9.2%	8.0	9.7	65.9
Utah	9.1%	1.9	2.8	16.2
Delaware	9.0%	0.8	1.0	6.6
Washington	8.5%	7.1	7.8	57.9
South Carolina	6.6%	3.8	5.2	31.4
North Dakota	6.3%	0.8	1.0	6.7
West Virginia	5.6%	1.5	1.8	12.5
Idaho	4.2%	0.7	1.0	6.1
Arizona	2.8%	1.3	2.1	10.9
Wyoming	1.8%	0.2	0.3	1.7
Montana	1.3%	0.1	0.2	1.2
Continental U.S. Total	13.6%	386.5	506.0	3,207.4

7 Estimated National PV Technical Potential for All Buildings

In this section, we present the combined PV technical potential estimates for small buildings (from Section 0) and medium and large buildings (from Section 6) to produce an estimate of the total national installed capacity and generation potential of rooftop PV. Because the medium and large building estimates are only available at the state level, the combined results are presented at that level. Table 6 shows potential installed PV capacity, rooftop area suitable for development, and annual generation (in terawatt-hours and as a percentage of total electricity sales in 2013) by state. Figure 17 and Figure 18 (at the end of this section) map the potential generation results.

The total nationwide technical potential of PV across all buildings is 1,118 GW of installed capacity and 1,432 TWh of annual energy generation, which equates to 39% of total national electric sales. This is significantly greater than a previous NREL estimate of 664 GW of installed capacity and 800 TWh of annual energy generation (Lopez et al. 2012).²⁰ The difference can be attributed to increases in module power density,²¹ improved estimation of building suitability, higher estimates of the total number of buildings, and improvements in PV performance simulation tools that previously tended to underestimate productivity.

Table 6 shows California with the greatest potential to offset use—PV on its rooftops could generate 74% of the electricity sold by its utilities in 2013. A cluster of New England states could generate more than 45%, despite these states' below-average solar resource. Washington, with the lowest population-weighted solar resource in the continental United States, could still generate 27%. The best-performing six states—in terms of potential PV generation as a percent of total state sales—all have significantly below-average household consumption, suggesting the role an energy-efficient residential sector could play in achieving a high penetration of energy from rooftop PV.

Wyoming has the lowest potential for offsetting statewide electricity sales with rooftop PV, at 14%, because it has the highest per-capita electricity sales of any state at 30.3 MWh/year/person (250% of the national average), driven by very high electricity use in the industrial sector (60% of retail electricity sales). Washington D.C. has the second-lowest potential to offset electricity sales, at 15%; lidar data indicate that this unique, almost entirely urban district has only 17.4 m² of developable roof area per capita, which is much lower than the average of 24.9 m² per capita throughout the rest of the lidar -covered regions.

Some states with below-average solar resource (such as Minnesota, Maine, New York, and South Dakota) have similar or even greater potential to offset total sales than states with higher-quality resource (such as Arizona and Texas). This highlights the observation that solar resource is only one of several factors that determine the offset potential.

²⁰ Lopez et al. (2012) draw most of their rooftop PV results from the analysis carried out by Denholm and Margolis (2008).

²¹ Our analysis assumed a panel power density of 160 W/m², whereas Lopez et al. (2012) assumed 135 W/m².

Florida can offset 47% of its total consumption, despite having an average household consumption that is 130% above the national average. This is largely explained by significantly below-average electricity consumption outside of the residential sector, which makes total per-capita state sales slightly lower than the national average, plus high-quality solar resource and a high percentage of buildings suitable for PV. In contrast, the other south Atlantic states range from a potential 23% to 35% of electricity offset, owing primarily to lower average suitability and higher per-capita electricity sales.

Table 6. Total Estimated Technical Potential (All Buildings) for Rooftop PV by State

State	Annual Generation Potential (% of sales)	Installed Capacity Potential (GW)	Annual Generation Potential (TWh/year)	Total Roof Area Suitable for PV Deployment (millions of m ²)
California	74.2%	128.9	194.0	961
Maine	60.0%	6.3	7.1	45
Vermont	60.0%	3.0	3.4	21
Rhode Island	56.6%	3.8	4.4	28
New Hampshire	53.4%	5.3	5.9	38
Connecticut	49.8%	12.8	14.8	95
Massachusetts	47.0%	22.5	26.0	165
Florida	46.5%	76.2	103.2	557
Michigan	45.9%	42.1	47.3	303
Colorado	44.0%	16.2	23.5	119
Oklahoma	44.1%	19.3	26.4	140
New Mexico	43.4%	6.1	10.0	45
Missouri	42.7%	28.3	35.6	204
Kansas	41.7%	12.5	16.6	90
Nevada	39.6%	8.7	13.9	67
New Jersey	40.4%	24.9	30.1	184
Wisconsin	40.1%	23.6	27.7	169
Maryland	38.7%	19.3	23.9	142
Minnesota	38.5%	23.1	26.4	168
South Dakota	38.7%	3.8	4.7	26
New York	37.4%	46.6	55.3	340
Illinois	37.0%	44.1	52.5	324
Ohio	35.3%	46.8	53.0	338
Iowa	35.5%	14.0	16.6	99
Texas	34.6%	97.8	131.2	715
North Carolina	34.9%	35.0	45.3	252

State	Annual Generation Potential (% of sales)	Installed Capacity Potential (GW)	Annual Generation Potential (TWh/year)	Total Roof Area Suitable for PV Deployment (millions of m²)
Pennsylvania	34.5%	43.6	50.4	316
Nebraska	34.1%	8.2	10.5	60
Utah	34.3%	7.2	10.4	52
Oregon	34.2%	14.1	16.3	101
Georgia	33.8%	34.6	44.1	251
Arizona	34.4%	16.3	26.1	114
Arkansas	33.3%	12.2	15.5	88
Virginia	32.4%	28.5	35.8	205
Tennessee	31.9%	24.4	30.9	175
Mississippi	31.2%	11.7	15.2	84
Delaware	31.0%	2.9	3.5	20
Louisiana	29.8%	20.1	25.6	146
Alabama	29.8%	20.4	26.2	147
Indiana	29.5%	26.3	31.1	188
Montana	28.0%	3.2	3.9	21
Washington	26.6%	22.8	24.7	164
Idaho	26.4%	4.7	6.4	33
Kentucky	25.2%	18.0	21.4	131
South Carolina	25.5%	15.2	20.0	108
North Dakota	24.6%	3.3	3.9	23
West Virginia	22.9%	6.3	7.2	45
Washington DC	15.1%	1.3	1.7	11
Wyoming	14.2%	1.7	2.4	12
Continental U.S. Total	38.6%	1,118	1,432	8,130

Figure 17. Total annual energy generation potential from rooftop PV for all building sizes

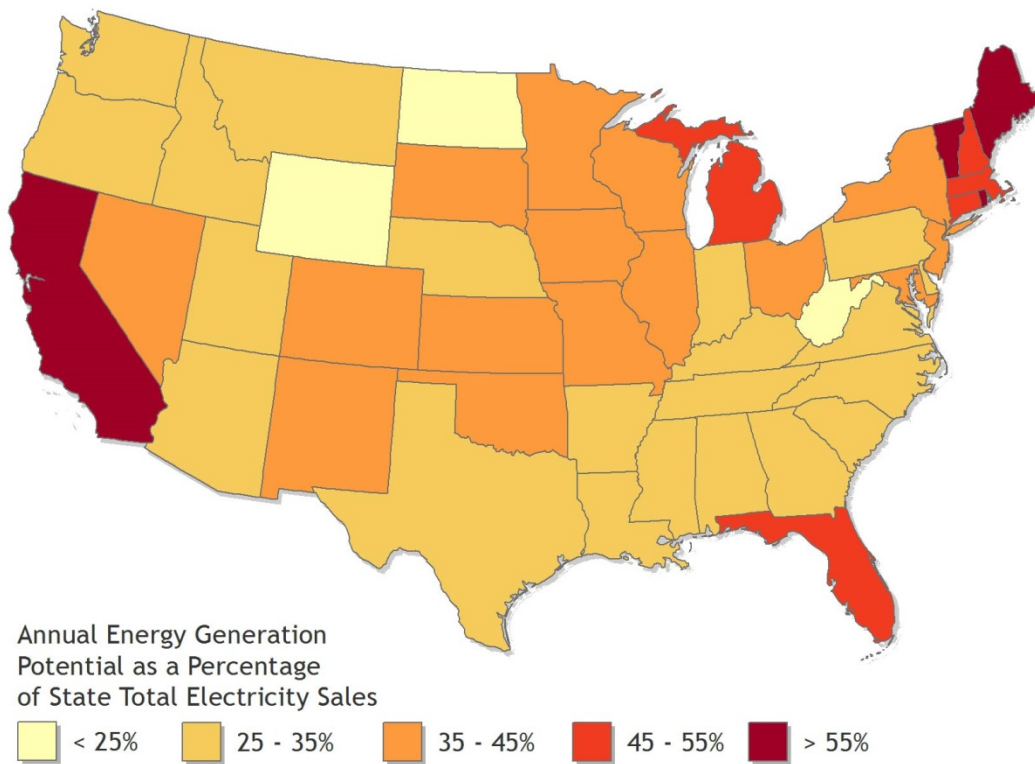


Figure 18. Potential rooftop PV annual generation from all buildings as a percentage of state total 2013 electricity sales

8 Conclusions and Future Research Directions

In this report, we considered the technical potential of PV systems installed on existing, suitable roofs within the continental United States. We evaluated the technical potential of rooftop PV within 47 cities—with a detailed analysis for 11 cities—by using a DHS lidar data set (covering 40% of the national population and 23% of the building stock) to estimate roof area suitable for PV deployment. We also leveraged the DHS lidar data to build models for predicting the total amount of suitable roof area in regions for which lidar data do not exist.

Using statistical models and ACS building counts, we made nationwide ZIP-code-level estimates of the number and characteristics of small buildings, showing broad regional trends in both the suitability and electric-generation potential of rooftops. Although only 26% of the total rooftop area on small buildings is suitable for PV deployment, the sheer number of buildings in this class gives small buildings the greatest technical potential. Small building rooftops could accommodate 731 GW of PV capacity and generate 926 TWh of PV energy annually, which represents approximately 65% of the total technical potential of rooftop PV.

Using statistical models and CBECS building counts, we made nationwide estimates of the number and characteristics of medium and large buildings on the census division level and then distributed them to the state level by population weight. Medium and large buildings have a total installed capacity potential of 386 GW and energy generation potential of 506 TWh/year, approximately 35% of the total technical potential of rooftop PV.

The total national technical potential of rooftop PV is 1,118 GW of installed capacity and 1,432 TWh of annual energy generation. This equates to 39% of total national electric-sector sales. Table 7 summarizes the total rooftop area suitable for PV and technical potential metrics, by building class for the nation. Because the relative magnitudes of the results are a strong function of the square footage used as a cutoff between divisions, these results should not be presented without that information.

