Research Report – UCD-ITS-RR-13-04

Updating the PECAS Modeling Framework to Include Energy Use Data for Buildings

February 2013

Giovanni Circella
Robert Johnston
Andrew Holguin
Eric Lehmer
Yang Wang
Michael McCoy

Institute of Transportation Studies ◦ University of California, Davis
One Shields Avenue ◦ Davis, California 95616
PHONE (530) 752-6548 ◦ FAX (530) 752-6572
www.its.ucdavis.edu

Public Interest Energy Research (PIER) Program

INTERIM PROJECT REPORT

Updating the PECAS Modeling Framework to Include Energy Use Data for Buildings

Prepared for: California Energy Commission

Prepared by: University of California Davis, Urban Land Use and Transportation Center

FEBRUARY 2013

Prepared by:

Primary Author(s):
Giovanni Circella*‡
Robert A. Johnston
Andrew J. Holguin
Eric W. Lehmer
Yang Wang
Michael McCoy†

* Principal Investigator, ‡ Corresponding Author
†Principal Investigator until October 2012.

University of California Davis
Urban Land Use and Transportation Center
One Shields Avenue
Davis, CA 95616
ultrans.its.ucdavis.edu

Contract Number: 500-10-033

Prepared for:

California Energy Commission

Dan Gallagher
Contract Manager

Dan Gallagher, for PIER; Zoe Elizabeth, for UCLA
Project Manager

Linda Spiegel
Office Manager
Energy Generation Research Office (EGRO)

Laurie ten Hope
Deputy Director
Energy Research and Development Division (ERDD)

Robert Oglesby
Executive Director

DISCLAIMER

This report was prepared as the result of work sponsored by the California Energy Commission. It does not necessarily represent the views of the Energy Commission, its employees or the State of California. The Energy Commission, the State of California, its employees, contractors and subcontractors make no warrant, express or implied, and assume no legal liability for the information in this report; nor does any party represent that the uses of this information will not infringe upon privately owned rights. This report has not been approved or disapproved by the California Energy Commission nor has the California Energy Commission passed upon the accuracy or adequacy of the information in this report.

ACKNOWLEDGEMENTS

We would like to acknowledge the support of the Los Angeles Department of Water and Power (LADWP) and Long Beach Gas and Oil (LBGO), which provided utility consumption data for this study. We also want to thank Stephanie Pincetl, Zoe Elizabeth and Sinnott Murphy from the University of California, Los Angeles, Paul Bunje from Los Angeles Regional Collaborative for Climate Action and Sustainability, and Lauren Rank from the Los Angeles County, whose support has been fundamental to gain access to electricity and natural gas consumption data in Los Angeles County. Sungbin Cho, from the Southern California Association of Governments (SCAG) provided useful insights and additional data on the distribution of the built floorspace in Los Angeles County. James Thorne, from the University of California, Davis, helped in the early stages of the research and provided the PRISM climate data. We also want to thank Patricia Mokhtarian and Nate Roth, from the University of California, Davis, for their useful advice during the development of the project and for their other contributions to the quality of the research. Finally, Ryan Boynton, also at UC Davis, and Luis Guzman, from the Universidad de Las Andes in Bogotá (Colombia), helped in the early stages of the data development and literature review for this project. The authors are responsible for any remaining errors in this report.

We would like also to acknowledge the technical advisory committee and the Institute of the Environment and Sustainability at the University of California Los Angeles for their valuable feedback and contributions with regard to the completion of this project. The technical advisory committee includes Sungbin Cho from Southern California Area Governments, Cristiano Facanha from ICF, Gordon Garry from the Sacramento Area Council of Governments, Monica Gilchrist from ICLEI, Kevin Gurney from Arizona State University, Beth Jines from the Los Angeles Department of Water and Power, Chris Kennedy from the University of Toronto, Jesse Langley from Southern California Edison, Josh Newell from the University of Michigan, Diane Pataki from UC Irvine, Lauren Rank from Los Angeles County, and Jennifer Stokes from UC Berkeley.

PREFACE

The California Energy Commission Public Interest Energy Research (PIER) Program supports public interest energy research and development that will help improve the quality of life in California by bringing environmentally safe, affordable, and reliable energy services and products to the marketplace.

The PIER Program conducts public interest research, development, and demonstration (RD&D) projects to benefit California.

The PIER Program strives to conduct the most promising public interest energy research by partnering with RD&D entities, including individuals, businesses, utilities, and public or private research institutions.

PIER funding efforts are focused on the following RD&D program areas:

- Buildings End-Use Energy Efficiency
- Energy Innovations Small Grants
- Energy-Related Environmental Research
- Energy Systems Integration
- Environmentally Preferred Advanced Generation
- Industrial/Agricultural/Water End-Use Energy Efficiency
- Renewable Energy Technologies
- Transportation

Updating the PECAS Modeling Framework to Include Energy Use Data for Buildings is the interim report for the Methodology to Establish Regional Energy Baselines project (contract number *500-10-033*) conducted by UCLA's Institute of the Environment and Sustainability (subcontract to UC Davis Urban Land Use and Transportation Center). The information from this project contributes to PIER's Transportation Research Area.

For more information about the PIER Program, please visit the Energy Commission's website at www.energy.ca.gov/research/ or contact the Energy Commission at 916-654-4878.

ABSTRACT

Building operations nowadays account for an important portion of total energy consumption. This study investigates the consumption of electricity and natural gas for building operations for several categories of residential and non-residential buildings. The purpose of the study is to update the Production Exchange Consumption Allocation System (PECAS) land use modeling framework to include energy components. The proposed approach is useful to serve as part of an urban metabolism framework, creating a methodology to account for environmental and energy balances of cities and complex regions. Annual electricity and natural gas consumption data from utility companies operating in Los Angeles County are used to build an energy database to study energy consumption in buildings, based on the analysis of almost 450,000 Energy Analysis Zones. Additional data on building stock, climate zones, geomorphological data, and sociodemographics are collected from multiple sources and integrated into the energy database. We conduct statistical analysis of utility data and estimate linear regression models to predict energy consumption in buildings. Electricity and natural gas consumption in residential and non-residential buildings are studied in relation to several variables, including building use type, building size, and climate zone. Energy profiles are created for several categories of buildings. Annual energy consumption is estimated for various types of residential units. Electricity and natural gas consumption per square foot of developed floorspace is estimated for various categories of non-residential buildings. We validate the results of the analyses through validity checks carried out using data from independent sources, including the California Residential Appliance Saturation Study (RASS) and the Commercial End-Use Survey (CEUS), given the limited amount of energy data provided by the utility companies, to date, and the lack of overlapping data for the consumption of both electricity and natural gas in the same zones. The results of the study are useful to update the PECAS land use modeling framework, and form part of the baseline study to estimate energy and greenhouse gas balances in an urban metabolism framework for the analysis of the environmental impacts of complex urban regions. The results also allow us to estimate the total energy consumption and greenhouse gas emissions for residential and commercial building operations through the application to the total residential and commercial building inventory in the region. These results are then useful for the evaluation of possible energy savings in buildings.

Keywords: Energy consumption, building operations, urban metabolism, land use modeling, energy mapping, energy efficiency, urban sustainability, sustainable communities

Please use the following citation for this report:

Circella, Giovanni, Robert A. Johnston, Andrew J. Holguin, Eric W. Lehmer, Yang Wang and Mike McCoy (University of California Davis Urban Land Use and Transportation Center). 2013. ***Updating the PECAS Modeling Framework to Include Energy Use Data for Buildings***. California Energy Commission. February 2013.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Urban metabolism (UM) is a comprehensive systems approach for assessing a city's sustainability. It measures the total energy, materials, and waste products that flow into and out of an urban area. As part of an urban metabolism analysis of the environmental impacts in complex regions, the Production Exchange Consumption Allocation System (PECAS) land use modeling framework is used to simulate the interactions among transportation, land use development and energy consumption.

The PECAS modeling framework is a complex land use transportation modeling framework that allows simulating land use development and the allocation of economic activities. It has the ability to assess and depict the interregional effects of major changes to land uses, economics, and transportation on the economy and the environment through the forecast of the interactions among economic activities, residential locations and travel behavior. Several updates were introduced in the PECAS modeling framework to properly account for energy consumption and greenhouse emissions associated with the various sectors of economic activities, land use and building operations.

The interim methodology report *Methodology to Establish Regional Energy Baselines* (Pincetl et al., 2012) provides a detailed description of the process that was followed to update the PECAS model to account for the energy consumption and greenhouse emissions associated with commodity flows and economic activities in a region. The approach is based on the integration of energy consumption data (for both direct, indirect, and lifecycle effects) that are associated with the production (and/or consumption) of each unit of economic output by economic sector (expressed in US$) in the PECAS modeling framework.

The remainder of this report describes the process that was developed in order to update the PECAS model also to account for energy consumption associated with building operations. This task was developed through the analysis of an extensive dataset for energy consumption in buildings, built using data for energy consumption from utility companies in Los Angeles County. Several types of buildings are studied to account for differential energy consumption, various land use types and building categories. A comprehensive database that includes almost 450,000 energy analysis zones (EAZs) is developed as part of the activities. The database is populated with energy consumption data obtained from the utility companies, information on the building stock in the region, climate and geomorphological data, and sociodemographics.

In the development of the research project, energy profiles are defined for single-family and multi-family residential buildings, as well as for several categories of non-residential buildings. The results of the analysis allow studying the variation of energy consumption for building operations in different floorspace types. Also in consideration of the limitations to the available data provided by the utility companies, in terms of both area of coverage and level of spatial aggregation, we run a series of validation checks using data from independent sources, such as the California Energy Commission's Residential Appliance Saturation Surveys (RASS) and the Commercial End Use Survey (CEUS).

The final results of the analysis confirm the impact of climate (i.e. California climate zones) as an important explanatory variable for energy consumption. Additionally, the impact of size and type of residential units is estimated, with a stronger impact on energy consumption associated with larger single family housing units, in particular for those that are equipped with a pool in their premises. Detailed analyses of per capita energy consumption (for electricity and natural gas) are carried out, and they highlight the role of building age, slope and aspects as important predictors of energy consumption. In particular, the age of the buildings is responsible for significant changes in energy consumption patterns in correspondence to the various thresholds for the implementation of Title 24 energy efficiency standards in California).

The analysis of energy consumption for non-residential buildings highlights many differences in energy consumption existing among various building categories. Estimated energy consumption for a square foot of floorspace types are computed for several categories and inform the modeling system on the consumption of electricity and natural gas associated with the different building types.

The analysis of energy consumption for building operations and the update of the land use modeling framework with energy data is fundamental in the definition of the baseline urban metabolism approach for the estimation of energy consumption and GHG emissions in complex regions. The application of the proposed approach to the building stock in the area of study allows researchers to estimate the distribution of energy consumption by category of building type, and it allows forecasts of energy consumption depending on the future expansion and modifications of the building stock. Thus, it yields more comprehensive knowledge on the formation of GHG emissions and on the energy impacts of various land use and other policy scenarios. The updated land use modeling framework provides an updated resource for MPOs struggling with SB 375 implementation, and it is useful to develop energy use and GHG baseline analyses and policy trend analyses in an urban metabolism framework.

CHAPTER 1: Introduction

Urban metabolism (UM) is a comprehensive systems approach for assessing a city's sustainability. Urban metabolism measures the total energy, materials, and waste products that flow into and out of an urban area. Urban metabolism analysis allows the identification and quantification of interactions between transportation, land use, water, waste and energy of the urban region to improve understanding of the material basis upon which a complex urban system depends. Cities concentrate energy and resources drawn from near and far for use in relatively compact spaces and UM is an important method for understanding the relationship between cities and the wider environment.

The first urban metabolism was conducted by Abel Wolman (1965), who estimated the water needs for a city of one million people to highlight the potential of resource scarcity over the longer term. Since then, over 50 studies have been conducted on cities across the globe. However, due to data availability limitations, these studies are typically performed at the scale of the whole city using average annual data or impute state or national data to localities. Obviously, this approach faces practical challenges and limits the ability of researchers to determine the specific metabolism of each place, key to understanding how places may differ and what their particular sustainability challenges might be. A second issue is that measuring flows across urban boundaries neglects some of the key metabolic processes within the city such as storage (water in aquifers, nutrients in waste dumps, materials in building stocks) (Kennedy et al., 2007), or local transportation, including water, energy and materials (Kierstead and Sivakumar, 2012). As Kierstead and Sivakumar point out, the distinction between aggregate urban metabolism data and the highly resolved demand data necessary for policy decision making into the twenty-first century presents a substantial research challenge.

As described in the interim report *Methodology to Establish Regional Energy Baselines (*Pincetl *et al*., 2012), this project advances urban metabolism research in three ways. First, researchers employ an analytical platform that enables quantification of the direct and embodied energy and emissions associated with economic activity occurring within the county. Second, researchers assess lifecycle, rather than solely direct, energy and emissions consumed and generated by the study area. Finally, researchers use significantly more granular data than in previous urban metabolism studies, enabling consideration of spatial patterns of energy use and waste production.

This pilot urban metabolism analysis of Los Angeles County addresses the problem of aggregate urban metabolism data and the need for highly resolved data through integrating methodologies. Researchers combine activity-based modeling of land use and transportation (using the California PECAS model) with highly disaggregated data on direct flows of electricity, gas, and water consumption collected from utilities serving Los Angeles County, economic input-output life-cycle analysis by industry sector, process-based life-cycle analysis of building materials for 32 building types, and hybrid life-cycle analysis of roadway and parking infrastructure. This last component takes into account existing stocks in cities that are "stored," as embedded energy that has thus far not been addressed in UM studies. The PECAS model provides the data integration and synthesis platform for this analysis. Flows of solid waste and air emissions are also added, with different scales of resolution reflecting data availability. This data is then overlaid with county parcel assessor information of land use type, building type, and building age, and employment data, socio-demographic data, and other information.

This report describes the approach that was developed to update the Production Exchange Consumption Allocation System (PECAS) modeling framework to include energy data and allow the estimation of energy consumption and greenhouse gas emissions in the proposed urban metabolism approach. The PECAS modeling framework is a complex land use and transportation modeling framework that allows the simulation of land use development and the allocation of economic activities. It has the ability to assess the interregional effects of major changes to land uses, economics, and transportation on the economy and the environment through the forecast of the interactions among economic activities, residential locations and travel behavior. Several updates were introduced in the PECAS modeling framework to properly account for energy consumption and greenhouse emissions associated with the various sectors of economic activities, land use and building operations.

Chapter 4 in the interim methodology report for this project (Pincetl *et al*., 2012) provides a detailed description of the process that was adopted to update the PECAS modeling framework to account for the energy consumption and greenhouse emissions associated with commodity flows and economic activities in a region. The approach is based on the integration of energy consumption data (for both direct, indirect, and lifecycle effects) that are associated with the production (and/or consumption) of each unit of economic output by economic sector of activities (expressed in US$) in the PECAS modeling framework.

The remainder of this report describes the research activities that were developed the update the PECAS modeling framework also to account for energy consumption associated with building operations. This research project investigates energy consumption patterns for residential and commercial buildings in Los Angeles County. The purpose of the research is to contribute to the energy baseline assessment for Los Angeles County, and to provide detailed information that can inform the land use modeling framework on energy consumption from buildings. The project is based on the analysis of energy consumption records provided by the utility companies that operate in Los Angeles County and on the generation of a comprehensive dataset that includes additional data from different sources that provide information on the building stock, land use patterns, geographic location, climate data, and sociodemographics.

Accurate accounting for energy consumption and greenhouse gas emissions from buildings is rather difficult for areas with large and complex land uses. Aggregate zonal estimates of energy use can be obtained from utilities, while individual building energy use by building type can be estimated from building characteristics. Previous studies have attempted to define energy baseline assessments for cities and regions. For example, a simple approach to generate GHG accounting would be to inventory the buildings in a zone and compute their estimated energy use from the literature. This approach, if useful to produce a "snapshot" of the estimated energy use in a zone based on past research and the building stock, is not sensitive enough to the effects of possible changes in land use, building technology, and building operations policies that are designed to reduce energy consumption and resulting greenhouse gas emissions.

Many studies have investigated energy consumption in buildings. Previous research has identified the main drivers of energy consumption, and has provided useful insights on the impact of technological solutions, energy costs, and building efficiency standards on energy consumption. These studies investigate the explanatory variables behind energy consumption in buildings and evaluate the potential for energy consumption reduction that could be achieved through the adoption of policies that favor energy efficiency and reduce environmental impacts from buildings.

This research project builds on previous experiences from the literature to develop an energy assessment for the building sector in Los Angeles County, one of the most populous and economically dynamic regions in the United States. The project aims at informing public agencies and decision-makers on the energy consumption and greenhouse gas emissions associated with the operation of various categories of existing buildings. The project investigates consumption of electricity and natural gas in the current stock of buildings in this region, at a highly disaggregated level of spatial details, through the definition of almost 450,000 Energy Analysis Zones in Los Angeles County. Through the estimation of econometric models, it provides energy profiles based on information on building age, location, land use, and sociodemographic variables. The results of the project are designed to serve as part of a land use modeling and urban metabolism approach for environmental analysis in the region.

