2020

Adoption of Residential Solar Energy: Exploratory Study Approach and Spatial Decision Support System

Ahmed Omar Alzahrani
Claremont Graduate University

Adoption of Residential Solar Energy: Exploratory Study Approach and Spatial Decision Support System

By
Ahmed Alzahrani

Claremont Graduate University
2019

APPROVAL OF THE DISSERTATION COMMITTEE

This dissertation has been duly read, reviewed, and critiqued by the Committee listed below, which hereby approves the manuscript of Ahmed Alzahrani as fulfilling the scope and quality requirements for meriting the degree of Doctor of Philosophy (Ph.D.) in Information Systems and Technology.

Brian Hilton, Chair

Claremont Graduate University

Professor

Lorne Olfman, Member

Claremont Graduate University

Professor

June Hilton, Member

Claremont Graduate University

Professor

Abstract

Adoption of Residential Solar Energy: Exploratory Study Approach and
Spatial Decision Support System

by
Ahmed Alzahrani

Claremont Graduate University: 2019

Stakeholders have recognized a need to employ and advance renewable-energy technologies. Renewable energy plays a crucial role in sustaining governments' energy resources security, mitigating environmental risks associated with extensive use of fossil fuels, and offering long-lasting energy resources for industrial, commercial, and residential needs for electrical power. This study focused on providing a better understanding of solar photovoltaic (PV) adoption at the residential level. Although many researchers investigated solar energy adoption in a search for what may motivate individuals to invest in such a technology, previous studies of solar energy adoption did not address spatial and nonspatial factors simultaneously. Geographic information system (GIS) technologies have emerged as powerful tools to study geographic characteristics that can be tested statistically and displayed on an interactive map to support decision makers. This dissertation investigated residential solar energy adoption by operationalizing homeowners' spatial and nonspatial factors. The dissertation features two artifacts that were created as part of the research. First, it includes an algorithm to build a predictive model, technically evaluating the model by applying it to a dataset of Los Angeles County as a case study. Second, this dissertation entailed developing a spatial decision-support system as an instantiation of the output model in the first artifact. The system is an interactive GIS-based web application that supports decision makers (e.g., policymakers, solar firms, utility companies and nonprofit organizations) in their decision-making process. A combined qualitative and quantitative methodological approach was used to evaluate the application's usability and user experience.

Dedication

To my dear parents, Omar and Jameilah

&

My lovely wife, Noura, and wonderful kids, Omar and Mansour

&

My sisters and brothers

Acknowledgments

I would like to express my deepest gratitude and appreciation to my committee chair, Dr. Brian Hilton, for his infinite support, guidance, and enthusiastic encouragement throughout my PhD coursework and the dissertation process. He taught me a couple of classes in GIS and I worked closely with him and learned much from his experience. D.Hilton is a great professor who has the ability to make any subject interesting and easy. I was honored to be a member of his lab, advanced GIS lab. Dr.Hilton's mentorship is a model for me to carry throughout my academic career.

In addition to my advisor, I would like to thank committee members, Dr. June Hilton and Dr. Lorne Olfman, for serving on my dissertation committee and for the great contributions you have made in my work. Dr. June taught a quantitative research method course and guided me through applying various statistical methods to multiple research paper. I also would like Warren Roberts for being for his support and help. Warren is an added value to the GIS lab at Claremont Graduate University.

Lastly, I would like to thank my friends at Claremont Graduate University for being supportive along the way of this journey. I would like to express my gratitude to my parents, Omar and Jamila, who raised me and my siblings to love science and be always the best that we can be, today they harvest what they planted. I would like to thank my brothers and sisters who believed in me. I also would like the too few people who stood by me when I decided to step out of my comfort zone when too many people doubted me at the time.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Renewable Energy

The concept of renewable energy is not new. In fact, the commitment to developing technologies that do not depend on fossil fuels has been a growing trend for last three decades. Leaders and scientists began introducing programs and inventions to help create a new energy generation. In the 1970s, the United States faced an energy crisis that was alarming to the government and that raised awareness of the danger of depending on fossil fuels in generating energy. For example, President Carter (1977) emphasized the need to create permanent energy resources, and in particular, solar power. Since then, the United States has been concerned about the future of its energy security as part of its national security.

Under the leadership of President Obama in 2009, the United States increased the production of solar power and wind capacity dramatically, increasing from 1.2 gigawatts (GW) in solar capacity to 31 GW and from 25 gigawatts GW in wind capacity to 75 GW by 2017 (Sargent, 2017). Energy security is a critical part of national-security stability of any nation. Hence, the U.S. government has been targeting efforts to reduce renewable-energy costs. For example, the U.S. government invested $90 billion in renewable energy with the American Recovery and Reinvestment Act of 2009 (The White House, 2016). Also, in 2015, Congress extended a tax cut to support renewable-energy industries, which was considered a major victory for solar and winds companies. This legislation allowed solar companies to claim federal tax cuts at 30% of solar array prices (Sweet, 2015). As a result, solar-power-installation costs dropped by about 50% since 2009, as shown in Figure 1 (Sargent, 2017; Weiner, 2016). Therefore, investments in and advancing the use of renewable energy is significant.

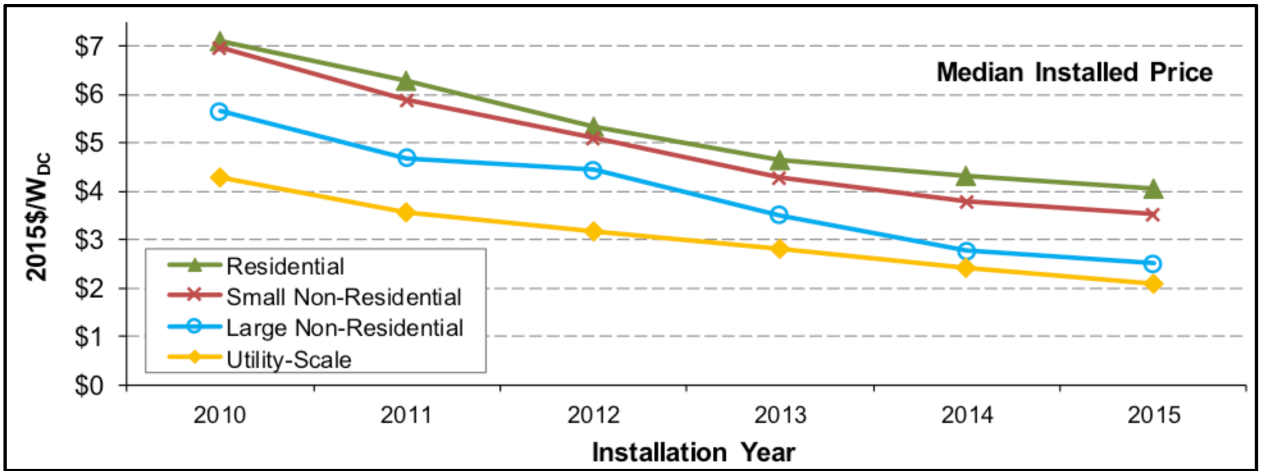


Figure 1. Installation cost timeline Residential, Small Non-Residential, Large Non-Residential, and Utility-Scale.

Source: *Median Installed Price of Solar in the United States Fell by 5–12% in 2015,* by J. Weiner, 2016, retrieved from https://newscenter.lbl.gov/2016/08/24/median-installed-price-solar-united-states-fell-5-12-2015/

The U.S. government is a major contributor to the global renewable-energy transformation, partnering with the European Union (EU), whose attitude toward renewable energy has been highly visible. The European Commission set various targets to "help combat climate change and air pollution, decrease its dependence on foreign fossil fuels, and keep energy affordable for consumers and businesses" (European Commission, n.d.). As a result, the European Commission set a roadmap for EU countries to have 20% of their overall energy mix as renewable energy by 2020 (Götz et al., 2016).

According to Eurostat, Directorate-General of the European Commission (n.d.), shares of renewable energy in the EU have already reached 17% of the 2020 target in 2016. The article listed 11 EU member states that have also reached the level of their national 2020 targets: Bulgaria, the Czech Republic, Denmark, Estonia, Croatia, Italy, Lithuania, Hungary, Romania, Finland, and Sweden (see Figure 2).

Similarly, as the world's largest energy consumer. China has committed to investing in renewable energy to meet the needs of its 1.4 billion (Worldometers, 2018) people. Despite all

these challenges, China has been developing sustainable technology on a large scale. China has a vast solar and wind manufacturing capacity; more than any other country in the world (Fialka, 2016). As a result, according to the Global Energy & CO2 Status Report from the International Energy Agency (2019), China's solar photovoltaic (PV) capacity exceeded 35 GW, which is equivalent to the total solar PV capacity of France and Germany combined, whereas India doubled its 8 GW solar PV capacity from 2016 to 2017.

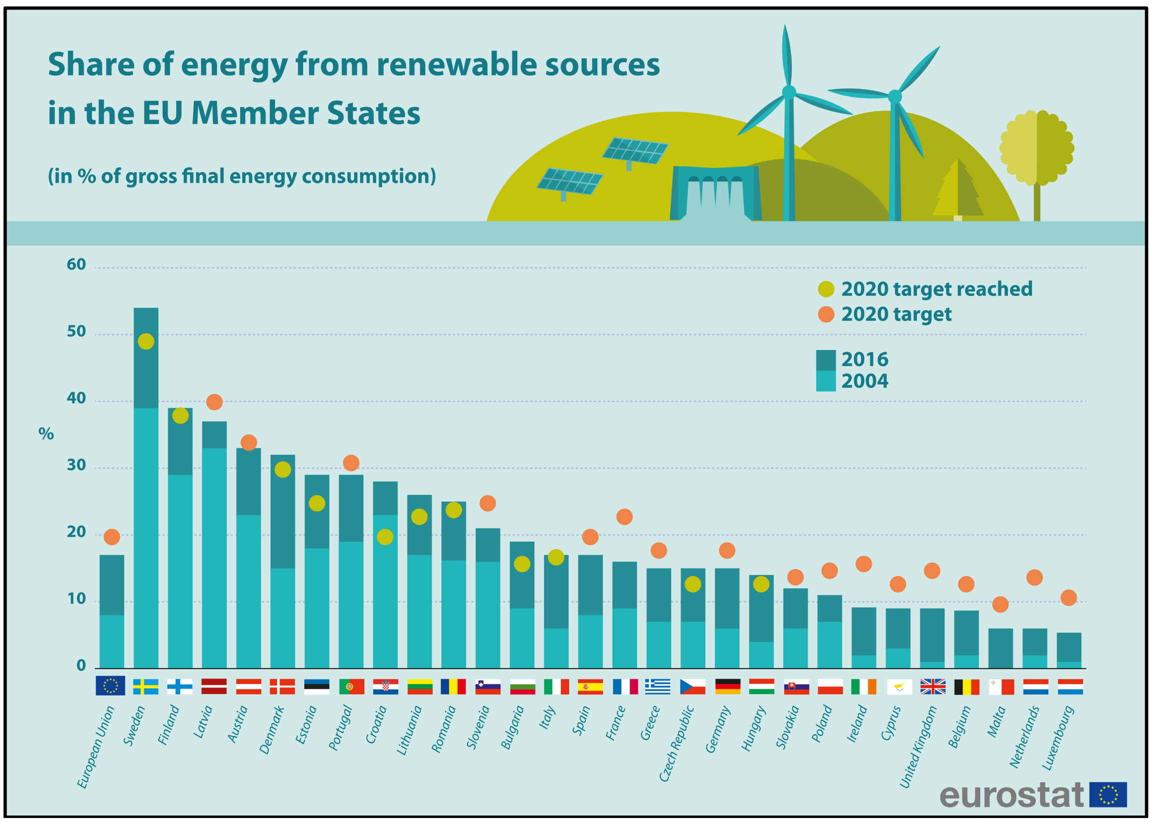


Figure 2. Shares of energy of renewable resources.
Source: *Eurostat Share of energy from renewable sources (in % of gross final energy consumption), Eurostat, 2017*, retrieved from https://ec.europa.eu/eurostat/statisticsexplained/index.php/Renewable_energy_statistics/)

Based on the initiatives and programs of these countries and others, concerns persist that governments are not striving to dedicate sufficient resources and technologies to improve renewable-energy capabilities to decrease dependence on coal and fossil fuels and promote the use of sustainable energy in commercial, residential, and industrial settings. Solar PV is the focus

of this dissertation research as it is the “most” renewable-energy technology available and individuals can adopt it on a small scale.

The motivation behind the incremental focus on renewable-energy diffusion is not merely driven by economics and politics. Global climate change has been a real threat and could lead to an environmental crisis as a result of the excessive use of fossil-fuel energy (Powell & Hill, 2009). Dincer (2000, p. 158) states that “Problems with energy supply and use are related not only to global warming, but also to such environmentalist concerns as air pollution, acid precipitation, ozone depletion, forest destruction, and emission of radioactive substances.” Moreover, rapid global population growth is increasing the complexity of environmentalist concerns, increasing the demand for greater energy consumption. According to the *World Population Prospects* (United Nations, 2017) the world’s population growth rate is 1.1% per year. In other words, by 2030, the world population will increase by 1 billion to be 8.6 billion; the United Nations projects the world population will reach 9.8 billion in 2050 (see Figure 3).

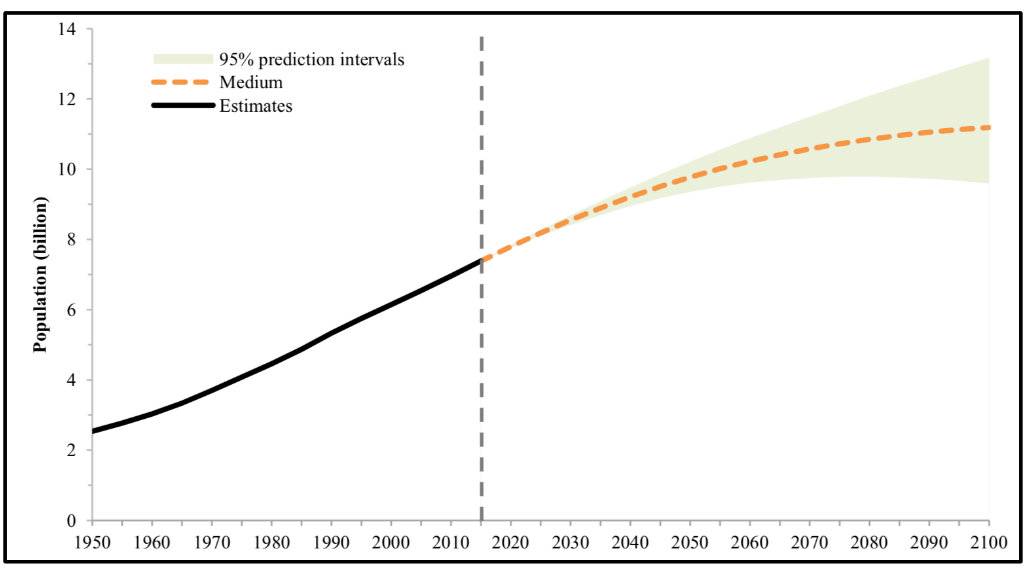


Figure 3 Population of the world: Estimates 1950–2015; projection: 2015–2100
Source: *World Population Prospects,* by United Nations, 2017, retrieved from https://population .un.org/wpp/

The National Aeronautics and Space Administration's (NASA's) department of Global Climate Change (National Aeronautics and Space Administration, 2018) reported that carbon dioxide (CO_2), released from human activities by burning fossil fuels, has reached the highest level in current history (see Figure 4). Additionally, NASA's Goddard Institute for Space Studies (2017) reported that the 2016 annual average global temperature was the warmest since 1880.

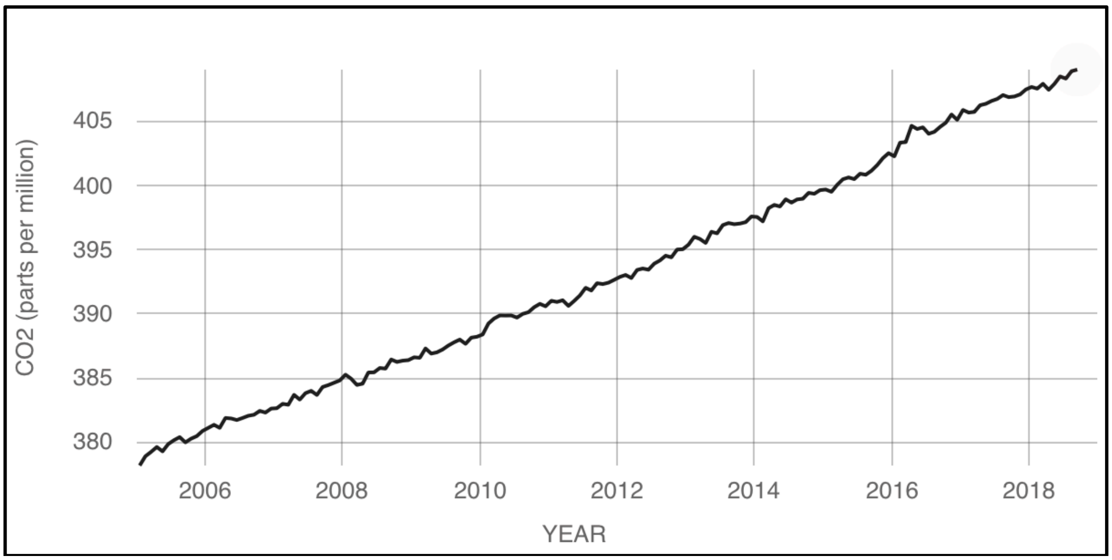


Figure 4. 1 Carbon dioxide parts per million 2006–2018.
Source: *Global Climate Change: Vital Signs of the Planet,* by National Aeronautics and Space Administration, 2018, retrieved from https://climate.nasa.gov

Mitigating this problem is becoming a necessity, and requires involvement from governments, business owners, and individuals. Many academics and governmental energy laboratories suggest the solution is to adopt clean energy by using renewable-energy technologies. In this research, I examine solar PV adoption behavior at the residential level by providing greater understanding of the factors that can help explain renewable-technology adoption.

Solar PV

Edmund Becquerel, in 1839, noticed that a material that was accidentally exposed to light generated electric current and in 1915, Einstein, in his special relativity theory, explained the

characteristics of light. These findings led to attempts to develop PV modules to generate electricity to power the satellites in NASA's space programs in the 1950s (Knier, 2008). Throughout the years, the technologies of solar PV advanced and became less costly. The global solar market reached almost 98 GW of new capacity in 2017, which is a significant growth compared to the global market in 2016 with 76 GW (International Energy Agency Photovoltaic Power Systems Programme, 2018).

Solar power is a valuable resource that produces an enormous amount of sustainable energy. In fact, "the amount of sunlight that strikes the earth's surface in an hour and a half is enough to handle the entire world's energy consumption for a full year" (Solar Energy Technology Basics, 2013, para 1). More importantly, people can install PV solar system technology on residential rooftops where it transforms sunlight into electricity without the need for transmission lines or other costly equipment (Kwan, 2012). Conveniently, solar panels come in various sizes and shapes, making them easy to install on residential, industrial, or commercial rooftop buildings.

To face the environmental risk of fossil-fuel consumption, governments and policymakers have introduced many programs to incentivize renewable electricity generation to finance home-installation projects and big industrial projects. In the United States, the federal government provides many incentive programs such as the investment tax credit, and states offer incentives in the form of solar renewable-energy credits. These programs vary among states. For example, California's Residential Property Assessed Clean Energy, funded with $1.8 billion, aligned with a PV-deployment increase at the residential level (Deason & Murphy, 2018). Policymakers usually evaluate the effectiveness of these incentives to ensure they achieve their renewable-energy transformation targets. However, some incentive programs were less

successful and did not achieve the anticipated results. For example, Matisoff and Johnson (2017) evaluated the value of over 400 state and utility incentive programs and concluded that the "Results suggest that approximately 67% of state and utility incentives, up to $1.9 billion over 11 years, were likely spent on incentives that did not increase residential solar PV installations" (p. 44). Similarly, Sarzynski, Larrieu, and Shrimali (2012) explained that "states offering tax incentives such as income, sales, or property tax incentives did not experience systematically stronger deployment of PV technology over the study period than states not offering tax incentives" (p. 555).