Table 7. Nationwide Estimated Suitable Area and Technical Potential by Building Class

Building Class (building footprint)	Total Suitable Area (billions of m ²)	Installed Capacity Potential (GW)	Annual Generation Potential (TWh/year)	Annual Generation Potential (% of national sales)
Small (< 5,000 ft²)	4.92	731	926	25.0%
Medium (5,000–25,000 ft²)	1.22	154	201	5.4%
Large (> 25,000 ft²)	1.99	232	305	8.2%
All Buildings	8.13	1,118	1,432	38.6%

These results are also sensitive to assumptions about module performance, which is expected to continue improving. For example, for this analysis we assumed a module power density of 160 W/m², corresponding to a module efficiency of approximately 16%. If a power density of 200 W/m² were assumed instead, corresponding to a module efficiency of approximately 20%, each of the technical potential estimates would increase about 25% above the values stated in this report. Furthermore, these are only estimates of the potential from existing, suitable roof

planes; they do not consider the immense potential of ground-mounted PV. Actual generation from PV in urban areas could exceed these estimates by installing systems on less suitable roof area, by mounting PV on canopies over open spaces such as parking lots, or by integrating PV into building facades.

Because our results are estimates of technical potential, they do not consider what would be required to use all the potential PV-generated energy. In practice, the integration of a significant quantity of rooftop PV into the national portfolio of generation capacity would require a flexible grid, supporting infrastructure, and a suite of enabling technologies.

8.1 Data for Public Use

The data set developed by NREL contains a significant amount of information specific to rooftops suitable for PV installation. The data were developed with the intention of providing a much-improved starting point for PV research regionally and nationwide. This base data set can be combined with numerous other data sets—such as data sets of insolation data, market data, or electricity consumption data—to provide more precise and actionable information than has been available previously.

There are many possible directions for future analyses, and it is the author’s hope that the data will be used creatively by municipalities, utility providers, solar energy researchers, and anyone else interested in using rooftop data to advance PV deployment. We have posted supporting data on the NSRDB Data Viewer website.²² Data layers include a regional summary and ZIP-code-level summary for all areas where lidar data are available as well as national coverage of ZIP code estimates as predicted by the methods described in this report. Detailed documentation of each step in our analysis, including scripts for running the GIS tools, are linked to in the metadata section of each layer. This information can be accessed by clicking the question mark icon next to each layer in the table of contents in the Data Viewer.

8.2 Future Work

Within the DHS lidar data set used in this analysis, approximately 30 cities have multiple years’ worth of data, and this type of multiyear data likely will continue to grow. Growth of these data could enable analysis to be performed over time to examine whether or how rooftop PV technical potential changes over time.

The estimates of PV electricity production from small buildings within this report were made on the ZIP-code level, and they could be aggregated to any useful administrative region. For example, capacity and electricity-production estimates could be aggregated for a given utility service territory to inform utilities of the quantity and spatial distribution of distributed generation potential within their territories. Information on the quantity and temporal characteristics of distributed generation within a service territory could help guide the design of new rate structures that accurately track services rendered, as the grid continues to modernize.

Using the DHS lidar data, we made building-level estimates of potential productivity from rooftop PV. Combining these results with modeled energy consumption within the same regions

²² See maps.nrel.gov/pv-rooftop-lidar.

would enable analysis of the ability of PV to offset consumption at the building level. This could provide information on how local distribution requirements could change under increasing levels of distributed PV adoption. If paired with a technology-diffusion model, such as NREL's Distributed Generation Market Demand Model (dGen), customer-adoption behavior could be incorporated, and a timetable for such changes could be estimated.

We could improve our model predictions by leveraging raw orthoimagery (satellite images) to classify potential areas for PV installations and thoroughly explore geostatistical methods on a national scale. Although the result of doing so cannot approach the accuracy of the lidar analysis described here, it may provide a finer-resolution estimate where fine-resolution orthoimagery is available but lidar data are not.

References

- Anderson, K.H., M.H. Coddington, and B.D. Kroposki. 2010. “Assessing Technical Potential for City PV Deployment using NREL's In My Backyard Tool.” Prepared for the 35th IEEE Photovoltaic Specialists Conference, June 20–25, Honolulu, HI.
- Barbose, Galen, and Naïm Darghouth. 2015. “Tracking the Sun VIII: An Historical Summary of the Installed Price of Photovoltaics in the United States from 1998 to 2014.” Berkeley, CA: Lawrence Berkeley National Laboratory.
- Barbose, G, N. Darghouth, S. Weaver, and R. Wiser. 2014. “Tracking the Sun VII: An Historical Summary of the Installed Price of Photovoltaics in the United States from 1998 to 2013.” Berkeley, CA: Lawrence Berkeley National Laboratory.
- Bright, J., and K. Burman. 2010. “Portland Ecodistrict Solar Site Assessment.” Golden, CO: National Renewable Energy Laboratory (for Solar Cities America).
- Brown, Austin, Philipp Beiter, Donna Heimiller, Carolyn Davidson, Paul Denholm, Jennifer Melius, Anthony Lopez, Dylan Hettinger, David Mulcahy, and Gian. Porro. 2015. “Estimating Renewable Energy Economic Potential in the United States: Methodology and Initial Results.” Golden, CO: National Renewable Energy Laboratory. NREL/TP-6A20-64503.
- Chaudhari, Maya, Lisa Frantzis, and Tom E. Hoff. 2004. “PV Grid Connected Market Potential Under a Cost Breakthrough Scenario.” Chicago: Navigant Consulting. 2004-117373.
- Compagnon, R. 2004. “Solar and Daylight Availability in the Urban Fabric.” *Energy and Buildings* 36(4): 321–28. doi:10.1016/j.enbuild.2004.01.009.
- Denholm, Paul, and Robert Margolis. 2008. “Supply Curves for Rooftop Solar PV-Generated Electricity for the United States.” Golden, CO: National Renewable Energy Laboratory. NREL/TP-6A20-44073.
- EIA (U.S. Energy Information Administration). 2009. “Residential Energy Consumption Survey (RECS).” Washington, DC: U.S. Energy Information Administration.
- EIA (U.S. Energy Information Administration). 2012. “Commercial Building Energy Consumption Survey (CBECS).” Washington, DC: U.S. Energy Information Administration.
- . 2014. “Monthly Electric Sales and Revenue Report with State Distributions Report.” EIA-826. Washington, DC: U.S. Energy Information Administration. ESRI (Environmental Systems Research Institute). 2014. ArcGIS Desktop: 10.1. Redlands, CA: Environmental Systems Research Institute.
- Frantzis, Lisa, Shannon Graham, and Jay Paidipati. 2007. “California Rooftop Photovoltaic (PV) Resource Assessment and Growth Potential by County.” Chicago: Navigant Consulting. CEC-500-2007-048.

Gilman, Paul, and Aron. Dobos. 2012. "System Advisor Model, SAM 2011.12.2. 2: General Description." Golden, CO: National Renewable Energy Laboratory. NREL/TP-6A20-53437.

Hofierka, Jaroslav, and Ján Kanuk. 2009. "Assessment of Photovoltaic Potential in Urban Areas using Open-Source Solar Radiation Tools." *Renewable Energy* 34(10): 2206–14. doi:10.1016/j.renene.2009.02.021.

Land, Charles E. 1971. "Confidence Intervals for Linear Functions of the Normal Mean and Variance." *Annals of Mathematical Statistics* 42: 1187–1205.

Lopez, Anthony, Billy Roberts, Donna Heimiller, Nate Blair, and Gain Porro. 2012. "U.S. Renewable Energy Technical Potentials: A GIS-Based Analysis." Golden, CO: National Renewable Energy Laboratory. NREL/TP-6A20-51946.

Melius, J., R. Margolis, and S. Ong. 2013. "Estimating Rooftop Suitability for PV: A Review of Methods, Patents, and Validation Techniques." Golden, CO: National Renewable Energy Laboratory. NREL/TP-6A20-60593.

Mondol, Jayanta Deb, Yigzaw G. Yohanis, and Brian Norton. "Optimising the Economic Viability of Grid-Connected Photovoltaic Systems." *Applied Energy* 86.7 (2009): 985-999.

Olsson, Ulf. 2005. "Confidence Intervals for the Mean of a Log-Normal Distribution." *Journal of Statistics Education* 13(1).

Ordonez, J., E. Jadraque, J. Alegre, and G. Martinez. 2010. "Analysis of the Photovoltaic Solar Energy Capacity of Residential Rooftops in Andalusia (Spain)." *Renewable and Sustainable Energy Reviews* 14: 2122–30.

Paidipati, J., L. Frantzis, H. Sawyer, and A. Kurrasch. 2008. "Rooftop Photovoltaics Market Penetration Scenarios." Golden, CO: National Renewable Energy Laboratory (work performed by Navigant Consulting). NREL/SR-581-42306.

Santos, T Teresa, Nuno Gomes, Miguel Brito, Sérgio Breire, Ana Fonseca, and José Antonio. 2011. "Solar Potential Analysis in Lisbon using lidar Data." Prepared for the 31st EARSeL Symposium and 35th General Assembly, May 30–June 2, Prague, Czech Republic.

Wilcox, S., and W. Marion. 2008. "Users Manual for TMY3 Data Sets." Golden, CO: National Renewable Energy Laboratory. NREL/TP-581-43156.

Zhang, Xilin, Rebecca Walker, Michael Salisbury, Rebecca Hromiko, and Jackson Schreiber. 2009. "Creating a Solar City: Determining the Potential of Solar Rooftop Systems in the City of Newark." Newark, DE: University of Delaware, Center for Energy and Environmental Policy.

Appendix A: Rooftop Tilt and Azimuth Model

In this appendix, we explore the distribution of rooftop planes in each tilt and azimuth category to understand the practical tilt and orientation of planes in the data set. Using that information, we develop a predictive model that can be used to infer the distribution of planes in an area where lidar data are not available.

Figure A-1 gives the fraction of buildings for each ZIP code grouped according to four tilt classes and five azimuth classes.²³ These classes, combined with a single class for flat roofs, create 21 different orientation bins that can be used to characterize rooftops. Most developable rooftop planes fall into the 28-degree tilt category, and steep rooftops (54 degrees) are an order of magnitude less common than the next category (41 degrees). Azimuths east, south, and west are most common, particularly among the 28-degree category of rooftops. These azimuths correspond to buildings aligned on a cardinal street grid. These observations appear to hold for small, medium, and large buildings alike.

Figure A-2 shows the fraction of flat rooftops in ZIP codes for each class of buildings. Nearly all large buildings provide flat planes. Approximately three in four medium buildings have suitable flat planes, while less than one in four small buildings have suitable flat planes. Figure A-3 groups the fraction of small buildings with flat planes according to the ZIP code locale type, showing that large cities have the most flat-roofed small buildings and that the fraction decreases in more urban areas. The same trend appears to be true of homogeneity when a chi-squared test is applied to the contingency distribution of building planes in each tilt and azimuth combination; 17% of ZIP codes labeled “small suburb” show statistically significant heterogeneity among their tilt/azimuth distribution, while only 7.2% of large cities are significantly heterogeneous.²⁴

²³ The five azimuth bins are 45 degrees each, covering 67.5 degrees to 292.5 degrees, as defined in Figure 6. The four tilt bins equally split the 50.5 degree range from 9.5 degrees to 60 degrees. Anything less than 9.5 degrees is considered flat.