The study builds on the previous experience developed by the researchers at the Urban Land Use and Transportation Center (ULTRANS) of the University of California, Davis on the treatment of spatial information and the development of land use and transportation modeling solutions to support informed decisions in planning. The study integrates data developed in previous projects carried out at UC Davis and produces information on energy consumption for building operations that is integrated in the PECAS land use modeling framework. The results of the project are useful for the definition of an energy baseline assessment for Los Angeles County. Additionally, they will be useful to inform land use models on the impact of changes in land use features on energy consumption in the area of study, as part of an ongoing modeling framework for economic activities, land use, transportation and energy use. In addition to providing information on energy consumption from buildings for Los Angeles County, the project also contributes to the current research in the field, through the development of a detailed modeling approach that studies energy use and environmental pollution effects from the building stock and that is of interest for many settings with complex land uses.

In the development of the project, researchers analyzed data from multiple sources to (1) create an inventory of the built floor space in Los Angeles County using information obtained from the Assessor's data; (2) generate a comprehensive database to study energy consumption in buildings that includes information on land use and the building stock, geographical location, climate data and sociodemographics; (3) analyze energy consumption using utility records obtained from the local utility companies that operate in the county; (4) estimate energy consumption models for various categories of existing buildings in the County; and (5) compute an assessment of energy consumption and the resulting greenhouse gas emissions in the region that can be used in an urban metabolism study.

The study is based on the integration of multiple sources of data and the development of energy consumption models from the analysis of utility records obtained from the local utility companies. The project analyzed data from Los Angeles County, and it was designed to support knowledge development on the impact of specific variables and selected building and environmental characteristics on energy efficiency. As part of this project, the researchers developed an innovative approach that integrates data collected from different sources and at different levels of spatial aggregation, and a comprehensive set of analytical tools that is able to investigate energy consumption in buildings in the area of study. Although these analytical tools and the spatial aggregation process are of general validity for the entire region of study, the completeness of the results from this study was somewhat limited by the availability of energy consumption records provided by local utility companies. Unfortunately, contrary to the expectations and previous contacts with utility companies and local administrators in the area of study, only two utility companies agreed and were able to provide good quality data on energy consumption in Los Angeles County, at a useful level of spatial disaggregation for this project. The Los Angeles Department of Water and Power (LADWP) provided data on the consumption of electricity in the city of Los Angeles, and the Long Beach Gas and Oil (LBGO) provided data on the consumption of natural gas in the Long Beach area.[1] Additional utility companies that operate in the area agreed to share energy consumption data, but no additional data were delivered in a useful format and in timely manner to allow use for this project, mainly as a result of possible concerns about privacy issues associated with the release of these data. The limits of the available data from utility companies limited the ability of the current project to cover the entire area of study and analyze in full depth the contributions to the formation of energy consumption behaviors.

Finally, an important source of limitation of the current study is associated with the lack of spatial overlap for the available data on the consumption of both energy and natural gas (the two main sources of energy that are used in buildings in the area of study). Such limitation, which derives from the availability of energy consumption data only for the areas respectively served by LADWP (electricity consumption) and LBGO (natural gas consumption) may cause distortions in the analysis of energy consumption in buildings[2] through the estimated energy consumption models.

[1] The limitation of the available data on energy consumption somewhat hampered the ability to develop comprehensive analysis of energy consumption in buildings in the entire area of study. The researchers made all necessary steps to develop alternative approaches that could reduce the disruptions caused by this issue on the quality of the research, as described in the following sections of this report. Despite the limits of the available data, the validity of the proposed analytical tools remains unchanged. The robustness of the results will increase when additional energy consumption data will become available.

[2] It might be responsible for the presence of "unobserved variables biases" in the estimation of energy consumption models, as discussed in Chapter 7 of this report.

Researchers originally planned, in the design of the study, to analyze energy consumption in buildings through the estimation of jointly estimated models for energy consumption of both electricity and natural gas. The lack of spatially overlapping data for the consumption of these two energy sources made it necessary to change the original research methodology. The estimation of independently estimated multiple linear regression models for the consumption of electricity and natural gas, as imposed by the limited data, may be responsible for the presence of eventual unobserved variable biases in the study. Natural gas and electricity are potential substitutes at least for some of the energy uses in buildings (e.g. heating in the residential sector), and the inability to estimate the consumption of both energy sources jointly may limit the validity, and robustness, of part of the results of the estimated models for energy consumption.

The latter sections of this document describe the steps that were developed by the researchers to deal with the issues associated with the limits of the available energy consumption data, to compare the results from the study with independent sources and to assess the impact of any eventual unobserved variable biases in the analysis of energy consumption in buildings. Chapter 7 also presents some simplified approaches for the analysis of energy consumption in buildings that were estimated using alternative datasets that do not involve the presence of unobserved variable biases, as in the case of data from the California Energy Commission's Residential Appliance Saturation Study (RASS) and Commercial End-Use Survey (CEUS). These simplified models of energy use, which miss some of the depth of the analysis that is allowed by the more detailed energy consumption models built on the analysis of energy utility records, can be used as benchmarks for the analysis of energy use in buildings for this project.

Despite the limits of the energy consumption data provided, to date, by the utility companies in Los Angeles County, the validity of the analytical tools developed as part of this study remains untouched. This research project embodies the first comprehensive study for the estimation of energy consumption in one of the most important and energy intensive metropolitan areas of the United States that is based on the analysis of utility records at a very fine level of spatial aggregation and that integrates many different sources of data for buildings, land use, climate data and sociodemographics. A great potential is associated with the full development of the proposed approach in future extensions of the project, when additional data on the consumption of electricity and natural gas will finally become available from the remaining utility companies that operate in the area. Using the analytical tools that have been developed in this research, it will be possible to estimate more comprehensive energy consumption models and investigate in more detail consumers' behavior associated with energy consumption. Besides, the application of the updated land use modeling framework, which includes the energy modeling component for building operations, will offer important insights on the forecasts for future energy and environmental impacts of land use development and of the policies developed to increase energy efficiency in buildings.

CHAPTER 2: Background

Several studies have investigated energy consumption in buildings. This section provides a summary overview of the relevant studies that have investigated the topic, including both studies that have focused on the analysis of energy consumption in residential buildings as well as in non-residential (and predominantly commercial) buildings. For each study, we reviewed the approach that was used, the available data, the level of spatial aggregation that was used in the analysis of energy consumption, and the main findings from the research. This literature review, even if not at all exhaustive of a field that is continuously evolving with many research projects being currently carried out, provides a brief overview of the main research streams that have focused on the analysis of energy consumption in buildings. It summarizes the background for the current research, and it identifies the supporting elements that led to the definition of the modeling approach for the analysis of energy consumption in buildings that is used in this study.

Many different approaches can be used to model the various aspects of energy supply and demand. Jebaraj and Iniyan (2006) conducted a review of the various emerging issues related to energy modeling, and discussed the various categories of models that have been developed to analyze certain aspects of energy production and consumption, including energy planning models, energy supply–demand models, forecasting models (e.g., commercial energy models, renewable energy models, etc.), optimization models, specific methodologies used to estimate energy consumption (e.g. using neural networks), and emission reduction models. All these models have received important attention in scientific research, given the importance of the topic, and the need for researchers and stakeholders to identify key variables that affect the demand for energy use, the sensitivity of energy demand (and supply) to perturbations in the economic, legislative and policy frameworks, and the environmental impacts associated with energy use.

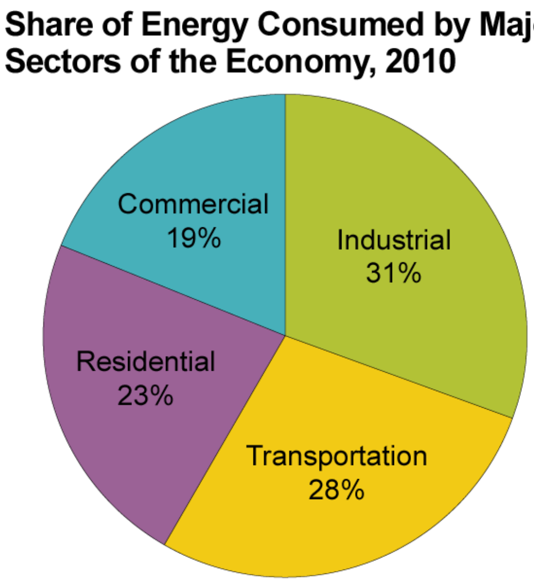
Final energy consumption is usually split into three main sectors: industry, transport, and others, including the service sector and residential buildings (Figure 1). Energy consumption in buildings other than residential dwellings constitutes a sizeable fraction of 'other' sectors. In developed countries, buildings account for 20-40% of the total final energy consumption.[3] The service sector, which covers all commercial and public buildings, includes many types of buildings (schools, restaurants, hotels, hospitals, museums, etc.) with a wide variety of uses and energy services - heating, ventilation and air conditioning (HVAC), domestic hot water (DHW), lighting, refrigeration, food preparation, etc.

In non-domestic buildings, the type of use and activities make a huge impact on the quality and quantity of energy services needed. Office and retail are the most energy intensive building categories typically accounting for over 50% of the total energy consumption for non-residential buildings. Other important building types that are responsible for a sizable portion of energy

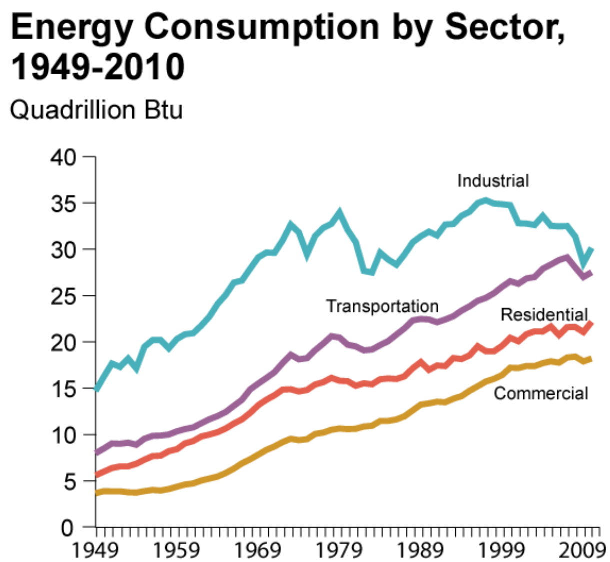
[3] In 2004, the EIA estimated the final energy consumption by building sectors: 18% commercial and 22% residential (in total 40% of final energy consumption).

consumption are hotels and restaurants, hospitals and schools.[4] According to the EIA data, HVAC is the main end use with a weight close to 50%, lighting follows with 15% and appliances with 10%. Building type is critical in how energy end uses are distributed and in their energy intensity: it is therefore essential to develop independent studies on energy consumption by building types.

Figure 1: Share of Energy Consumption by Major Economic Sectors (left) and Trends in Energy Consumption in 1949-2010 (right)



Source: U.S. Energy Information Administration, *Annual Energy Review 2009*, and *Monthly Energy Review (June 2011)*, preliminary 2010 data.



Source: U.S. Energy Information Administration, *Annual Energy Review 2009*, Table 2.1a, and Monthly Energy Review (June 2011), preliminary 2010 data.

In the Unites States, offices account for 17% of total non-residential area and about 18% of the energy use, equivalent to a 3.2% of the total consumption. Moreover, the amount of artificial lighting required, IT equipment use and air-conditioned area have steadily increased over time. Three key energy end uses, respectively HVAC, lighting and appliances, add up together to about 85% of the total energy bill for these buildings.

Early behavioral models for the estimation of energy use in residential buildings have been proposed since several decades ago, while the discussion on technological innovations and their impact on energy conservation started to be an important topic in this research field (Darley, 1978). Van Raaij and Verhallen (1983a, 1983b) developed a behavioral model of residential energy use, which included different end uses for which energy is consumed in a household. Several more comprehensive, and accurate, studies have followed, accounting for the various end uses for which energy is consumed in a household and the rationale behind the adoption of specific energy consumption patterns (Keirstead, 2006; Lopes et al., 2012).

[4] In 2003, the EIA estimated the energy use in the US commercial sector by building type: Retail 32%, Offices 18%, Hotels and restaurants 14%, Schools 13%, Hospitals 9%, Leisure 6%, and Others 9%.

Several studies have attempted to model energy use in residential buildings as a dependent variable of land use and building characteristics. The physically-based model BREHOMES (Building Research Establishment Housing Model for Energy Studies) has been used to model the energy use of the UK housing stock. Shorrock and Dunster (1997) provide a description of the model. The model was useful to develop scenarios for energy use and carbon dioxide emissions in future years. It provided information on a reference scenario, which represents what is likely to happen if current trends continue, and compared it with the possible outcomes from a number of energy efficiency measures (based on the technology available at the time the study was developed). This study also highlighted the importance of the time at which the outcomes of the proposed policies are achieved, as a central point in the evaluation of the future scenarios.

A related approach, which focuses on energy use at the level of an individual building is illustrated by the Hong Kong Building Environmental Assessment Method (HK-BEAM). Lee, Yik and Burnett (2007) describe the building energy performance assessment method implemented in the latest versions of the HK-BEAM model. This model is based on the energy budget approach, which has been formulated to rate energy performance of a wide range of new and existing buildings. This flexible approach can cater to a wide range of building types and allows energy performance trade-offs among various components in a building. The energy budget for an assessed building is the predicted annual energy use for a 'baseline' building. The baseline building model has the same shape and dimensions, and comprises the same mix of areas and types of premises as the assessed building (except for window-to-wall ratio adjustment to meet the relevant regulatory requirement). It also incorporates a range of standard (default) characteristics such that the model represents a building that has a level of energy performance that barely meets the relevant regulatory requirements or meets only basic design quality. The predicted annual energy use of the assessed building will be based on its specific design characteristics.

In a recent study, Miller (2011) applies parametric energy simulation modeling to assess the impact of different urban forms, and combination of building types, on the consumption of energy in buildings in Vancouver (Canada). The study identifies different energy patterns that can be explained by the specific combinations of urban forms and building types, isolating them from the effects of building construction standards and occupant behavior.

The analysis of energy consumption in different building categories is the object of the study from Pérez-Lombard et al. (2008): this study analyzes the available information concerning energy consumption in buildings, in particular as it relates to heating, ventilation, and air conditioning systems. The authors discuss energy consumption data for residential and non-residential buildings in different countries, comparing different energy end uses. They also discuss the difficulty in accessing information on energy use in non-domestic buildings, where the type of use and specific activities make a huge impact on the quality and quantity of energy services needed. Office buildings are identified together with retail as buildings responsible for very large energy consumption and CO2 emissions. Finally, the authors conclude that it is essential to make available comprehensive building energy information, in order to allow energy consumption analysis and forecast, and plan efficient energy policies. The assessment of energy balance in individual buildings is also the topic of the investigation from Hani and Koiv (2012), who analyze the thermal and electrical energy consumptions for different types of residential, educational and other public buildings. The study considers the impact of climate variables on energy consumption (as an effect of cooling vs. heating days) and separates the components of energy consumption by energy end use in the building.

Tso and Yau (2007) performed a study comparing three methods of predicting electricity consumption in buildings: regression analysis, decision tree, and neural networks. They found that decision tree and neural networks performed better in different seasons but the difference in error between the three methods were minimal, indicating that, as a predictive tool, linear regression is a valid method, and usually the easiest to develop.

Regression analysis has been the most popular modeling technique in predicting energy consumption. The least-squares method is generally used for estimation purposes in the multiple-regression model. Once regression coefficients are obtained, a prediction equation can then be used to predict the value of a continuous output (target) as a linear function of one or more independent inputs. The popularity of the regression models may be attributed to the interpretability of model parameters and ease of use. However, the major conceptual limitation of all regression techniques is that one can only ascertain relationship but cannot identify causal mechanisms among variables. Moreover, the estimation of linear regression models is based on the assumptions of normality and of independent distributions of the explanatory variables, which are often violated in many empirical studies.

For neural network models, feedforward network is the simplest and most popular type of network. Training a neural network is the process of setting the best weights on the inputs of each of the units and backpropagation (backprop) is the most common method for computing the error gradient for a feedforward network. Neural networks perform well in applications when the functional form is nonlinear. They are especially useful for prediction problems where mathematical formulae and prior knowledge on the relationship between inputs and outputs are unknown. A disadvantage in using neural network for a regression analysis is that it does not provide p-values for testing the significance of the parameter estimates. Moreover, a preliminary step of feature selection before learning is needed. Artificial neural networks with hidden layers are better as classifiers for problems involving nonlinear decision hyper-surfaces, but are much harder to interpret.

In decision tree modeling, an empirical tree represents a segmentation of the data that is created by applying a series of simple rules. These models generate set of rules that can be used for prediction through the repetitive process of splitting. The most common tree methods include chi-squared automatic interaction detection, classification and regression trees. A major advantage of the decision tree over other modeling techniques is that it produces a model that can represent interpretable rules or logic statements. The explanation capability that exists for trees producing axis parallel decision surfaces is an important feature. Besides, classification can be performed without complicated computations and the technique can be used for both continuous and categorical variables. Furthermore, decision tree model results provide clear information on the importance of significant factors for prediction or classification. However, decision tree induction generally does not perform as well as neural networks for nonlinear data, and it is susceptible to noisy data. The technique is more suitable for predicting categorical outcomes and, unless visible trends and sequential patterns are available, decision trees are less appropriate for application to time series data.

The use of alternative analytical methods has not been popular in the energy consumption literature. While the regression analysis method is supported by statistical theories as producing good estimates according to certain statistical properties (for instance, being the best linear unbiased estimator), other approaches such as decision tree and neural network are useful in developing predictive models in other fields. In the past decade, advancements in database management and improvements in computing speed have led to new ways of conducting data analysis. Data mining is now receiving attention and is being recognized as a newly emerging analysis tool. When searching for a predictive model, common practice in data mining is to develop various models using different approaches, then select a final model after comparing their accuracies according to some model selection criteria.