Research Problem and Questions

This research study outlines a plan to shed light on the importance of the use of employing renewable-energy technologies, explaining the consequences associated with inadequate solar adoption rates, and clarifying the benefits connected with an overall increase in solar PV adoption. This research has value for various stakeholders (e.g., utility companies, policymakers, service providers homeowners, and researchers). For example, electric utilities are under pressure from the North American Electric Reliability Council, to improve grid reliability, which is a major challenge, especially with the incremental demand for electricity. According to the U.S. Department of Energy, employing renewable-energy technologies such as solar PV would effectively contribute to grid reliability and resiliency. Thus, this study can help electric utilities better identify locations with high expected adoption so they can more effectively allocate resources to react to electric power fluctuations in the grid. Similarly, policymakers can better understand what drives households to adopt solar PV, which in turn, would help them better design regulations and evaluate existing incentive programs. Specifically, this research provides valuable insights to solar panel marketers in California, particularly to improve their

market-targeting efforts. California is the leading solar market in the United States. In 2017, California ranked first nationally, with about 934 installers/developers, 293 manufacturers, and 16.68% of Californian state electricity production from solar (Solar Energy Industries Association, 2019).

To address the challenges facing policymakers, utility companies, homeowners, and others involved in solar PV adoption, this study investigated expenditures, demographics, political preferences, and geographic and environmental factors and their influence on residential solar PV adoption. Although researchers have studied residential solar PV adoption, they have not statistically tested the overall influence of the noted factors on the number of existing solar PV installations by census tract at the residential level. Moreover, extant research showed limited use of GIS in examining spatial factors related to solar PV adoption. This research aims to exploit the capabilities of the geographic information system (GIS) in addressing the research problem (e.g., creating spatial indices, visualization, spatial analysis, or building a decision-support system). Thus, to solve this research problem, and to address the identified literature gaps, this research answers the following questions:

- **Question 1**: How might factors such as expenditures, demographics, political preferences, and geographic and environmentalist indicators predict solar adoption at the residential level?
- **Question 2**: How might the derived solar adoption algorithm be instantiated into a Spatial Decision Support System (SDSS) to provide stakeholders with locational adoption characteristics?

Research Contribution

In this research study, I sought to provide two contributions. First, I created an algorithm that explains the many aspects that provide a better understanding of solar PV adoption and the essential characteristics from social, environmental, spending, and geographical perspectives. This information contributes to the knowledgebase of the energy, environmental, consumer, and behavioral domains. Additionally, this research contributes to the information system (IS) literature and particularly energy informatics and as a response to the call of Watson, Boudreau, & Chen (2010), who introduced Green IS to address the need to focus on environmental–sustainable-development issues. To the best of my knowledge, this is the first attempt to provide an empirical evidence-based approach to explain solar adoption at the residential level, incorporating spatial and nonspatial indicators. Second, this research extends the knowledgebase of decision-support-system literature as an instantiation of a spatial solar adoption index. In addition, this research contributes by adding practical value to stakeholders, utility companies, policymakers, service providers, homeowners, and researchers through the development of an SDSS that allows them to use the output of the algorithm to visualize and interact with a map of areas with high/low potential for solar adoption. This SDSS will enable users to perform interactive scenario-based roleplaying that can serve their purposes in making better decisions.

Chapter 2: Literature Review

Related and Existing Research

Many researchers investigated the various factors that may explain why people adopt solar technologies and particularly solar PV. Scholars from many disciplines have shed light on solar power adoption; for example, researchers from energy (renewable energy), engineering, social studies, consumer behavior, and information systems have used various research methods, answered various research questions, reported their findings and limitations, and suggested some future guidelines for other promising research areas.

Wang, Guan and Wu (2017) used the Heckman model approach to study factors that affect renewable-energy technology, particularly solar water heaters. Wang et al. used Heckman's model as a theoretical background to describe whether people would adopt such technology and what usage level should they adopt. Wang et al. gathered quantitative data from participants in Jiangxi Province, China, to predict homeowners' (n = 972) willingness to adopt solar water heaters.

Wang et al. (2017) found that "the awareness of solar water heater (SWH) technology did not significantly impact willingness to adopt, though rural residents with better knowledge about SWHs use them more, once they have adopted SWHs" (p.1317). Although the study lacked an explanation of the geographic factors that influence renewable-energy adoption, the importance of including spatial data, which would provide better understating of solar systems' adoption by recognizing spatial patterns, would add value to the research findings.

Similarly, Kwan (2012) conducted a study, building a regression model that predicts solar PV installations in residential areas by zip code across the United States. Kwan included variables such as education level, income, authenticity, incentives, unit density, number of units,

cost of electricity, and political preferences. Additionally, Kwan presented a map predicting the spatial distribution of residential solar PV shares of zip codes. Many demographic variables had a statistically significant impact on solar adoption. However, people's perspectives toward environmentalist matters, such as their perspective on global warming or green energy issues, was missing from Kwan's model. The present study addressed this missing information.

For this study, I also considered a more granular unit of analysis than the zip code, namely the census tract, for many reasons. Census tracts are well-defined geographic areas and align coterminously to county boundaries. Also, census tracts provide more statistical homogeneity because census tracts average 4,000 in population, whereas a single zip code may exceed 10,000 (ProximityOne, 2019).

Moreover, to explore reasons that may contribute to a person's reason for adopting renewable-energy technologies, Mills and Schleich (2009) assessed the effects of geographic, residential, and household characteristics on the adoption of solar-thermal water and space-heating technologies in Germany. The authors used a quantitative research approach to conduct a study in which they used descriptive analysis, comparing means between three groups (n = 12,331): (a) households adopting neither solar-water-heating nor solar-space-heating systems, (b) households adopting solar-water-heating systems, and (c) households adopting solar-space-heating systems. Mills and Schleich obtained datasets based on a mail survey of private-sector household energy consumption conducted in December of 2002. Research findings included that low energy cost-saving limited the diffusion of residential solar-thermal technologies. The researchers advised that studying wealthy-household attributes and preferences can lead to a better explanation of why households may adopt energy-saving technologies. In this

study, I added many variables centered on spending habits on various commodities, explained in detail in a different section of the research.

Palm and Eriksson (2018) suggested that to increase the number of PV installations, people need to better understand PVs. Palm and Eriksson claimed that households face an information barrier that affects adoption of PVs. Thus, they conducted an in-depth qualitative analysis, interviewing members of 58 households in three studies from 2013 to 2016. Participants were nonadopters of PVs, adopters of PVs, and a mix of adopters of PVs and nonadopters of PVs. Palm and Eriksson found different levels of household-adoption rates with respect to information absorption and that people act in various ways. For example, nonadopters tended to find information on PVs to be too complicated and were less likely to adopt PVs. In contrast, environmentally engaged adopters looked for the correct information to help learn and understand PVs. In sum, when organizing households into various groups to detect the most suitable information that fit their needs, it was most effective to emphasize the tendency to adopt PV technologies (Palm & Eriksson, 2018). This finding confirmed that households that care about the environment and have higher education levels align positively with renewable-energy technologies (Brechling & Smith, 1994; Weber & Perrels, 2000).

Similarly, in an in-depth study across the state of Wisconsin, Schelly (2014) aimed to examine what motivates homeowners (*n* = 48) to adopt residential solar electric technology, including concerns about environmental motivations, economic considerations, and demographic characteristics that influence the adoption and diffusion of innovations. Study findings suggested that, based on the characteristics of Wisconsin homeowners, environmentalist values and economic value return are not sufficient to motivate adoption (Schelly, 2014). Qualitative research methods provided rich insight and an explanation of what Wisconsin homeowners

believed motivated them to adopt residential solar electric technology; however, these finding cannot be generalized. Generalization is the extent to which findings apply to other groups (Robson & McCartan, 2016). Thus, in this study, I developed an algorithm that could be operationalized using data from various counties, states, and cities. This type of generalization can be classified as an empirical statistical generalization, as Tsang (2014) described.

Theoretical Framework

This research draws its theoretical grounding from the energy informatics (EI) framework proposed by Goebel et al. (2014; see Figure 5). To illustrate, the EI framework can guide scholars from computer science (CS) and IS on "how the respective research questions are broken down into specific research projects and how EI researchers have made contributions based on their academic background" (Goebel et al., 2014). Based on this conclusion, the purpose of this research fits under the Energy Efficiency goal and Smart Grid category which targets renewable energy in residential buildings as highlighted in Figure 5.

The research also drew on diffusion of innovations theory (DOI) to provide a better understanding of solar PV adoption diffusion characteristic. DOI explains how, over time, an idea or product diffuses in populations or social systems (Rogers, 2003). The theory categorizes the population into five groups: innovators, early adopters, early majority, late majority, and laggards, with each category distinct from the other. For example, innovators are people who are first to try the innovation and are willing to take risks; in contrast, the late majority would not try the innovation until the majority has effectively adopted it. DOI theory helped provide an explanation by considering the nearby influence of adopters, by examining the adoption density in the study area. The research also drew on the theory of planned behavior to capture how an individual's values, norms, and beliefs drive positive behavior toward the environment (aligned

with Ajzen, 1991). The theory has five components that collectively represent a person's actual control over their behavior: attitudes, behavioral intention, subjective norms, social norms, and perceived power.

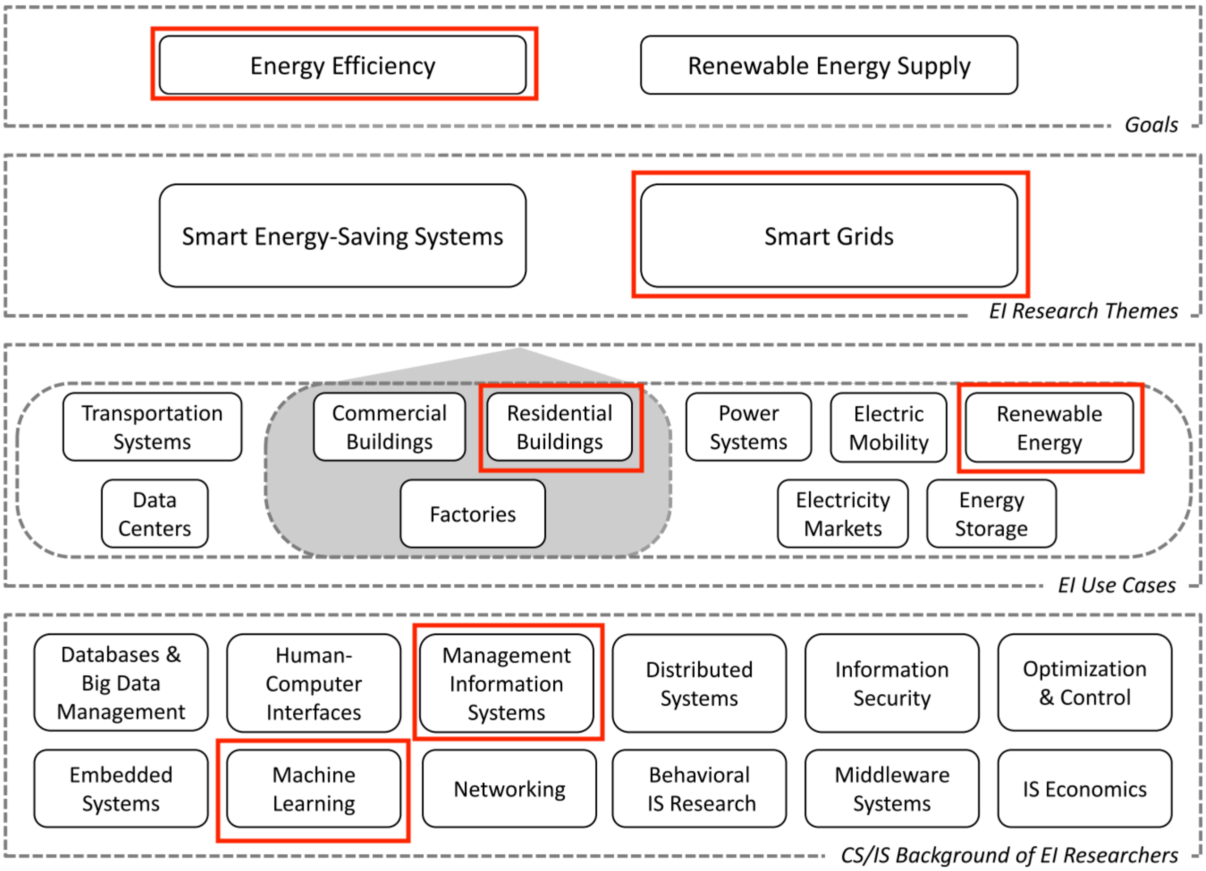


Figure 5. Energy informatics research scope

Note. Source: "Energy Informatics," by C. Goebel, H.-A. Jacobsen, V. Razo, C. Doblander, J. Rivera, J. Ilg, … J. Lässig, 2014, *Business & Information Systems Engineering, 6,* p. 26, https://doi.org/10.1007/s12599-013-0304-2

Chapter 3: Methodology Plan

Research Approach

In this dissertation, I addressed the problem of renewable-energy adoption, particularly solar PV. Due to different characteristics and motivational goals in solar adoption on the commercial and industrial levels, this research focused on the factors related to homeowners, that is, at the residential level. Thus, I suggest a framework that indicates the factors that may explain renewable-energy technology adoption (see Figure 6). This framework categorizes variables that share similar attributes. For example, I gathered demographic characteristics under key demographic indicators (see Table 1) whereas I grouped values and norms indicators toward environmental issues and perspectives under environmentalist indicators (see Tables 2 through 5). Correspondingly, I grouped all other indicators to represent a factor, shown in Figure 6. Tables 2 through 5 summarize factors derived from previous research and from different disciplinary fields. Here, I used exploratory research in the sense that I added and tested new factors to previously validated variables that affect solar adoption rates. For example, I used GIS to identify new variables and test their contributions to solar adoption, as explained in the artifacts section.

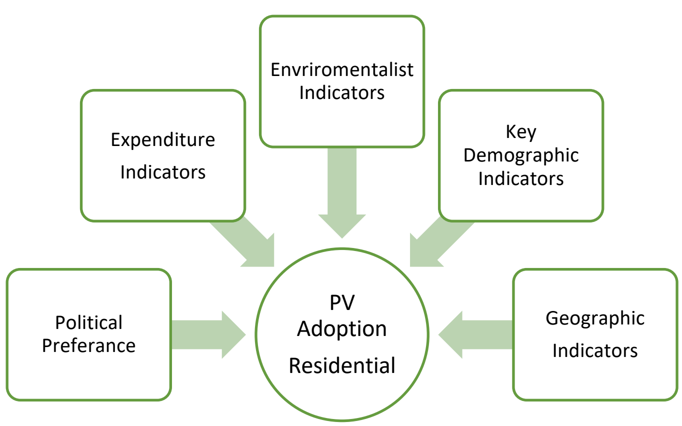


Figure 6. Solar adoption factors framework.

Table 1. *Key Demographic Indicators*

Number	Variables	Resource
1	Income	Kwan, 2012; Mills & Schleich, 2009
	Average household income	
	Median household income	
	Household income less than $15,000	
	Household income $15,000–$24,999	
	Household income $25,000–$34,999	
	Household income $35,000–$49,999	
	Household income $50,000–$74,999	
	Household income $75,000–$99,999	
	Household income $100,000–$149,999	
	Household income $150,000–$199,999	
	Household income $200,000 or greater	
2	Home value	Kwan, 2012
	Average home value	
	Median home value	
	Home value less than $50,000	
	Home value $50,000–$99,999	
	Home value $100,000–$149,999	
	Home value $150,000–$199,999	
	Home value $200,000–$249,999	
	Home value $250,000–$299,999	
	Home value $300,000–$399,999	
	Home value $400,000–$499,999	
	Home value $500,000–$749,999	
	Home value $750,000–$999,999	
	Home value $1,000,000 or greater	
3	Educational attainment	Lutzenhiser, 1993; Mills & Schleich, 2009; Weber & Perrels, 2000
	Associate's degree	
	Graduate/Professional degree	
	GED/Alternative credential	
	High school diploma	
	Some college/No degree	
4	Marital status	
	Population Age 15+: Divorced	
	Population Age 15+: Married	
	Population Age 15+: Never Married	
	Population Age 15+: Widowed	
5	Housing units	Graziano & Gillingham, 2014
	Owner-occupied housing units	
	Renter-occupied housing units	
	Vacant housing units	
	Total housing units	
	Total households	
6	Population	Graziano & Gillingham, 2014
	Total population	
	Population density (population per square mile)	
	Family population	
	Average family size	
7	Net worth	
	Median net worth	
	Average net worth	
8	Race/Ethnicity	
	White population	
	Black/African American population	
	Hispanic population	
	American Indian/Alaska Native population	
	Asian population	
	Pacific Islander population	
	Other race population	

Table 2. *Expenditure Indicators*

Variables	Resource
Expenditure	Kwan, 2012
Spending on electricity: Average	
Spending on apparel and services: Average	
Spending on dining out: Average	
Spending on education: Average	
Spending on entertainment/Recreation: Average	
Annual budget expenditures: Average	

Table 3. *Political Preference Indicators*

Variables	Resource
Political party	Hsu, 2018; Kwan, 2012
Democratic: Average 2016 election voters	
Republican: Average 2016 election voters	
Other	

Table 4. *Geographic Indicators*

Variables	Resource
Yearly sunlight KW Per hour: Average	Dorvlo & Ampratwum, 1998; Guerrero, Hang, & Uceda, 2008; Robinson, Stringer, Rai, & Tondon, 2013; Tsoutsos, Gekas, & Marketaki, 2003; Zhang, Vorobeychik, Letchford, & Lakkaraju, 2016
Yearly sunlight KW Per hour: Total	
Nearby adoption affect: Index	
Carbon dioxide concentration	

Table 5. *Environmentalist Indicators*

Variables	Resource
Attitudes and Perspectives toward	Dastrup, Zivin, Costa, & Kahn, 2012; Hsu, 2018; Kwan, 2012; Zhang, Vorobeychik, Letchford, & Lakkaraju, 2016
Helping to preserve nature	
The government's focus on environmental issues	
An interest in how to help the environment	
Global warming	
Environmentally conscious compared to most people	
Buying from companies with environment record	
Paying more for environmentally safe products	
Valuing green products over convenience	

Data Acquisition, Tools, and Techniques

I sourced factors of key demographics indicators, spending indicators, and environmentalist indicators (see Table 5) from the datasets of ESRI's ArcGIS Business Analyst. The choice of ArcGIS Business Analyst was based on the available spatial data associated with each variable, and the accuracy of the data. ESRI data has the lowest total precision error compared to other vendors in the United States (Cropper, McKibben, Swanson, & Tayman, 2018).

Regarding geographic indicators, the data are publicly available on Project Sunroof Data Explorer, powered by Google. Also, I gathered the political-preference-indicator data from the election data archive Dataverse, which is available for public consumption. The most granular data available for the number of solar installation variable is the census tract; hence, I used the census tract as the unit of analysis for this study. The study area is Los Angeles County, chosen due to data availability and to the number of tracts. Los Angeles County includes 2,317 census tracts ranging from 6037101110 to 6037980033.

Proposed Artifact

Design Science Research

Design science research (DSR) is a research methodology that proposes prescriptive artifacts to improve information technology (IT) performance (Hevner, 2007; Hevner & Chatterjee, 2010). The main activity is iteratively designing and building the artifacts, which are in the form of constructs, models, methods, instantiations, and theory. Most importantly, DSR reflects on lessons learned and contributes to the IT body of knowledge. DSR has three iterative and interconnected cycle components: rigor, relevance, and design. The rigor cycle focuses on the idea that the research has to be informed and grounded in basic knowledge by looking to previous research on the specific topic, relies on a kernel theory that guides the research approach, and uses rigorous research-evaluation methods (quantitative, qualitative, or mixed-method approaches) to assess the proposed artifact. The rigor cycle connects to the design stage in an iterative process. The relevance cycle connects to the design such that the researcher detects research opportunities and adds value by solving the relevant problem. The design cycle is the core component where the researcher builds and evaluates an artifact in an iterative process.