²⁴ The chi-squared test was performed with factored building plane counts and $\alpha = 0.05$.

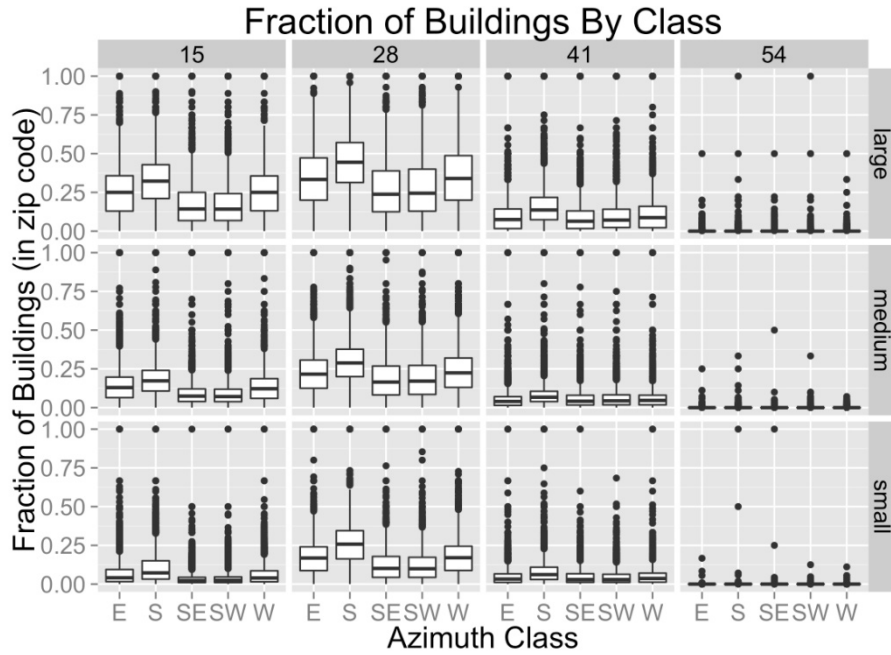


Figure A-1. Fraction of suitable rooftop planes in each tilt and azimuth category for small, medium, and large buildings

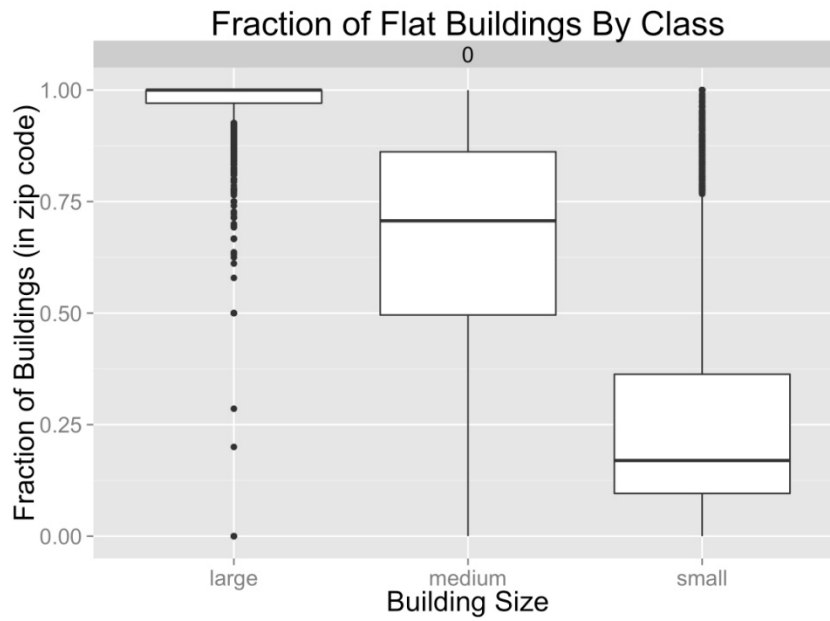


Figure A-2. Fraction of flat suitable rooftop planes for small, medium, and large buildings

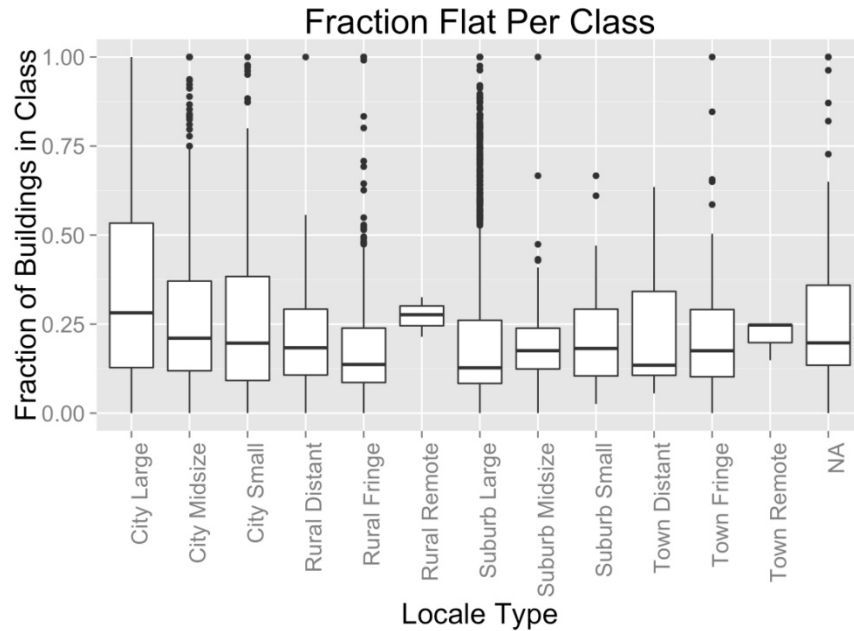


Figure A-3. Fraction of small buildings with flat planes grouped by locale type

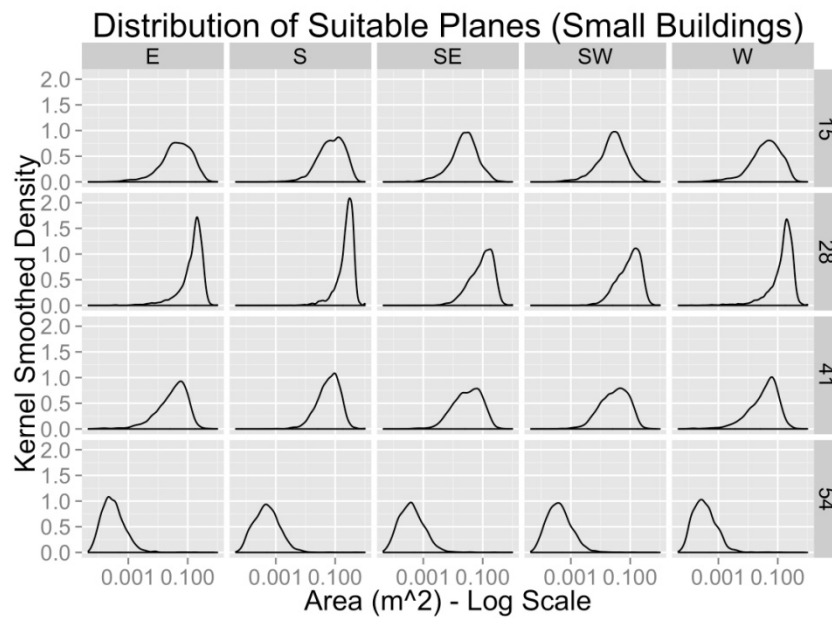


Figure A-4. Probability distribution functions for each combination of azimuth and tilt for small buildings

Figure A-4 shows the probability density function for each combination of tilt and azimuth for small buildings alone. This plot is log-scaled and makes clear the long tail on each distribution (also apparent in the box plots in Figure A-1) and lognormal shape. Using this observation, we can formulate average “characteristic” matrices by using the lognormal mean:

$$T = \begin{bmatrix} t_{11} & \cdots & t_{1l} \\ \vdots & \ddots & \vdots \\ t_{k1} & \cdots & t_{kl} \end{bmatrix}$$

$$t_{ij} = e^{\mu + \frac{\sigma^2}{2}}$$

$$s. t. \sum t_{ij} = 1$$

where T is the “characteristic” 2-D matrix describing the average distribution of a number of matrices. T is normalized to 1 so that the fractions form a complete distribution.

Using this method, national average distribution can be calculated for each building class, providing insight into the distribution of azimuth and tilt in each class (Table A-1, Table A-2, and TableA-3).

Table A-1. National Average Tilt/Azimuth Distribution for Planes on Large Buildings

	Flat	E	SE	S	SW	W
0°	0.93					
15°		0.01		0.01		0.01
28°		0.01		0.01		0.01
41°						
54°						

Table A-2. National Average Tilt/Azimuth Distribution for Planes on Medium Buildings

	Flat	E	SE	S	SW	W
0°	0.74					
15°		0.02	0.01	0.04	0.01	0.02
28°		0.03	0.02	0.05	0.02	0.03
41°				0.01		
54°						

Table A-3. National Average Tilt/Azimuth Distribution for Planes on Small Buildings

	Flat	E	SE	S	SW	W
0°	0.26					
15°		0.03	0.02	0.06	0.02	0.03
28°		0.09	0.06	0.18	0.06	0.09
41°		0.02	0.02	0.03	0.01	0.02
54°						

Confidence interval estimates can be generated for each fitted mean using the Cox method (Olsson 2005; Land 1971):

$$t_{ij} \pm z \sqrt{\frac{\sigma^2}{n} + \frac{(\sigma^2)^2}{2(n-1)}}$$

where z is the distribution parameter for the desired confidence level.

As a final method to understand sources of variability in this data set, we perform a principle components analysis on the 21 categorical fractions (20 azimuth and tilt combinations, plus fraction flat) for small buildings. Table A-4 provides the loadings for the first three components, which cumulatively describe 90% of the total variance observed. In the table, the dominant loadings have been bolded. The first component is dominated by the fraction of flat roofs. The second component most strongly measures the degree to which a ZIP code has southwest and southeast azimuths (most rare in the data), and the third component measures the fraction of south-facing azimuths (next most rare).

Table A-4. Principle Component Loadings for Tilt/Azimuth

	Component 1	Component 2	Component 3
Flat	0.94	-0.09	0.07
E, 15°	-0.02	-0.12	-0.20
E, 28°	-0.14	-0.20	-0.33
E, 41°	-0.04	-0.02	-0.08
E, 54°			
SE, 15°	-0.01	0.10	0.02
SE, 28°	-0.09	0.51	0.17
SE, 41°	-0.03	0.17	0.01
SE, 54°			
S, 15°	-0.02	-0.25	0.13
S, 28°	-0.22	-0.48	0.72
S, 41°	-0.06	-0.05	0.05
S, 54°			
SW, 15°	-0.01	0.11	0.03
SW, 28°	-0.08	0.50	0.15
SW, 41°	-0.03	0.16	0.01
SW, 54°			
W, 15°	-0.02	-0.10	-0.18
W, 28°	-0.14	-0.22	-0.46
W, 41°	-0.04	-0.03	-0.10
W, 54°			

Appendix B: Small Building Suitability Model

This appendix describes our model for determining how many small buildings in a given area are suitable for PV installation. It combines an analysis of rooftop suitability (fraction of suitable rooftops in a given area) with a method for estimating the total number of small buildings. When combined, the total expected number of small *suitable* buildings can be determined.