Kalogirou and Bojic (2000) adopt an innovative method for the estimation of energy consumption in buildings. The energy consumption of the building is studied depending on the thickness of the masonry, the building insulation and the season. Simulated data for a number of cases are used to train an artificial neural network (ANN) in order to generate a mapping between the easily measurable inputs and the desired output, i.e., the building electricity consumption (in kWh).

Short-term weather patterns are an additional factor that can cause substantial spikes in household electricity consumption. The lack of comprehensive information on household characteristics, however, can make the development of accurate models difficult. As a result, several groups (Mihalakakou, et al., 2002; Beccali, et al. 2008) have used artificial neural network (ANN) models for short-term prediction of electricity demand.

Mihalakakou et al. (2002) use a neural network model, for estimating the energy consumption time series of a residential building in Athens using several climatic parameters as inputs. These parameters are hourly values of the energy consumption, for heating and cooling purposes in buildings. The primary objective of the study is to examine the ability of neural network systems to estimate the hourly values of energy consumption for a residential building. The second objective is to examine the feasibility of the neural network system in predicting future values of energy consumption using as inputs "multi-lag" predicted values of ambient air temperature and total solar radiation time series. The authors found that the neural network approach is able to estimate building energy consumption rather successfully for both the warm and the cold period of the year.

Beccali et al. (2008) present a forecasting model for the short-time prediction of the household electricity consumption related to a suburban area of Palermo, Italy. An Elman artificial neural network (ANN) model predicts the household electric energy demand of the investigated area and evaluates the influence of Heating, Ventilation, Air Conditioning (HVAC) equipment on the overall consumption. The model estimates the electricity consumption for each hour of the day, starting from weather data and electricity demand related to the hour before the hour of the forecast. The model was designed to predict, one hour ahead, the intensity of the electric current supplied to a sub-urban area of the town of Palermo, characterized by the sole presence of household users.

Artificial neural network models have also been used to model urban heat islands, and their impact on energy demand (Kolokotroni et al., 2010; Gobakis et al., 2011). Kolokotroni et al. (2010) describe a method for predicting air temperatures within the Urban Heat Island at discrete locations, based on input data from one meteorological station in London. The paper describes a method that can be applied to other cities using historical air temperature data, in many cases available through air pollution networks or meteorological stations. The authors use London as a case-study to describe the method and its applications. The model is capable of predicting site specific hourly air temperature within the Urban Heat Island based on input data from one meteorological station for the time the prediction is required, and historic measured air temperatures within the Greater London Area. Gobakis et al. (2011) adopt a similar approach, using artificial neural networks and learning paradigms for predicting the intensity of urban heat islands in Athens. They present several variations on the neural networks architectures, and evaluate the feasibility of predicting urban heat island phenomena using a limited data series. The Athens case study was used to demonstrate the feasibility and accuracy of the overall approach. The methodology presented showed that the urban heat island intensity can be predicted quite accurately for at least a 24-h prediction horizon using a limited set of data.

Yu et al. (2010) develop a building energy demand predictive model based on the decision tree method. In the study, the method is applied to estimate residential building energy performance indexes by modeling building energy use intensity levels. The results demonstrate that the use of decision tree method can classify and predict building energy demand levels accurately, identify and rank significant factors of building energy use intensity automatically. One of the advantages of this method, as suggested by the authors, is associated with its ability to predict categorical variables and generate accurate predictive models with interpretable flowchart-like tree structures that enable users to quickly extract useful information on the studied phenomena.

In a recent energy consumption study, Howard, et al. (2012) build a model to estimate the building sector energy end-use intensity (in KWh/m2 of floor area) for space heating, domestic hot water, electricity for space cooling and electricity for non-space cooling applications in New York City. The model assumes that such end use is primarily dependent on building function and not on construction type or the age of the building. The modeled intensities are calibrated using 5-digit ZIP code level data reported by the New York City Mayor's Office of Long-Term Planning and Sustainability on the annual electricity and natural gas, steam, or fuel oil consumption for 191 ZIP codes. End-use ratios are derived from the Residential Energy Consumption Survey (RECS) and CBECS's Public Use Microdata. The results provide the ability to estimate the end-use energy consumption of each tax lot in New York City. Annual end-use energy consumption intensities are developed by performing a robust multiple linear regression to obtain electricity and total fuel intensities for eight different building functions:

residential 1–4 family, residential multi-family, office, store, education, health, warehouse and other commercial. The electricity and total fuel intensities are apportioned into base electric, space heating, water heating, and space cooling end uses by ratios derived from the RECS and CBECS end use estimation. The base electric end use includes energy consumed for appliances, lighting, ventilation, and refrigeration. The annual end-use intensities are applied to building floor area across New York City to determine the spatial distribution of energy consumption for the four primary end uses. Weather is found to have a large impact on energy consumption from year to year indicated by the high correlation between the consumption of fuel oil, natural gas, and to some extent steam, with heating degree days. The study uses information from the New York City Department of City Planning on the NYC building stock stored in a geo-rectified database, PLUTO. The study is able to project energy consumption for different building types using the total building floor area for each tax lot available in PLUTO for eight different building categories: commercial, residential, office, retail, garage, storage, factory, and other.

Recent applications of energy use models have integrated energy consumption models in integrated urban models for the simulation of land use, transportation and economic development. These models can simulate both the short-term and long-term decisions of firms and households that directly affect urban energy consumption. Chingcuanco and Miller (2012) integrated a model of energy use for residential space heating demand in the ILUTE model. The model combines a bottom-up approach to aggregate individual uses with a logit-type discrete choice model that simulates the heating fuel and equipment choice. The model is developed and estimated using household microdata for the City of Toronto, Canada. Energy consumption for heating purposes in the individual dwelling units is then computed with the HOT2000 software. The resulting residential space heating model component is added to the ILUTE model as the first step towards the creation of an integrated energy-land use model that can study energy consumption in cities.

To support the modeling of energy demand, and the viability of alternative and distributed generation systems, accurate load profiles are required for different types of buildings. Armstrong et al. (2009) review the efforts to synthetically generate electric and thermal load profiles in Quebec, for three targets single-family detached households – low, medium and high consumers – based on a limited amount of available information. Although the synthetic Canadian profiles proved useful to simulate a residential cogeneration system, and compared favorably to simulation results with measured data, there is still room for improving the realism of the synthetic profiles. The current generated profiles include only seasonal variations for lighting. According to the authors, the method of generating domestic load profiles could easily be applied for different target households, or even different countries.

The form of an urban area can affect energy use in other ways, such as through its impact on the types and amount of transportation that it generates. The empirical results regarding urban size, density, and distance to the city center can be related to the concept of the compact city. Høyer and Holden (2003) found that the extent of environmentally harmful household consumption varies substantially with the physical/structural conditions in housing areas. The authors' research supports the assumption that compact urban structures would lead to reductions in the overall ecological footprints of households. This is because shorter distances between houses and services results in less travel, and at the same time because dense and concentrated types of housing use less energy for heating and other technical equipment. Steemers (2003) has similar findings, and also establishes the relative magnitudes of building energy use in comparison to transport. An important finding from this study is that dense cities are generally low energy cities. The results shows that for residential buildings, the energy implications of compact densification are balanced between the benefits from reduced heat losses and the non-benefits of reduced solar and daylight availability. Specific results can vary, however, and the author found that for naturally ventilated office buildings, increasing urban density increased energy use because of the reduced availability of daylight. The results from both studies emphasize the importance of physical urban planning, and demonstrate how housing and land-use planning can be important tools in achieving sustainable levels of consumption.

Many governments and public agencies have introduced important changes in regulations for new buildings, in order to reduce the energy consumption and the GHG footprint of new developments. For European Union (EU) countries, this shift has been highly supported by EU policies to increase efficiency in buildings, as supported by the Directive 2002/91/CE on the energy performance of buildings from 2002. The Department for Communities and Local Government in the UK issued a report in 2006, in which it outlined a plan for achieving zero carbon buildings for new homes in UK within a decade. To achieve this target, they set out a package of measures that includes innovative standards and regulations for new technologies. This was followed by the report "Building Regulations: Energy efficiency requirements for new dwellings", from 2007, where likely changes to the building regulations were described. The report intended to provide an early indication of the changes that were likely needed to meet future targets for energy efficiency. These changes to the building regulations are part of a larger initiative to reduce energy consumption and GHG production throughout the UK (Department of Trade and Industry 2007). The intention of the UK government is to significantly reduce energy use in buildings as an important element in its climate change strategy, and in its approach to securing energy supplies in the future.

An evaluation of the technical feasibility of achieving major reductions in CO2 emissions by the year 2050 has been described by Johnston (2003) and Johnston et al. (2005). These papers describe the development of an energy use and CO2 emission model of the UK housing stock, which is capable of being used to explore a range of possible future scenarios. The model is used to explore the technological feasibility of achieving CO2 emission reductions within the housing stock under a number of different illustrative scenarios, with the objective of achieving emission reductions in excess of 80% within this sector by the middle of the century.[5] The model uses a bottom-up approach for forecasting energy and CO2 emissions, and tends to focus on the energy sector alone, using highly disaggregated, physically-based, engineering-type models to represent in detail the energy demand and supply sectors. Although the model has some weaknesses, such as a concentration on the residential sector – leaving aside the business and

[5] Reductions of this order are likely to be required across the industrialized countries in order to stabilize the atmospheric CO2 concentration and global climate.

industrial sectors – and a relatively simple energy supply model, it is has significant value as a policy tool. It represents a detailed and scientifically defensible attempt to project the delivered energy use and CO2 emissions attributable to the UK housing stock through the middle of this century. Johnston et al. (2005) concludes that despite increases in the total number of households, and increasing standards of thermal comfort, it is possible to achieve CO2 emission reductions in excess of 80% within the UK housing stock by 2050. However, achieving these sorts of reductions will require strategic shifts in both energy supply and demand side technology. Overall, initiating substantial changes in the energy performance of new and existing buildings is likely to require changes to energy use and greenhouse gas regulations. Bell (2004) reviews existing regulatory energy provision and CO2 performance in the UK and Europe, and discusses ways in which it could be modified, or new mechanisms developed to have a greater impact on the performance of existing buildings.

In the United States, several measures have been implemented to increase energy efficiency in buildings. A wide range of interventions at federal and state level have contributed to increase energy efficiency of appliances, on one side, and improve building standards, on the other, with both effects contributing to an increase in the efficiency of energy use in buildings. In California in particular, an important milestone in this field was the approval of the Regulations Establishing Energy Conservation Standards for New Residential and New Nonresidential Buildings ("Title 24") in 1978 (California Energy Commission, 1978). A series of more recent generation of building energy efficiency standards has followed, with the latest set of energy efficiency standards adopted in 2008. A new set of 2013 standards will continue to improve upon the current 2008 Standards for new construction of, and additions and alterations to, residential and nonresidential buildings, and should be effective starting on January 1, 2014. The building energy efficiency standards have produced significant effects in improving energy efficiency in buildings in the State of California. Additional support in the direction of increased energy efficiency and the creation of zero net energy solution came from the 2008 California Long-Term Energy Efficiency Strategic Plan. Moreover, significant contributions to the increase of energy efficiency in the State are associated with the efficiency programs currently promoted by the Investor Owned Utilities (IOUs) in the state, which have contributed to spread low energy consumption technologies and have incentivized California residents to adopt technological solutions that reduce energy use in their residences.[6]

The impact of these policies has contributed to smooth the demand for energy use, in particular during the daily and seasonal peaks, as reported in the updated forecasts provided by the utilities and contained in the California Revised California Energy Demand Forecast 2012 - 2022 (Kavalec et al., 2012) and the 2011 California Energy Commission Integrated Policy Report (California Energy Commission, 2011).

---

[6] Example of the efficiency programs promoted by the IOUs in California include education programs, energy audits, analysis of energy use, measures of infiltration and free energy improvement upgrades promoted by the IOU among their customers. These programs have been mainly addressed to reduce peak-energy use after the record energy consumption registered in the 2000s. A similar pattern of energy efficiency programs is also provided by the main Municipally Owned Utilities (MOUs) in California.

The actual effectiveness of policies designed to increase energy efficiency in reducing energy consumption in buildings is an important research topic that is increasingly studied in the literature. For example, Scott (2011) evaluates the energy consumption from the residential sector through the development of a structural equation model (SEM) designed to investigate the relations among physical, demographic and behavioral characteristics of dwellings and their occupants. Structural equation models are useful to deal with the limitations of multiple linear regression models, and they can be used to study causality issues among variables in a dataset. In the study, Scott discusses the causality issues that can be investigated with a SEM approach, and concludes that according to the UK data used in the research, homes with a propensity to consume more energy are also those that have higher energy efficiency standard rates.

In view of the changes that are introduced in newer buildings, as an effect of regulations for increased energy efficiency, Pérez-Lombard et al. (2009) provide a summary of the benchmarking, labeling and rating concepts used for building energy certification schemes, and that helps the comparison of energy efficiency plans and requirements in different context. Similarly, Marszal et al. (2011) discuss the need for a standardization of the definitions and calculation methodologies used in energy consumption analysis, as they apply to the Zero Energy Building (ZEB) concept, which can help to mitigate CO2 emissions and reduce energy use in the building sector. They conclude that the most important issues which should be given special attention before developing a new ZEB definition are:

- The metric of the balance
- The balancing period
- The type of energy use included in the balance
- The type of energy balance
- The accepted renewable energy supply options
- The connection to the energy infrastructure
- The requirements for the energy efficiency, the indoor climate, and in case of gird connected ZEB, for the building–grid interaction

Among the many other technological solutions that have been proposed to increase energy efficiency in buildings and reduce the resulting GHG emissions is the addition of a green roof (Castleton et al., 2010). The greatest benefits from this solution seem to be realized in older buildings with poor existing insulation (as current building regulations require high levels of insulation, green roofs are seen to hardly affect annual building energy consumption). In their review, Castleton, et al. discuss the current state of knowledge on the potential benefits that green roofs offer in relation to building energy consumption, and also discuss the issues involved in retrofitting older buildings.

Research on the effects of energy efficiency policies on the consumption of energy in commercial buildings has analyzed several possible scenarios for the reduction of energy consumption in this category of buildings. For instance, in 2007, researchers at the National

Renewable Energy Laboratory (NREL) conducted a study to assess the technical potential for achieving net zero-energy commercial buildings (Griffith, et al. 2007). The simulation of building energy use needs to cover all interactions among systems, components, occupants' activities and weather. NREL used EnergyPlus as the modeling tool to assess alternative scenarios as it accounts for the complicated interactions among climate, internal gains, building form, HVAC systems and renewable energy systems. The analysis framework for the study used detailed energy performance simulations for a large number of individual building models, which were intended to represent the entire commercial sector at the national level. The simulation used distributed computing to assess what would happen if an aggressive set of Zero-Energy Building (ZEB) technologies and practices were applied to the buildings, under several different scenarios. This study focused on energy use and energy consumption costs, but did not consider life-cycle environmental and economic performance for the entire building.

Many sources of data provide information on energy consumption in buildings. In addition to datasets collected for specific research projects, and for energy efficiency programs conducted by utility companies, energy authorities usually monitor energy consumption in both residential and commercial buildings through periodic surveys. The U.S. Energy Information Administration has collected many useful datasets, such as the Residential Energy Consumption Survey (RECS), Commercial Buildings Energy Consumption Survey (CBECS) and several reports on energy consumption in the industrial sector. These studies usually estimate energy consumption per square foot of buildings and are useful to inform environmental studies in a specific geographic area.

In the State of California, the California Energy Commission (CEC) administers the Residential Appliance Saturation Study (RASS) and the Commercial End-Use Survey (CEUS) to periodically collect information on energy use, respectively, in residential units and commercial buildings. The CEC also has databases of electricity and natural gas consumption for the residential and nonresidential sectors by county and for six sectors by utility area. The California Energy Commission sources apply to residential, commercial and industrial buildings and take into account climate zones and other locally and regionally unique modifiers. Some of these sources include building characteristics that are of interest in land use modeling studies, such as the California PECAS model developed at UC Davis. Other sources include the International Council for Local Environmental Initiatives (ICLEI), the World Resources Institute (WRI), the World Business Council for Sustainable Development (WCSB), the Greenhouse Gas Regional Inventory Protocol (GRIP) and the California Climate Action Registry (CCAR).

In this project, the available sources of data and the relevant experiences available in the literature were reviewed to inform the energy consumption study for Los Angeles County on the standards, processes and calculations to use in assessing the energy use of the surveyed floorspace types and to update the PECAS land use modeling framework with these energy components. The following sections of this report describe the process that was used to develop the analysis on energy consumption in buildings, and discuss the outcomes from the analysis.

CHAPTER 3: Building and Floorspace Inventory for Los Angeles County

An inventory of buildings in Los Angeles County was prepared using information from the Los Angeles County Assessor's data. The purpose of this study is to define an inventory of energy consumption and GHG emissions from buildings that is suitable for use in modeling approaches that study urban metabolism and that will allow the analysis of future scenarios of development and the effects of the adoption of energy efficiency policies. To do this, it is necessary to develop a methodology to account for the total amount of developed floorspace in the region. This floorspace inventory is then useful to compute the resulting energy consumption and GHG emissions from buildings. To accomplish this task, the information on the building stock was analyzed and classified using categories that are compatible with land use modeling approaches in the State of California, so that they can be easily integrated in modeling applications to compute energy consumption and the resulting environmental impact associated with the building stock.