In this research study, I developed two artifacts that use the three-component DSR cycle. First, I built a predictive model that operationalizes the proposed framework (see Figure 6). The factors were grounded in the literature and based on existing related work to test how spatial and nonspatial attributes explain solar energy adoption rates. Solar adoption is a highly relevant research opportunity that has value for organizations and households. I used various statistical analyses to build the predictive model, which was evaluated quantitatively by testing the model's goodness of fit. The second artifact was an instantiation of the first artifact that served as an

SDSS for stakeholders to visualize and explore locations of low- and high-solar energy adoption potential. I evaluated the instantiation effectiveness by conducting a mixed-methods study.

Artifact 1: PV Installations per Tract

The goal was to build a prediction algorithm that explains the independent variables' influence on the dependent variable, the number of existing PV installations per tract (Tables 1–5). Regression analysis was one of the more practical ways of understanding how the independent variables, individually and collectively, predict the dependent variable (see Table 6).

Analysis plan. I initiated this process by following the steps below:

1. Preparing data and generating all the variables.

Table 6. *Artifact 1: Sample Response Calculation for One Tract of Global Warming (N = 3,225)*

Items	Not a serious threat	A minor problem	A problem	A serious threat
Code	1	2	3	4
Responses	376	654	1,011	984
Calculated value (eq. 1)	0.12	0.43	1.00	1.30

Note: Tract number 06029003305.

- Environmentalist Indicators: The study area, Los Angeles County, has about 2,400 census tracts; each census tract has 31 items (see Table 7). For each person who responded to these items, I extracted their set of responses (values). However, the plan is to have one indicator for the entire tract to measure "environmental friendliness." To do that, I categorized environmental items into eight factors: (a) Preserve_Nature, (b) Government_Focus, (c) Help_Enviroment, (d) Global_Warming, (e) Environmental_Consciousness, (f) Company_Environmental_Record, (g) Buy_Environmentally_Safe Products and (h) GreenProducts_vs_Convenience. Then, to reduce the number of items to have only one score per variable by census tract, I coded the items per variable from 1 to 4,

as shown in Table 8. Then, I applied the following formula to each variable to so that each represents one variable per census tract:

$$\left(\frac{total\ responses\ per\ single\ item}{overall\ repsonses}\right) * coded\ score\ per\ item. \tag{1}$$

Depending on the numbers of items for each variable, the maximum number represents the majority of responses for that tract, considered its environmental-friendliness attitude. For example, Table 6 provides a sample response calculation of one tract for Global_Warming. As a result, the perspective of tract number 06029003305 regarding global warming is that “Global warming is a serious threat” because the majority of responses is represented by the calculated score of 1.30.

Table 7. *Environmentalist Items*

1.	Helping to preserve nature is not important
2.	Helping to preserve nature is of average importance
3.	Helping to preserve nature is very important
4.	Government shouldn`t focus more on environmental issues
5.	Government should focus slightly more on environmental issues
6.	Government should focus somewhat more on environmental issues
7.	Government should focus more on environmental issues
8.	I am not interested in how to help the environment
9.	I am slightly interested in how to help the environment
10.	I am somewhat interested in how to help the environment
11.	I am interested in how to help the environment
12.	Global warming is not a serious threat
13.	Global warming is a minor problem
14.	Global warming is a problem
15.	Global warming is a serious threat
16.	I am not more environmentally conscious than most people
17.	I am slightly more environmentally conscious than most people
18.	I am somewhat more environmentally conscious than most people
19.	I am more environmentally conscious than most people
20.	A company`s environmental record is not important when I am considering buying its products
21.	A company`s environmental record is slightly important when I am considering buying its products
22.	Company`s environmental record is somewhat important when buying
23.	Company`s environ record important when buying
24.	I rarely will pay more for an environmentally safe product
25.	Occasionally pay more for environmentally safe product
26.	Frequently pay more for environmentally safe product
27.	Usually pay more for environ safe product
28.	I rarely value green products over convenience
29.	I occasionally value green products over convenience
30.	I frequently value green products over convenience
31.	I usually value green products over convenience

- Geographic Indicators: Carbon dioxide concentration represents the potential carbon dioxide abatement of the solar capacity that meets the technical potential criterion, which is used as an indication for installation suitability in a certain census tract.

These data are publicly available on Project Sunroof's Data Explorer, powered by Google. However, I extracted nearby adoption-rate affects as a weighted value using a density-based clustering tool in ArcGIS 10.6.1. The tool is used to find clusters of point features, based on their spatial distribution. I calculated the average number of installations by tract, identifying tracts with an above-average number of installations. I used a clustering tool to highlight nearby tracts with an above-average number of installations as one cluster. Hotspot clustering indicates nearby tracts were likely to have a nearby adoption affect.

Table 8. *Environmentalist Variables and Scores*

Variable \ Code	1	2	3	4
Preserve_Nature	Helping to preserve nature is not important	Helping to preserve nature is of average importance	Helping to preserve nature is very important	
Govt_Focus	Govt shouldn`t focus more on environmental issues	Govt should focus slightly more on environ issues	Govt should focus somewhat more on environ issues	Govt should focus more on environmental issues
Help_Enviroment	Am not interested in how to help the environment	Am slightly interested in how to help the environ	Am somewhat interested in how to help the environ	Am interested in how to help the environment
Global_Warming	Global warming is not a serious threat	Global warming is a minor problem	Global warming is a problem	Global warming is a serious threat
Environ_Conscious	Am not more environ conscious than most people	Am slightly more environ conscious than most	Am somewhat more environ conscious than most	Am more environ conscious than most people
Company_Environ_ Record	Company`s environ record not important when buying	Company`s env rec slightly important when buying	Company`s env rec somewhat important when buying	Company`s environ record important when buying
Buy_Environ_ SafeProducts	Rarely pay more for environ safe product	Occasionally pay more for environ safe product	Frequently pay more for environ safe product	Usually pay more for environ safe product
GreenProducts_vs_ Convenience	Rarely value green products over convenience	Occasionally value green products over convenience	Frequently value green products over convenience	Usually value green products over convenience

- Political Preference Indicator: These data are from the election-data archive, Dataverse. Dataverse contains the 2016 presidential voting outcomes by precinct,

which is a county subdivision used as a voting district. A precinct does not exactly correspond to a census tract; thus, I geo-spatially joined these data to the corresponding census tract. In order to make the correspondence between census tract and voting result map, ArcMap was used to overlay the voting map by precinct with census tract maps of Los Angeles County. Because I assigned all variables by tract, I assigned a political-preference value to each tract. I identified each tract as to whether it is Democratic or Republican, based on its number of votes.

- Demographic Indicators: I gathered data for these indicators from ArcMap Business Analyst. I applied the preparation steps outlined in this section.
- Expenditure Indicators: Similarly, I gathered data of various expenditure types through ArcMap Business Analyst.

2. Data preparation: I used MS Access 2016 for data organizing and cleaning, and to perform simple calculations. I integrated spatial and nonspatial data into one table to perform further processing steps. After conducting a spatial analysis on the variables, I extracted it into an Excel file for use in additional applications, such as SPSS. The data-cleaning process included checking data normality to ensure a linear relationship between the variables and, if necessary, by deleting outliers and records with a missing value. In the case of environmentalist indicators, I used SPSS to test data reliability to ensure data consistency. I set Cronbach's alpha at 0.7 or higher to be acceptable. I tested correlation between environmentalist indicators and did not accept variables with correlation higher than 0.8 to control for multicollinearity. As part of the analysis plan, I used stepwise multiple regression in SPSS to test the independent variables (environmentalist variables) on the dependent variable (the

number of existing installations). Afterwards, I tested all other variables shown in Figure 5 for multicollinearity and then created a prediction model. I report and discuss betas, F-statistics, the multiple correlation coefficients (R), and multiple coefficients of determination (R^2).

Evaluation criteria. Design evaluation is a crucial component in DSR and must show utility and efficacy of the proposed artifact. Thus, to demonstrate the quality of the output model, I examined how well the data fit the model. A well-fitting regression model is one that predicts values close to observed values. I performed a model goodness of fit assessment using "studentized" residuals. In this case, 95% of standardized residuals would lie between ± 2. This technical evaluation relied on SPSS to perform the model goodness of fit, and I also needed to perform further model-fit assessments, such as standardized residuals and Cook's distance.

Artifact 2: Residential Solar Energy Adoption

I instantiated the Residential Solar Energy Adoption (RSEA) algorithm as an artifact in the form of an SDSS, the instantiation was a GIS- and web-based application that enables stakeholders (e.g., utility companies, policymakers, service providers, homeowners, and researchers) to navigate across locations on a map interactively. I term this artifact RSEA-SDSS. To illustrate, the RSEA-SDSS maps highlight locations with high- and low- solar energy adoption potential. Additionally, the web application offered different types of visualization results (e.g., charts, tables, and maps), and decision-support functions to actively inform stakeholders and to help them make better decisions regarding plans, incentive programs, and marketing strategies. For example, users can select a tract with high potential solar energy adoption and visualize the variables that contributed most to the prediction. In addition, they can compare it to a nearby area with a different adoption level.

Evaluation criteria. In order to understand and evaluate RSEA-SDSS's effectiveness and utility, I used a mixed-method strategy with a concurrent triangulation design, which included a descriptive quantitative analysis using the System Usability Scale (SUS) and the User Experience Questionnaire (UEQ); also, I collected feedback through end-user interviews.

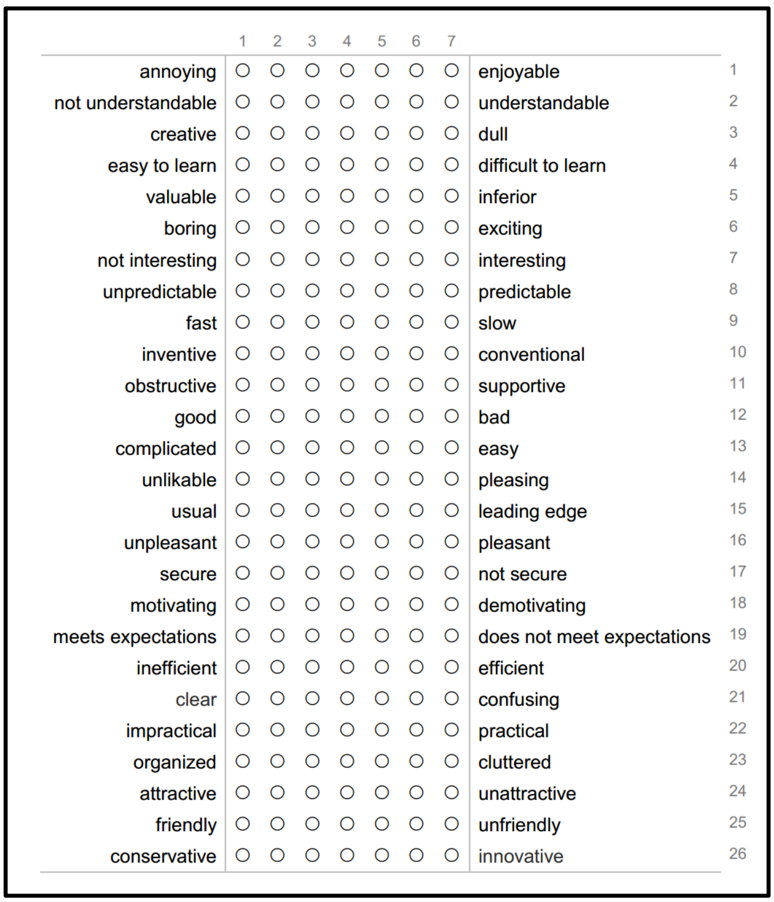
The SUS is a widely used validated instrument (Bangor, Kortum, & Miller, 2008). Brooke (1996) designed SUS to cover: (a) Effectiveness: the ability of a user to complete system tasks with high quality output, (b) Efficiency: the level of resources consumed to perform the tasks, and (c) Satisfaction: the user subjective reaction to using the system. SUS is a 10-item questionnaire that has five response options (strongly agree to strongly disagree) as follows:

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

The UEQ is a valid measurement of user experience (Laugwitz, Held, & Schrepp, 2008). Researchers have widely used the UEQ scale to measure user experience with a product. The main objective of the UEQ is

to allow a quick assessment done by end users covering a preferably comprehensive impression of user experience. It should allow the users to express feelings, impressions, and attitudes that arise when experiencing the product under investigation in a very simple and immediate way. (Schrepp, Hinderks, & Thomaschewski, 2017a)

As in Figure 7, the UEQ is a semantic differential scale that consists of 26 items that capture six components as outlined in Appendix A (Schrepp, Hinderks, & Thomaschewski, 2017a, p. 2).



	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

Figure 7. The User Experience Questionnaire (UEQ).

After receiving approval for this study from the Claremont Graduate University Institutional Review Board, I designed a use-case scenario for a participant to navigate through the artifact and perform specific tasks. I planned to recruit at least 30 participants, who could be

considered potential stakeholders, to perform the study evaluation. The potential participants included employees from solar energy firms in California, using the database from the *gosolarcalifornia.com* website; members of *Sustainable Claremont*, which is a nonprofit organization that engages people in education and action to create a more sustainable community; and representatives from utility companies. The aim of seeking participants from these categories was to recruit representative user populations to give expert feedback and assessment. I sent an e-mail to each potential participant explaining the purpose of the study and the scenarios to be performed on the websites. If they agreed to participate, I sent a follow-up e-mail with the questionnaire link.

Interviews offer the potential for rich, in-depth insights that a questionnaire may not offer. Accordingly, I interviewed some experts who have in-depth knowledge of solar energy adoption to evaluate the proposed artifact's effectiveness. I drew participants from a convenience sample of individuals who used the artifact. Interviews were semi-structured to allow for emerging questions. The interviews were transcribed, coded, and analyzed. I report and discuss common themes. During the interviews, I asked the questions outlined in Appendix B

Modeling Solar Energy Adoption at the Residential Level

Artifact 1: PV Installations per Tract aimed to operationalize the proposed framework, shown in Figure 5 (see page 14), to provide a predictive model regarding solar adoption at a certain location. I conducted the analysis by utilizing two methods: multiple regression and a machine-learning forest-based classification and regression technique. Multiple regression is an advanced statistical technique that examined the effect of multiple independent variables on one dependent variable. Forest-based classification and regression serves a similar purpose by

creating bootstrap samples of training data and random-feature selection in tree induction. I set up the modeling process through the following steps:

Step 1. Data Preparation

This step covers data generation, data transformation, and data aggregation for all of the independent variables.

- **Environmentalist Indicators:** As indicated earlier, I extracted 31 items to represent each census tract. Then, I categorized environmentalist items into eight variables and coded the items as shown in Table 5. Using the formula described in Chapter 3, I calculated a single environmentalist attitude for that tract. See Table 6 for a sample.

- **Geographic Indicators:** These data are publicly available on Project Sunroof Data Explorer, powered by Google. For the nearby effect variable, I used the spatial analysis clustering tool in ArcGIS 10.6.1 named Cluster and Outlier Analysis. The tool identifies statistically significant hot spots, cold spots, and spatial outliers; and assigns weighted scores to the targeted features, accordingly. This process involves measuring the Anselin Local Moran's I Index, *z* score and *p*-value. The *z* scores and *p*-values are measures of the statistical significance between two neighboring features to determine the similarly and dissimilarity between them. In measuring the nearby tract effects with solar installations, the Los Angeles tracts polygon was the input feature class while the existing installation served as the input field. Because the analysis was performed on polygon features, the spatial relationship was specified as Contiguity_Edges_Corners; that is, polygon features that share boundaries, share a node, or overlap influence computations for the target polygon feature (see Figure 8). The computation output map of the areas shows significant clustering and non-

significant clustering as demonstrated in Figure 9; results are summarized in Table 9. Then, I created a nearby-effect variable and coded each tract based on the cluster level assigned to it.

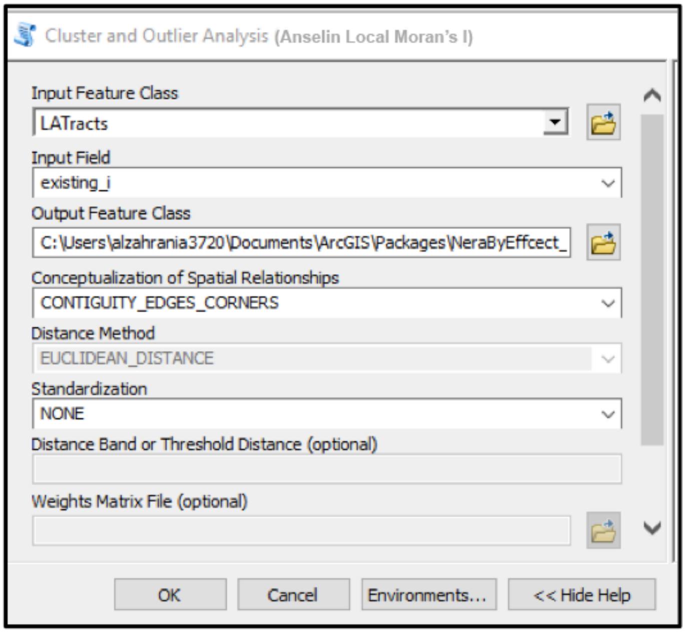


Figure 8. Cluster analysis.

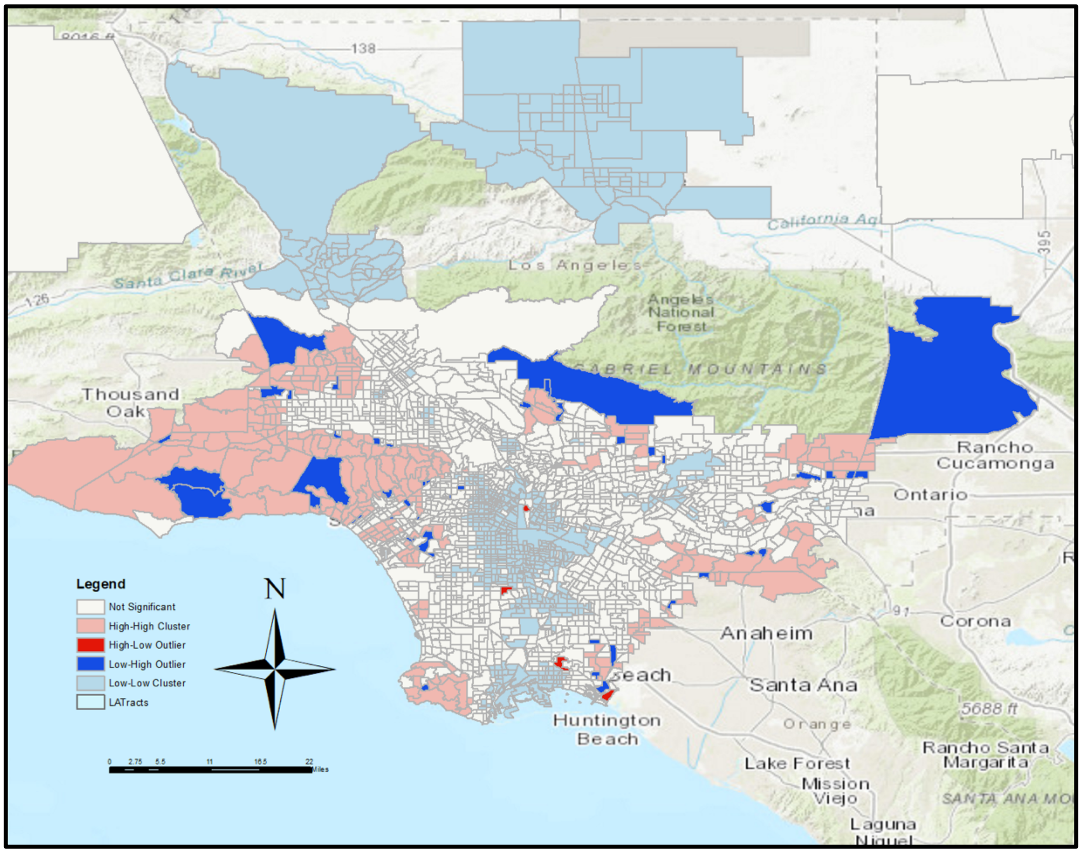


Figure 9. Significant clustering and non-significant clustering result map.