FigureB-1 and FigureB-2 show the distribution of fraction of suitable rooftops at the state level. Small building suitability at the state level appears to be normally distributed. On average, per-state suitability is 79% (median: 79.39%, minimum: 58.70%, maximum: 92.93%, standard deviation: 7.21%).

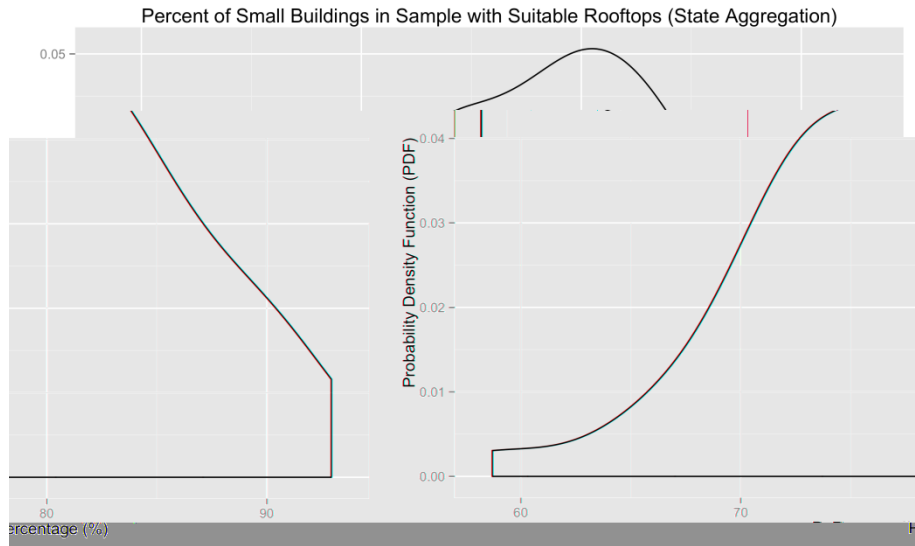


Figure B-1. Distribution of suitable percentage among small buildings (within states)

Figure B-2. Suitable percentage among small buildings (within states) and 95% confidence intervals

Aggregating instead by ZIP codes reveals finer-scale variability, particularly among urban and rural regions. Looking at ZIP codes in the aggregate as seen in FigureB-3, we see a skewed distribution, with a tail to the left. To address this tail, we find that ZIP-code-level statistics can be more harmoniously described by partitioning using National Center for Education Statistics (NCES) locale description and U.S. census division. FigureB-4 gives the percentage of small suitable buildings in each ZIP code, grouped according to these variables, and TableB-1 gives the complete statistics for every combination of categories.

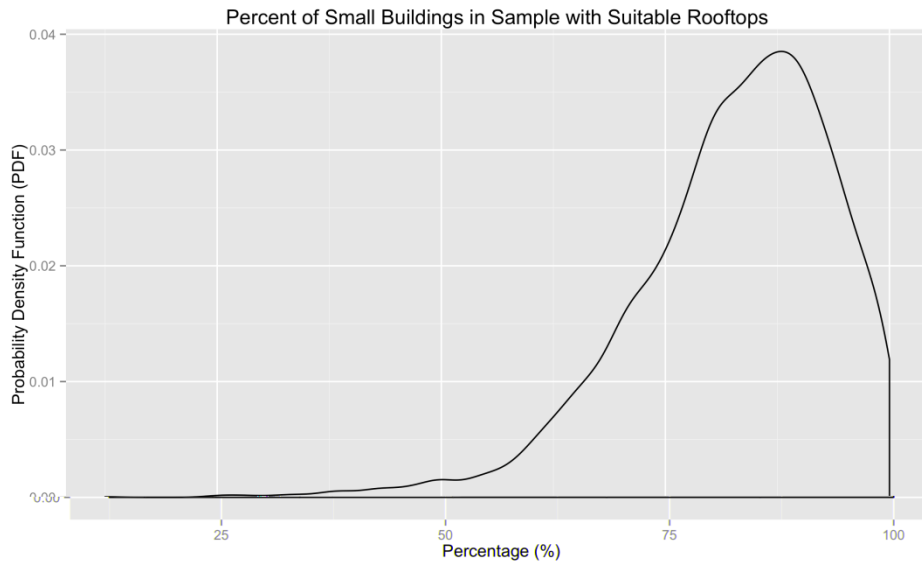


Figure B-3. Distribution of the percentage of small suitable buildings by ZIP code

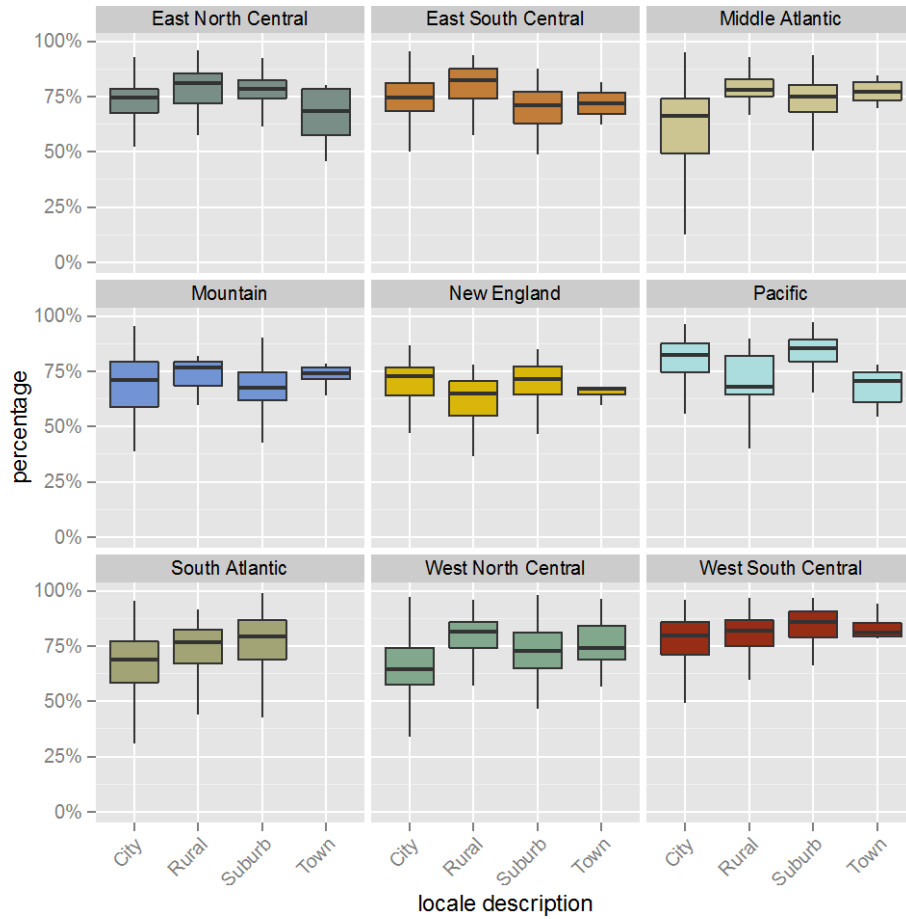


Figure B-4. Percentage of small buildings containing a flat area with at least 10 m² suitable for PV, by census division and aggregated locale description

Table B-1. Suitable Percentage of Small Buildings by Locale/Division

Locale Type	Census Division	Mean	Median	Min	Max	N
City Large	East North Central	80.32	81.80	12.50	100.00	222.00
City Midsize	East North Central	79.36	78.38	65.57	100.00	65.00
City Small	East North Central	81.06	81.70	71.01	100.00	39.00
Rural Distant	East North Central	87.51	88.03	61.54	100.00	15.00
Rural Fringe	East North Central	84.67	88.16	52.63	100.00	61.00
Suburb Large	East North Central	84.06	84.62	47.68	100.00	286.00
Suburb Midsize	East North Central	76.55	79.67	69.48	80.50	3.00
Suburb Small	East North Central	100.00	100.0	100.0	100.00	1.00
Town Fringe	East North Central	81.78	84.62	50.00	97.06	17.00
City Large	East South Central	84.16	83.08	70.72	100.00	46.00
City Midsize	East South Central	81.93	84.82	47.34	100.00	53.00
Rural Distant	East South Central	85.76	84.30	74.30	97.39	8.00
Rural Fringe	East South Central	88.13	91.11	63.92	100.00	18.00
Suburb Large	East South Central	72.20	74.12	33.33	96.28	29.00
Suburb Midsize	East South Central	92.89	92.89	91.12	94.67	2.00
Town Distant	East South Central	87.72	87.72	87.72	87.72	1.00
Town Fringe	East South Central	88.88	88.31	78.90	100.00	4.00
City Large	Middle Atlantic	82.37	84.08	40.52	100.00	273.00
City Midsize	Middle Atlantic	77.00	77.15	45.45	93.75	22.00
City Small	Middle Atlantic	84.81	84.62	70.69	100.00	54.00
Rural Distant	Middle Atlantic	84.15	84.37	66.67	100.00	11.00
Rural Fringe	Middle Atlantic	84.95	85.31	66.67	100.00	56.00
Suburb Large	Middle Atlantic	80.62	81.65	25.00	100.00	523.00
Suburb Midsize	Middle Atlantic	75.00	75.00	75.00	75.00	1.00
Suburb Small	Middle Atlantic	82.67	81.63	80.00	87.43	4.00
Town Fringe	Middle Atlantic	84.34	84.54	67.44	92.59	9.00
City Large	Mountain	77.10	80.39	46.07	100.00	89.00
City Midsize	Mountain	71.54	73.71	55.02	89.02	22.00
City Small	Mountain	79.75	84.82	61.54	94.26	9.00
Rural Distant	Mountain	76.41	77.40	58.55	85.31	8.00
Rural Fringe	Mountain	88.93	90.04	78.77	98.68	5.00
Rural Remote	Mountain	66.61	66.61	61.79	71.44	2.00
Suburb Large	Mountain	74.25	74.20	33.33	100.00	147.00
Suburb Small	Mountain	65.83	68.73	42.11	78.08	11.00