The California Production, Exchange, Consumption, Allocation System (PECAS) modeling system adopts the developed floorspace categories reported in Table 1 to classify the available building types in the State of California.

Table 1: Non-agricultural, developed floorspace types in the California PECAS model

Floorspace types	
1	Light industrial space
2	Heavy industrial space
3	Warehouse space
4	Highway retail space
5	Downtown retail space
6	Mall and big box retail space
7	Neighborhood retail space
8	Low density office space
9	High density office space
10	Developed amusement parks space
11	Hospital space
12	Secondary education space
13	Primary K-12 education space
14	Religious space
15	Government operations space
16	Military space
17	Fishing dock space
18	Depot space
19	Rural luxury residential

20	Rural economy residential
21	Acreage luxury residential
22	Acreage economy residential
23	Single family detached luxury residential
24	Single family detached economy residential
25	Joined luxury residential
26	Joined economy residential
27	Low-rise luxury residential
28	Low-rise economy residential
29	High-rise luxury residential
30	High-rise economy residential
31	Urban “mobile home” residential
32	GQ (Group Quarters) residential

Due to the difficulties in tracking energy consumption (and the associated GHG emission) patterns for several floorspace types reported in Table 1, and the similar difficulties in crosswalking the building information from the Assessor’s data to the PECAS categories, the researchers further aggregated these floorspace types into a shorter list of floorspace categories for energy consumption purposes. The floorspace types that were considered for the inventory of the building stock in Los Angeles County are reported in Table 2.

Table 2: Building/Floorspace types used in the definition of the Los Angeles County building inventory

Building categories	*PECAS floorspace types*
Apartment residential	Low-rise luxury residential Low-rise economy residential High-rise luxury residential High-rise economy residential
Developed amusement park space	Developed amusement parks space
General commercial	Neighborhood retail space Downtown retail space Highway retail space
Government operations space	Government operations space
GQ residential	GQ residential
Heavy industrial space	Heavy industrial space
High density office space	High density office space
Hospital space	Hospital space
Joined residential	Joined econ residential Joined lux residential
Light industrial space	Light industrial space
Low density office space	Low density office space
Mall and big box retail space	Mall and big box retail space

Mixed use space	*(combination of various floorspace types)*
Parking	*(not included in PECAS)*
Primary k-12 education space	Primary k-12 education space
Religious space	Religious space
Secondary education space	Secondary education space
Single family detached residential	Single family detached economy residential Single family detached luxury residential Acreage economy residential Acreage luxury residential Rural econ residential Rural luxury residential
Single family detached with pool	*(included in previous SF categories above)*
Urban mobile home residential	Urban mobile home residential
Warehouse & distribution space	Warehouse space Depot space

Assessor's data

The Los Angeles County Assessor Parcel Dataset provides a list of parcels located in the County of Los Angeles for property tax purpose. The dataset contains information on the characteristics of each building (up to five buildings on each parcel), which are of valuable interest for the definition of a building inventory for Los Angeles County, as well as for the analysis of relationships of the building characteristics with the energy use.

The Los Angeles County Assessor Parcel Dataset was developed by the Office of the Assessor in the Los Angeles County. Assessor's parcel data were provided by Los Angeles County in a text file format covering the entire parcels (for 2008 property tax purpose) in Los Angeles County. The County Assessor Parcel Dataset is a master dataset for this study, as it provides information on the building stock in Los Angeles County and provides many explanatory variables used in the statistical analysis for the estimation of energy consumption models. Each record of this parcel dataset represents a single parcel with geographic information attached, including street location, size, and building attribution (up to five buildings on the same parcel).

The dataset was provided to the University of California, Davis in the condition of 'as is' from the Los Angeles County Office of Assessor. The spatial resolution is at the parcel level. However, records with typos and missing values are present in the dataset. The researchers corrected and interpolated them as much as possible for the purposes of the analyses for this project. Records with missing location information were removed from the analyses since these data cannot be used to retrieve ZIP+4 information from the U.S. Postal Services website, and therefore cannot be matched to any energy consumption data available for the project.[7]

[7] The removed records consist of approximately 5% of the entire parcel dataset, which is considered reasonable given the approximately 2.3 million records in the parcel inventory.

Data processing

The original Los Angeles County Assessor's Roll dataset came as a fixed length text file nearing 2.66 GB in size. To facilitate its import into SQL Server this file needed to be transformed into a comma separated value (CSV) file. This process began by discovering from the documentation what fields were included in the file and their column positions. Then a python script was developed to read the fixed length file line by line and split the line into elements for each field. Address parts were combined together into one field for simplicity. After running this script and trying to import the fixed length text file into SQL Server it was found that there were errors in the Assessor's file. Some records (lines) did not contain all fields: this made lines have varying number of columns and therefore impossible to import correctly. Also some invalid characters were incorporated into the file. This caused SQL Server to throw an error during the process of data import.

A second python script was developed to determine how many columns were in each line, as well as an additional script that identified the invalid characters. These errors were then manually corrected in the original file through a text editor. After these fixes in the original fixed length file, it was possible to parse the data using the transformtocsv.py script into a ".csv" format file. This file was then imported into a table of the SQL Server database, and the field specifications were changed to match the attribute types in the data.

Once the data table had been successfully imported into the SQL Server database, it had to be transformed into a more useful format: this was done using a series of SQL statements. This allowed the use of the information contained in the Assessor's dataset in combination with the other datasets described in the following sections of this report, and allowed the researchers to easily merge the information from the Assessor's data into the energy database developed at the Energy Analysis Zone level of spatial aggregation, as described in the following sections of this report.

The Assessor's database was developed and stored in SQL Server to ensure the relationship among variables and records. Information from the Assessor's data was also spatially joined to a GIS parcel shapefile of Los Angeles County. The Assessor Identification Number (AIN) recorded in both the roll parcel text file and the GIS parcel shapefile was used as spatial reference, allowing easy relationships among variables. The projected coordinate system is Albers Conical Equal Area whose European Petroleum Survey Group (EPSG) code is 9822 and the geographic coordinate system is North American Datum 1983 whose EPSG code is 6269.

Building Inventory

The information from the Assessor's data was processed to obtain an inventory of the building stock in Los Angeles County. The Los Angeles County Assessor Dataset consists of 126 variables, including street location, size, value, and building attributions. Their descriptions are elaborated in the Record Layout and Field Definitions document (pages 9 -31) prepared by the Office of Assessor of Los Angeles County. The property use code and building type fields respectively identify the land use of each parcel and classify each building contained in the parcel.

There are 1,121 different use codes in the Los Angeles County Assessor's data. The researchers crosswalked the use code field from the parcel database to the 21 floorspace types listed in Table 2 according to a functional crosswalk list of floorspace types that was developed as part of this research (and that is available on request from the researchers). Table 3 summarizes the number of parcels, the number of buildings and the total amount of square feet by each floorspace type in Los Angeles County.

Table 3: Distribution of floorspace types in the Los Angeles County Assessor's data

Floorspace type	*No. of Parcels*	*No. of Buildings*	*Sum of Sq. Ft.*
Apartment residential	67,955	95,339	827,435,126
Developed amusement park space	64	81	5,266,180
General commercial	37,861	45,311	448,102,968
Government operations space	25,377	1,385	15,040,709
GQ residential	1,570	2,136	29,741,686
Heavy industrial space	1,416	2,998	117,307,348
High density office space	1,231	1,735	219,590,994
Hospital space	585	1,098	47,672,473
Joined residential	452,684	558,581	766,031,088
Light industrial space	33,762	45,422	528,756,911
Low density office space	17,138	19,778	206,922,884
Mall and big box retail space	1,266	1,793	127,114,088
Mixed use space	12,989	20,193	85,038,073
Parking	15,783	15,060	221,881,094
Primary k-12 education space	2,467	2,968	28,213,383
Religious space	5,410	8,417	71,662,943
Secondary education space	245	521	19,005,819
SFD residential	1,248,123	1,242,569	2,002,020,745
SFD residential with pool	246,059	245,831	602,266,635
Urban MH residential	2,403	2,995	8,071,276
Warehouse & distribution space	13,374	18,490	566,879,657
Total Developed Floorspace	**2,187,762**	**2,332,701**	**6,944,022,080**
Vacant	133,009	2,626	35,927,690
*(Null)**	*1,609*	*561*	*3,099,793*
Total (Non Agricultural)	**2,322,380**	**2,335,888**	**6,983,049,563**
Agricultural and Park Space	*53,979*	*1,755*	*15,136,251*
Total (including Agricultural)	**2,376,359**	**2,337,643**	**6,998,185,814**

*The "*null*" field refers to parcels with invalid or missing code that could not be matched to any floorspace type

The Assessor's database contains information on up to five buildings for each parcel. If a parcel contains more than five buildings, the information is truncated, and the information reported by the Office of the Assessor of the Los Angeles County for the additional buildings is omitted. This constitutes a potential loss of information in the building inventory. To try to compensate for this error, the researchers analyzed the distribution of the number of buildings in each parcel, by floorspace type, to establish a method to estimate the number of parcels with six or more buildings in the dataset. The proportion of parcels with six or more buildings was then calculated from the number of parcels with information for five buildings in the Assessor's data.[8]

Another additional issue is associated with the purpose for which the Assessor's data are created and maintained. Assessor's data are mainly created and used for collecting information on property taxes. Accordingly, they do usually contain accurate information on private residential and non-residential buildings that are subject to property taxes. They often do not contain very updated information for public buildings and other "non-property tax" buildings though.

Table 4: Distribution of floorspace types and adjustments introduced in the building inventory for Los Angeles County

Floorspace type	*No. of Parcels with missing Sq. Ft.*	*Percentage of total Parcels*	*Adjusted Sum of Sq. Ft.*
Apartment residential	617	0.9%	835,607,112
Developed amusement park space	19	29.7%	7,723,730
General commercial	1,077	2.8%	462,507,541
Government operations space	24,381	96.1%	35,523,283*
GQ residential	70	4.5%	31,387,392
Heavy industrial space	148	10.5%	141,715,444
High density office space	53	4.3%	229,852,255
Hospital space	29	5.0%	55,574,937
Joined residential	3,700	0.8%	772,371,343
Light industrial space	1,020	3.0%	547,576,935
Low density office space	193	1.1%	209,378,398
Mall and big box retail space	137	10.8%	144,934,874
Mixed use space	378	2.9%	87,648,920
Parking	1,284	8.1%	241,595,694
Primary k-12 education space	732	29.7%	183,716,885**
Religious space	100	1.8%	73,282,749

[8] Given the way the assessor's data is structured, a parcel with a high number (larger than 5) of buildings is reported with information for five buildings. Therefore, parcels with six of more buildings have to be searched among the subset of parcels with information for five buildings in the dataset. This significantly restricts the amount of parcels that might include a large number of buildings, and therefore restricts the possible interval of values, and the error associated with it, for the number of parcels with six or more buildings.

Secondary education space	54	22.0%	39,241,821*
SFD residential	6,556	0.5%	2,012,644,137
SFD residential with pool	310	0.1%	603,038,633
Urban MH residential	273	11.4%	30,157,566
Warehouse & distribution space	326	2.4%	584,560,624
Total developed floorspace in Los Angeles County			**7,330,040,272**

Notes: *Adjusted through computation of information for Federal and State buildings from government sources; **Estimated using the California PECAS floorspace synthesizer modeling framework

After an examination of the parcel record dataset for Los Angeles County, we determined that the information reported for public buildings (administrative, educational, religious and other non-property tax buildings) was not very accurate (and the building stock and the amount of floorspace for these types of buildings were probably largely underestimated). Unfortunately, there are not a lot of sources of information that could allow access to more reliable sources of data to complement building information for these categories. For this reason, the researchers tried to analyze the Assessor's data, in order to find information that could help assess the order of size of the error associated with the estimation of the amount of floorspace by each category found in the data, and try to correct it.

Table 4 reports the distribution of the number of parcels with missing information for the size (in square feet) of the buildings contained in each parcel. As expected, the number of parcels with missing information for the size of the buildings is particularly high for non-property tax buildings. In particular, they reach very high percentages of the total number of parcels in dataset for primary education space (K-12 schools), secondary education (colleges and higher education institutions), and for the government buildings. Information related to the amount of floorspace for government buildings is missing for more than 96% of the parcels belonging to this category.

In order to correct for the described issues, the researchers developed a set of adjustment factors that were applied to scale the amount of floorspace (in square feet) by category. These factors included both a term that compensated for the eventual presence of more than six buildings in a parcel[9] and a term that compensated for the proportion of parcels of each floorspace type that do not contain information on the building size. The latter term of the factors was further corrected in order to attenuate large corrections: for instance, in the case of government buildings, the researchers assumed that probably only smaller buildings were usually left out of the Assessor's data, and therefore corrected the scaling factor to 60% of the original factor that was proposed in earlier versions of the building inventory.

[9] This attempt to correct the amount of floorspace for the eventual presence of more than five buildings in a parcel was developed in a rather conservative way, in order not to over-inflate the number of square feet in the building inventory. This means that, probably, the issue was only partially corrected in the dataset. However, given the lack of more detailed and reliable information in this field, it is difficult to assess the exact order of magnitude of this error.

The researchers also compared the information on the amount of developed floorspace used for public buildings to other available data sources. In particular, they requested information on the Public Building Inventory from the Professional Services Branch of the Department of General Services of the State of California. The DGS data provided information on 1,889 public buildings that are located in Los Angeles County, including 627 buildings that are considered university facilities in the various campuses of the University of California and the California State University. The information for these buildings, which accounted for more than 30 million sq.ft. in total size in Los Angeles County, was used to update the information on secondary education from the Assessor's data. The remaining 1,261 public building records provided information for buildings prevalently owned by the State of California and a few County buildings, for a total of 12,260,346 sq. ft. in Los Angeles County.[10] Additional information were obtained from the General Service Administration (GSA) of the U.S. Federal Government. GSA data included information for 136 Federally-owned buildings located in Los Angeles County, for a total floorspace of 6.8 million sq. ft. The DGS and GSA public building records were used to update the estimates for the floorspace inventory for public buildings. Additional adjustments were included to account for City (not included in the DGS records), County (only partially included in DGS records) and other public buildings located in the area of study.

Moreover, the researchers compared the results from the building inventory with the amount of floorspace for each building category predicted by the floorspace synthesizer built as part of the California PECAS model. The comparison identified some categories (in particular, single family detached homes, primary and secondary education, GQ residential and apartment residential buildings) for which the building inventory obtained from the Assessor's data largely differ from the results of the floorspace synthesizer. In particular, for the "primary K-12 education space", data from the Assessor's dataset appeared particularly low from a comparison to the other source. For this floorspace category, the researchers adopted the estimate for the sum of square feet that was estimated and validated by data from the California Department of Education as part of the development of the PECAS floorspace synthesizer. For all other floorspace types, as the official Assessor's data[11] were, overall, considered a more reliable source of information on the building stock in Los Angeles County than any modeling approach, the researchers decided to rely on the data obtained from the Assessor's dataset.

The researchers also compared the numbers reported in the adjusted building inventory from Table 4 with other sources to verify the reliability and consistency of the data. In particular, the total building inventory (in square feet) for industrial areas (sum of heavy and light industrial areas) was found to be rather consistent with the estimates for industrial floorspace developed by real estate operators (http://www.grubb-ellis.com/Forecast2012/PDFs/Los-Angeles_IND_2012_1Q.pdf, last accessed on December 4, 2012;

---

[10] Unfortunately, also in the DGS data on public buildings, information on the building size was missing for 274 of the 1261 records (21.7% of the total).

[11] At least in theory, Assessor's data contain real information, and not modeled data, on existing buildings.

http://www.cushwake.com/cwmbs3q12/us_3q12.html, last accessed on December 4, 2012; http://www.colliers.com, last accessed on December 4, 2012).

Estimates for the retail and commercial space, as well as for the low-rise and high-rise office space from the building inventory built using the Assessor's data appeared to be higher than estimates reported by real estate operators (http://www.grubb-ellis.com/Forecast2012/PDFs/Los-Angeles_OFF_2011_4Q.pdf, last accessed on December 4, 2012; http://www.cushwake.com/cwmbs3q12/us_3q12.html, last accessed on December 4, 2012; http://www.colliers.com, last accessed on December 4, 2012). However, given the official source of information that was used in the case of the Los Angeles County Assessor's data, the researchers did not reduce the amount of floorspace for these building categories.

The final numbers of the adjusted sum of square feet of developed space in Los Angeles County, by each category of building type, are reported in the last column to the right of Table 4. It is implied that the proposed inventory of the building stock for Los Angeles County is not expected to be perfectly exact, given the difficulties in acquiring accurate information on the amount of floorspace by each building category. Moreover, data from the Assessor's dataset are referred to 2008, and do not include later changes in Los Angeles County real estate development.

Still, the inventory provides a reliable enough basis, computed from the observed data from the Assessor, which can be used in the assessment of a baseline energy study for Los Angeles County. As previously discussed, the reliability of the results for a specific building type category varies, in particular between buildings subject to property tax and non-property tax buildings, with the former categories of buildings better reported in the Assessor's data than the latter ones. Even with the limitations here discussed, this building inventory and the resulting energy baseline study is useful to inform researchers on the trends in energy use in buildings. Further, they can be used as the basis in modeling projects for the development of forecasts for the estimation of future energy consumption from the building stock, for instance under specific assumptions on policies that will increase energy efficiency in some specific sectors and for specific categories of buildings.