Table 9. *Coding Census Tracts Based on Cluster Level*

Code	Cluster level	Number of tracts	Interpretation
5	High–High (HH)	252	Statistically significant cluster of high number of solar installations from a value that one would expect from a random distribution
4	High–Low (HL)	4	High number of solar installations surrounded by low numbers
3	Low–High (LH)	46	Low number of solar installations surrounded by high numbers
2	Low–Low (LL)	622	Statistically significant cluster of low number of solar installations from a value that one would expect from a random distribution
1	Not significant	1,446	No statistically significant cluster from a value that one would expect from a random distribution

- **Political-Preference Indicator**: Data accrued from Election Data Archive Dataverse. Data contain the 2016 presidential voting outcomes by precinct, which is a county subdivision used as a voting district. A precinct does not exactly correspond to census tract boundaries; thus, I used ArcGIS 10.6 to covert voting data to feature points and

spatially joined them with Los Angeles tract layers to correspond with the census tract. The correspondence creation used ArcMap to overlay voting map by precinct with census tracts map of Los Angeles County. After all votes of the 2016 presidential election were generated to Los Angeles census tract boundaries, I compared votes for Hillary Clinton to votes for Donald Trump. I labeled tracts with a majority voting for Clinton Democratic with blue color and similarly, I labeled those with a majority voting for Trump Republican, as shown in Figure 10. The total number of Democratic tracts were 2,106, whereas the area had only 80 Republican tracts. The total tracts with null values were 219 so I deleted them from further analyses. As a result, the political-preference indicator in the analysis shows that the total Republican tracts represent only less than 10% of the total number of tracts and caused no effect in the statistical-significance calculation. Therefor, the Political-Preference Indicator will be eliminated due to weak data variance.

- **Demographic Indicators:** I gathered data of the indicators listed in Table 1 from ArcMap Business Analyst, as described in the data-acquisition section. Then, I renamed the selected variables with more detailed names to ease addressing the variables in subsequent steps. For example, the total households with income level between $35,000 and $50,000 variable was renamed from HINC35_CY to Income_35K50K. The demographic attributes are listed in Table 10.

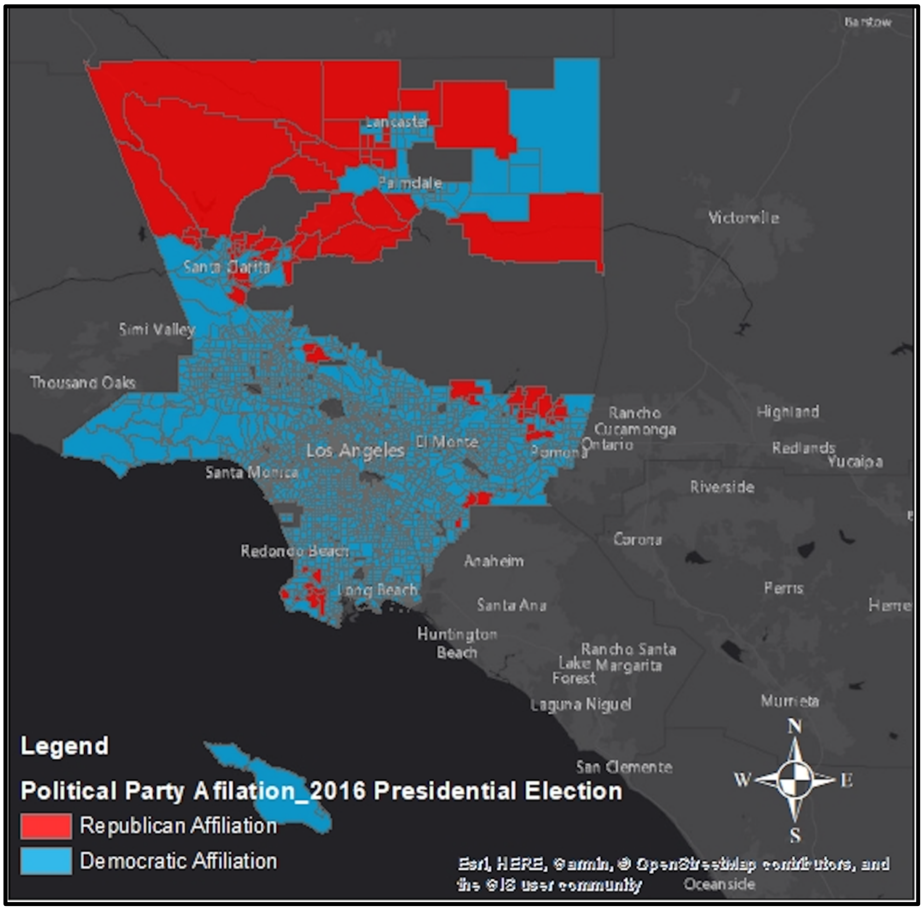


Figure 10. Democratic and Republican affiliation in Los Angeles County.

Table 10. *Demographic and Expenditure Attribute Variables*

	Demographic Variables		
AMERICINDIAN	Income_150K200K	OTHRACE	Value_250K300K
ASIAN	Income_15K25K	OWNERUnits	Value_2M
BLACK	Income_200K	PACIFIC	Value_300K400K
DIVORCD	Income_25K35K	POPDENS	Value_400K500K
EdeAlteernativeCred	Income_35k50K	RENTERUnits	Value_500K750K
EduASSCDEG	Income_50K75K	TOTPOP	Value_50k
EduBACHDEG	Income_75K100K	VACANTUnits	Value_50K100K
EduGRADDEG	Income_Avrg	Value_1.5M2M	Value_750K1M
EduHighDeploma	Income_Less15k	Value_100K150K	Value_Avrg
EduSomeColge	Income_Median	Value_150K200K	Value_Median
HomeValueAvrg	MARRIED	Value_1M1.5M	WHITE
HouseHoldSizeAvrg	NEVMARR	Value_200K250K	WIDOWED
Income_100K150K			
	Expenditure Variables		
Annual_Bdget_Exp	SpendinApperalAvrg	SpendinDiningAvrg	SpendEduAvrg
Electricity_Avrg	SpendinDining		

- **Expenditure Indicators:** I gathered data of various expenditure types, listed in Table 2, through ArcMap Business Analyst by census tracts of Los Angeles County. Then, I renamed the selected variables with more detailed names to ease addressing the variables in subsequent steps. For example, the average spending variable on electricity was renamed from *X3063_X* to *Electricity_Avrg.* The expenditure attributes are listed in Table 10.

I used MS Access 2016 to organize, clean and perform simple calculations on the data. I integrated all variables into one table to perform further processing steps in SPSS for analysis and modeling, explained in the next section. After removing tracts from the table due to missing values, the totals items ready for the next step were n = 2,245.

Step 2. Evaluating Regression Assumptions

Regression analysis requires assumptions such as linearity, distribution normality, and no multicollinearity to ensure data are adequate for the analysis. For this step, I used SPSS to examine the variables against the assumptions as follows:

- Because environmentalist indictor data is from household responses, I used the reliability test on all environmentalist indictors with a Cronbach's alpha coefficient test to measure the consistency of data. The reliability test results showed $\alpha = .763$ indicating an acceptable result ($\alpha \geq 0.7$).
- In the data file, income, home value and electricity-spending variables were available as an average and a median. To select variables for the subsequent step, I tested distribution normality to examine which variable would have a better normal curve and was not severely skewed. Table 10 shows the skewness scores for the median and average values of the three variables. Income_Median, Value_Avrg and

Electricity_Avrg were selected as they are closer to 0 meaning they have lower skewness, as highlighted in Table 11.

Table 11. *Measures of Skewness*

Income	Income_Median	Income_Avrg
Skewness	1.288	1.598
Std. error of skewness	.051	.051
Home value	Value_Median	Value_Avrg
Skewness	1.942	1.648
Std. error of skewness	.051	.051
Electricity spending	Electricity	Electricity_Avrg
Skewness	1.874	1.535
Std. error of skewness	.051	.051

Note. Highlighted values indicate lower-skewness variables

- Detecting multicollinearity is crucial in multiple regression analysis. Multicollinearity occurs when two independent variables are highly correlated with each other, that is, when the Pearson coefficient correlation exceeds .80 ($r > .80$) meaning they are behaving similarly.

First, I examined the correlation matrix for all environmentalist variables. Table C1 (Appendix C) shows that the majority of the variables have a significant correlation with one another ($r < 0.70$, $p < .05$). There was no significant correlation between Environ_Conscious and Global_Warming ($r = 0.03$, $p > .05$), Buy_Environ_SafeProducts and Global_Warming ($r = 0.07$, $p > .05$), and Buy_Environ_SafeProducts and Company_Environ_Record ($r = - 0.01$, $p > .05$). Thus, I did not enter the variables Buy_Environ_SafeProducts and Environ_Conscious in the regression analysis. Table C3 shows the correlation matrix of the remaining variables with significant correlation between all variables ($r < 0.70$, $p < .05$). To conclude, the variables Preserve_Nature, Govt_Focus, Help_Environment, Global_Warming,

GreenProducts_vs_Convenience and Company_Environ_Record have significant bivariate relationships among one another; therefore, I entered them in the regression analysis.

Second, I examined the correlation matrix for demographic variables. I tested the correlation of all variables related to home value. Table C2 shows that the majority of the variables have a significant correlation with one another ($r < 0.70$, $p < .05$). However, I detected high correlation between Value_100k150K and Value_150k200k ($r = 0.742$, $p < .05$), and Value_1M1.5M and Value_1.5M12M ($r = 0.717$, $p < .05$). Thus, I combined Value_100k150K and Value_150k200k to a new variable named Value_100K200K_Comp and correspondingly combined Value_1M1.5M and Value_1.5M12M to a new variable named Value_1M2M_Comp. Accordingly, the variables Value_50k, Value_50K100K, Value_100K200K_Comp, Value_200K250K, Value_250K300K, Value_300K400K, Value_400K500K, Value_500K750K, Value_750K1M, Value_1M2M_Comp, and Value_2M had significant bivariate relationships with an acceptable Pearson correlation coefficient; thus, I entered them in the regression analysis.

Income-level variables are demographic variables. I calculated a correlation matrix to examine multicollinearity assumptions. Tables C4 and C5 show some variables significantly correlate ($r < 0.70$, $p < .05$). High correlation emerged between Income_Less15k and Income_15K25K ($r = 0.743$, $p < .05$), Income_15K25K with Income_25K35K ($r = 0.769$, $p < .05$), Income_50K75K with Income_75K100K ($r = 0.730$, $p < .05$), Income_100K150K with Income_150K200K ($r = 0.816$, $p < .05$). I combined Income_Less15k, Income_15K25K, and Income_25K35K to a new variable, Income_Less5k35k_Comp. I combined Income_50K75K and Income_75K100K to a new variable, Income_50k100k_Comp. Lastly, I combined Income_100K150K and Income_150K200K to a new variable, Income_100k200k_Comp.

Variables Income_Less5k35k_Comp, Income_35k50K, Income_50k100k_Comp, and Income_100k200k_Comp have significant bivariate relationships with acceptable Pearson correlation coefficients (see Tables C4 and C5); thus, I entered them in the regression analysis.

Similarly, correlation relationships between the variables TOTHouseUnits, TOTPOP, VACANTUnits, RENTERUnits, POPDENS, OWNERUnits were significant ($r > 0.70$, $p < .05$). However, TOTHouseUnits highly correlated with TOTPOP ($r = 0.711$, $p < .05$) and RENTERUnits ($r = 0.73$, $p < .05$; see Table C6). Therefore, I excluded TOTHouseUnits from the regression analysis. Similarly, multicollinearity diagnosis shows that all marital-status variables—NEVMARR, MARRIED, WIDOWED—significantly correlated with an acceptable Pearson correlation coefficient ($r < 0.70$, $p < .05$). I examined educational level variables—EduHighDeploma, EdeAlteernativeCred, EduSomeColge, EduASSCDEG, EduBACHDEG, EduGRADDEG—to ensure the multicollinearity assumption was not violated. Table C7 shows high Pearson correlation coefficients between EduSomeColge and EduASSCDEG ($r = 0.733$, $p < .05$), resulting in substituting them with a new variable, EduBach_EduSomeColg__Comp.

Lastly, the spending variables—Electricity_Avrg, SpendinApperalAvrg, Annual_Bdget_Exp, SpendinDiningAvrg, and SpendEduAvrg—had multicollinearity ($r > 0.90$, $p < .05$), as shown in Table C8. As a result, I combined all spending variables under one new variable named GeneralSpending. Then, I tested all independent variables with the dependent variable—Existing_Installs_Count—to ensure the correlation relationship persisted. The correlation matrix showed a significant correlation with Existing_Installs_Count; however no significant correlation emerged between Existing_Installs_Count and Preserve_Nature ($r = 0.041$, $p > .05$) , and Existing_Installs_Count and Value_400K500K ($r = 0.029$, $p > .05$). Thus, I

did not enter Preserve_Nature and Value_400K500K in the regression analysis. Table 12 summarizes all the variables I entered for further analysis.

Table 12. *List of the Variables That Met the Regression Assumptions*

AMERICINDIAN	Global_Warming	MARRIED	Value_200K250K
ASIAN	Govt_Focus	NearBy_IndexCode	Value_250K300K
BLACK	GreenProducts_vs_Convenience	NEVMARR	Value_2M
Company_Environ_Record	Help_Enviroment	OWNERUnits	Value_300K400K
DIVORCD	HouseHoldSizeAvrg	PACIFIC	Value_500K750K
EdeAlteernativeCred	Income_100k200k_Comp	POPDENS	Value_50k
EduASSCDEG	Income_200K	RENTERUnits	Value_50K100K
EduBach_EduSomeColg_Comp	Income_35k50K	VACANTUnits	Value_750K1M
EduGRADDEG	Income_50k100k_Comp	Value_100K200K_Comp	WHITE
EduHighDeploma	Income_Less5k35k_Comp	Value_1M2M_Comp	WIDOWED
GeneralSpending			

Step 3. SPSS Modeling Analysis

I processed the regression using the stepwise method. In the stepwise method, researchers enter variables based on the size of their partial correlation coefficient, entering the variable with the largest correlation first and then the variable with second largest correlation coefficient; the process ends if there is no improvement in R^2. The regression process did not consider a h hierarchical regression because there was no order of the variables entering the regression process. I processed the regression twice. In the first run, I aimed to assure that Income_Median and HomeValueAvrg are significant predictors, before entering all home-value level and income-level variables. Table 13 shows that Income_Median and HomeValueAvrg are significant predictors. As a result, I repeated the regression process by entering all the variables of income levels and home-value levels.

Table 13. *Results of the First Multiple Regression Analysis—Coefficients*

Model	Unstandardized coefficients		Standardized coefficient		
	B	Std. error	Beta	*t*	Sig.
(Constant)	-35.174	7.943		-4.428	.000
PACIFIC	-0.036	0.016	-.035	-2.285	.022
WIDOWED	0.010	0.005	.040	2.150	.032
BLACK	0.002	0.001	.043	2.745	.006
Govt_Focus	-12.326	5.693	-.044	-2.165	.030
Carbon_Offset_Metric_Tons	0.000	0.000	.091	5.983	.000
Global_Warming	22.196	4.379	.103	5.069	.000
RENTERUnits	-0.005	0.001	-.103	-4.114	.000
Income_Median	0.000	0.000	.135	3.712	.000
OWNERUnits	0.008	0.002	.142	4.576	.000
Value_Avrg	1.285E-5	0.000	.164	5.753	.000
NearBy_IndexCode	4.370	0.352	.206	12.405	.000
EduGRADDEG	0.016	0.002	.240	7.081	.000

Note. a. Dependent variable: Existing_Installs_Count

After including income levels and home-value levels, I performed the analysis and report the results shown in Tables 14 through 16. The model summary in Table 14 shows that I entered 17 steps in the regression equation. The model summary indicates that Adjusted R^2 = 0.56. The 15 significant predictors shown in Table 15 accounted for 56% of the variance in the dependent variable: Existing_Installs_Count. Table 16 shows that the model significantly predicts the dependent variable, solar installations count (F = 197.414, p < .001). Accordingly, the regression equation would help compute the solar installation likelihood for a census tract in Los Angeles County. Thus, I generated the equation as follows:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \ldots + b_iX_i\,,$$

where y = the dependent variable, X_i = the independent variables, a = constant or intercept, and b_i is the untenderized regression coefficient of X_i

Solar Installations = -44.99 + (.003 * *GeneralSpending*) + (4.533 * *NearBy_IndexCode*)

$$+ (.024 * Value\_750K1M) + (.007 * OWNERUnits) + (-.040 * Value\_200K250K)$$

$$+ (20.100 * Global\_Warming) + (0.000473 * Carbon\_Metric\_Tons) + (-.061 * Value\_50K100K)$$

$$+ (.003 * BLACK) + (-.008 * EduASSCDEG) + (-.024 * Value\_250K300K) + (.027 * Value\_2M)$$

$$+ (-.037 * PACIFIC) + (.024 * Value\_1M2M\_Comp) + (.023 * Value\_500K750K)$$

Table 14. *Results of the Multiple Regression Analysis—Model Summary*

Model	*R*	R^2	Adjusted R^2	Std. error of the estimate
1	.652[a]	.425	.425	19.376
2	.682[b]	.465	.465	18.688
3	.711[c]	.506	.505	17.974
4	.720[d]	.518	.517	17.746
5	.729[e]	.532	.531	17.503
6	.735[f]	.541	.540	17.334
7	.740[g]	.548	.547	17.201
8	.745[h]	.555	.553	17.074
9	.749[i]	.560	.559	16.974
10	.751[j]	.564	.562	16.914
11	.752[k]	.565	.563	16.890
12	.753[l]	.567	.564	16.862
13	.754[m]	.568	.565	16.844
14	.754[n]	.569	.566	16.827
15	.755[o]	.570	.567	16.814
16	.756[p]	.571	.568	16.796
17	.755[q]	.571	.568	16.797

Table 15. *Results of the First Multiple Regression Analysis—ANOVA*[a]

Model		Sum of squares	*df*	Mean square	*F*	Sig.
17	Regression	835479.221	15	55698.615	197.414	.000[4]
	Residual	628894.093	2229	282.142		
	Total	1464373.310	2244			

Note. a. Dependent variable: Existing_Installs_Count.

Table 16. *Results of the Second Multiple Regression Analysis—Coefficients*

Model	Unstandardized coefficients		Standardized coefficient		
	B	Std. error	Beta	*t*	Sig.
(Constant)	-44.999	5.999		-7.501	.000
Value_200K250K	-0.040	0.013	-.060	-3.152	.002
Value_50K100K	-0.061	0.020	-.049	-3.107	.002
Value_250K300K	-0.024	0.012	-.042	-1.942	.052
EduASSCDEG	-0.008	0.004	-.040	-1.947	.052
PACIFIC	-0.037	0.015	-.035	-2.386	.017
BLACK	0.003	0.001	.071	4.613	.000
Global_Warming	20.100	3.413	.093	5.889	.000
Carbon_Offset_Metric_Tons	0.000	0.000	.105	7.014	.000
Value_2M	0.027	0.005	.115	5.122	.000
OWNERUnits	0.007	0.003	.123	2.483	.013
Value_750K1M	0.024	0.004	.127	5.820	.000
GeneralSpending	0.003	0.001	.152	5.391	.000
Value_500K750K	0.023	0.005	.169	4.833	.000
Value_1M2M_Comp	0.024	0.004	.178	5.753	.000
NearBy_IndexCode	4.533	0.344	.214	13.192	.000

Note. a. Dependent variable: Existing_Installs_Count.