Locale Type	Census Division	Mean	Median	Min	Max	N
Town Distant	Mountain	72.89	72.89	71.71	74.07	2.00
Town Fringe	Mountain	78.97	77.28	72.26	91.07	8.00
Town Remote	Mountain	86.31	86.31	86.31	86.31	1.00
City Midsize	New England	74.71	80.92	28.90	94.96	45.00
City Small	New England	78.86	79.08	71.59	85.57	18.00
Rural Distant	New England	71.31	70.81	44.44	83.11	12.00
Rural Fringe	New England	67.69	71.48	38.46	89.39	14.00
Rural Remote	New England	89.58	89.58	89.58	89.58	1.00
Suburb Large	New England	74.97	78.38	31.09	100.00	122.00
Suburb Midsize	New England	77.25	77.02	70.00	84.46	6.00
Town Distant	New England	73.80	73.52	59.63	87.95	7.00
Town Fringe	New England	74.12	71.81	66.94	84.52	6.00
City Large	Pacific	87.29	89.16	49.48	100.00	244.00
City Midsize	Pacific	87.10	89.94	46.41	98.17	93.00
City Small	Pacific	89.86	91.09	57.14	100.00	38.00
Rural Distant	Pacific	80.13	80.13	76.04	84.21	2.00
Rural Fringe	Pacific	80.14	77.89	64.72	94.72	12.00
Suburb Large	Pacific	89.47	91.48	53.85	100.00	186.00
Suburb Midsize	Pacific	90.20	90.49	80.34	96.58	7.00
Suburb Small	Pacific	80.10	84.89	60.00	92.34	10.00
Town Distant	Pacific	82.15	82.15	82.15	82.15	1.00
Town Fringe	Pacific	74.83	78.35	60.00	88.86	12.00
City Large	South Atlantic	80.40	81.12	49.44	100.00	155.00
City Midsize	South Atlantic	80.92	80.53	27.21	100.00	185.00
City Small	South Atlantic	77.02	83.12	41.82	100.00	4.00
Rural Distant	South Atlantic	76.77	81.11	43.28	100.00	16.00
Rural Fringe	South Atlantic	79.51	81.95	28.30	100.00	44.00
Suburb Large	South Atlantic	85.66	87.80	25.29	100.00	534.00
Suburb Midsize	South Atlantic	87.28	88.95	50.00	100.00	42.00
Suburb Small	South Atlantic	85.47	85.44	74.93	95.06	18.00
Town Distant	South Atlantic	71.43	71.43	71.43	71.43	1.00
City Large	West North Central	76.50	75.38	55.51	100.00	133.00
City Midsize	West North Central	72.67	71.94	48.61	100.00	46.00
City Small	West North Central	77.58	77.51	62.02	100.00	29.00
Rural Distant	West North Central	84.18	84.88	57.14	100.00	21.00

Locale Type	Census Division	Mean	Median	Min	Max	N
Rural Fringe	West North Central	85.04	88.51	37.68	98.99	45.00
Suburb Large	West North Central	79.14	80.20	54.46	98.74	146.00
Suburb Midsize	West North Central	77.62	85.79	57.42	89.65	3.00
Suburb Small	West North Central	82.95	84.78	71.03	93.53	6.00
Town Distant	West North Central	84.92	84.92	84.02	85.83	2.00
Town Fringe	West North Central	83.16	86.74	48.78	98.59	14.00
Town Remote	West North Central	75.37	75.37	73.00	77.73	2.00
City Large	West South Central	88.90	89.24	61.66	100.00	207.00
City Midsize	West South Central	82.30	82.68	56.90	100.00	66.00
City Small	West South Central	84.79	85.71	66.67	97.55	18.00
Rural Distant	West South Central	84.88	84.94	75.43	96.46	8.00
Rural Fringe	West South Central	87.11	90.31	25.84	100.00	44.00
Suburb Large	West South Central	90.17	91.99	40.00	100.00	101.00
Suburb Small	West South Central	86.96	89.53	63.04	95.30	7.00
Town Distant	West South Central	94.99	96.38	91.25	97.34	3.00
Town Fringe	West South Central	89.65	89.53	80.91	98.25	10.00

B.1 Predictive Regression Model

As a basis for prediction in a multiple regression model, we gathered geospatial data from the U.S. National Land Cover Database (NLCD 2006), NCES (2006), and U.S. Census Bureau. At the outset of model building, we expected different regions of the country may have different characteristics with respect to PV suitability. We also expected, based on our understanding of the shading process, building density is likely well correlated with suitability.

In the search for the best-performing model for this data set, we tried combinations several variables. The variables that appear well correlated with PV suitability are:

Locale Description: NCES (2006) locale description (e.g., Small Town, Remote Rural) of each ZIP code area.

Census Division: The U.S. Census divides the country into nine divisions: East North Central, East South Central, West North Central, West South Central, Middle Atlantic, South Atlantic, New England, Pacific, and Mountain (FigureB-5).

Land Cover Classification Percentage: NLCD (2006) percentages for each ZIP code. Those that appeared to be particularly useful and are included in our model include: Developed Open Space, Developed Low Intensity, Developed Medium Intensity, Developed High Intensity, Forest, and Scrub.

Population Density: 2010 population density per square mile from the U.S. Census Bureau was used.

Location: We found meaningful correlations between northing (or latitude) and suitability, while easting (or longitude) did not appear significantly correlated. In this work, we use projected coordinates²⁵ for easy comparability and convenient (meters-based) scale.

Lidar Coverage: When fitting a regression model, substantial improvements were obtained by weighting ZIP codes by their coverage in the lidar data. Both the total number of buildings detected and the fraction of the area covered were tested, with the former proving more useful.

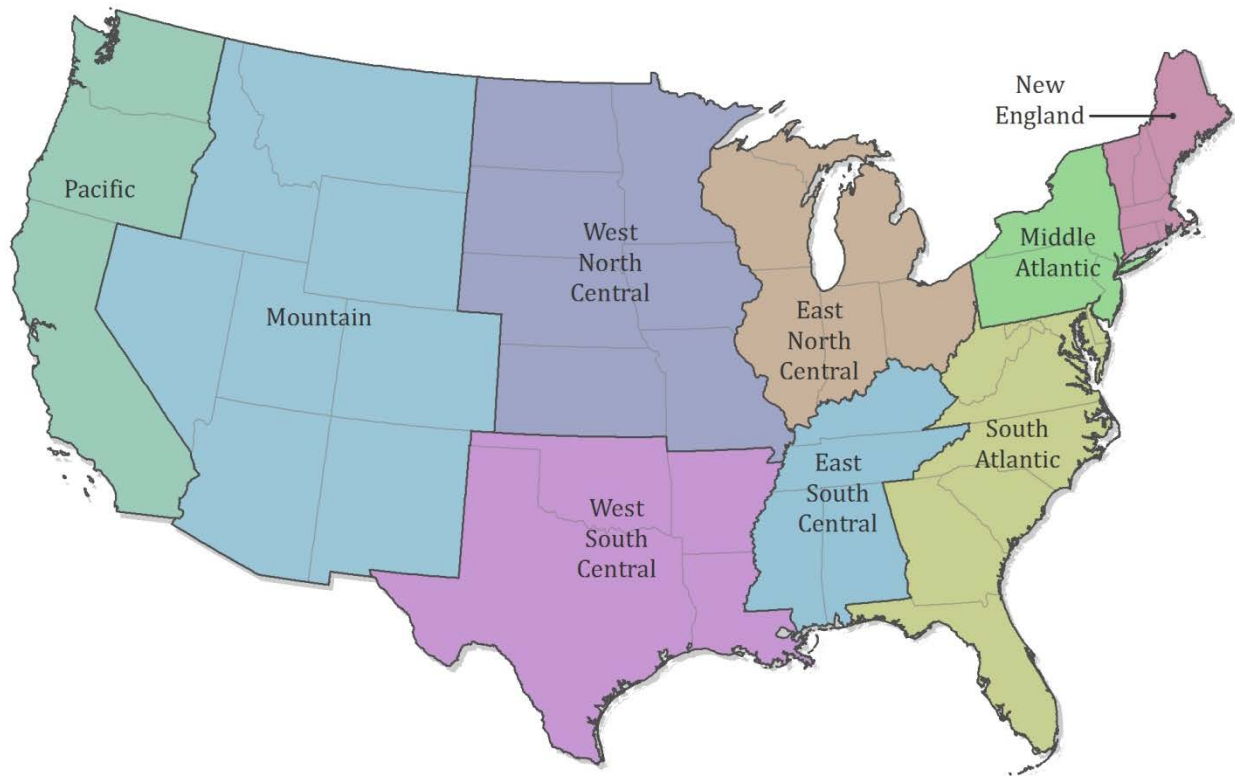


Figure B-5. Census divisions

Additional variables that were considered but not ultimately used owing to low correlation with small building rooftop suitability or for having too coarse of a resolution include the Residential Energy Consumption Survey, Commercial Building Energy Consumption Survey (CBECS), American Housing Survey, U.S. Economic Census, and Zillow building and sales data.

We constructed several models using combinations of the variables above, and we “trained” each model using a randomly selected subset of 75% of the original lidar data set. We then assessed the predictive accuracy of our models using the remaining 25% of the data by predicting the percentage of buildings that are small and suitable for PV and comparing those predictions to their actual values. We found that reducing our best-performing model by removing the three

²⁵ Common Web Mercator (EPSG 3857) was used for the spatial projection.

least useful variables provided a simplified version that sacrificed little in terms of accuracy. This reduced model served as our final model. About 60% of our final model’s predictions were within 10% of the observed lidar values, and about 80% of the predictions were within 20% of the observed values.

Table B-2 shows fit statistics for a subset of the small building models evaluated. In each case the root-mean-square error (RMSE) is evaluated via 100-fold cross validation, withholding 25% of the training data to test against. The Basic Model fits against only locale description and census division. The *Large* model adds land cover and canopy information (classes 10, 20, 40, and 50 and their subclasses) along with population density. The States Model adds the U.S. state name, for a modest improvement, but has the consequence of causing undesirable transitions at state boundaries. Finally, the Spatial Model incorporates the northing of the ZIP code centroid being fitted.

Table B-2. Model Tuning and Performance, Final (Reduced) Model Marked with an Asterisk

Model	Degrees of Freedom	AIC	BIC	RMSE
Basic	81	-4,585.853	-4,067.163	0.142
Large	86	-5,016.133	-4,465.425	0.135
States	64	-5,399.268	-4,989.439	0.130
Spatial	90	-5,981.261	-5,404.939	0.121
Weighted	90	-6,515.968	-5,939.646	0.041
Reduced*	86	-6,430.093	-5,879.386	0.042

Although these models perform well (RMSE 12%–14%), substantial additional gains can be obtained with the Weighted Model, in which a least squares regression is augmented with weights from the number of buildings detected during the lidar GIS processing. This weighted least squares (WLS) approach has the effect of up-weighting dense urban areas with good lidar coverage in our data set while down-weighting ZIP areas with few buildings or poor lidar coverage. For instance, Table B-3 provides statistics for four outliers common to the prior models, whose poor coverage seemingly affects the quality of the fit. For simplicity, we also present a Reduced Model in which the three least useful variables have been removed. This reduced model performs nearly as well, with 4.2% RMSE. TableB-4 and TableB-5 provide analysis of variance (ANOVA) results for these models and their parameters.