CHAPTER 4: Spatial Aggregation: Energy Analysis Zones

One of the main challenges in the development of this project relates to the need for the definition of a common level of spatial aggregation for the development of the analysis of energy consumption in buildings. This task became necessary for the need to treat information available from different sources and aggregated at different geographic scales. In particular, building information was obtained from the parcel dataset of the Assessor's data for Los Angeles County. This dataset contains a rich source of information on building location, size, age and technology, all information available at the parcel level data.

Unfortunately, energy consumption data for electricity and natural gas in the area of study were not available at the billing address level. Utility consumption data were provided at the Zip+4 level of spatial aggregation. This level of aggregation is not particularly user-friendly and not easily treatable in terms of geographical information. Zip+4 areas are not uniquely defined geographic areas, but rather a level of functional aggregation of addresses defined by the United States Postal Service (USPS) to facilitate mail delivery. As such, it allows rather straightforward aggregation of data on the side of utility companies: they can easily sum up the energy consumption data from the individual billing addresses in a region based on the Zip+4 field they have on records.[12]

The treatment of information aggregated at the Zip+4 level is not easy when integrated with spatial data from other sources, which are usually provided at standardized levels of spatial aggregation (e.g. parcels, blocks, block groups, census tracts, etc.). Currently, no GIS layer of Zip+4 areas exist for the United States. Moreover, Zip+4 areas may considerably vary in size, with smaller areas located in more densely built areas, and even spatially overlapping Zip+4 areas often found in large buildings in the Central Business Districts or in densely populated areas of a city.[13]

The spatial aggregation problem for the analysis of energy consumption in buildings was solved for this project through the definition of Energy Analysis Zones (EAZs). EAZs are defined in a way that allows the treatment of all data contained in the energy use dataset that was built for this study, regardless of the original scale in which each variable was measured.

---

[12] Each valid U.S. address where mail is delivered is usually associated with zip+4 information. The aggregation of billing address data at the Zip+4 level is therefore a rather straightforward process in a database that includes complete street addresses of the customers. However, the aggregation of other types of spatial information, as those measured at the parcel level, to Zip+4 areas is not similarly straightforward, and poses serious spatial and computational difficulties that are discussed in this section of the report.

[13] Example of spatially overlapping Zip+4s can be found in large residential or commercial buildings, where they refer to aggregations of suites located on different floors (in a large commercial building), or aggregation of apartments (or condos) in a large apartment complex or residential building.

The following subsections of this chapter describe the process that was developed to generate the Energy Analysis Zones.

ZIP+4

Energy use data for this project was supplied by utility companies at the ZIP+4 level, for the LADWP and LBGO service areas. To compare the energy data provided by the utility companies with parcel-level information, a correspondence had to be developed between the ZIP+4 codes used for the energy data and the addresses from the Los Angeles County Assessor's parcel dataset.

The Assessor's parcel data could be joined to a spatial dataset of parcel polygons. This provided a valid geography and a well-defined level of spatial aggregation for all data available at the parcel level for the research. Unfortunately, a spatial approach to developing the ZIP+4-to-parcel relationships was hindered by the lack of a corresponding ZIP+4 spatial dataset.

ZIP+4 codes are defined by the U.S. Postal Service (USPS) using roads and address ranges, and the USPS does not attempt to produce a ZIP+4 spatial dataset of the aggregated parcel polygons which would make up the areal extent of each ZIP+4. Some point approximations are available from private sources, but researchers could not access any source that could provide a complete spatial dataset sufficient for developing ZIP+4-to-parcel relationships using a spatial approach.

Therefore we developed a non-spatial approach for merging the different levels of aggregation and matching address records from the Assessor's parcel database to the ZIP+4 codes provided by the utility companies. This task was carried out using the address information available from the Los Angeles County Assessor's dataset. The *situs address* for each parcel record was used to query the U.S. Postal Service's ZIP Code Lookup webpage through a script developed as part of the research project. For each record, the available address information was submitted, and a standardized version of the address was returned. This standardized address data was then gleaned to extract any valid ZIP+4 codes associated with the address. This process was implemented in Python scripts, to automate the extraction and processing of each address in the dataset. The raw outputs were stored in text files, which were later transferred into spreadsheets for review. Finally, they were loaded into SQL Server where the ZIP+4 codes were attached to the rest of the parcel data. Many records[14] in the Assessor's parcel dataset lacked valid addresses, and as a result, could not be directly cross-walked to any Zip+4 record nor linked to the energy use records.

By matching addresses from the Assessor's dataset to ZIP+4 codes, it was possible to develop a table of the required ZIP+4-to-parcel relationships. This allowed energy data summarized by ZIP+4 to be attached to parcel information, and then analyzed together. Also, as a result of this effort, it is possible to extract a spatial representation of ZIP+4 zones within Los Angeles

---

[14] 214,669 parcels (approximately 9% of the total number of parcels) in the dataset did not have a valid *situs* address in the assessor's dataset, and therefore they could not be queried in the process of the Zip+4 information extraction.

County, which was built through the aggregation of parcel polygons associated with each Zip+4.

Figure 2 shows an example of the correspondence between parcels and Zip+4s in downtown Los Angeles: each parcel might be associated to one or more Zip+4s.[15] Similarly, a Zip+4 may contain multiple parcels, with some of these also associated with other Zip+4s.

Figure 2: Parcels and Zip+4s in downtown Los Angeles

Processing of Parcel Roll Records

There are 2,376,361 roll records in the Assessor's Roll database. The vast majority of these records include complete information for the following variables:

- parcel number (AIN)
- street number

[15] Each parcel might contain one or more buildings, and each large building might contain multiple residential or commercial units, which are not necessarily all associated with the same Zip+4.

- street name
- prefix/suffix
- apartment/unit number
- city name
- 5-digit zip code

Through querying the USPS website with a Python script, researchers were able to match most of the parcels in the dataset to a valid Zip+4 code. However, there are records with address data in a valid street address format that did not return USPS ZIP+4 information.

When processing the information from the Assessor's dataset, some data cleansing and transformation were necessary, e.g. the removal of records with missing address information. Of the original 2,376,361 roll records in the Assessor's database, 2,118,065 parcels (about 89.1% of the total) were matched with a valid ZIP+4 using the USPS ZIP Code Lookup webpage. These parcels were linked to a total of 649,457 unique ZIP+4 codes.

At this point, the tabular data with ZIP+4 attributes were joined to the Los Angeles County Assessor's Parcel GIS dataset, which contains parcel geometries. Within the GIS parcel dataset there are 2,382,897 records; however, some AINs are duplicated (e.g. a building annex is geographically separated, but it still shares the same AIN with its main building). A GIS dissolve process was performed on the AIN column to get 2,382,017 unique AINs from the parcel GIS dataset. After this process, there was still a discrepancy in the number of AIN between the GIS-based parcel dataset and the table-based roll dataset. This was due to the presence of AINs that are contained in one file but not in the other. This resulted in a reduced number of coincident AINs when the two sources of information are joined.

There are four possible types of spatial relationships between the AIN codes and the ZIP+4s, which can be summarized as follows:

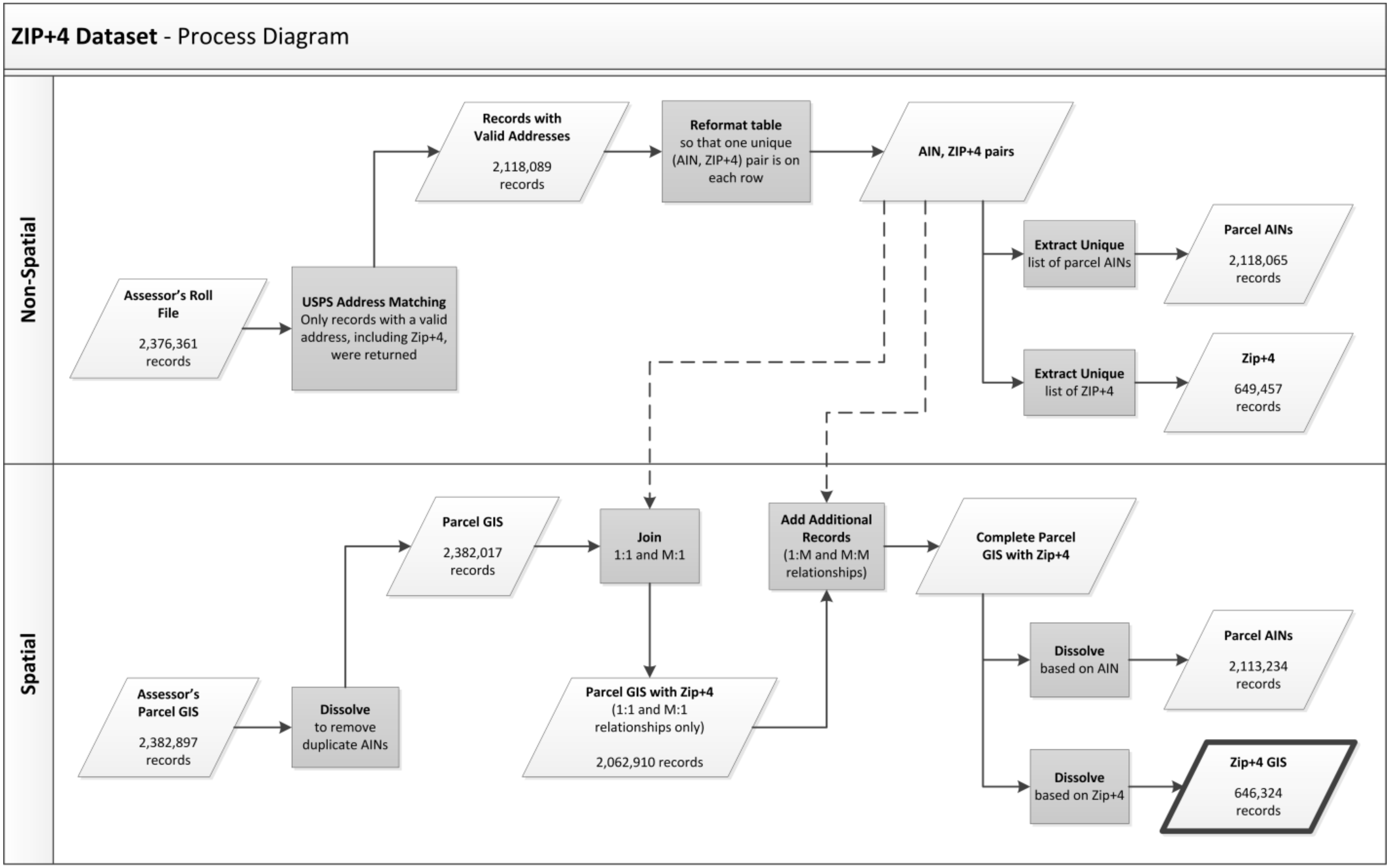
- 1 AIN to 1 ZIP+4
- Many AINs to 1 ZIP+4
- 1 AIN to Many ZIP+4s
- Many AINs to Many ZIP+4s

The majority of USPS queried outputs belong to the first two categories: either one parcel (AIN) is uniquely associated with one ZIP+4, or many parcels (AINs) are linked to the same ZIP+4. There were 2,067,298 queries in the merged dataset (parcel; Zip+4) that fitted in either one of these two categories. Of these, the number of cases that have a spatial geometry in the GIS layer is 2,062,910. These records can be easily joined to the parcel GIS dataset[16]: the geometry of the parcels is dissolved to become a great portion of the ZIP+4 GIS layer.

---

[16] GIS analyses for this project were developed in the ESRI ArcMap environment

Figure 3: Process Diagram for Zip+4 and Parcel Data



AIN = Assessor's Identification Number

For the two remaining relationships between AIN and ZIP+4, a Python script was developed by the researchers at the University of California, Davis to spatially represent the remaining cases of "1 AIN to Many ZIP+4s" and "Many AINs to Many ZIP+4s". In the case of one parcel (AIN) linked to many ZIP+4s, the script automatically created duplicate GIS polygons with identical shapes on top of the existing shape. It then assigned each of the unique ZIP+4 codes to one of the resultant polygons. For example, if there are 10 ZIP+4 associated with a certain AIN (in the case of large apartment complex), the script generated 9 extra polygons on top of the pre-existing polygon (thus generating 10 polygons with the same shape) and then assigned each of the unique ZIP+4 codes associated with this AIN to one of those 10 polygons.

The treatment of the "Many AINs to Many ZIP+4" was more difficult: the automatic script identified for each parcel (AIN) all ZIP+4 values with which this parcel was associated. It then queried the dataset to search for additional AINs that were associated with any of these ZIP+4s initially linked to the first AIN. The algorithm keeps searching for AINs and ZIP+4 iteratively, until it identifies all possible combinations (parcel; ZIP+4) that are linked to each other.

The result of this process is a GIS layer with both AIN and ZIP+4 codes. This GIS dataset contains the 2,113,234 AINs that can be geographically represented. Since nearly 5,000 AINs (.002%) are lost in the spatial join process, slightly fewer ZIP+4 than those contained in the original dataset are contained in this adjusted dataset. The final number of Zip+4s in this dataset is 646,324. By doing a GIS dissolve process on the ZIP+4 code, a ZIP+4 GIS dataset was created, representing all 646,324 unique ZIP+4 codes. Figure 3 summarizes the process that was used for the generation of the Zip+4 dataset.

The following possible sources of errors are associated with the process of generation of the Zip+4 GIS dataset:

- Geographically, the ZIP+4 zones are based upon the Los Angeles County parcel dataset geometry. Therefore, they inherit this dataset's spatial precision and accuracy. Los Angeles County parcels are digitized to a high quality, and the Zip+4 dataset shares the same high quality.

- As mentioned before, the Assessor's dataset was created and is maintained for tax purpose only. It is not an exhaustive list of buildings in Los Angeles County. Therefore, due to the process that was used for the creation of the Zip+4 dataset, we do not have a complete record of Zip+4 codes for all properties in LA County. The quality of the dataset is significantly higher in the areas predominantly occupied by buildings subject to property tax.

- 214,669 parcels could not be attributed with a ZIP+4 code because they do not have a valid situs address. In addition, 43,627 address records either did not return a ZIP+4 from the U.S.P.S. website or returned only a five digit ZIP code. These records were not used in the resulting dataset.

- The addresses used to build this dataset were obtained from the Los Angeles County Assessor's dataset from 2008. The ZIP+4 attributes gathered from the U.S. Postal Service website are 2012 data collected at the time the project was developed. This mismatch in the years the data are referred to may generate some inconsistencies in the data, and it might be responsible for part of the missing ZIP+4 results from the USPS website queries.

Energy Analysis Zones

A new level of spatial aggregation was created so that energy consumption could be analyzed with data aggregated from several different sources. A complex spatial relationship between ZIP codes and parcels exists and a common denominator needed to be found so that energy use could be analyzed spatially. This process led to the definition of the Energy Analysis Zones (EAZs).

EAZs are defined from the overlap of ZIP+4s and parcels. They were generated through a series of database queries, which selected all parcels that shared common ZIP+4 designations, and all ZIP+4 codes that share common AIN (parcel) designations, in an iterative process.

Figure 4: Creation of EAZ from the Zip+4 and Parcel Data

AIN	EAZ	ZIP+4
5544011033	**1**	900279
4332026022	**2**	900358
4332026022	**3**	900359
6032012015	**4**	900446
5123004429	**5**	900891
5123004433	**5**	900891
4330004036	**6**	902129
4330004037	**6**	902129
4330004038	**6**	902129
4330004038	**6**	902130
4330004040	**6**	902130
4330004041	**6**	902130
4330004042	**6**	902130
4330004042	**6**	902131
4330004044	**6**	902131
4051003005	**7**	902501

4051003005	**7**	902502
4051006002	**8**	902503
4051006002	**8**	902504

First a unique list of ZIP+4 codes was created and looped through. Parcels matching a ZIP+4 were selected and then the set of AINs (unique key for each parcel) from the parcels were re-queried for the ZIP+4s that belonged to them. Then again the list of ZIP+4s was queried for AINs. This iterative process continued until the number of records returned from the AIN list was equal to the number of records returned from the ZIP+4 list. The final set was then given a unique EAZ number that was written to the parcel dataset, for the parcels in the set of AINs. Figure 4 shows an example of the relationship AIN - ZIP+4 that led to the creation of the EAZ system.

After generating the complete set of ZIP+4 to AIN relationships in the database, the newly created EAZ numbers were joined to the spatial parcel dataset. The parcel polygons were then dissolved on common EAZ numbers, to create the spatial EAZ dataset. The final number of EAZs in the dataset is 448,380.

Researchers ran a number of quality checks to verify the quality and completeness of the EAZ system and the correspondence of the parcel-Zip+4 matches to spatially contiguous areas and functional aggregations of parcels. By looking up parcel addresses on the USPS ZIP code lookup page, 2,118,065 parcels were successfully matched to valid ZIP+4 codes and integrated into an EAZ. Of the remaining 258,296 parcels, 43,627 had some level of information available in the Situs Address field. In an attempt to match some of these parcels to a ZIP+4 code, researchers tried using a proprietary address verification service, provided by the company SmartyStreets. This process ultimately matched about 4,500 additional parcels to ZIP+4 codes. These parcels, however, were not added to the current version of the EAZ system, as many of them did not fit in the areas that are currently covered by the available energy data provided by the utility companies. Therefore, the computational burden to update the EAZ system was not justified by eventual increases in the quality of the results of the energy consumption study.[17]

The following Figures 5, 6 and 7 show some examples of how parcel geographies are associated with Zip+4 codes (after querying street addresses through the USPS website) and how these correspondences are aggregated in the Energy Analysis Zones.