Step 4. ArcGIS Modeling Analysis (Forest-based Classification and Regression)

In a similar manner, in this step I aimed to develop a model based on all predictors used in the previous step to predict the number of solar installations by census tract using a forest-based classification and regression algorithm in ArcGIS Pro. The tool is a supervised machine-learning method that uses an adaptation of Breiman's (2001) random-forest algorithm. This algorithm creates an ensemble forest of a number of decision trees. Each tree has its own prediction, and the final prediction is determined by the collective prediction of the entire forest as a part of a voting scheme of all trees (Breiman, 2001). Notably, combining multiple trees in a forest to make a prediction addresses the overfitting problem associated with a single tree.

I built the forest regression using 90% of the data for training and extracted 10% of the data for validation. I included all the variables in the model. I used the selected forest-regression model with the following parameters: 100 was the number of trees, 5 was the maximum branch, 29 was the tree depth, and I had 2,246 observations. The forest-based regression analysis showed that solar installations was a continuous dependent variable explained by 20 independent variables (see Table 17). The importance score is calculated using Gini coefficients. Gini coefficients are "a measure of how each variable contributes to the homogeneity of the nodes and leaves in the resulting random forest" (Dinsdale, n.d., para. 5). In other words, Gini importance counts the number of times a variable is responsible for a split and the impact of that split divided by the number of trees. The model resulted in R^2 = 0.637 after validating the testing data extracted from the training model by comparing the observed values for each test feature with the predictions for those features based on the trained model.

Table 17. *Top Important Variables*

Variable	%	Variable	%	Variable	%
NearBy_IndexCode	14	OWNERUnits	3	MARRIED	1
Value_1M2M_Comp	13	Value_2M	3	GreenProducts_vs_Convenience	1
GeneralSpending	10	EduGRADDEG	3	Value_400K500K	1
Income_200K	8	POPDENS	2	WHITE	1
Carbon_Offset_Metric_Tons	7	Income_100K200K_Comp	2	Edubach_Edusomecolg_Comp	1
Value_750K1M	4	HouseHoldSizeAvrg	1	RENTERUnits	1
Value_500K750K	4	Govt_Focus	1		
Training data: Regression diagnostics					
R^2	.946				
p-value	.000				
Standard error	.004				
*Predictions for the data used to train the model compared to the					
Validation data: Regression diagnostics					
R^2	.637				
p-value	.000				
Standard error	.027				

The analysis here mainly answers the research question, how might factors such as expenditures, demographics, political preferences, and geographic and environmentalist indicators, predict solar adoption at the residential level? To achieve the best explanation of the data, I used two techniques: First, using the multiple regression method within SPSS, version 20 and second, by utilizing a machine-learning, forest-based regression, algorithm within ArcGIS Pro 2.3.

After collecting, aggregating, and merging the data variables, I carefully tested the regression assumptions and the data normalization. I examined correlation between variables and deleted variables with high correlation. I addressed variables with high correlations by combining correlated variables or removing one of the highly correlated variables. I calculated reliability using Cronbach's alpha for environmentalist indicator responses. I used modeling techniques to predict the dependent variable, which is the number of solar installations per census tract. Table 18 shows a comparison of the results of the output model coefficient of determination.

I deployed the output model of the forest-based regression as an interactive GIS tool to be used by the user using the ESRI web-application builder. The second model achieved a better explanation of the variance in the data, resulting in a more accurate prediction of the number of solar installations. In the next chapter, I discuss the design, functionality, and evaluation of Artifact 2: Residential Solar Energy Adoption.

Table 18. *Mutiple Regression Versus Foreset Based Regression Analysis Results*

Algorithm	Multiple regression	Forest based regression
Coefficient of determination	$R^2 = .568$	$R^2 = .637$
Significant predictors	General spending per household (total average of electricity, dining, education, apparel)	Nearby solar installations
	Nearby solar installations	Houses with a market value of 1M to 2M
	Houses with market value of 750K to 1M	General spending per household (total average of electricity, dining, education, apparel)
	Units that are owned (not rented not vacant)	Income of 200K and more
	Houses with a market value of 200k to 250K (negative relationship)	Carbon dioxide concentration
	Houses with a market value of 50k to 100K (negative relationship)	Houses with market value of 750K to 1M
	The perspective towards whether Global warming is a threat	Houses market value of 400K to 500k
	Carbon dioxide concentration	Education: BSc degree/some degree
	Education level: associate's degree (negative relationship)	Houses market value of 500K to 750K
	Houses with a market value of 1M to 2M	Owned houses with market value of 2M and more
	Houses with market value over 2M	Education level: Graduate degree
	Houses with market value of 500K to 750K	Average income 100k to 200K
	Houses with a market value of 250k to 300K (negative relationship)	Household with positive attitude towards that the Government should focus more on environmental issues
	Black ethnicity	The perspective that people who value green products over convenience
	Pacific ethnicity (negative relationship)	White ethnicity
		Population density per 1 mile

Chapter 4: Residential Solar Energy Adoption SDSS Artifact (RSEA-SDSS)

The RSEA algorithm is an instantiation artifact in the form of an SDSS. RSEA is a GIS web-based application that enables stakeholders (e.g., utility companies, policymakers, service providers, homeowners, and researchers) to interactively navigate through locations on a map. For example, there are maps that highlight locations with high and low solar energy adoption potential. The prediction output rests on the collective contribution of the statistically significant variables mentioned in the analysis section. In addition to the map visualization, the web application shows information of the variables for each census tract. Stakeholders can use this information to make better decisions regarding plans, incentive programs, and marketing strategies. For example, the user can select a tract with a high potential for solar energy adoption and be able to visualize the variables, on click popup window, that contributed most to the prediction and compare that outcome to a nearby area with a different adoption level.

I developed the system using ESRI's ArcGIS Web AppBuilder for Developers. The RSEA-SDSS is optimized for different platforms: computers, tablets, and mobile devices. I created a story map that includes extended information about the research purpose, Los Angeles County history and demographics, the prediction-calculation results, the interactive map, and the evaluation survey.

Description of the RSEA-SDSS

The RSEA-SDSS's home page shows the Los Angeles County map divided into census tracts (Figure 11). The solar adoption prediction has five classes or natural breaks. The points identified by ArcMap determined the breaks that best group similar values and maximize the differences between classes (Figure 12). The representation of the level of adoption is shown by

the greenness color of each tract, where dark green designates 153 and more in predicted solar installations and light green designates 12 or fewer predicted solar installations.

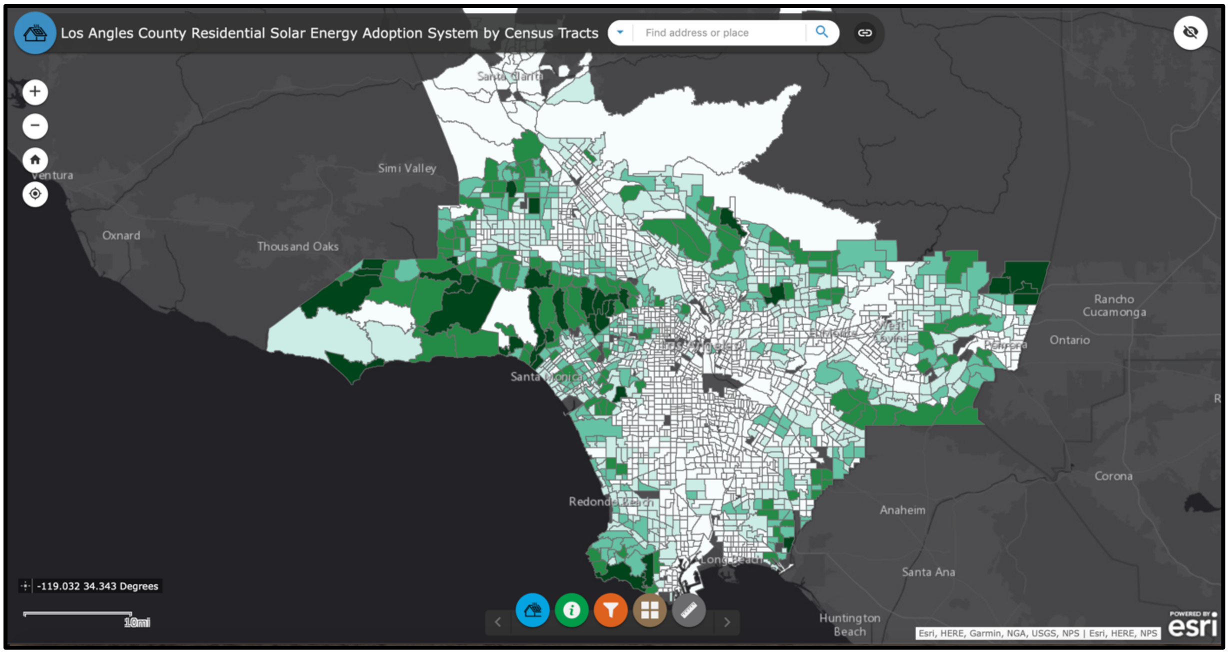


Figure 11. Los Angeles County map-output prediction map per census tract.

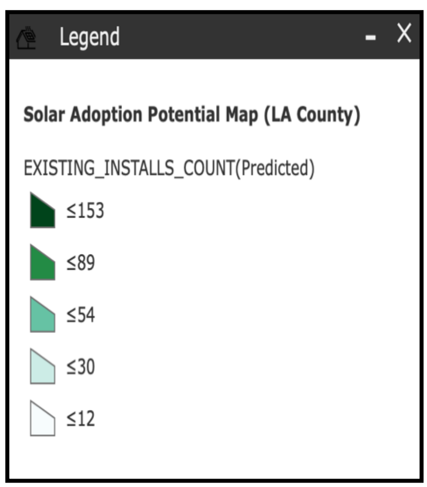


Figure 12. Residential Solar Energy Adoption legend.

Moreover, the RSEA-SDSS provides a filter function (see Figure 13). This function filters tracts with high predicted solar adoption (90 and above), moderate predicted solar adoption 30 to 60), and low predicted solar adoption (30 and less). Additionally, the popup (see

Figure 14) appears when the user clicks on a specific tract on the map. The RSEA-SDSS provides all information associated with the selected tract: number of predicted installations, house ownership, expenditure characteristics, houses with market values, income, education, ethnicity, attitude toward the environment, and general demographics, however this information is for visualization and the tool doesn't embed the information.

Other functions are available in the system. The Basemap function provides a basemap gallery. Users who prefer imagery can select the imagery basemap as the basis for the map (see Figure 15). The information icon displays a detailed framework on which the algorithm was built and the factors that were found to be significant. The system offers a search bar to locate a specific address or place. For example, users can use the map to look up their home address and identify their tract number and their solar adoption prediction level.

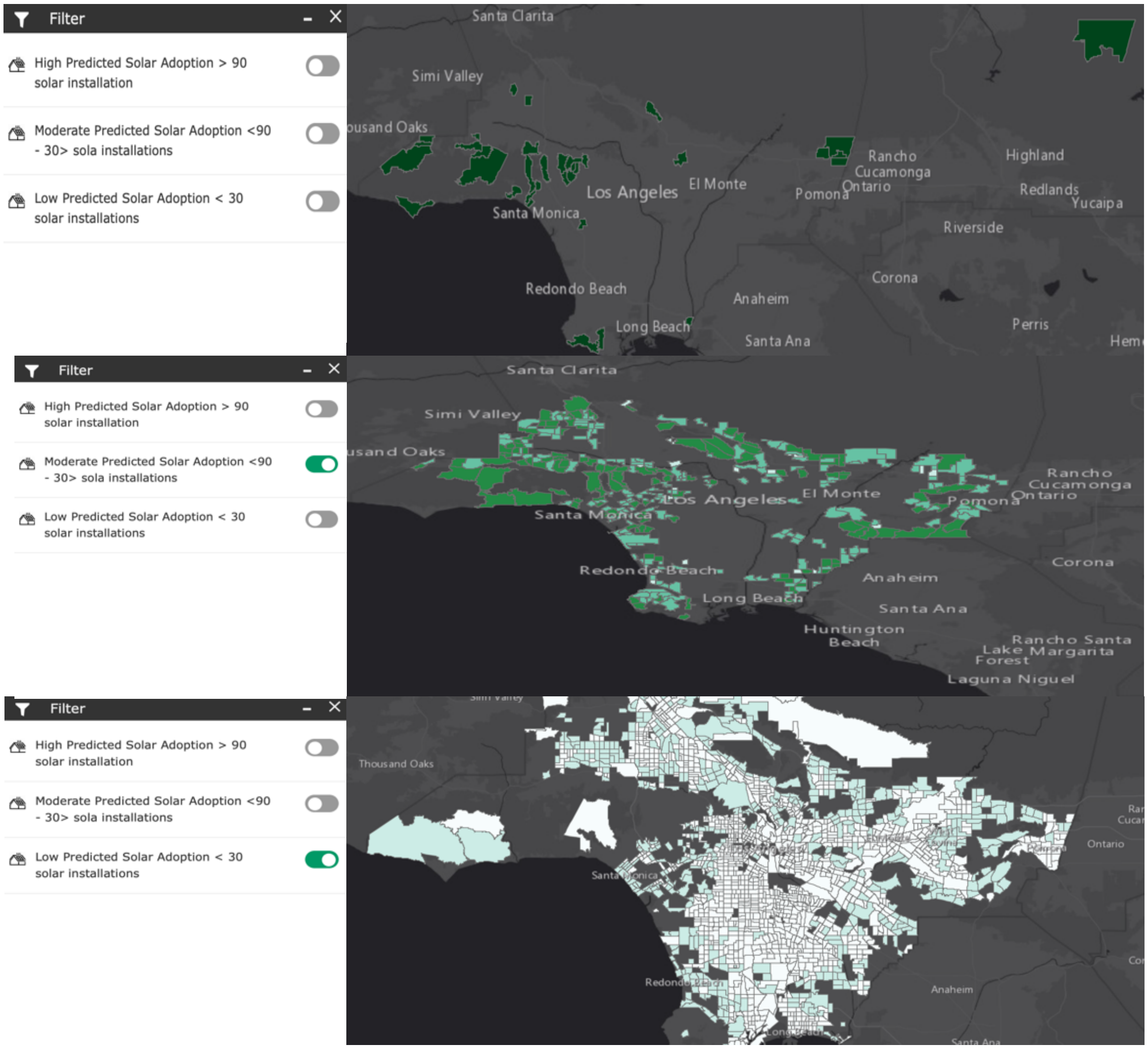


Figure 13. The level of solar adoption filter function.

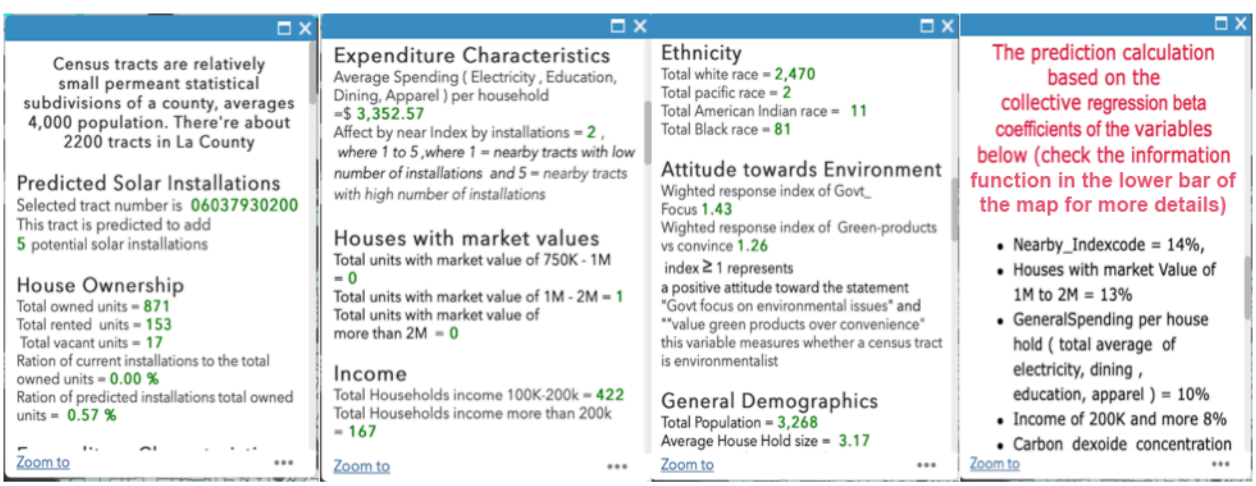


Figure 14. Popup function.

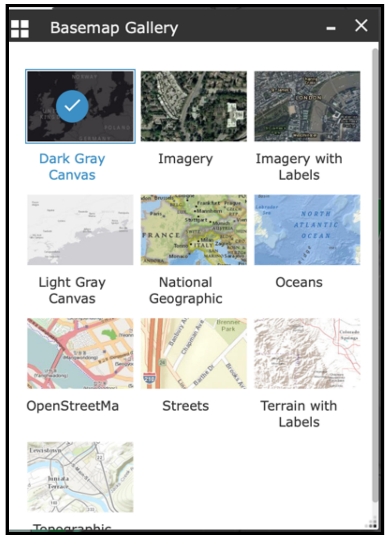


Figure 15. Changing Basemap function.

RSEA-SDSS Evaluation

To evaluate the RSEA-SDSS, I used a mixed-method strategy concurrent triangulation design by conducting a descriptive quantitative analysis through the SUS and UEQ, as well as collecting qualitative feedback from end-user interviews. Chapter 3 provided a brief description of each quantitative measure. Both questionnaires were approved as exempt by the Claremont Graduate University Institutional Review Board (see Appendix D and E).

First, I designed the survey using Qualtrics to be distributed as a pilot study. I recruited 14 participants from Claremont Graduate University from the IST371, Introduction to GIS Solution Development. After I presented the research purpose and explained RSEA-SDSS, I asked students to navigate through the system and take the survey. The pilot study helped me make necessary changes to the survey. I created a story map to describe the research goal, explain the model results, explore the solar map for navigation, and request participants take the quantitative questionnaire for RSEA-SDSS (see Appendix F for story map).

The population frame included participants from utility companies, solar installation firms, renewable-energy activists, nonprofit organizations, and government employees. I

distributed the surveys by sending introductory e-mails to solar firms followed by e-mails with a link to a story map for those who were interested in participating (see Appendix G for e-mail sample). I used the database form *gosolarcalifornia.com* to recruit marketing and sales employees in solar firms but had difficulty finding an adequate number of participants. Therefore, I also requested participation by distributing flyers at renewable-energy-related events (see Appendix H for the flyer).

Second, I conducted a qualitative study to evaluate RSEA-SDSS to gain a subjective interpretation of the system's utility in improving decision making for solar energy adoption at the residential level. I conducted interviews with participants who consented to participate and were knowledgeable about the solar installation industry and renewable-energy domain. Interviewees were employees from a utility company, GIS specialists, solar activists, and solar firm marketers. The interviews followed a semi-structured approach (see Appendix B for the prepared interview questions).

Quantitative Results

Based on the data gathered from participants, the RSEA-SDSS is useable and achieved a positive evaluation on the UEQ. Fifty-four participants evaluated the system. I received 10 incomplete responses; thus, the total sample size of those who completed the survey was 44. Scores ranged from 1 (strongly disagree) to 5 (strongly agree). As directed by Bangor et al. (2008), I calculated the SUS score by subtracting 1 from each odd item's value and subtracting 5 from each even item's value. After that, I summed all numbers and multiplied by 2.5. The answer range allowed from each participant was from 0 to 100 (not percentage), where 0 was the worst score and 100 indicated perfect usability. The overall SUS score is the mean of users' scores. Based on participants' responses, the RSEA-SDSS scored 69.6 on the SUS. An above-

average usability score of 68 is considered an acceptable score for most IT artifacts (Sauro, 2011). Thus, the RSEA-SDSS is considered to be a good usable system.

As noted above, 44 participants evaluated the RSEA-SDSS. The UEQ consists of 26 items. that captures the following scales: attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty (see Appendix A). For each item, participants checked the circle that most closely reflected their impression. The range of scales is between -3 (horribly bad) and +3 (extremely good). Therefore, interpretation of the scale means is that values between -0.8 and 0.8 represent a neutral evaluation of the corresponding scale, values > 0.8 represent a positive evaluation and values < -0.8 represent a negative evaluation (Schrepp et al., 2017b). The UEQ offers an Excel-tool to examine and detect response errors or misunderstandings. I used that tool and identified three responses of the 44 as inconsistent responses; as a result, I deleted them from the analysis. Before continuing the evaluation analysis, I examined reliability using Cronbach's alpha. Table 19 confirms that the reliability tests of constructs (attractiveness, perspicuity, efficiency, dependability, stimulation and novelty) are all above 0.7, the recommended threshold.