Table B-3. Examples of Outliers in Unweighted Regression

ZIP Code	City	State	Total Buildings	lidar Coverage (%)	Population Density (pop./mi ²)	True Suitability (%)
20560	Washington	DC	3	100	7.5	0
63045	St Louis	MO	171	100	69.1	7
90095	Los Angeles	CA	8	100	15,165.0	0
50061	Des Moines	IA	261	28	87.3	79

Table B-4. ANOVA on Weighted Model

Source	Df	SS	MS	F	P-value
Locale Description	11	0.26	0.02	13.15	0.0000
Census Division	8	2.29	0.29	162.59	0.0000
NLCD21	1	0.28	0.28	157.93	0.0000
NLCD22	1	0.01	0.01	2.86	0.0910
NLCD23	1	0.00	0.00	0.89	0.3453
NLCD24	1	0.65	0.65	369.04	0.0000
NLCD11	1	0.02	0.02	13.93	0.0002
NLCD40	1	0.34	0.34	190.98	0.0000
NLCD50	1	0.01	0.01	7.59	0.0059
Population Dens.	1	0.01	0.01	2.88	0.0898
Northing	1	0.68	0.68	386.60	0.0000
Locale Desc. * Census Division	60	0.47	0.01	4.43	0.0000
Residuals	4,374	7.72	0.00	—	—

Table B-5. ANOVA on Reduced (Final) Model

Source	Df	SS	MS	F	P-value
Locale Description	11	0.26	0.02	12.89	0.0000
Census Division	8	2.29	0.29	159.35	0.0000
NLCD21	1	0.28	0.28	154.78	0.0000
NLCD24	1	0.53	0.53	296.69	0.0000
NLCD11	1	0.04	0.04	23.44	0.0000
NLCD40	1	0.32	0.32	176.76	0.0000
Northing	1	0.69	0.69	381.41	0.0000
Locale Desc. * Census Division	60	0.44	0.01	4.41	0.0000
Residuals	4,378	7.88	0.00	—	—

The proportion of predictions within 10% and 20% of their actual values using the final full (weighted) model was 59.4% and 83.3%, respectively. Similarly, the proportion of predictions within 10% and 20% of the actual values using the final reduced model was 59.4% and 83.1%, respectively. Therefore, we suggest using the reduced model—henceforth referred to as the final model—as the basis for future predictions and analyses because it does not sacrifice predictive capability and it is more parsimonious.

B.2 Spatial Correlation

To understand whether prediction errors might be correlated in space, we computed Moran's I on the underlying process and the regression residuals. As is standard, we computed Moran's I using the network of eight nearest neighbors. The underlying process has an accumulated Moran's I of 0.56 with p-value $\ll 0.05$, suggesting a significant spatial autocorrelation. By comparison, the post-fit residuals have an aggregate Moran's I of 0.42 with p-value $\ll 0.05$. FigureB-6 provides correlograms plotting Moran's I as a function of lag distance for both the actual values and the regression residuals. Lag distance is defined as the number of "hops" on the nearest neighbor network.

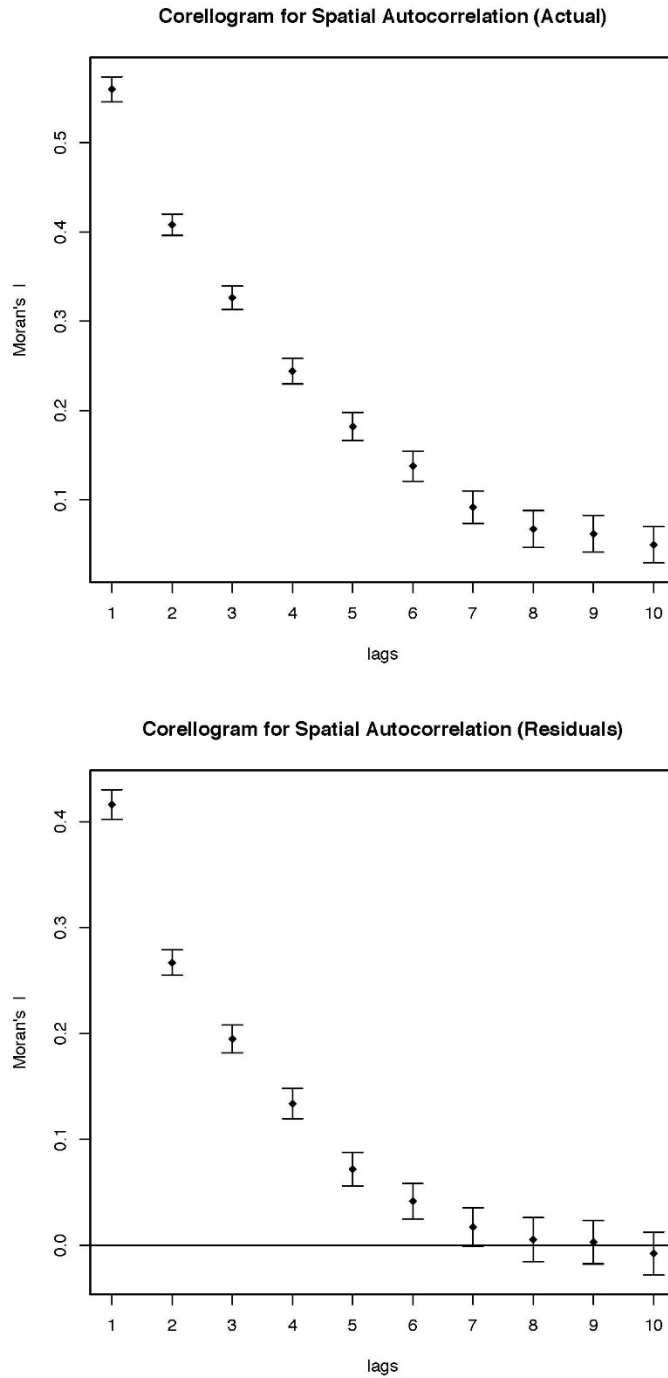


Figure B-6. Spatial autocorrelation of percentage of buildings that are small and suitable for PV, with data aggregated into ZIP codes, both before and after regression fitting

These results suggest that, while the regression can capture some spatial relationships in the data, there is still substantial spatial dependency that might be used to better fit the model. There is also a relatively strong spatial autocorrelation for ZIP codes within one to three lags, but the autocorrelation falls off at greater distances. Based on this result, we explore geostatistical modeling as a potential extension to the national regression model.

B.3 Regression Kriging

As an optional extension to the linear regression model, we explored regression Kriging as a method for further increasing predictive accuracy. Kriging is a geostatistical interpolation originally developed for mapping in resource extraction and mine valuation. Today Kriging is finding a foothold in diverse applications mainly owing to its favorable properties as an interpolation method; (1) it is an exact method, so the map of predictions will always map to ground truth at locations where data exist and (2) it provides a notion of uncertainty around predictions by reporting the Kriging variance at each estimated point. In regression Kriging, the residual error from a regression model is the basis for training a geostatistical model that attempts to make an additional fine-tuning correction. Although regression Kriging is a powerful method, it is also computationally costly. In the present study, we limit our analysis to Colorado, leaving the task of performing a nationwide prediction to later work.

The first step in geostatistical modeling involves characterizing the semivariogram function. The semivariogram is a plot of half the difference between points separated by some distance. Based on our analysis of spatial autocorrelation, we understand that there is the highest degree of spatial dependency with 2–3 lags. For this reason, we focus our semivariogram model on urban areas by truncating the model at a maximum of 40 km between points. In the Colorado data, this appears to be a reasonable heuristic that maximizes generalizability while preserving model fidelity. FigureB-7 (top) shows an empirical semivariogram for Colorado using an exponential semivariogram model. The semivariogram model is used by the Kriging process to find optimal weights when predicting additional points.

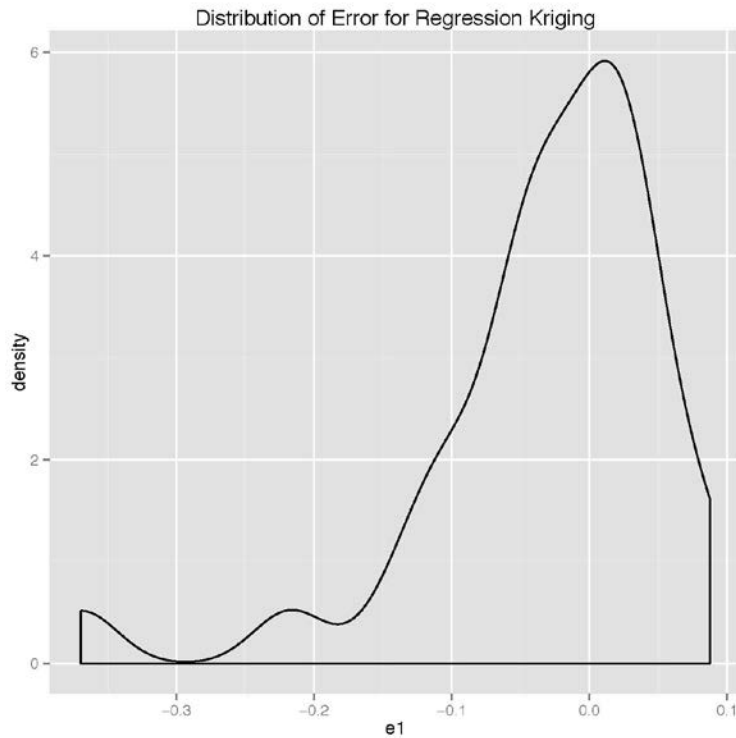
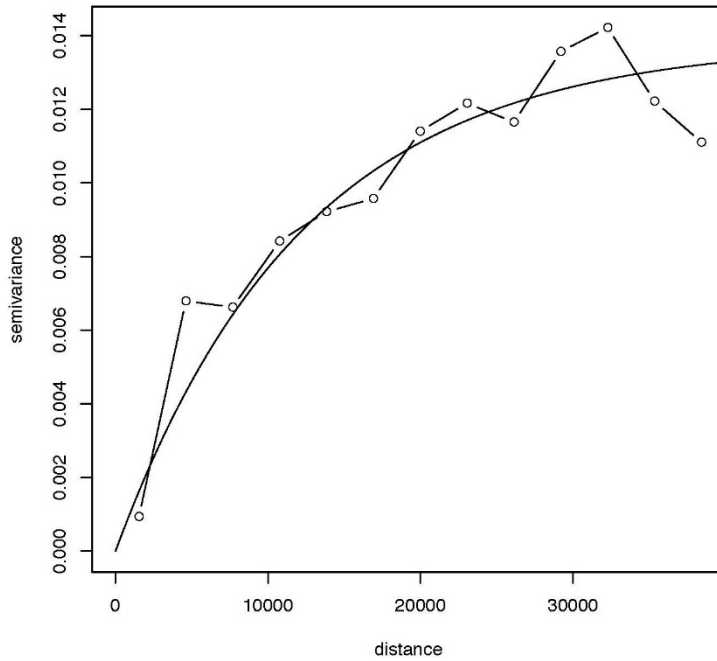


Figure B-7. Fitted semivariogram using an exponential semivariogram model (top) and residual error of regression Kriging model (bottom)

To evaluate this method, we perform a standard cross-validation where 25% of the data are withheld while the model is trained on the remaining 75%. The distribution of residual error of the Kriging-corrected regression model is given in FigureB-7 (bottom). We see an overall improvement of approximately 2%, from 12% error in the basic regression model to 10%.