---

[17] An additional reason not to include these parcels in the current EAZ system is that the information used to match these parcels to Zip+4 codes is based on a different source than the rest of the database. This might generate an additional source of errors in the dataset, without significant gains in terms of additional records added to the database. In future extensions of the research, when data from more utility companies will be available, the quality of the information obtained from this different source will be checked more thoroughly, and the additional parcels will be added to the Energy Analysis Zone System.

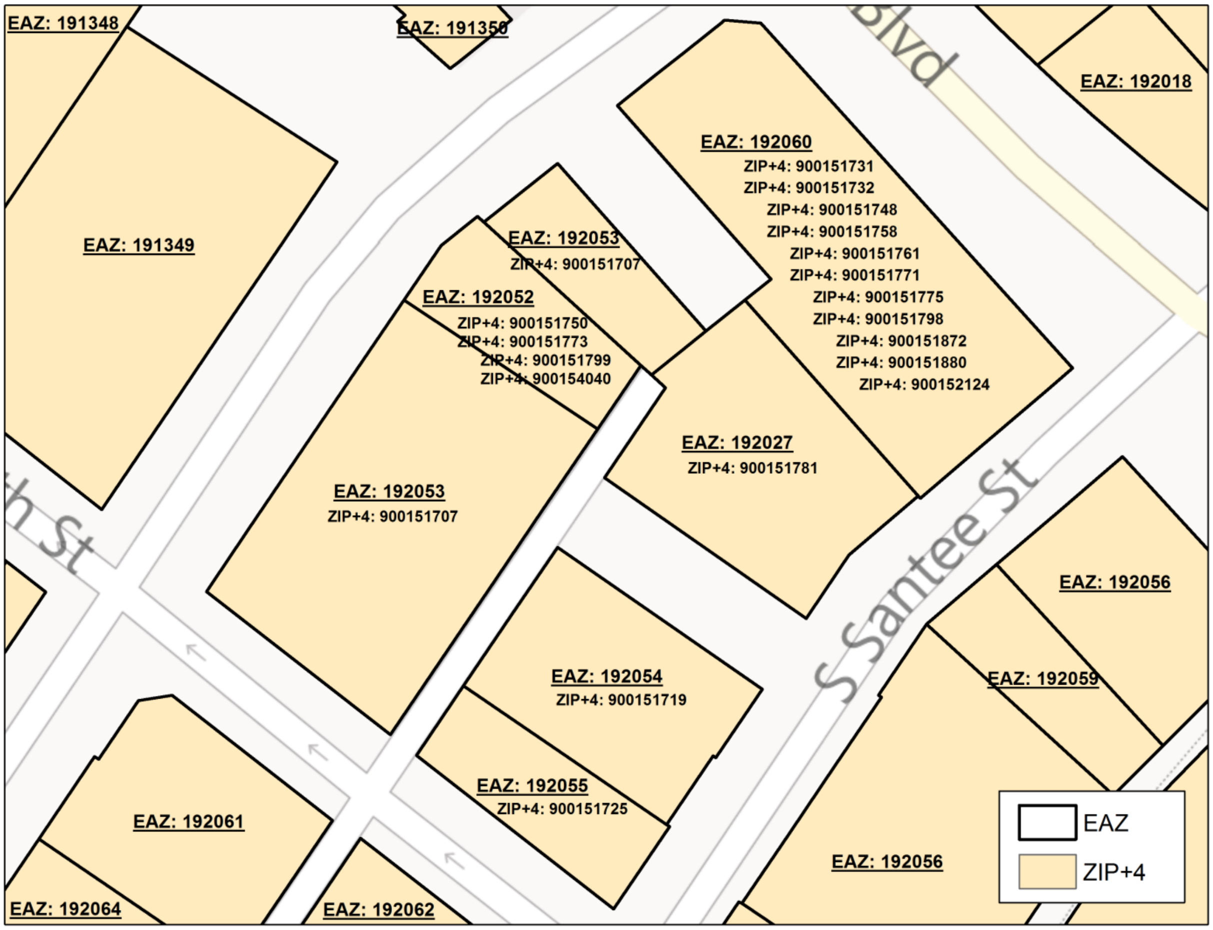
Figure 5: Spatial overlap of Zip+4 areas and parcels

Figure 6: Spatial overlap of EAZs and Parcels

As mentioned earlier in this section of the report, four different relationships might regulate the correspondence between parcels and Zip+4s. Each of the four different cases that have been described contribute to create the Energy Analysis Zones, which might aggregate a rather variable number of Zip+4 codes and parcels, depending on the location[18], and the specific relationships between Zip+4s and parcels.

[18] Neighborhoods with more uniform land uses and regular urban form tend to have EAZs that include a smaller number of parcels and very few Zip+4s.

Figure 7: Spatial overlap of EAZs and Zip+4s



Overlaps of EAZs with other geography systems

Energy Analysis Zones can be overlaid spatially with other geographic units, for the purpose of comparing and analyzing data available in different units of geographic aggregation. EAZs created for this study are based on the aggregation of parcels and Zip+4s. Given the way the EAZs are created, they nest very well in the county and city boundaries in Los Angeles County. Appendix A contains a table with the distribution of the 448,380 Energy Analysis Zones in the various cities inside Los Angeles County.

Researchers also overlaid the Energy Analysis Zones with other levels of spatial aggregation that are of interest for this research project, and in particular with census tracts and census block groups (important levels of aggregation at which sociodemographic data are aggregated by the U.S. Census Bureau). Almost all EAZs (about 99% of the total) nest perfectly in the 2,346 census tracts in Los Angeles County. Figure 8 shows an example of the overlap of the Energy Analysis Zones with the census tracts. Similarly, Figure 9 shows the spatial overlap of the Energy Analysis Zones with the census block groups in the LA County.

Figure 8: Spatial overlap of Energy Analysis Zones and Census Tracts

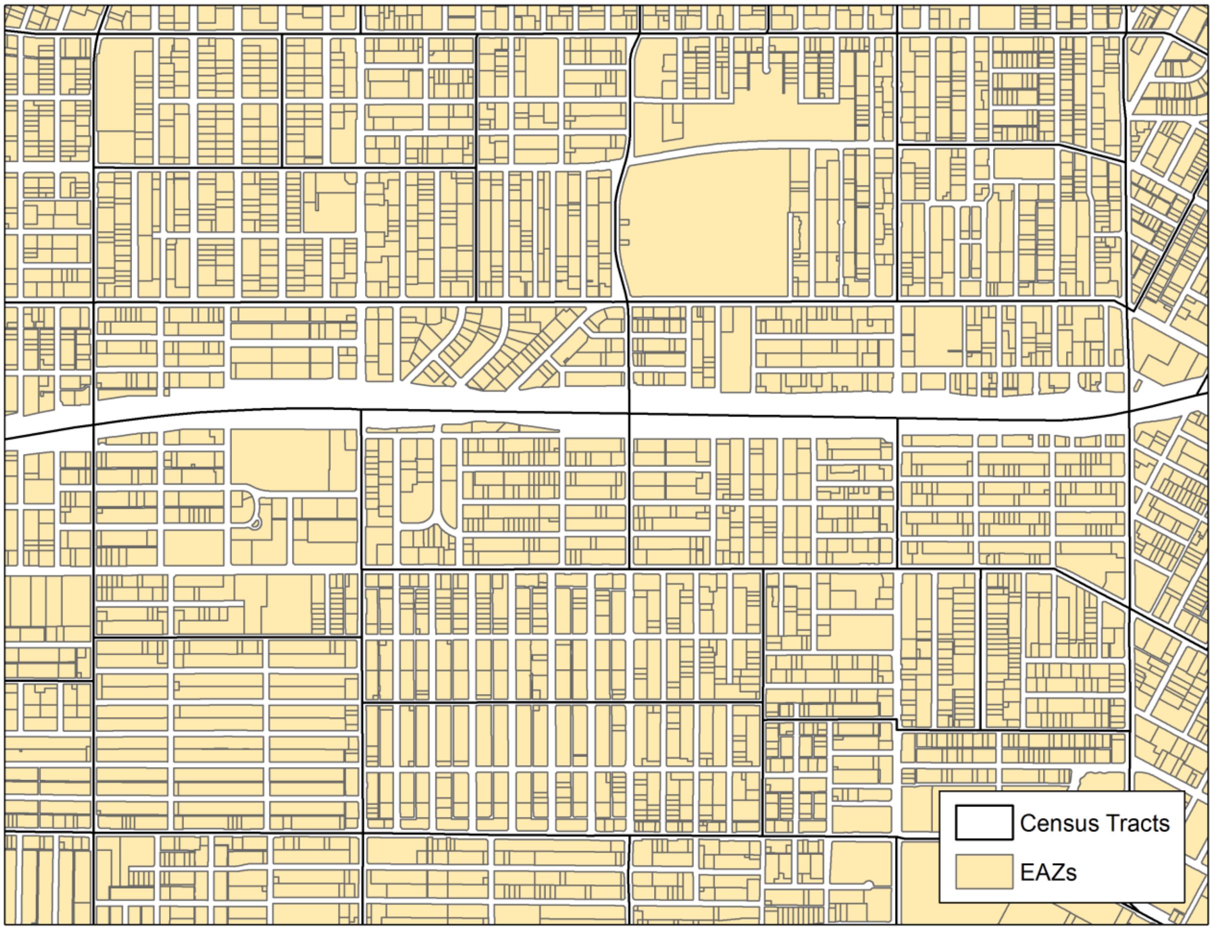
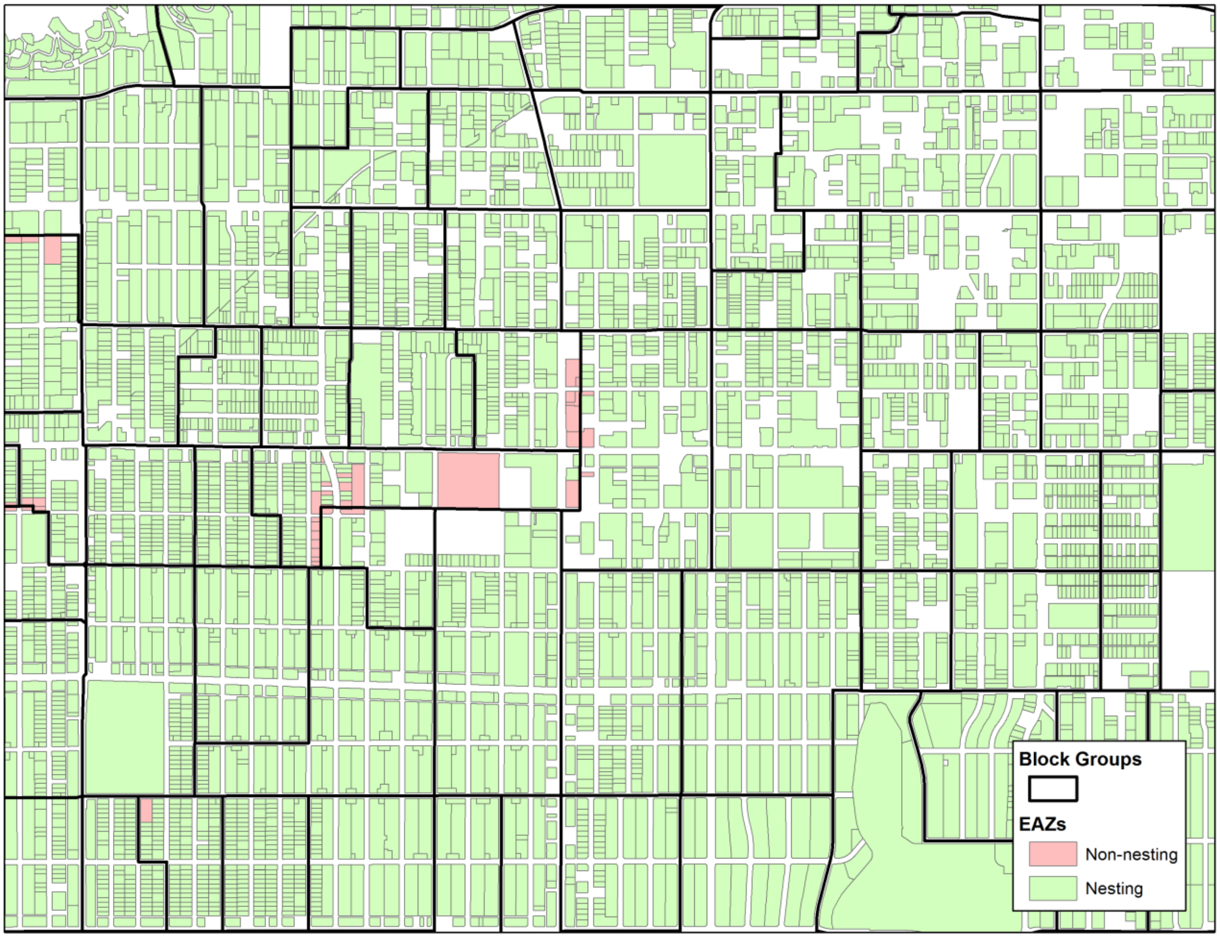


Figure 9: Spatial overlap of Energy Analysis Zones and Block Groups



Appendix A contains additional information on the spatial overlap of the Energy Analysis Zones developed for Los Angeles County and other levels of spatial aggregations, as the Traffic Analysis Zones (TAZs) and the Land Use Zones (LUZs), developed at the University of California, Davis, respectively for the analysis of transportation demand in the California Statewide Travel Demand Model and the distribution of land use activities in the PECAS model (ULTRANS, 2011).

CHAPTER 5: Utility Data

Utility companies provide the vast majority of electricity and natural gas to residents and commercial and industrial establishments within Los Angeles County through centralized distribution systems. For both electricity and natural gas, there are a small number of utility providers. The two largest electric utilities serve 92 percent of grid-supplied demand in Los Angeles County, whereas the two largest natural gas utilities serve almost 100 percent of grid-supplied demand.

Tables 5 provides a list of all utility companies that provide electricity with service territories either partly or entirely located within Los Angeles County, ordered by the quantity supplied within the county in 2010.

Table 5: Electricity utility companies servicing Los Angeles County

Utility	Usage (GWh, 2010)	Customers (2010)
Southern California Edison	31,877	1,730,792
Los Angeles Department of Water and Power	22,944	1,449,174
Burbank Department of Water and Power	1174	50,100
Pasadena Department of Water and Power	1144	62,130
City of Vernon	1138	1129
Glendale Department of Water and Power	1076	84,118
Azusa Light and Water	239	15,326
City of Cerritos	45	52
City of Industry	32	106

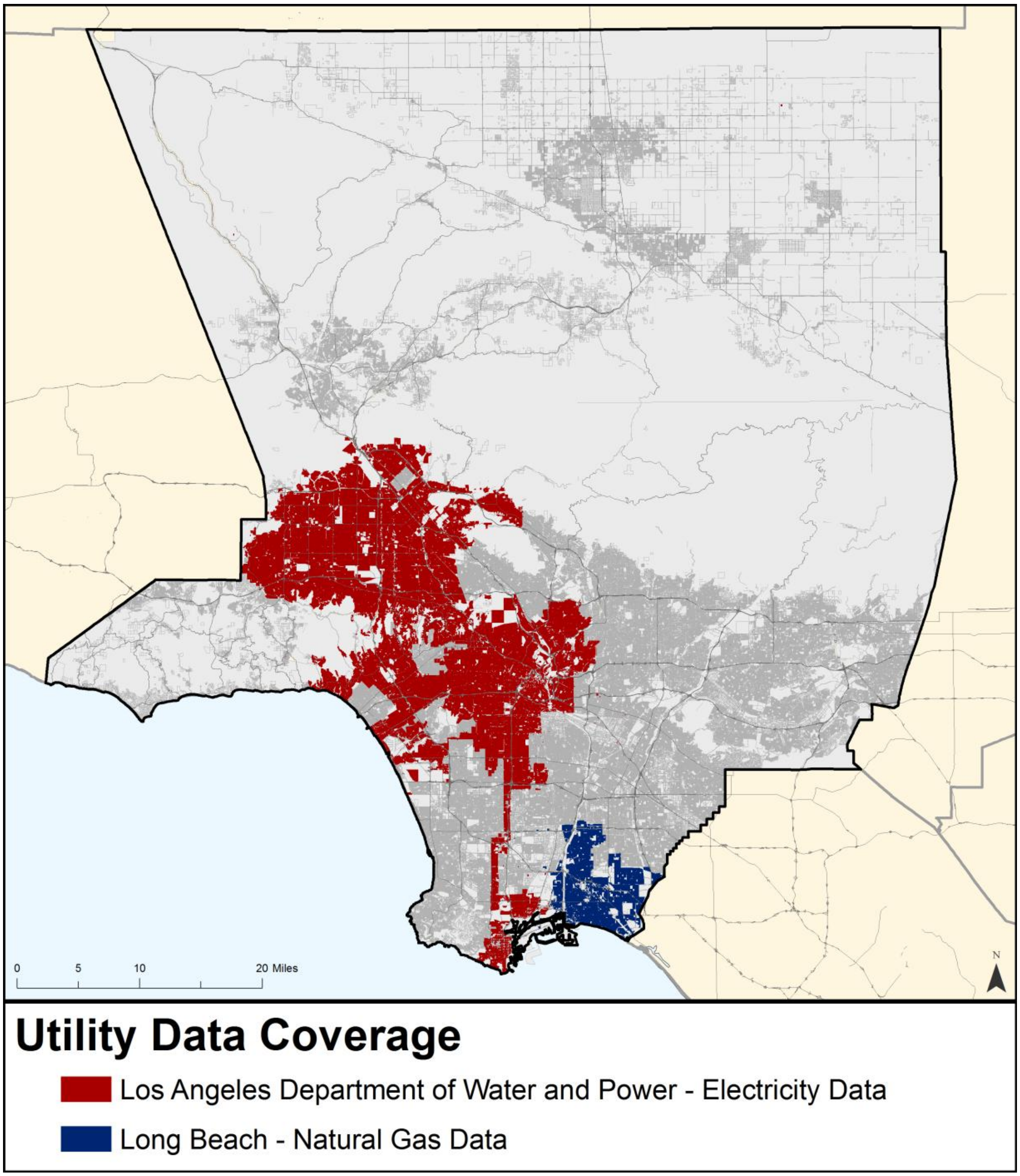
Source for electricity calculations: data provided by utilities and the CEC: http://www.energy.ca.gov/maps/serviceareas/Electric_Service_Areas_Detail.pdf

Electricity and natural gas utilities are either municipally owned or investor owned. The California Public Utilities Commission (CPUC) regulates investor-owned utilities (IOUs) providing electricity, natural gas, water, or telecommunications services within the state. The CPUC enacts decisions and rulemakings that guide IOU operations—for example, directing IOUs providing electricity to invest in energy efficiency and conservation. The CPUC also sets out provisions regarding disclosure of customer data. While disclosure of customer data to third parties is generally not permitted, several exceptions exist, including:

- A customer consents to the release of their data;
- Data are aggregated so that customers' individual identities are not disclosed;

- Data are disclosed for a primary purpose being carried out under contract with or on behalf of the utility, including for utility system, grid, or operational needs, or for the implementation of demand response, energy management, or energy efficiency programs;
- Data disclosure is otherwise permitted or required under state or federal law, or is required by an order of the CPUC.