The mean score of attractiveness was 1.93, which confirms that the RSEA-SDSS is likable and enjoyable. Similarly, the RSEA-SDSS scored a mean of 1.73 for perspicuity, which reflects that the system is clear and understandable. In addition, the mean score for dependability was 1.46. Moreover, results showed that the RSEA-SDSS achieved a positive impression in stimulation with a mean score of 1.68. Lastly, responses revealed that novelty gained a positive evaluation with a mean of 1.07, which is the lowest of all the constructs. For detailed results, mean values of the 26 items are listed in Appendices I and J.

Table 19. *Cronbach's Alphas for UEQ*

Constructs	Cronbach's alpha (alpha)	Means	Variance
Attractiveness	0.93	1.79	0.82
Perspicuity	0.88	1.73	1.30
Efficiency	0.80	1.78	0.78
Dependability	0.75	1.46	1.00
Stimulation	0.87	1.68	1.00
Novelty	0.69	1.08	1.10

To compare the utility of the RSEA-SDSS with other systems, the UEQ offers a comparison tool with the data in the dataset benchmark. This method provides a highly efficient way to measure the quality of new products where no results from previous evaluations exist for comparison (Schrepp et al., 2017a). The benchmark data set contains data from 18,483 people from 401 studies on different products (business software, web pages, web shops, and social networks). As shown in Figure 16 and Table 20, only 10% of results in the benchmark are better than the RSEA-SDSS and 75% are worse; only novelty, with a 25% benchmark, had better results and 50% had worse results. All in all, the quantitative study results clearly demonstrated the utility of the RSEA-SDSS with respect to its usability and user experience.

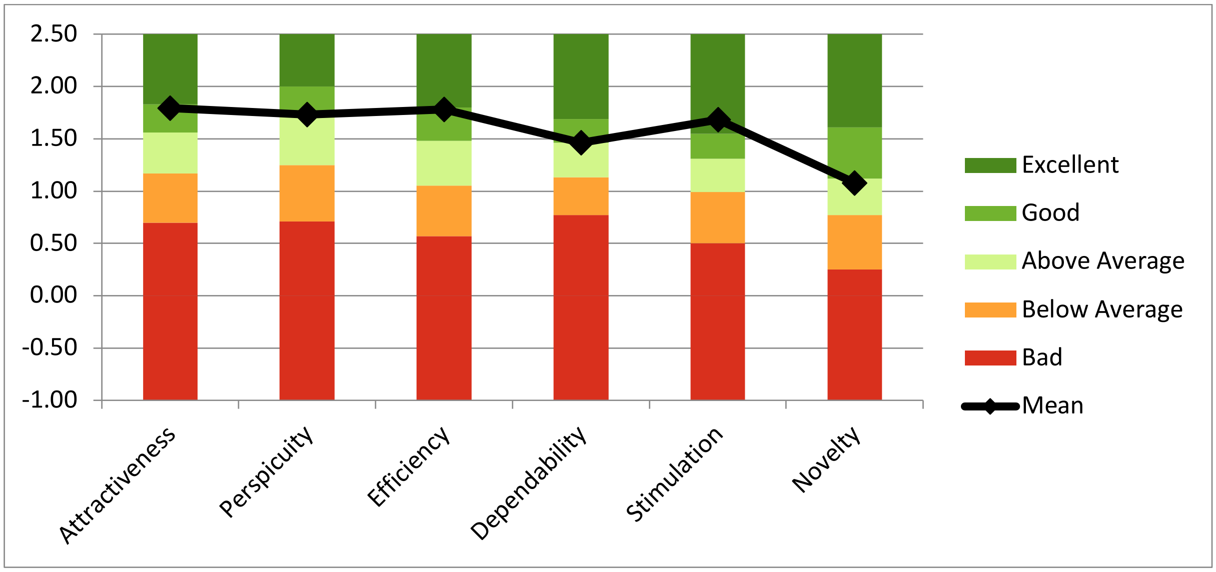


Figure 16. The RSEA-SDSS evaluation in comparison to the benchmark.

Table 20. *The RSEA-SDSS Evaluation in Comparison to the Benchmark*

Scale	Mean	Comparison to benchmark	Interpretation
Attractiveness	1.79	Excellent	10% of results better, 75% of results worse
Perspicuity	1.73	Good	10% of results better, 75% of results worse
Efficiency	1.78	Excellent	10% of results better, 75% of results worse
Dependability	1.46	Above Average	10% of results better, 75% of results worse
Stimulation	1.68	Excellent	10% of results better, 75% of results worse
Novelty	1.08	Good	25% of results better, 50% of results worse

Qualitative Results

I conducted interviews with seven interviewees from various stakeholders: a utility company, solar firms, nonprofit organizations, and government sectors. I reached theoretical saturation, where no more interviews were needed. Four interviews took place online, and three were in-person. Interviews were recorded and transcribed. The data analysis method was thematic coding. Researchers commonly use this analysis approach when analyzing and categorizing qualitative data into summaries of shared meaning patterns called themes (Braun,

Clarke, Hayfield, & Terry, 2018). The analysis approach resulted in two main themes: usefulness and improvements.

Usefulness

As noted above, I used mixed-method research to evaluate the usefulness of the RSEA system. Qualitative data offers deep insights to the utility of using a system that numeric data does not necessarily provide. Participants clearly confirmed the usefulness of the RSEA. They viewed the tool as helpful for many tasks in decision-making. For example, Participant 1 thought the system would help better direct an incentives program.

> Very important for the decision makers to keep enhance their incentives programs to target those people, but more importantly to focus on people with demographics that not associated with solar adoption to design a customized program to encourage investing more in solar.

Participant 6, a government employee for a local county, stated, “the tool highlights key behavioral-economics findings whose implications can help policymakers understand which factors, aside from purely economic ones, may affect individuals’ adoption behavior.”

The tool would help utility companies in their tasks and “add value to utility companies to consider areas with high installations in their consideration to grid impact.”

In a similar manner, Participant 2, who works for a local utility company, added,

> Looking through the map that these dark greens areas are of high adoption possibly and from there we can see what circuits are in that region and depending on that we can see if we have had any thermal, overloading, voltage issues in that region that would help in our forecasting plan or maybe respond by installing energy storage or upgrade the conductor or something like that in order to mitigate future complications.

Likewise, Participant 3 specified the departments that would benefit from using the system: “This system would be of a value to these departments: energy procurements, corporate financial planning and forecasting, and distribution system planning.” Participant 4, a local nonprofit origination member, highlighted the potential use for the system as follows:

> Primary industry would be the installations companies that are interested where to market their services. … Secondary market, I would think that startups who try to identify gaps in the industry in those regions that do not have providers. … community-based organizations, like Sustainable Claremont that are working to facilitate solar adoption, where these groups can locate areas with opportunity with adoption. … nonprofit organizations would come and do some education work about incentives programs to help with the adoption.

Moreover, Participants 2, 3, 4, and 6 indicated the usability of the system, as quoted below, respectively:

“I found it easy to navigate. Response time was good.”

“I think it is pretty straight forward and I did not have issues navigating.”

“I think the tool is great. I wish I had it a couple years ago.”

“It is an interesting project, I can see potential in this project.”

Improvements

In contrast, participants suggested ideas and changes to the system. Thus, I created a theme merged from the textual data that includes all participants’ insights. First, participants from Southern California Edison (SCE) suggested connecting SCE’s Distributed Energy Resource Interconnection Map with RSEA. This connection would allow developers and

planners to better locate the energy-distribution resources in tracts with high/low adoption for further action. Participant 5, a data scientist/senior engineer, recommended,

> It would be valuable if the system has an API that can be integrated to another system. For instance, the integration may have to transfer the information from the map to the feeder to the circuits level map for each tract level. Another suggestion is to use the [Distributed Energy Resource Interconnection Map] which can also be integrated with your system to create a new layer to visualize the solar adoption along with the circuits and substation located in each census tract.

I will consider this suggestion in future development.

Furthermore, a few participants valued the idea of time-framed prediction. For instance, Participant 3 added, “I think it would be useful to have a forecast for a given time frame, i.e., forecast for 2020, 2022, 2024, etc.” In a similar manner, Participant 7, a researcher in the energy domain and former utility employee, suggested to “explore the payment/incentive changes over time for solar programs. Has it had any effect on the outcome? Changes happened over time? Did it accelerate or reduce the diffusion rates?” I may not consider this suggestion because inadequate information and data are available about it. Lastly, two participants proposed minor ideas about adding the option to overlay a street view and better organize the information in the popup list to improve readability. I will address these ideas in future versions of the system.

Chapter 5: Discussion

This dissertation aimed to provide a decision-support system to identify locations that would likely adopt solar energy at the residential level. To achieve this goal, I designed two novel artifacts based on a solar adoption framework grounded in existing literature and a Residential Solar Energy Adoption Spatial Decision Support System (RSEA-SDSS). The first artifact is a predictive algorithm that operationalizes solar adoption framework components: environmentalist indicators, geographic indicators, political-preference indicators, demographic indicators, and expenditure indicators in Los Angeles County. To achieve high accuracy, I used two analytical techniques for the predictive analysis: multiple regression and forest-based regression. The second artifact is a working system in the form of a web-based application that supports decision making for potential stakeholders such as those in the utility industry, solar firms, and policymakers. I implemented the RSEA-SDSS based on modeling results from the first artifact. I adopted DSR to deliver a practical solution for an existing problem by adding to the knowledge base and following rigorous research methodologies in the evaluation process.

Prior researchers noted the importance of employing renewable-energy technologies, explaining the consequences associated with inadequate solar adoption rates, and clarifying the benefits connected with an overall increase in solar PV adoption with respect to the overall influences of the noted factors on the solar adoption framework. The first question in this study sought to determine how expenditures, demographics, political preferences, and geographic and environmentalist indicators predict solar adoption at the residential level. After examining spatial and nonspatial factors, a significant finding was that 20 factors explain solar adoption. Perhaps the most compelling finding is that the geographic indicator of “nearby index” was the most influential factor. In addition, the factor of population density per mile is also important in

determining solar adoption. The nearby index considers the clustering of the number of nearby previously installed systems as an indication of spatial patterns of diffusion. This finding broadly supports the work of other research studies, linking technology diffusion with social interaction and visibility. The result is in line with DOI (Rogers, 2003), averring that an idea or product diffuses in populations or social systems.

Researchers reported a positive relationship between household income and education level with adoption of energy-saving technologies. Study results here showed that households with an annual income of $100,000 and above, with members who hold a higher education degree, are likely to adopt solar energy more than households with less income and no college degree. Results also indicated that the probability of adopting solar energy increases for households with total annual income greater than $200,000 and a graduate degree. This finding is consistent with those of Brechling and Smith (1994) and Mills and Schleich (2009), who found an association between the level of education, income, and adoption of energy-saving technologies.

With respect to the first research question, I found that expenditures on electricity, education, and dining out, for example, are crucial in explaining solar energy adoption. For example, households with more spending on electricity, education, apparel, dining out, and entertainment/recreation tend to adopt solar energy. Correspondingly, solar energy adoption is more likely in households that own rather than rent their residences. Results also showed that houses' market value contributes to solar energy adoption. To illustrate, houses with market values between $1 and $2 million are more likely to adopt solar energy than houses with market values in the $500,000 to $1 million range. A possible explanation for this might be that households' financial capabilities influence their attitude toward solar energy adoption. These

results support previous research into this area, linking expenditures, economic factors, and solar energy adoption.

Moreover, I found very little in the literature on the question of how environmental values may motivate people to adopt solar energy technologies. Thus, I set out to assess the importance of environmental values and attitudes toward adopting solar energy technologies. Study results indicated that households' positive attitude toward a green environment might determine their adoption of solar energy. Specifically, results highlighted two environmental values: (a) household members who think government should focus more on environmental issues, and (b) household members who value green products over convenience. These results did not contribute as much to solar energy adoption as other factors. Likewise, marital status and ethnicity made a weak contribution to determining solar energy adoption. Findings showed that members of the White race and married households had a greater likelihood of adopting solar energy than others. I was unable to demonstrate that political affiliation aligned with solar energy adoption, due to the inadequacy of datasets, as explained in the analysis section.

With respect to the second research question—How might the derived solar adoption algorithm be instantiated to provide stakeholders with locational adoption characteristics?—results of the quantitative and qualitative studies confirmed the utility, usefulness, and positive user experience of the RSEA-SDSS. The implementation of the RSEA-SDSS, in consideration of the collective important factors, spatial and nonspatial, from the solar adoption framework made a novel contribution to the knowledge base within the information systems, decision support systems, and renewable energy fields. Most importantly, this finding enhances the EI framework proposed by Goebel et al. (2014), that aims to guide scholars for future work in "how EI researchers have made contributions based on their academic background" (Goebel et al., 2014,

para.1). Thus, the dissertation would fit under the Energy Efficiency goal and Smart Grid category that targets renewable energy in residential buildings.

RSEA-SDSS improves the decision-making process for various potential stakeholders in utility, solar installations policymaking, and nonprofit renewable-energy domains. Furthermore, RSEA-SDSS demonstrated its utility in identifying locations with solar energy adoption level—high, moderate, and low—as well as in providing detailed mapped information related to the desired location. Hence, RSEA-SDSS provides a practical contribution by offering an artifact that can be used to improve the decision making for various stakeholders to ultimately increase the amount of residential adoption of solar energy.

For instance, utility companies aim to optimize its resources allocation processes of upgrading infrastructures projects in locations where solar installations numbers are higher than other locations. Thus, the outcome information of RSEA-SDSS will be of a value to projects managers in utilities company to improve their proactive practices to mitigate negative complications in locations with high installation.

Similarly, policymakers would improve decision related to renewable energy incentives programs. To illustrate, the information from RSEA-SDSS could be used to evaluate the effectiveness of these existing programs or by enhance the decision making regarding where and to whom these incentives programs should directed to. Additionally, providers of solar energy service would also benefit from RSEA-SDSS to improve their directing marketing campaign by better identify and focus market niche which would essential to create and enhance their competitive advantage over other comparators in the industry. In general, the finding of the study would contribute positively on practice and policy which ultimately to increase the solar energy diffusion.

Conclusion, Limitations, and Future Work

The diffusion of solar energy among residents in the United States has not been growing as desired. The purpose of this dissertation was to design two artifacts to better understand the implications associated with adopting solar energy, and then help increase solar energy adoption at the residential level. First, I built an algorithm to explain solar energy adoption by operationalizing spatial and nonspatial factors such as demographics, environment, expenditures, and geographical factors; and used Los Angeles County as the case study location.

Second, I introduced an SDSS to support decision making to increase solar energy installation diffusion. The RSEA-SDSS is an interactive GIS-based and web-based system that enables decision makers (e.g., policymakers, solar firms, utility companies, and nonprofit organizations) to help reach their goals in planning new programs, evaluating current incentive programs, targeting marketing efforts, mitigating electricity overload risks, identifying locations to raise awareness of current environmental risks, and others. For this dissertation, I adopted DSR as the methodological approach by providing a practical solution to a highly relevant problem by designing and evaluating two artifacts. The results of the evaluations showed their utility.

Although the study explained 63.7% of the variance in predicting solar energy adoption by residents, it has certain limitations in terms of training the model with datasets using a single county in California. The limitation also extends to considering only California as the resource to train the predictive model, which limits the generalizability of the findings. Additionally, the dissertation is limited by the lack of data on the exact locations of current solar installations. The precise location of a solar installation would have resulted in better examination of the diffusion-effect innovation.

Future research may seek to include more data from different counties in different states to validate and enhance the current predictive model. Such data would make examining political-affiliation preferences on solar energy adoption possible. Likewise, future work may examine the expected payback investment time on homeowners' decisions to adopt solar energy technology. Finally, as suggested in the qualitative portion, future work may improve the SDSS design by adding a satellite layer with customizable transparency for visualization, integrating the electric grid, providing a substation map, and better organizing the popup information.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human–Computer Interaction, 24,* 574–594. https://doi.org/10.1080/10447310802205776

Braun, V., Clarke, V., Hayfield, N., & Terry, G. (2018). Thematic analysis. In P. Liamputtong (Ed.), *Handbook of research methods in health social sciences* (pp. 1–18). https://doi .org/10.1007/978-981-10-2779-6_103-1

Brechling, V., & Smith, S. (1994). Household energy efficiency in the UK. *Fiscal Studies, 15*(2), 44–56. https://doi.org/10.1111/j.1475-5890.1994.tb00196.x

Breiman, L. (2001). Random forests. *Machine Learning, 45,* 5–32. https://doi.org/10.1023 /A:1010933404324

Brooke, J. (1996). SUS—A quick and dirty usability scale. In P. W. Jordan, B. Thomas, B. A. McLelland, & B. A. Weerdmeester (Eds.), *Usability evaluation in industry, 189*(194), 4–7. London, England: Taylor and Francis.

Carter, J. (1977). The president's proposed energy policy. *Vital Speeches of the Day, 43,* 418–420. Retrieved from https://alternativeenergy.procon.org/sourcefiles/Carter%27s_Speech .pdf

Cropper, M., McKibben, J., Swanson, D., & Tayman, J. (2018). *Vendor accuracy study*. Retrieved from https://www.esri.com/~/media/Files/Pdfs/library/brochures/pdfs/vendor-accuracy-study.pdf

Dastrup, S. R., Zivin, J. G., Costa, D. L., & Kahn, M. E. (2012). Understanding the solar home price premium: Electricity generation and “green” social status. *European Economic Review, 56,* 961–973. https://doi.org/10.1016/j.euroecorev.2012.02.006

Deason, J., & Murphy, S. (2018). *Assessing the PACE of California residential solar deployment: Impacts of Property Assessed Clean Energy programs on residential solar photovoltaic deployment in California, 2010–2015.* Retrieved from https://emp.lbl.gov /publications/assessing-pace-california-residential

Dincer, I. (2000). Renewable energy and sustainable development: A crucial review. *Renewable and Sustainable Energy Reviews, 4*(2), 157–175. https://doi.org/10.1016/S1364-0321 (99)00011-8

Dinsdale, E. (n.d.). *Supplemental material.* Retrieved April 2, 2019, from https://dinsdalelab .sdsu.edu/metag.stats/code/randomforest.html

Dorvlo, A. S., & Ampratwum, D. B. (1998). Summary climatic data for solar technology development in Oman. *Renewable Energy, 14,* 255–262. https://doi.org/10.1016 /S0960-1481(98)00075-5

European Commission. (n.d.). Energy strategy and energy union. Retrieved October 27, 2018, from /energy/en/topics/energy-strategy-and-energy-union/2020-energy-strategy

Fialka, J. (2016). Why China is dominating the solar industry. *Scientific American.* Retrieved October 27, 2018, from https://www.scientificamerican.com/article/why-china-is -dominating-the-solar-industry/

Goebel, C., Jacobsen, H.-A., Razo, V., Doblander, C., Rivera, J., Ilg, J., … Lässig, J. (2014). Energy informatics. *Business & Information Systems Engineering, 6,* 25–31. https://doi.org/10.1007/s12599-013-0304-2

Götz, M., Lefebvre, J., Mörs, F., Koch, A. M., Graf, F., Bajohr, S., … Kolb, T. (2016). Renewable power-to-gas: A technological and economic review. *Renewable Energy, 85,* 1371–1390. https://doi.org/10.1016/j.renene.2015.07.066

Graziano, M., & Gillingham, K. (2014). Spatial patterns of solar photovoltaic system adoption: the influence of neighbors and the built environment. *Journal of Economic Geography, 15,* 815–839. https://doi.org/10.1093/jeg/lbu036

Guerrero, J. M., Hang, L., & Uceda, J. (2008). Control of distributed uninterruptible power supply systems. *IEEE Transactions on Industrial Electronics, 55,* 2845–2859. https://doi.org/10.1109/TIE.2008.924173

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems, 19*(2), Art. 4.