FigureB-8 compares the maps generated with this method with the predictions of the final regression model. Error is substantially reduced from the prior model in locations where data are available because regression Kriging matches the prediction values at those points. While these data do not permit us to approximate the value of regression Kriging in rural areas, accounting for and explicitly fitting the spatial dependency in residuals does appear to increase fidelity in the predictions within urban areas. Because this model builds on the regression model, it can be optionally applied when higher-resolution predictions are desired for a particular city or set of metropolitan areas, particular when partial lidar data may be available, which allows for per-urban-area model training and interpolation.

Figure B-8. Performance of Regression Kriging model for Colorado

Figure shows actual values (top left), regression-Kriging predictions (bottom left), regression predictions (bottom right), and error for regression Kriging (top right).

B.4 Prediction of the Number of Buildings

As a final consideration, to predict the number of suitable buildings in a given region, we must have a notion of how many buildings are present (suitable or not). Because these data are not readily available, we make a similar prediction using a regression model:

$$n_i = \text{population} + \text{nlcd}_{21} + \text{nlcd}_{23} + \text{nlcd}_{811} + \text{northing} + \text{acs}_1 + \text{pop density}$$

where *locale* is the NCES locale, *region* is the census division, *nlcd* is the fraction of land cover for the given classes, *northing* is the Universal Transverse Mercator northing (converted from degrees latitude), and *acs*₁ is the number of single-unit detached housing according to the 2011 ACS. This model was evaluated against the data set using 10-fold cross validation with 25% of data withheld for testing (Tableb-6). We censored 1,922 ZIP codes with less than 90% coverage during fitting, because this produced a more harmonious fit.

Table B-6. Predictive Accuracy of Predicted Variables for National Potential Model

Model	RMSE	2.5 th Percentile Error	97.5 th Percentile Error	Median Error	Proportion within 10%	Proportion within 20%
n_i	870	-1,939	1,676	69	49%	72%

In practice, owing to the large variability, predicting the number of buildings poses challenges for fitting. Nonlinear methods for prediction (e.g., Multiple Adaptive Regression Splines or Support Vector Machines) do not appear to perform appreciably better. In this work, we proceed with the given model for n_i , treating its errors directly and accepting that future work may improve this component of the estimate (e.g., with finer estimates of building counts from commercial data).

Appendix C: Rooftop Plane Area Model

To estimate the suitable area per building we must create model-based estimates for both the number of planes per building and the size of each of those planes. As input to a data-driven model, we perform a uniform random sample of one million small buildings across all ZIP codes where lidar data are available. This results in a sample of 1.86 million planes, which form the basis for our model derivation (Figure C-1).

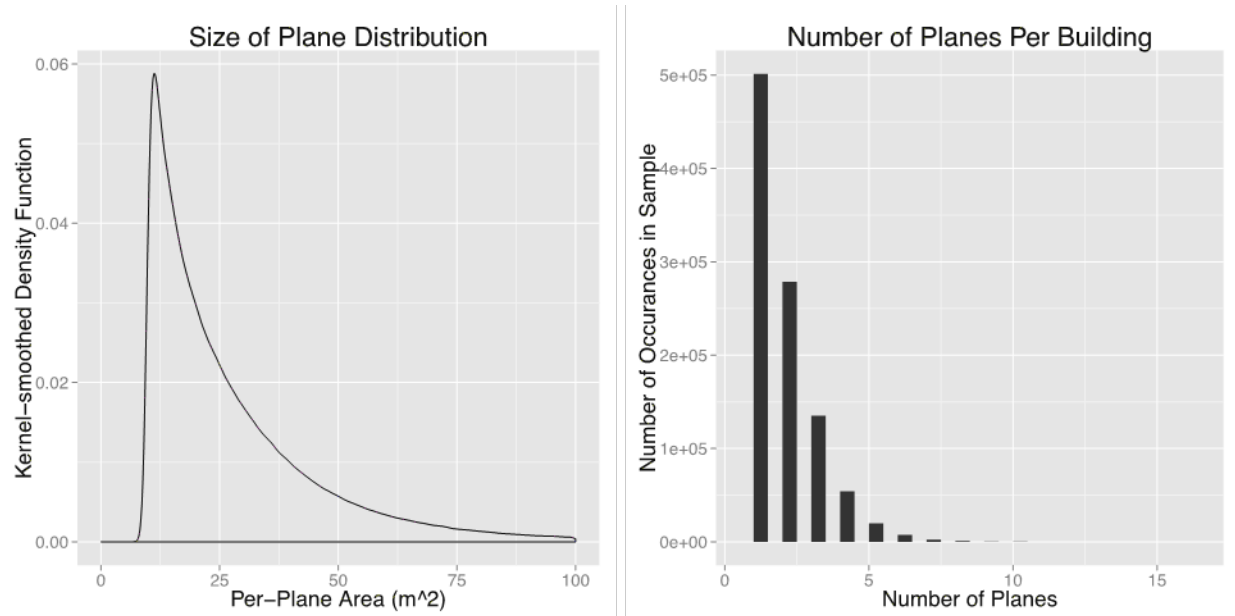


Figure C-1. Distributions of plane sizes and plane numbers for small buildings

Figure C-1 shows that, for small buildings, the size of each plane and the number of planes per building are well characterized by an exponential distribution. Total suitable flat area size for small buildings is a Weibull-distributed random variable with a shape parameter of 1.4 (standard error = 0.01) and a scale parameter of 62.0 (standard error = 0.47). For the number of planes per building, 99.6% of small buildings have six or fewer planes, and 50% have only one plane. We use these fractions directly, normalized to 1.0 in the model (Table C-1).

Table C-1. Building Planes: Small Buildings

1 plane	2 planes	3 planes	4 planes	5 planes	6 planes
50%	28%	14%	5%	2%	1%

The per-plane area is similar for small buildings with two to six planes. However, small buildings with only one plane generally have a larger single plane. To account for this, we build exponential fits for one-plane buildings separately from multiple (six or fewer) plane buildings. Figure C-2 shows the fits obtained, which are exceptionally close to the observed values.

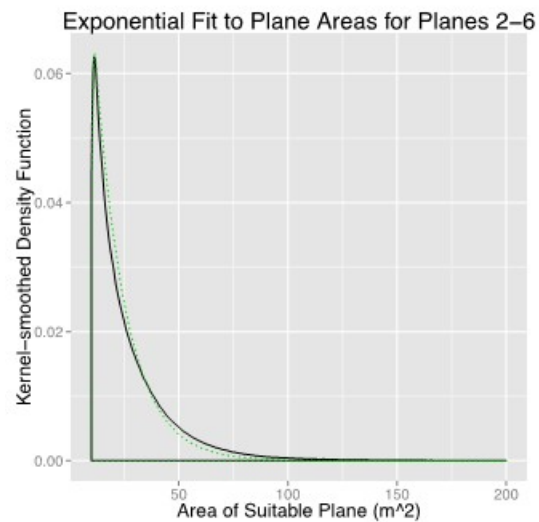
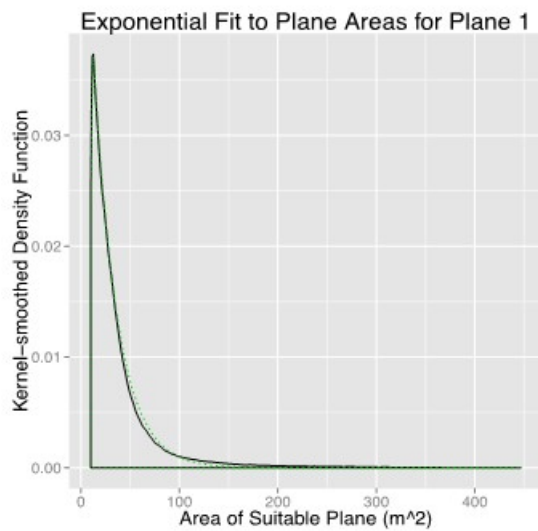
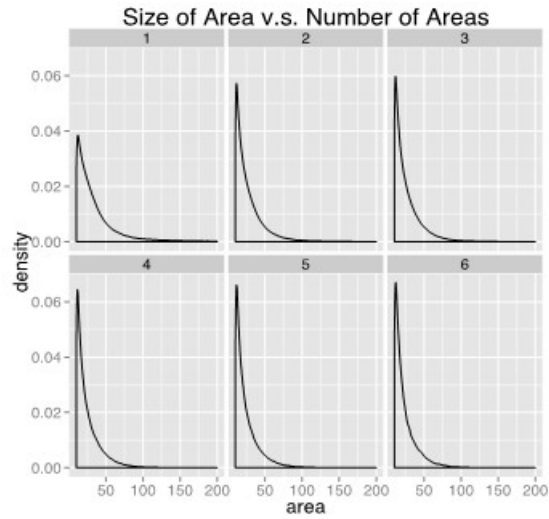


Figure C-2. Exponential fits

The fitted probability density function is:

$$p(x) = re^{-rx} + \theta$$

where r is the fitted rate of decay of the exponential process and θ is an optional location parameter, which we set to $\theta = 10$, the minimum plane size in our data. For small buildings, we find $r_1 = 0.042$ and $r_{2-6} = 0.071$. The mean of each distribution can be calculated directly:

$$s_i = \frac{1}{r} + \theta$$

Lower and upper bounds for the mean area in each category are computed as the 95th percentile range (2.5% to 97.5%):

$$s_i = \left\{ \frac{\log(0.975)}{r}, \frac{\log(0.025)}{r} \right\} + \theta$$

An identical method is used to obtain fits for medium and large buildings. For medium buildings, 98.2% of buildings have 14 or fewer suitable planes. As before, a truncated distribution is used such that the smallest category is at least 1% of all buildings (TableC-2).

Table C-2. Building Planes: Medium Buildings

1 plane	2 planes	3 planes	4 planes	5 planes	6 planes	7 planes
21%	14%	13%	12%	10%	8%	6%
8 planes	9 planes	10 planes	11 planes	12 planes	13 planes	14 planes
5%	3%	2%	2%	1%	1%	1%

In the medium building class, the distribution for single-plane buildings starts farther to the right, which is accommodated by setting $\theta_1 = 300$ (i.e., the minimum practical size of a single-plane rooftop is 300 m²):

$$r_1 = 0.002, \theta_1 = 300$$

$$r_{2-14} = 0.046, \theta_{2-14} = 10$$

Large buildings have as many as 32 planes in our model (TableC-3).