Figure 10: Utility Data Coverage



Through the support of the researchers at the University of California, Los Angeles, and of Los Angeles County officials, researchers collected direct consumption data for flows of electricity and natural gas. Unfortunately, the data collection for these flows proved to be a very time-consuming process: although utilities universally collect these data, they were not immediately willing to provide them with the requested parameters. This significantly delayed the access to spatially and temporally disaggregated data from the utility companies.

By the time the project was developed, energy consumption data were provided by utility companies only for some geographic areas within Los Angeles County. In particular, the Los Angeles Department of Water & Power (LADWP) provided data on electricity consumption within the City of Los Angeles, and the Long Beach Gas & Oil Department (LBGO) provided data on natural gas consumption within the City of Long Beach (Figure 10). Unfortunately, the service areas of these two utility companies do not spatially overlap, which generates some problems in the analysis of the energy consumption patterns, as discussed in the following Chapter 7, which focuses on the estimation of energy consumption models for buildings with the use of these utility data.

Electricity data

The Los Angeles Department of Water and Power provided electricity consumption data for their entire area of service in the City of Los Angeles aggregated at the Zip+4 level. The data provided by this utility company included total annual consumption of electricity (in KWh/Zip+4 area) for all customers located in each of the 254,910 Zip+4s in the LADWP area of service. Electricity consumption data did not separate for different uses (e.g. residential vs. commercial) and covered all calendar years 2005 to 2010.

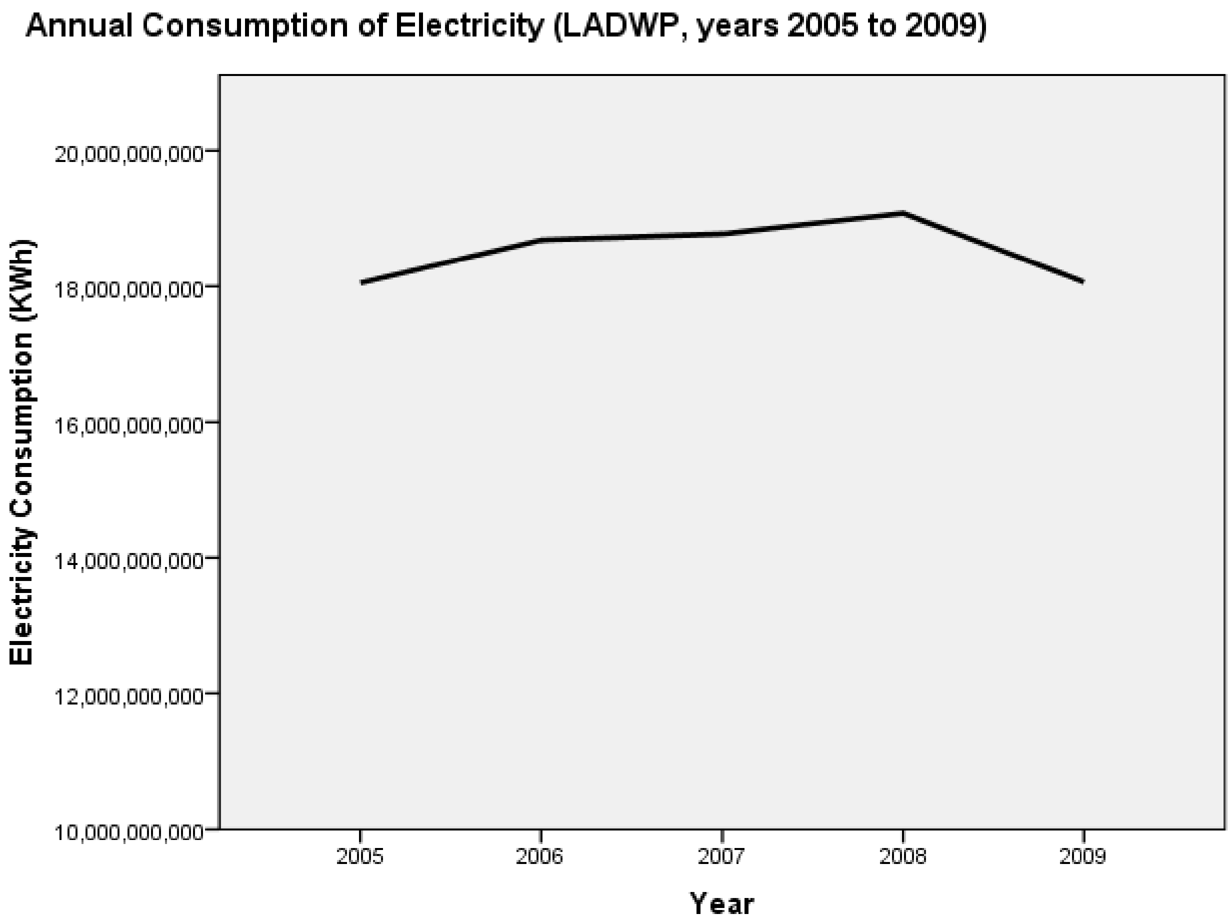
The electricity consumption data that were received are referred to 254,910 Zip+4s that are included in the LADWP area of service. For the purposes of the analysis of this project, researchers aggregated annual electricity consumption data at the Energy Analysis Zone level, using the crosswalk between Zip+4s and EAZs that was created in the process of spatial aggregation described in Chapter 4 of this report. The final energy dataset contains information on electricity consumption for 150,743 EAZs.

Figure 11 shows the total annual electricity consumption in all EAZs in the LADWP dataset for the years 2005 to 2009. Annual electricity consumption data for 2010 were discarded because the data proved to be significantly incomplete. After verification with LADWP representatives, it was established that this was due to the process of temporal aggregation that was used by the utility company and the time in which these data were compiled, as the data included only the first three quarters in year 2010. As Figure 11 shows, the total electricity consumption for all EAZs in the LADWP area of service vary between 18,050 GWh (in 2005) and 19,078 GWh (in 2008), and it approximately includes 90% of the total electricity consumption of LADWP users.[19]

[19] Electricity consumption of some users may not be included in the database used for studying energy consumption in buildings in this project, either because (1) it was not possible for LADWP to aggregate electricity consumption for these customers at the Zip+4 level of spatial aggregation, or (2) it was not

Total electricity consumption in the area of study increased at an average annual growth rate of 1.87% from 2005 until 2008. The electricity consumption significantly decreased (of more than 5%) from 2008 to 2009. This result might be explained by the economic crisis, and it is consistent with the reduction in energy use observed for LADWP and other energy providers and the conclusions reported in the Revised California Energy Demand Forecast for 2012-2022 (Kavalec et al., 2012).[20] Additional effects might be explained by weather effects, which often significantly affect both winter and summer energy consumption, respectively for heating and cooling purposes.

Figure 11: Total annual electricity consumption in the Energy Analysis Zones served by LADWP



It is also important to stress that the total consumption of electricity reported in the energy dataset does not include the total consumption of electricity consumed by all LADWP

---

possible to match these Zip+4 areas with the corresponding parcels (and building information) to generate the Energy Analysis Zones used in the study.

[20] According to the report, electricity consumption declined during 2009, and in particular became lower than the previously forecasted California Energy Demand (CED) for 2009 due primarily to the economic downturn.

customers in Los Angeles County, but accounts for about 90% of it. Some customers may not be included in the dataset, as a result of the process of aggregation of the data at the Zip+4 level done by the utility company before releasing the data to the researchers.[21]

The geographic specificity of electricity consumption data allows researchers to explore relationships between consumption outcomes and a number of explanatory variables such as land use, income, and socio-demographic characteristics using a variety of data sources and statistical approaches. Researchers also generated maps of electricity consumption to provide a spatial representation of usage patterns in the county. Researchers generated both static maps and supported the development of a web-based interactive map that was developed by the colleagues at the University of California, Los Angeles.

Natural Gas

At the time the project was developed, data for the consumption of natural gas in Los Angeles County were only available for the Long Beach Gas and Oil (LBGO) utility company. Original data were provided by LBGO in several files, which were broken out by facility type and consumption level, and summarized by ZIP+4 codes. The researchers merged the files together and imported the resulting table into a SQL Server database where they were joined to the other additional data for further analysis. The data were joined to the corresponding EAZ using the ZIP+4 code, and then aggregated by EAZ and time period, to produce summary values.

[21] Actual growth rate in the annual electricity consumption in the City of Los Angeles and in the complete LADWP service area might slightly differ from the data presented above, for the aforementioned reasons.