Hevner, A., & Chatterjee, S. (2010). *Design research in information systems: Theory and practice* (Vol. 22). London, England: Springer Science & Business Media.

Hsu, J. H.-Y. (2018). Predictors for adoption of local solar approval processes and impact on residential solar installations in California cities. *Energy Policy, 117,* 463–472. https://doi.org/10.1016/j.enpol.2018.03.008

International Energy Agency Photovoltaic Power Systems Programme. (2019). *A snapshot of global photovoltaic.* Retrieved November 3, 2018, from The International Energy Agency (IEA) website: http://www.iea-pvps.org/?id=266

Knier, G. (2008, August 6). *How do photovoltaics work?* Retrieved November 3, 2018, from https://science.nasa.gov/science-news/science-at-nasa/2002/solarcells

Kwan, C. L. (2012). Influence of local environmental, social, economic and political variables on the spatial distribution of residential solar PV arrays across the United States. *Energy Policy, 47,* 332–344. https://doi.org/10.1016/j.enpol.2012.04.074

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. In A. Holzinger (Ed.), *Proceedings of the 4th Symposium workgroups Human–Computer Interaction and Usability Engineering of the Austrian computer Society* (pp. 63–76). London, England: Springer. https://doi.org/10.1007/978-3-540-89350-9_6

Lutzenhiser, L. (1993). Social and behavioral aspects of energy use. *Annual Review of Energy and the Environment, 18*(1), 247–289. https://doi.org/10.1146/annurev.eg.18.110193.001335

Matisoff, D. C., & Johnson, E. P. (2017). The comparative effectiveness of residential solar incentives. *Energy Policy, 108,* 44–54. https://doi.org/10.1016/j.enpol.2017.05.032

Mills, B. F., & Schleich, J. (2009). Profits or preferences? Assessing the adoption of residential solar thermal technologies. *Energy Policy, 37,* 4145–4154. https://doi.org/10.1016/j.enpol.2009.05.014

National Aeronautics and Space Administration, Goddard Institute for Space Studies. (2017). Retrieved from https://www.giss.nasa.gov

National Aeronautics and Space Administration. (2018). *Global climate change: Vital signs of the planet.* Retrieved from https://climate.nasa.gov

Office of Energy Efficiency & Renewable Energy. (2013). *Solar energy technology basics.* Retrieved October 23, 2018, from https://www.energy.gov/eere/solar/articles/solar-energy-technology-basics

Palm, J., & Eriksson, E. (2018). Residential solar electricity adoption: How households in Sweden search for and use information. *Energy, Sustainability and Society, 8,* 14. https://doi.org.10.1186/s13705-018-0156-1

Powell, E. E., & Hill, G. A. (2009). Economic assessment of an integrated bioethanol–biodiesel–microbial fuel cell facility utilizing yeast and photosynthetic algae. *Chemical Engineering Research and Design, 87,* 1340–1348. https://doi.org/10.1016/j.cherd.2009.06.018

ProximityOne. (2019). *10 reasons to use census tract versus ZIP code geography & demographics.* Retrieved from http://proximityone.com/tracts_zips.htm

Robinson, S. A., Stringer, M., Rai, V., & Tondon, A. (2013). GIS-integrated agent-based model of residential solar PV diffusion. *Proceedings of the 32nd USAEE/IAEE North American Conference* (pp. 28–31). Cleveland, OH, US: International Association for Energy Economics.

Robson, C., & McCartan, K. (2016). *Real world research.* Hoboken, NJ, US: John Wiley & Sons.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York, NY, US: Simon and Schuster.

Sargent, R. (2017). President Obama's renewable energy legacy [Web log]. Retrieved October 27, 2018, from https://environmentamerica.org/blogs/environment-america-blog/ame/president-obama%E2%80%99s-renewable-energy-legacy

Sarzynski, A., Larrieu, J., & Shrimali, G. (2012). The impact of state financial incentives on market deployment of solar technology. *Energy Policy, 46,* 550–557. https://doi.org/10.1016/j.enpol.2012.04.032

Sauro, J. (2011). *A practical guide to the system usability scale: Background, benchmarks & best practices.* Denver, CO, US: Measuring Usability.

Schelly, C. (2014). Residential solar electricity adoption: What motivates, and what matters? A case study of early adopters. *Energy Research & Social Science, 2,* 183–191. https://doi.org/10.1016/j.erss.2014.01.001

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017a). Construction of a benchmark for the User Experience Questionnaire (UEQ). *International Journal of Interactive Multimedia and Artificial Intelligence, 4*(4), 40–44. https://doi.org/10.9781/ijimai.2017.445

Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017b). Design and evaluation of a short version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence, 4*(6), 103–108. https://doi.org/10.9781/ijimai.2017.09.001

Solar Energy Industries Association. (2019). *Solar state by state.* Retrieved November 30, 2018, from https://www.seia.org/states-map

Sweet, C. (2015). Wind, solar companies get boost from tax-credit extension. *The Wall Street Journal.* Retrieved from https://www.wsj.com/articles/wind-solar-companies-get-boost-from-tax-credit-extension-1450311501

Tsang, E. W. (2014). Case studies and generalization in information systems research: A critical realist perspective. *Journal of Strategic Information Systems, 23*(2), 174–186.

Tsoutsos, T., Gekas, V., & Marketaki, K. (2003). Technical and economical evaluation of solar thermal power generation. *Renewable Energy, 28,* 873–886. https://doi.org/10.1016/S0960-1481(02)00152-0

United Nations. (2017). *World population prospects.* Retrieved from https://population.un.org

Wang, X., Guan, Z., & Wu, F. (2017). Solar energy adoption in rural China: A sequential decision approach. *Journal of Cleaner Production, 168,* 1312–1318.

Watson, R. T., Boudreau, M.-C., & Chen, A. J. (2010). Information systems and environmentally sustainable development: Energy informatics and new directions for the IS community. *MIS Quarterly, 34*, 23–38. https://doi.org/10.2307/20721413

Weber, C., & Perrels, A. (2000). Modelling lifestyle effects on energy demand and related emissions. *Energy Policy, 28,* 549–566. https://doi.org/10.1016/S0301-4215(00)00040-9

Weiner, J. (2016, August 24). *Median installed price of solar in the United States fell by 5–12% in 2015.* Berkeley Lab. Retrieved October 27, 2018, from https://newscenter.lbl.gov/2016/08/24/median-installed-price-solar-united-states-fell-5-12-2015/

The White House, Office of the Press Secretary (2016, February 25). *FACT SHEET: The Recovery Act made the largest single investment in clean energy in history, driving the deployment of clean energy, promoting energy efficiency, and supporting manufacturing* [Press release]. Retrieved October 27, 2018, from https://obamawhitehouse.archives.gov/the-press-office/2016/02/25/fact-sheet-recovery-act-made-largest-single-investment-clean-energy

Worldometers. (2018). *China population*. Retrieved October 27, 2018, from http://www.worldometers.info/world-population/china-population/

Zhang, H., Vorobeychik, Y., Letchford, J., & Lakkaraju, K. (2016). Data-driven agent-based modeling, with application to rooftop solar adoption. *Autonomous Agents and Multi-Agent Systems, 30,* 1023–1049. https://doi.org/10.1007/s10458-016-9326-8

Appendix A: Items of User Experience Questionnaire

Scale	Items
Attractiveness	annoying/enjoyable
	bad/good
	unlikable/pleasing
	unpleasant/pleasant
	unattractive/attractive
	unfriendly/friendly
Efficiency	slow/fast
	inefficient/efficient
	impractical/practical
	cluttered/organized
Perspicuity	not understandable/ understandable
	difficult to learn/easy to learn
	complicated/easy
	confusing/clear
Dependability	unpredictable/ predictable
	obstructive/supportive
	not secure/secure
	does not meet expectation/meet expectation
Stimulation	inferior/valuable
	boring/exiting
	not interesting/ interesting
	demotivating/ motivating
Novelty	dull/creative
	conventional/inventive
	usual/leading edge
	conservative/innovative

Appendix B: Interview questions

1. What do you think about the system?
2. Would you use the system for decision making?
3. Why/why not do you think that the system will help you better make a decision with regard to solar adoption plans/regulations/property living locations?
4. What do you think of the data visualization? Coloring? Ease of navigation?
5. What functions did you find mostly useful in the system?
6. What do you think is missing in the system?
7. How easy could you find a piece of information you want to know about a certain area?
8. How did you find the system response time during the navigation process?
9. What features/functions should be added to help the decision-making process?

Appendix C: Correlation Matrices

Table C1. *First Correlation Matrix for the Environmentalist Variables (N = 2,245)*

		Preserve_Nature	Govt_Focus	Help_Environment	Global_Warming	Company_Environ_Record	Environ_Conscious	Buy_Environ_SafeProducts	GreenProducts_vs._Convenience
Preserve_Nature	Pearson correlation	1	.350**	.160**	.521**	.546**	.372**	.325**	.444**
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.000
Govt_Focus	Pearson correlation	.350**	1	-.258**	.642**	.405**	-.054*	-.103**	.246**
	Sig. (2-tailed)	.000		.000	.000	.000	.010	.000	.000
Help_Environment	Pearson correlation	.160**	-.258**	1	-.398**	.207**	.654**	.372**	.504**
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.000
Global_Warming	Pearson correlation	.521**	.642**	-.398**	1	.216**	.030	.007	.276**
	Sig. (2-tailed)	.000	.000	.000		.000	.153	.748	.000
Company_Environ_Record	Pearson correlation	.546**	.405**	.207**	.216**	1	.126**	-.010	.243**
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.644	.000
Environ_Conscious	Pearson correlation	.372**	-.054*	.654**	.030	.126**	1	.551**	.633**
	Sig. (2-tailed)	.000	.010	.000	.153	.000		.000	.000
Buy_Environ_SafeProducts	Pearson correlation	.325**	-.103**	.372**	.007	.010	.451**	1	.338**
	Sig. (2-tailed)	.000	.000	.000	.748	.644	.000		.000
GreenProducts_vs._Convenience	Pearson correlation	.444**	.246**	.504**	.276**	.243**	.533**	.338**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	

Note. ** Correlation is significant at the .01 level (2-tailed); * Correlation is significant at the .05 level (2-tailed).

Table C2. *Second Correlation Matrix Examined for the Environmentalist Variables (N = 2,245)*

		Preserve_ Nature	Govt_Focus	Help_ Environment	Global_ Warming	GreenProducts_ vs._Convenience	Company_ Environ_ Record
Preserve_Nature	Pearson correlation	1	.350**	.160**	.521**	.444**	.546**
	Sig. (2-tailed)		.000	.000	.000	.000	.000
Govt_Focus	Pearson correlation	.350**	1	-.258**	.642**	.246**	.405**
	Sig. (2-tailed)	.000		.000	.000	.000	.000
Help_Environment	Pearson correlation	.160**	-.258**	1	-.398**	.504**	.207**
	Sig. (2-tailed)	.000	.000		.000	.000	.000
Global_Warming	Pearson correlation	.521**	.642**	-.398**	1	.276**	.216**
	Sig. (2-tailed)	.000	.000	.000		.000	.000
GreenProducts_vs._ Convenience	Pearson correlation	.444**	.246**	.504**	.276**	1	.243**
	Sig. (2-tailed)	.000	.000	.000	.000		.000
Company_Environ_ Record	Pearson correlation	.546**	.405**	.207**	.216**	.243**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000

Note. ** Correlation is significant at the .01 level (2-tailed); * Correlation is significant at the .05 level (2-tailed).

Table C3. *First Correlation Matrix for Income Level Variables (N = 2,270)*

		Value_50K	Value_ 50K100K	Value_ 100K150K	Value_ 150K200K	Value_ 200K250K	Value_ 250K300K	Value_ 300K400K	Value_ 400K500K	Value_ 500K750K	Value_ 750K1M	Value_ 1M1.5M	Value_ 1.5M2M	Value_2M
Value_ 100K150K	Pearson correlation	.418**	.628**	1	.742**	.424**	.232**	.128**	-.020	-.082**	-.104**	-.105**	-.081**	-.060**
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.335	.000	.000	.000	.000	.004
Value_ 150K200K	Pearson correlation	.383**	.498**	.742**	1	.597**	.325**	.197**	.022	-.099**	-.134**	-.129**	-.100**	-.073**
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.000	.300	.000	.000	.000	.000	.001
Value_ 200K250K	Pearson correlation	.301**	.262**	.424**	.597**	1	.577**	.352**	.161**	-.095**	-.196**	-.187**	-.156**	-.113**
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.000	.000	.000	.000	.000	.000
Value_ 250K300K	Pearson correlation	.219**	.170**	.232**	.325**	.577**	1	.602**	.282**	-.034	-.231**	-.236**	-.202**	-.140**
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000	.107	.000	.000	.000	.000
Value_ 300K400K	Pearson correlation	.184**	.110**	.128**	.197**	.352**	.602**	1	.650**	.209**	-.236**	-.285**	-.260**	-.188**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000		.000	.000	.000	.000	.000	.000
Value_ 400K500K	Pearson correlation	.100**	.015	-.020	.022	.161**	.282**	.650**	1	.602**	-.049*	-.207**	-.229**	-.173**
	Sig. (2-tailed)	.000	.484	.335	.300	.000	.000	.000		.000	.019	.000	.000	.000
Value_ 500K750K	Pearson correlation	.013	.003	-.082**	-.099**	-.095**	-.034	.209**	.602**	1	.435**	.026	-.097**	-.107**
	Sig. (2-tailed)	.537	.871	.000	.000	.000	.107	.000	.000		.000	.221	.000	.000
Value_ 750K1M	Pearson correlation	-.094***	-.062**	-.104**	-.134**	-.196**	-.231**	-.236**	-.049*	.435**	1	.592**	.319**	.110**
	Sig. (2-tailed)	.000	.003	.000	.000	.000	.000	.000	.019	.000		.000	.000	.000
Value_ 1M1.5M	Pearson correlation	-.101**	-.078**	-.105**	-.129**	-.187**	-.236**	-.285**	-.207**	.026	.592**	1	.717**	.411**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.221	.000		.000	.000
Value_ 1.5M2M	Pearson correlation	-.085**	-.066**	-.081**	-.100**	-.156**	-.202**	-.260**	-.229**	-.097**	.319**	.717**	1	.664**
	Sig. (2-tailed)	.000	.002	.000	.000	.000	.000	.000	.000	.000	.000	.000		.000
Value_2M	Pearson correlation	-.062**	-.048*	-.060**	-.073**	-.113**	-.140**	-.188**	-.173**	-.107**	.110**	.411**	.664**	1
	Sig. (2-tailed)	.003	.023	.004	.001	.000	.000	.000	.000	.000	.000	.000	.000	
Value_ Avrg	Pearson correlation	-.225**	-.192**	-.241**	-.275**	-.350**	-.392**	-.425**	-.289**	-.015	.427**	.673**	.733**	.713**

Note. ** Correlation is significant at the .01 level (2-tailed); * Correlation is significant at the .05 level (2-tailed).

Table C4. *Second Correlation Matrix for Income-Level Variables (N = 2,245)*

		Income_ less15K	Income_ 15K25K	Income_ 25K35K	Income_ 35K50K	Income_ 50K75K	Income_ 75K100K	Income_ 100K150K	Income_ 150K200K	Income_200K
Income_Less15k	Pearson correlation	1	.743**	.605**	.519**	.352**	.117**	-.025	-.100**	-.117**
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.232	.000	.000
Income_15K25K	Pearson correlation	.743**	1	.769**	.641**	.403**	.131**	-.102**	-.216**	-.265**
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.000	.000
Income_25K35K	Pearson correlation	.605**	.769**	1	.726**	.446**	.200**	-.027	-.188**	-.266**
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.207	.000	.000
Income_35K50K	Pearson correlation	.519**	.641**	.726**	1	.658**	.398**	.183**	.015	-.156**
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.000	.477	.000
Income_50K75K	Pearson correlation	.352**	.403**	.446**	.658**	1	.730**	.517**	.344**	.101**
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.000	.000
Income_75K100K	Pearson correlation	.117**	.131**	.200**	.398**	.730**	1	.770**	.556**	.294**
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000	.000
Income_100K150K	Pearson correlation	-.025	-.102**	-.027	.183**	.517**	.770**	1	.816**	.569**
	Sig. (2-tailed)	.232	.000	.207	.000	.000	.000		.000	.000
Income_150K200K	Pearson correlation	-.100**	-.216**	-.188**	.015	.344**	.556**	.816**	1	.775**
	Sig. (2-tailed)	.000	.000	.000	.477	.000	.000	.000		.000
Income_200K	Pearson correlation	-.117**	-.265**	-.266**	-.156**	.101**	.294**	.569**	.757**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	

Note. ** Correlation is significant at the .01 level (2-tailed); * Correlation is significant at the .05 level (2-tailed).

Table C5. *Correlation Matrix for Income-Level Variables (N = 2,245)*

		Income_35K50K	Income_Less 5K35K_Comp	Income_ 50K100K_Comp	Income_ 100K200K_ Comp
Income_35K50K	Pearson correlation	1	.672**	.584**	.121**
	Sig. (2-tailed)		.000	.000	.000
Income_Less 5K35K_Comp	Pearson correlation	.672**	1	.335**	-.106**
	Sig. (2-tailed)	.000		.000	.000
Income_50K100K_Comp	Pearson correlation	.584**	.335**	1	.620**
	Sig. (2-tailed)	.000	.000		.000
Income_100K200K_Comp	Pearson correlation	.121**	-.106**	.620**	1
	Sig. (2-tailed)	.000	.000	.00	

Note. ** Correlation is significant at the .01 level (2-tailed); * Correlation is significant at the .05 level (2-tailed).

Table C6. *Correlation Matrix for Some Demographic Variables (N = 2,270)*

		TOTHouse Units	TOTPOP	VACANT Units	RENTER Units	POPDENS	OWNER Units
TOTHouseUnits	Pearson correlation	1	.711**	.551**	.713**	.008	.471**
	Sig. (2-tailed)		.000	.000	.000	.688	.000
TOTPOP	Pearson correlation	.711**	1	.200**	.368**	-.015	.535**
	Sig. (2-tailed)	.000		.000	.000	.486	.000
VACANTUnits	Pearson correlation	.551**	.200**	1	.507**	.093**	.000
	Sig. (2-tailed)	.000	.000		.000	.000	.983
RENTERUnits	Pearson correlation	.713**	.368**	.507**	1	.462**	-.274**
	Sig. (2-tailed)	.000	.000	.000		.000	.000
POPDENS	Pearson correlation	.008	-.015	.093**	.462**	1	-.558**
	Sig. (2-tailed)	.688	.486	.000	.000		.000
OWNERUnits	Pearson correlation	.471**	.535**	.000	-.274**	-.558**	1
	Sig. (2-tailed)	.000	.000	.983	.000	.000	

Note. ** Correlation is significant at the .01 level (2-tailed); * Correlation is significant at the .05 level (2-tailed).

Table C7. *Correlation Matrix for Education-Level Variables (N = 2,245)*

		EduHigh Deploma	Edu Alternative Cred	EduSome Colge	Edu ASSCDEG	EduBACH DEG	EduGRAD DEG
EduHighDeploma	Pearson correlation	1	.516**	.515**	.400**	-.129**	-.291**
	Sig. (2-tailed)		.000	.000	.000	.000	.000
EduAlternativeCred	Pearson correlation	.516**	1	.453**	.300**	-.068**	-.162**
	Sig. (2-tailed)	.000		.000	.000	.001	.000
EduSomeColge	Pearson correlation	.515**	.453**	1	.773**	.397**	.256**
	Sig. (2-tailed)	.000	.000		.000	.000	.000
EduASSCDEG	Pearson correlation	.400**	.300**	.733**	1	.516**	.368**
	Sig. (2-tailed)	.000	.000	.000		.000	.000
EduBACHDEG	Pearson correlation	-.129**	-.068**	.397**	.516**	1	.838**
	Sig. (2-tailed)	.000	.001	.000	.000		.000
EduGRADDEG	Pearson correlation	-.291**	-.162**	.256**	.368**	.838**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	

Note. ** Correlation is significant at the .01 level (2-tailed); * Correlation is significant at the .05 level (2-tailed).