Table C-3. Building Planes: Large Buildings

1 plane	2 planes	3 planes	4 planes	5 planes	6 planes	7 planes	8 planes
14%	7%	6%	5%	5%	5%	5%	4%
9 planes	10 planes	11 planes	12 planes	13 planes	14 planes	15 planes	16 planes
4%	4%	3%	3%	3%	3%	3%	3%
17 planes	18 planes	19 planes	20 planes	21 planes	22 planes	23 planes	24 planes
2%	2%	2%	2%	1%	1%	1%	1%
25 planes	26 planes	27 planes	28 planes	29 planes	30 planes	31 planes	32 planes
1%	1%	1%	1%	1%	1%	1%	1%

$$r_1 = 0.0005, \theta_1 = 2,000$$

$$r_{2-32} = 0.055, \theta_{2-14} = 0$$

The large buildings are least well modeled using this method, largely because there is no simple dichotomy between buildings with one plane and buildings with two or more planes. As FigureC-3 demonstrates, one-plane buildings are clearly exponential; however, buildings with 2–10 planes appear bimodal. By assuming that all buildings in the model have a *first* plane selected from the one-plane distribution, we account for this multimodality. One goal for future work is to fit a continuous model that accounts for the relationship between size of panel and number of panels.

To understand how this model compares to assumptions of suitable rooftop area used in prior studies, we compared the expected average suitable rooftop size from our model for medium and large buildings to the average sizes from a previous NREL report on economic potential (Brown et al. 2015). In that study, medium buildings were assumed on average to have 1,487 m² of suitable rooftop area, while in our model the mean is 929 m². Also, that study assumed large buildings had on average 5,471 m² of suitable rooftop space, while our model predicts 4,178 m². FigureC-4 and FigureC-5 show these figures overlaid as vertical lines on the actual observed distribution of suitable building areas in the lidar data set. For this data set, the nonparametric modeling approach used here seems to better account for the heavy-tailed nature of the underlying distribution and it better models the true central tendency in the data.

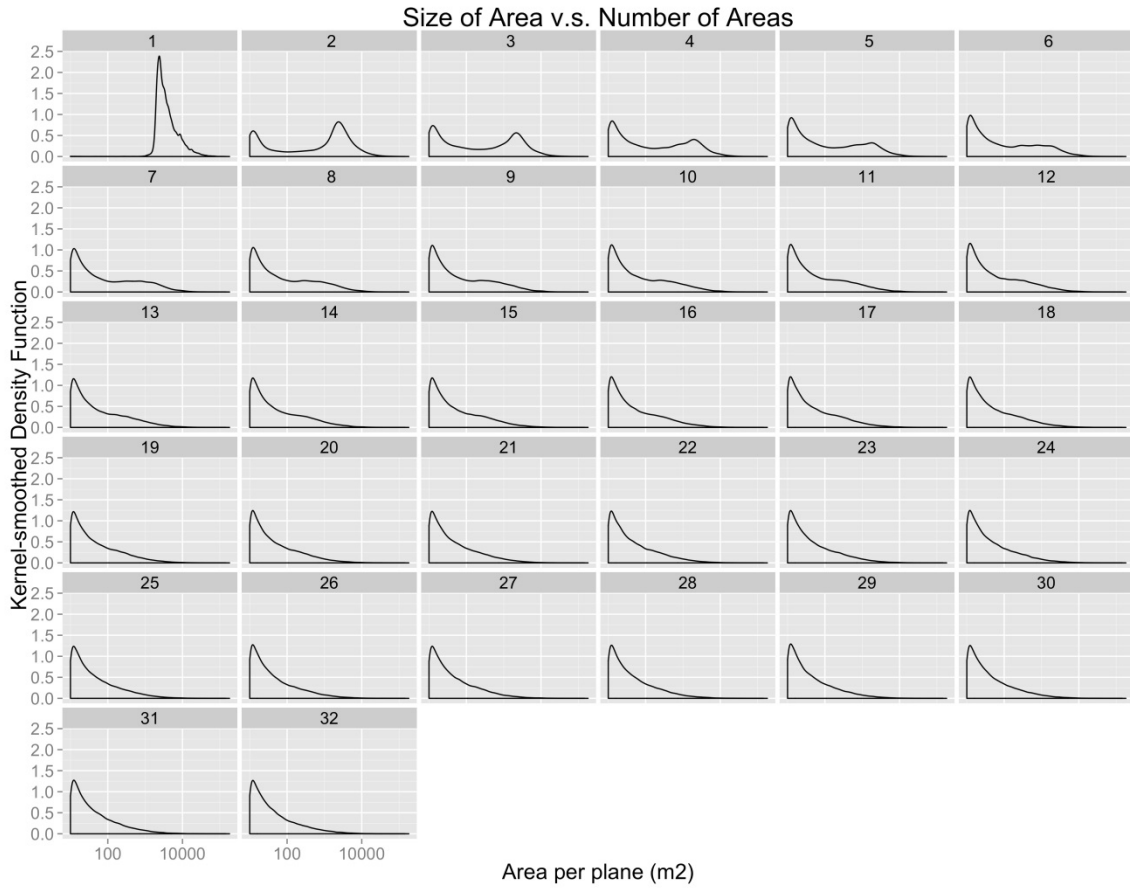


Figure C-3. Observed size of suitable planes for buildings with 1–32 planes in lidar data

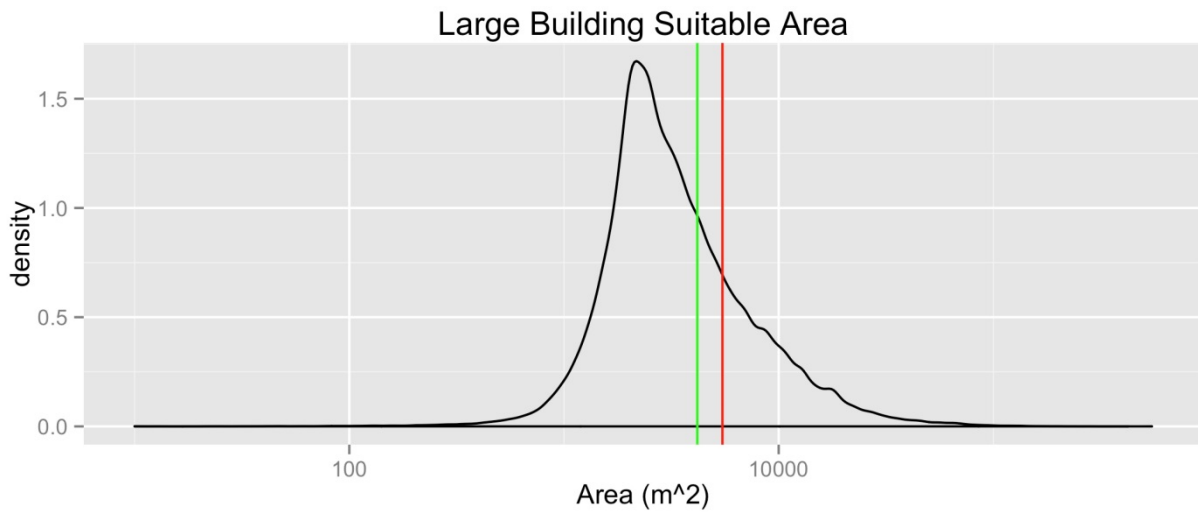


Figure C-4. Distribution of large building suitable areas (total usable rooftop area) in lidar data compared to model-calculated mean (green line) and assumed constant value in prior NREL studies (red line)

Plot is in log-scale for better visibility.

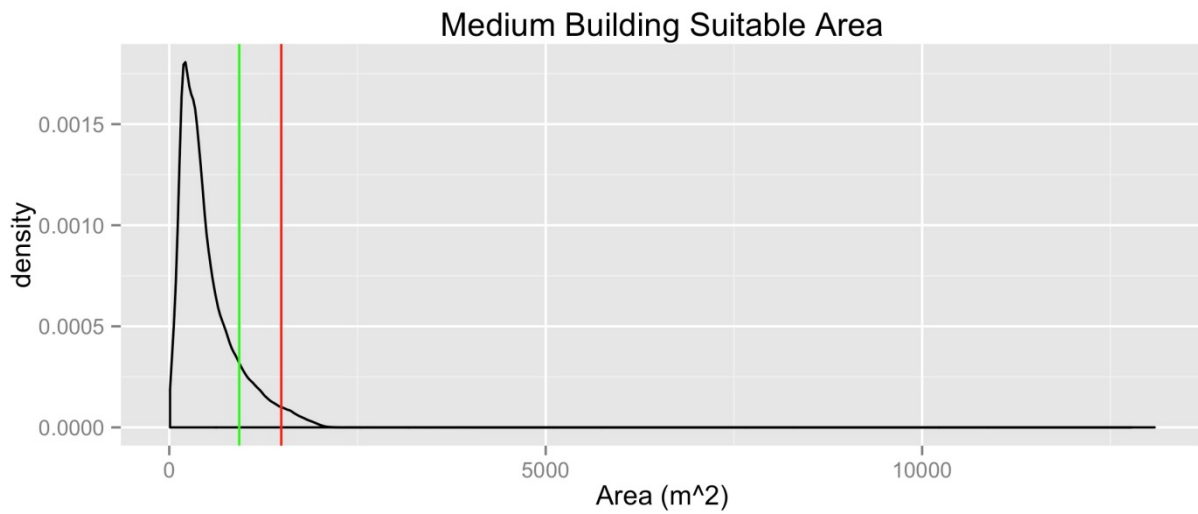


Figure C-5. Distribution of medium building suitable areas (total usable rooftop area) in lidar data compared to model calculated mean (green line) and assumed constant value from prior NREL studies (red line)

- For medium and large buildings, a constant value of 1 is used (assuming that 100% of medium and large buildings have at least one suitable plane)

f_i : Expected fraction of buildings with exactly i suitable planes

- Derived from per-class (i.e., small, medium, and large) random sample corresponding to the actual fractions seen in the data.

s_i : Expected size of suitable plane space in square meters for the i^{th} suitable plane (see Appendix C)

- Defined as the mean of a fitted exponential distribution using a per-class random sample.
- The first plane ($i = 1$) is fitted separately from successive planes ($i = 2..m$) because substantial variation exists between these classes.
- $s_i = \frac{1}{r} + l$, where l is the location parameter (threshold) and r is the return rate of the fitted exponential.

t_{ij} : Expected fraction of rooftop planes on a surface with tilt class i and azimuth class j (see Appendix A)

- For small buildings, per-locale-type average values are used from a lognormal fit of data in each azimuth and tilt class combination.
- For medium and large buildings, national average lognormal fitted values are used.
- In both cases, the matrix is normalized such that $T = \sum t_{ij} = 1$.
- $t_{ij} = e^{\mu + \sigma^2/2}$ where μ and σ^2 are the fitted parameters of each lognormal distribution, as described in Appendix A.

In the equation for a_i , $\|n_i p_i\|$ is the predicted number of buildings, $\overrightarrow{l_{ijk}}$ is the number of planes in each planes-per-building category, and the dot product of $\overrightarrow{l_{ijk}}$ and \vec{s} gives the total suitable predicted area in azimuth/tilt combination j/k for the i^{th} geography.