Table C8. *Correlation Matrix for Expenditure Variables (N = 2,245)*

		Annual_Bdget_ Exp	Electricity_ Avrg	SpendinApperal Avrg	SpendinDining Avrg	SpendEduAvrg
Annual_Bdget_Exp	Pearson correlation	1	.911**	.998**	.997**	.992**
	Sig. (2-tailed)		.000	.000	.000	.000
Electricity_Avrg	Pearson correlation	.991**	1	.989**	.989**	.981**
	Sig. (2-tailed)	.000		.000	.000	.000
SpendinApperalAvrg	Pearson correlation	.998**	.989**	1	.999**	.987**
	Sig. (2-tailed)	.000	.000		.000	.000
SpendinDiningAvrg	Pearson correlation	.997**	.989**	.999**	1	.983**
	Sig. (2-tailed)	.000	.000	.000		.000
SpendEduAvrg	Pearson correlation	.992**	.981**	.987**	.983**	1
	Sig. (2-tailed)	.000	.000	.000	.000	

Note. ** Correlation is significant at the .01 level (2-tailed); * Correlation is significant at the .05 level (2-tailed).

Appendix D: Survey of Residential Solar Energy Adoption Spatial Decision Support System (RSEA-SDSS)

Part I: General Information

1. Have you finished reviewing the Residential Solar Energy Adoption System (RSEA-SDSS)?

   □ Yes □ No

2. To what extent are you consider your knowledge of (RSEA-SDSS)?

Completely Un-experienced						Completely Experienced
1	2	3	4	5	6	7

3. Would you like to participate in an interview? □ Yes □ No

Part II: Survey Questions:

II.I. For each of the following statements, mark one box that best describes your reactions to the (RSEA-SDSS)

		Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1	I think that I would like to use this system frequently	1	2	3	4	5
2	I found the system unnecessarily complex	1	2	3	4	5
3	I thought the system was easy to use	1	2	3	4	5
4	I think that I would need the support of a technical person to be able to use this system	1	2	3	4	5
5	I found the various functions in this system were well integrated	1	2	3	4	5
6	I thought there was too much inconsistency in this system	1	2	3	4	5
7	I would imagine that most people would learn to use this system very quickly	1	2	3	4	5
8	I found the system very cumbersome to use	1	2	3	4	5
9	I felt very confident using the system	1	2	3	4	5
10	I needed to learn a lot of things before I could get going with this system	1	2	3	4	5

II.II. Please tick the circle which most closely reflect your impressions about using (RSEA-SDSS)

	1	2	3	4	5	6	7		
annoying	○	○	○	○	○	○	○	enjoyable	1
not understandable	○	○	○	○	○	○	○	understandable	2
creative	○	○	○	○	○	○	○	dull	3
easy to learn	○	○	○	○	○	○	○	difficult to learn	4
valuable	○	○	○	○	○	○	○	inferior	5
boring	○	○	○	○	○	○	○	exciting	6
not interesting	○	○	○	○	○	○	○	interesting	7
unpredictable	○	○	○	○	○	○	○	predictable	8
fast	○	○	○	○	○	○	○	slow	9
inventive	○	○	○	○	○	○	○	conventional	10
obstructive	○	○	○	○	○	○	○	supportive	11
good	○	○	○	○	○	○	○	bad	12
complicated	○	○	○	○	○	○	○	easy	13
unlikable	○	○	○	○	○	○	○	pleasing	14
usual	○	○	○	○	○	○	○	leading edge	15
unpleasant	○	○	○	○	○	○	○	pleasant	16
secure	○	○	○	○	○	○	○	not secure	17
motivating	○	○	○	○	○	○	○	demotivating	18
meets expectations	○	○	○	○	○	○	○	does not meet expectations	19
inefficient	○	○	○	○	○	○	○	efficient	20
clear	○	○	○	○	○	○	○	confusing	21
impractical	○	○	○	○	○	○	○	practical	22
organized	○	○	○	○	○	○	○	cluttered	23
attractive	○	○	○	○	○	○	○	unattractive	24
friendly	○	○	○	○	○	○	○	unfriendly	25
conservative	○	○	○	○	○	○	○	innovative	26

Part III: About yourself and your organization

1. Your gender ◻ Male ◻ Female

2. Your highest level of education

◻ Less than high school ◻ Undergraduate degree

◻ High school degree ◻ Graduate degree

◻ College degree ◻ Other

3. Your age

□ 18–25 □ 56–65

□ 26–35 □ 66–75

□ 36–45 □ 76–85

4. In which industry is your organization operating?

□ Information Technology □ Education

□ Financial services □ Energy □Manufacturing

□ Government □ Nonprofit □Real State □Health care

□ Utility □ Other, please specify ______________

5. Have you worked in Information Technology domain?

□ Yes □ No

6. If yes, how many years have you worked in renewable energy domain? _______

7.How many years have you worked for your current organization?

□ less than 1 year □ 11 - 15 years

□ 1 – 5 years □ more than 15 years

□ 6 – 10 years

8. Your job title is ______________________________________

10. Number of employees in your organization

□ Fewer than 500 □ 5,000–10,000

□ 500–999 □ More than 10,000

□ 1,000–4,999

Thank you for participation in this study.

Appendix E: Institutional Review Board Approval

Claremont Graduate University
Institutional Review Board

Dear Ahmed,

The IRB manager has determined protocol **3395 Adoption of Residential Solar Energy: Exploratory Study Approach and Spatial Decision Support System (SDSS)** is not human subjects research and does not require further IRB review or oversight.

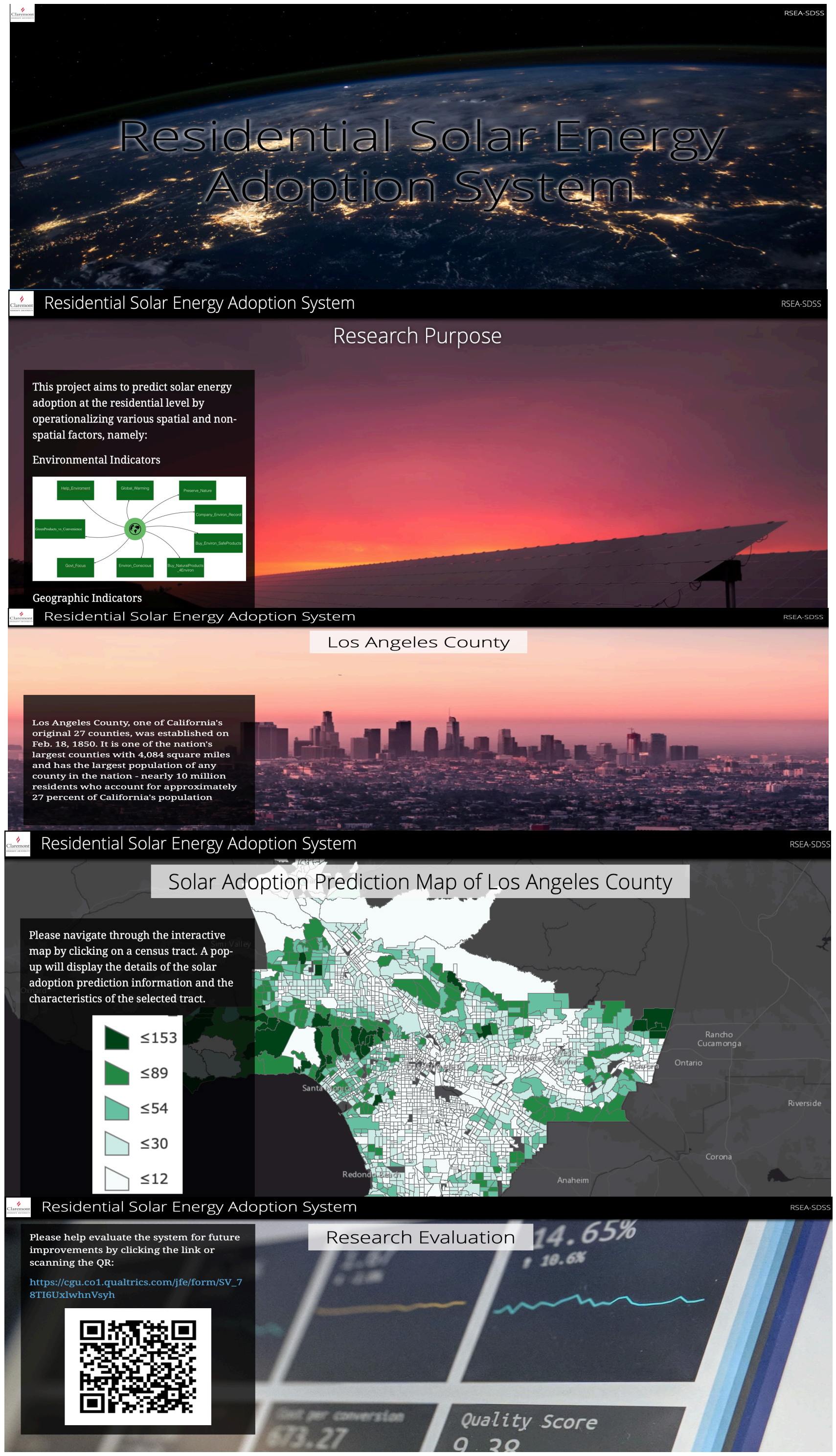
Please note that changes to your protocol may affect this determination. Please contact me directly to discuss any changes you may contemplate.

Thanks,

James Griffith,
IRB Manager
james.griffith2@cgu.edu

150 East Tenth Street ● Claremont, California 91711-6160
Tel: 909.607.9406

Appendix F: Snaps of RSEA-SDSS story map



Appendix G: Email sample for evaluation participation request

Dear Invitee,

My name is Ahmed Alzahrani. I am a doctoral student at Claremont Graduate University 's Information Systems and Technology Program. I am kindly requesting your participation in a doctoral research study that I am conducting titled: Adoption of Residential Solar Energy: Exploratory Study Approach and Spatial Decision Support System. The intention is to assess the effectiveness and utility of the algorithm used to predict solar energy adoption in residential areas in LA county.Participation is completely voluntary, and you may withdraw from the study at any time. The study is completely anonymous; therefore, it does not require you to provide your name or any other identifying information.

To begin the study, open the story map and navigate through the system and then click the survey link at the end. Your participation in this research will be of great importance to assist solar energy installations firms, policymakers, and utility companies in understanding the factors associated with solar installations in LA county. Ultimately, the system would benefit the stakeholders such as yourself in targeting your efforts to increase solar energy adoption. Thank you for your time and participation! Simply click in the links to go to directly to the map or the survey. If the link does not work, copy and paste the URL int the address bar of your Internet browser. Your participation in this research is strictly voluntary.

The website for the map is: https://agis.maps.arcgis.com/apps/Cascade/index.html?appid=81bcddf50b3c429ba87f2a01b1dcd3d1

The website for the survey is: https://cgu.co1.qualtrics.com/jfe/form/SV_78TI6UxlwhnVsyh

Sincerely,
Ahmed Alzahrani
Doctoral Candidate, GIS advanced lab, Center of Information and Technology
Claremont Graduate University
150 E 10th St,
Claremont, CA 91711

Appendix H: Prediction of Solar Adoption Solution in LA County

1.Scan the QR code to access the interactive map

OR
On ArcGIS online: Residential Solar Energy Adoption System

2. Scan the QR code top open the survey

OR
Find the link at the end go the story map

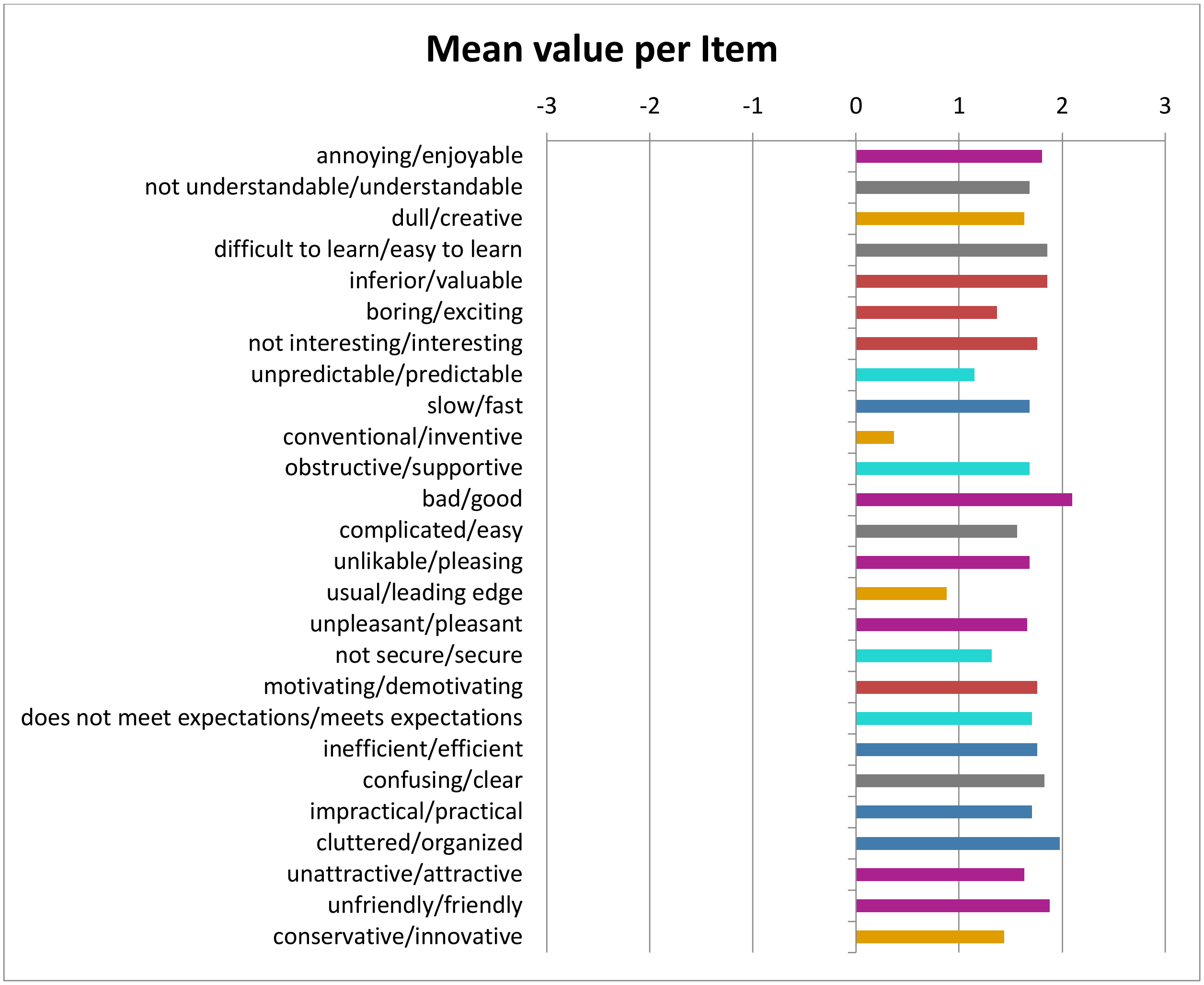
My name is Ahmed Alzahrani. I am a doctoral student at Claremont Graduate University 's Information Systems and Technology Program, under Dr.Brian Hilton supervision. I am kindly requesting your participation in a doctoral research study that I am conducting titled: **Adoption of Residential Solar Energy: Exploratory Study Approach and Spatial Decision Support System**. The intention is to assess through a short survey the effectiveness and utility of the algorithm used to predict solar energy adoption in residential areas in LA county. Participation is voluntary and completely anonymous. Scan the QR code to open the story map, navigate through the map and kindly click the survey link at the end for feedback. Hopefully, this research would assist solar energy installations firms, policymakers, and utility companies in understanding the factors associated with solar installations in LA county, which ultimately would help increasing the adoption of solar technology.

Doctoral Candidate, GIS advanced lab, Center of Information and Technology
Tel: 909 224 9928
Email: ahmed.alzahrani@cgu.edu
Claremont Graduate University
150 E 10th St,
Claremont, CA 91711

Appendix I: Descriptive Results the User Experience Questionnaire

Item	Mean	Variance	Std. Dev.	No.	Left	Right	Scale	
1	1.8	1.3	1.1	41	annoying	enjoyable	Attractiveness	
2	1.7	1.7	1.3	41	not understandable	understandable	Perspicuity	
3	1.6	1.5	1.2	41	creative	dull	Novelty	
4	1.9	1.6	1.3	41	easy to learn	difficult to learn	Perspicuity	
5	1.9	1.4	1.2	41	valuable	inferior	Stimulation	
6	1.4	1.6	1.3	41	boring	exciting	Stimulation	
7	1.8	1.4	1.2	41	not interesting	interesting	Stimulation	
8	1.1	2.3	1.5	41	unpredictable	predictable	Dependability	
9	1.7	1.4	1.2	41	fast	slow	Efficiency	
10	0.4	2.7	1.7	41	inventive	conventional	Novelty	
11	1.7	1.4	1.2	41	obstructive	supportive	Dependability	
12	2.1	0.8	0.9	41	good	bad	Attractiveness	
13	1.6	2.1	1.4	41	complicated	easy	Perspicuity	
14	1.7	1.3	1.1	41	unlikable	pleasing	Attractiveness	
15	0.9	2.4	1.5	41	usual	leading edge	Novelty	
16	1.7	0.9	1.0	41	unpleasant	pleasant	Attractiveness	
17	1.3	1.8	1.3	41	secure	not secure	Dependability	
18	1.8	1.1	1.1	41	motivating	demotivating	Stimulation	
19	1.7	1.4	1.2	41	meets expectations	does not meet expectations	Dependability	
20	1.8	1.1	1.1	41	inefficient	efficient	Efficiency	
21	1.8	1.7	1.3	41	clear	confusing	Perspicuity	
22	1.7	1.3	1.1	41	impractical	practical	Efficiency	
23	2.0	1.2	1.1	41	organized	cluttered	Efficiency	
24	1.6	1.5	1.2	41	attractive	unattractive	Attractiveness	
25	1.9	1.0	1.0	41	friendly	unfriendly	Attractiveness	
26	1.4	2.0	1.4	41	conservative	innovative	Novelty	

Appendix J: the UEQ Mean Value per Item



Appendix K: Excluded Variables from the Regression Output Model

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics Tolerance
17	Govt_Focus	-.036[r]	-1.803	.072	-.038	.488
	Help_Enviroment	.006[r]	.341	.734	.007	.564
	Company_Environ_Record	.014[r]	.880	.379	.019	.739
	GreenProducts_vs_Convenience	-.010[r]	-.611	.542	-.013	.660
	NEVMARR	.005[r]	.310	.757	.007	.637
	MARRIED	-.030[r]	-1.103	.270	-.023	.255
	WIDOWED	.022[r]	1.210	.226	.026	.606
	DIVORCD	.012[r]	.652	.514	.014	.528
	WHITE	.003[r]	.131	.896	.003	.411
	AmericIndian	.017[r]	.996	.319	.021	.626
	ASIAN	-.029[r]	-1.726	.084	-.037	.701
	EduHighDeploma	.002[r]	.077	.938	.002	.435
	EdeAlteernativeCred	.014[r]	.864	.388	.018	.691
	EduGRADDEG	.011[r]	.396	.692	.008	.268
	VACANTUnits	-.011[r]	-.722	.470	-.015	.846
	RENTERUnits	-.028[r]	-1.480	.139	-.031	.539
	POPDENS	-.027[r]	-1.459	.145	-.031	.567
	Income_less5k35k_comp	-.034[r]	-1.732	.083	-.037	.503
	Income_35k50K	-.008[r]	-.410	.682	-.009	.576
	Income_50k100k_comp	-.006[r]	-.315	.753	-.007	.459
	Income_100k200k_comp	-.029[r]	-1.021	.307	-.022	.243
	Income_200K	-.068[r]	-1.702	.089	-.036	.121
	Value_300K400K	-.043[r]	-1.142	.253	-.024	.136

a. Dependent Variable: existing_installs_count