Figure 12: EAZs with 2008 natural gas consumption data from LBGO

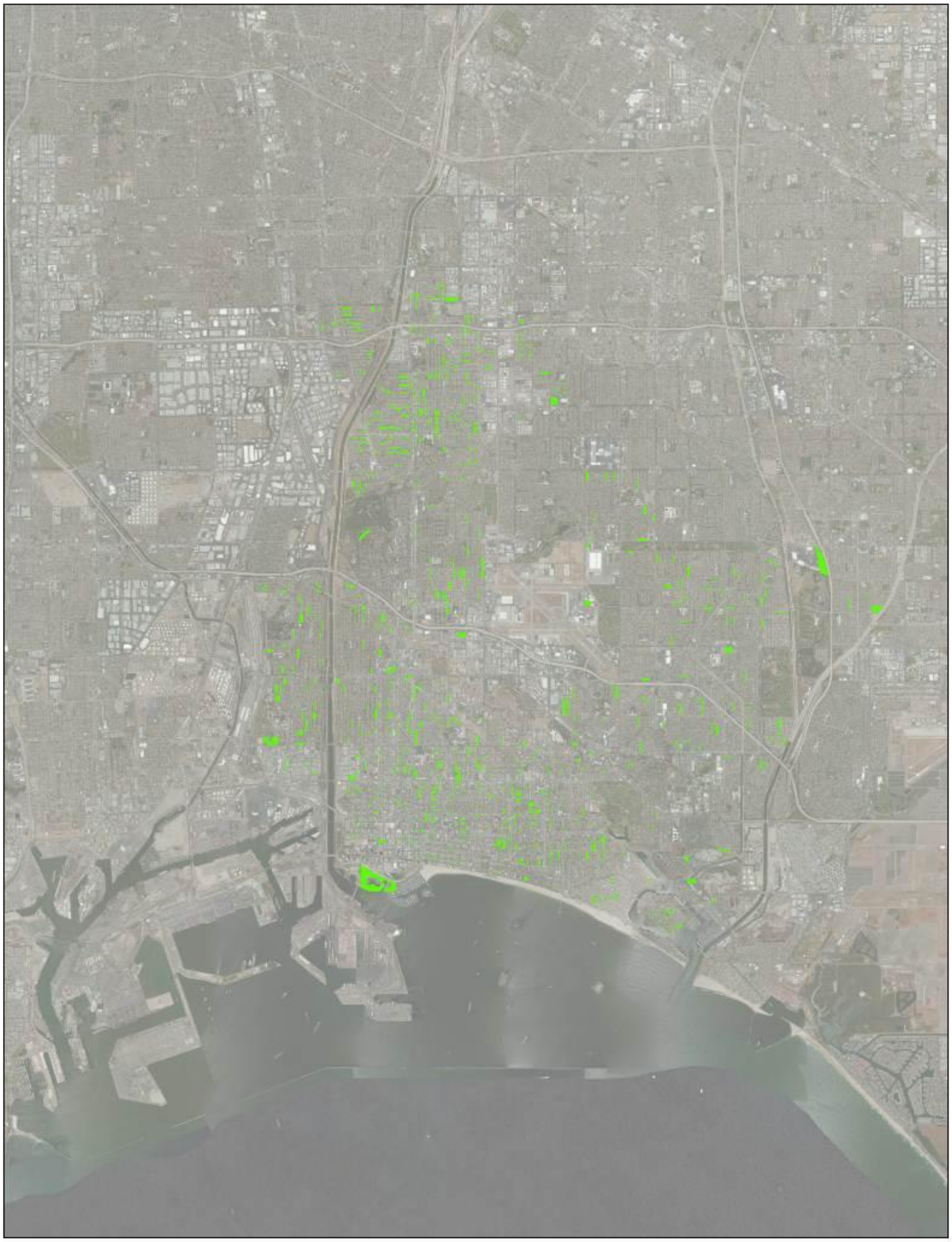


Figure 13: EAZs with 2009 natural gas consumption data from LBGO

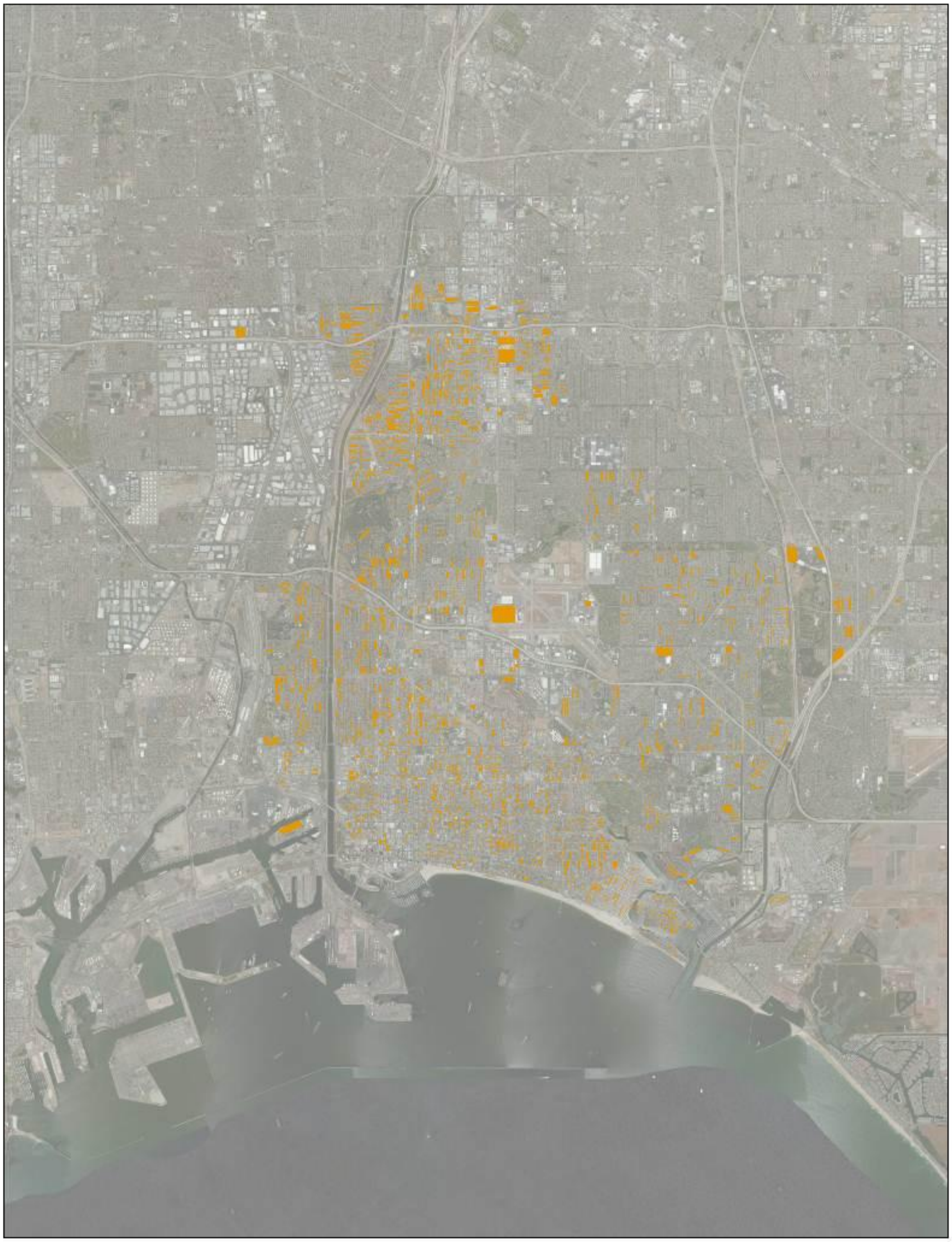
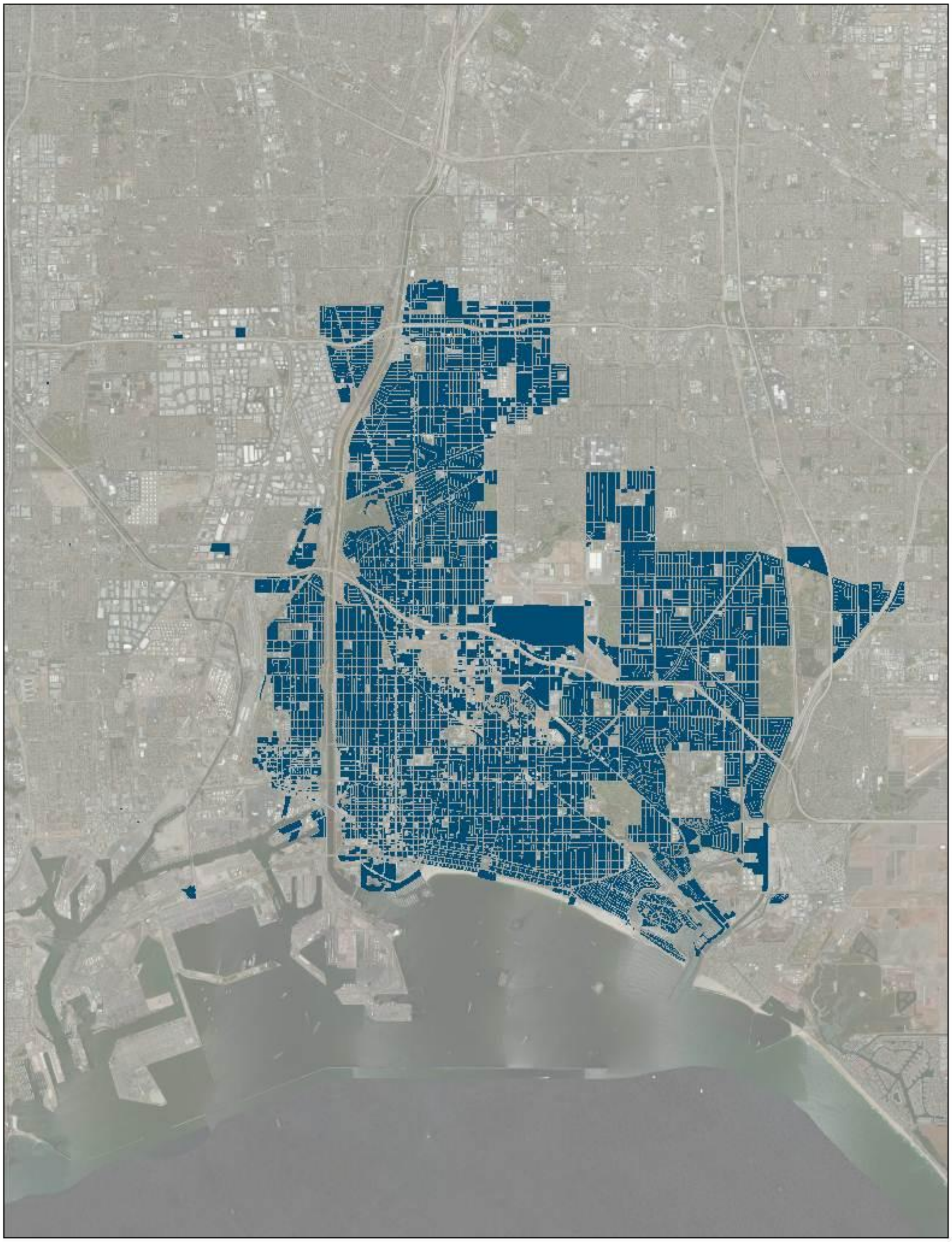


Figure 14: EAZs with 2011 natural gas consumption data from LBGO



However, from the analysis of the data that were provided by the utility company, it resulted that the data for the consumption of natural gas were largely incomplete for all years before

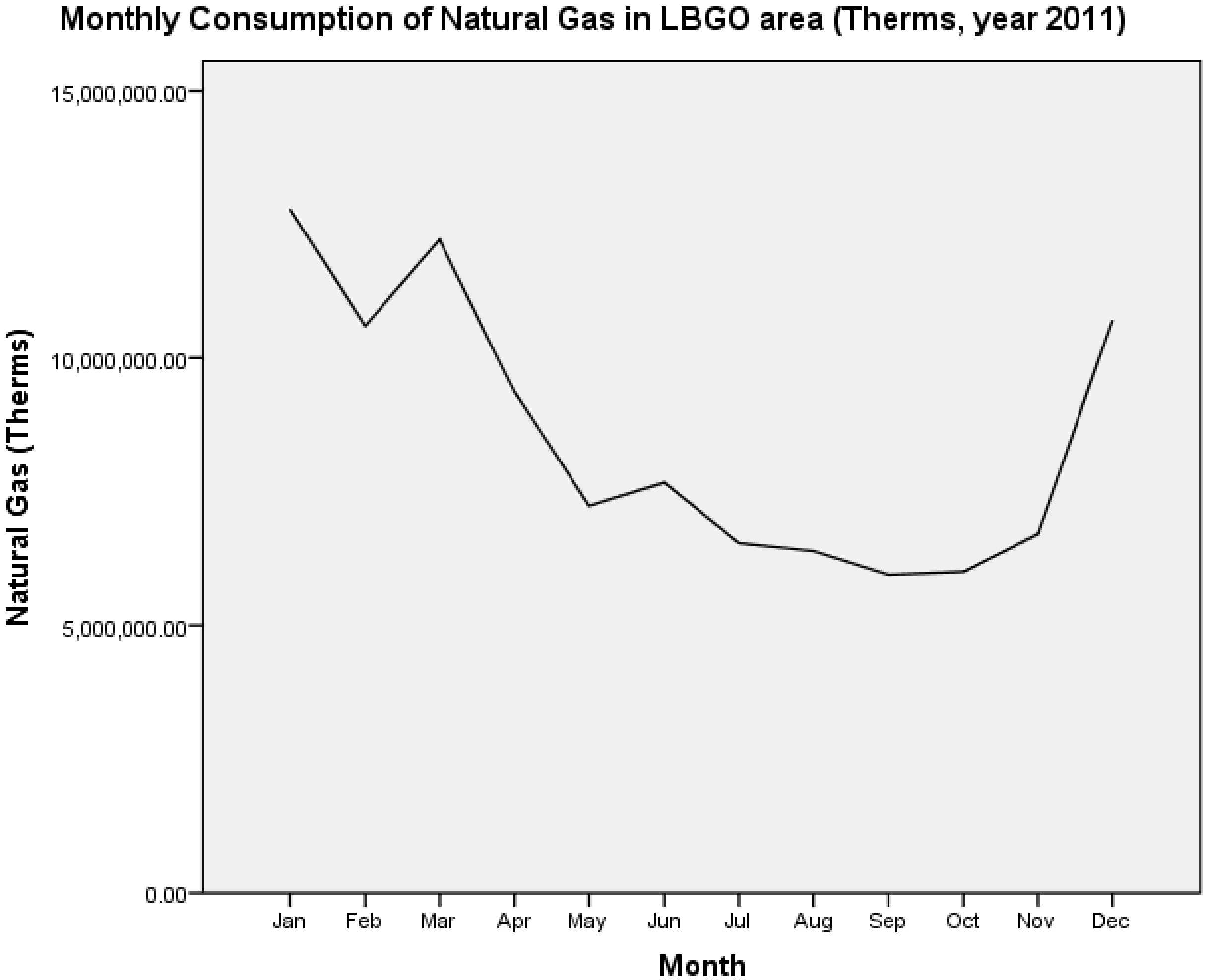
2011. Figures 12, 13 and 14 summarize the areas of coverage for the LBGO natural gas consumption data respectively for 2008, 2009, 2011. Monthly natural gas consumption data were complete only for year 2011, and for the first months of year 2012.

The incomplete data issue was caused by the LBGO billing system, which only retains 24 months of billing data for each account number in its database. The system is designed to overwrite new billing information over older records for all customers that do not move their location or their "tenant" status. This causes older records to be largely incomplete, and generate missing gaps in between continuous records for all records that are referred to more than 24 months before the data were released by LBGO.[22] In the context of this data request, this means that all records outside the 24 months before the data request are largely incomplete. As data were received by LBGO in April 2012, the natural gas consumption dataset is supposed to be largely incomplete for all records before May 2010.

As a consequence of the incompleteness of the data on natural gas consumption data until 2010, the researchers selected 2011 as the year for the analysis of natural gas consumption data. This decision, which was forced by the data availability from the utility company, generated one additional problem for the development of energy consumption models (and their interpretation) in terms of the consistency of the data used in the study. Not only do these natural gas consumption data describe a different, non-overlapping geographical area than the electricity consumption data, and therefore it is not possible to control for the possible substitution effects for the use of the two energy sources in buildings. The data do not overlap on a temporal scale either: electricity data refer to 2008, natural gas data to 2011. This limitation to the validity of the results of the study is further discussed in the following sections of the report, and is the object of further investigation and comparison with data on energy consumption in buildings from other sources. Figure 15 shows the variation of natural gas consumption in the entire LBGO service area during the year 2011.

[22] A one-digit "tenant code" is included as part of each account number: therefore, if for example a housing unit keeps turning among different uses over every 24 months or less, the associated account number is changed before the billing system can no longer store additional monthly usage data. In this way, a continuous record of usage for that unit is built over many years. However, if the housing unit turns over less frequently than every 24 months, new months will replace old months in the billing system database and gaps in the record will be introduced. LBGO also confirmed that apartment units turn over much more frequently than houses in their service territory, so in addition of being largely incomplete, residential data from before May 2010 will be skewed toward the usage patterns of multi-family dwellings.

Figure 15: Total monthly natural gas consumption in the Energy Analysis Zones served by LBGO



Natural gas data were delivered by LBGO as monthly record for Zip+4 areas. The data were aggregated at the Energy Analysis Zone level of spatial aggregation for the purposes of studying the building energy consumption in this study.

Future availability of energy consumption data for Los Angeles County

While researchers have acquired and used the data described in this section of the report in the development of the baseline energy analysis, additional negotiations are under way with the utility companies operating in the Los Angeles County for the provision of additional data that will improve the coverage and certainty of the results of this study. At the time the project is developed, researchers at UC Davis, together with colleagues at UCLA are continuing to work with utility providers to acquire these data. This process follows the efforts of the Los Angeles County Office of Sustainability, which spent more than a year working to collect similar data before the beginning of this project.[23] In particular, researchers request data for each electricity and natural gas utility's full service territory within Los Angeles County with a higher level of details with the following parameters:

- Spatial granularity: by service address;
- Temporal granularity: by billing cycle (monthly or bimonthly);
- End-user granularity: by tariff (e.g. residential, commercial, industrial).

The request for spatially and temporally more disaggregated data is motivated by the interest to develop a more disaggregated modeling analysis on energy consumption patterns in buildings to inform the environmental sustainability studies in Los Angeles County.

Researchers initiate data collection by identifying and contacting appropriate utility staff, including customer service representatives, account managers, and general managers. The high spatial granularity requested by researchers was one of the main issues for some utilities, primarily because of implications for customer confidentiality. As identified above, the CPUC does not allow release of customer data except in certain situations—one such situation being that data are aggregated such that individual consumption signatures are not identifiable. Since its promulgation by the CPUC, IOUs have used a threshold referred to as the "15/15 rule"—where any data released must be composed of at least 15 customer accounts, with no one account comprising more than 15 percent of total usage—to determine whether data are sufficiently aggregated. In meetings with UCLA and the IOUs, the CPUC has affirmed that this threshold is only a guideline rather than a steadfast rule, but it has not issued a formal clarification to this effect or provided further guidance.[24] To address this limitation, researchers at UC Davis developed a computer routine to flexibly aggregate IOU service address data to data points that just satisfy 15/15, thereby maximizing spatial granularity. The IOUs did not initially support researchers' efforts to devise a collaborative solution to data provision under the 15/15 guideline. Their willingness to work with researchers developed in response to researchers at UCLA building strong relationships with CPUC staff and commissioners and with the Governor's Office of Planning and Research (OPR). Researchers argued the social benefits and need for access to consumption data to generate effective policy responses to environmental challenges such as climate change.

---

[23] This process entailed significant time spent by the researchers and by the colleagues at the University of California, Los Angeles for meetings with utility staff and management, as well as with local and state government representatives and CPUC staff and commissioners to generate pressure for release of the data.

[24] The primary challenge this poses for research is that the IOUs were only willing to provide data satisfying the 15/15 guideline through a one-size-fits-all aggregation. For residential data, this may equate to the ZIP code, while for commercial or industrial usage the data often only satisfy 15/15 when aggregated to the individual city. In both cases, this granularity is insufficient to provide the level of certainty required for the energy consumption study.

As of this writing, Southern California Edison was working with researchers to implement researchers' computer algorithm. In contrast, Southern California Gas Company (SCG) told researchers it would continue to abstain from providing customer data until required to do so by law or by an order from the CPUC. SCG maintains that their customer database is a business asset owned by their shareholders. The high spatial granularity of researchers' data requests similarly generated concern from municipally owned utilities (MOUs). However, MOUs' rules generally allow them to work with researchers to provide the requested data as long as there is a clear and significant benefit to the public and the utility, such as improving energy efficiency and conservation programs, and as long as customer data remains strictly protected. Energy consumption data from other MOUs operating in the area of study were not provided to the researchers to date. However, the process of obtaining access to these data is under development and close to a future successful conclusion. Although it was not possible to include these additional data in the development of the analysis of this current project, the researchers organized the research activities so that, when these data will become available, it will be possible to analyze them in future extensions of the project using the database structure and methodology approach already developed for this study.[25]

[25] The definition of the Energy Analysis Zones and the structure of the energy consumption database developed for this project were prepared for the entire Los Angeles County, and are designed to use data from all utility companies in the County. In the current study, however, given the current limitations on the amount of data provided by the utility companies, the estimation of energy consumption models was carried out for the areas covered by the energy consumption data available at the time of development of the project. The following sections of this report discuss the validity and extension of the results from this study to the other areas currently not covered by energy consumption data.

CHAPTER 6: Input Data

Several different sources of data were used to create the energy consumption database that is used in this project. These sources cover a wide variety of physical and socioeconomic variables, and were compiled into a single database to allow the study of energy consumption in buildings at the level of Energy Analysis Zones (EAZs). There are 449,539 EAZs in the final energy database, 448,380 of which can be represented spatially in a GIS dataset. The database includes more than 1,000 explanatory variables, which measure different characteristics of the land use, building stock, natural environment, and sociodemographics in the area of study. The sources for these data include the Los Angeles County Assessor's Property Database, demographic data from the U.S. Census Bureau – American Community Survey, geomorphologic information on slope and aspect, information related to the building climate zones as defined by the California Energy Commission, and climate data from previous research studies developed at the University of California, Davis. This section of the report describes the data that researchers used in the analyses, with necessary details on the data sources that were accessed, the transformations that were applied for data processing and analysis, and the level of spatial aggregation at which the data were available.

Building Information from the LA County Assessor's Property Database

The Los Angeles County Assessor's Property database contains information on individual parcels, and on the building stock in Los Angeles County, including information on square footage, construction type, and value. The data received from the Assessor were imported into a Microsoft SQL Server database, following the methodology described in the Chapter 3 of this report. The following list summarizes some of the main fields that were included in the original dataset, or that were attached by researchers to the parcel records, for inclusion in the energy consumption database:

AIN (Assessor's ID Number) – Identifies individual parcels within the Assessor's database. It consists of a Mapbook number (4 digits), page number (3 digits), and parcel number (3 digits).

LUZ (PECAS Land Use Zone) – Zone that the parcel majority resides in; derived by GIS overlay of parcels and zones. This zone is used by the PECAS model to determine commodity flows between the geographic areas.

TAZ (CSTDM Traffic Analysis Zone) – Zone where the parcel majority resides in; derived by GIS overlay of parcels and zones. These zones are used by the CSTDM (California Statewide Travel Demand Model).

ZIP+4 (United States Postal Service 9 digit zip code) – This information was attached to parcel records via address matching with USPS records.

Land Year – Year of current land value

Land Value – Value of land, excluding improvements

Improvement Year – Year of current improvement value

Improvement Value – Value of improvements (structures) on parcel.

Situs Address – The street address of the parcel; used to identify ZIP+4 codes.

Zone Code – Zoning classification given to the parcel by local jurisdictions or cites. The first two characters represent the city code. The 3rd character represents the type of zoning, such as agricultural, commercial, industrial, or residential. The 4th through 15th place characters represent the zoning of the parcel.

Use Code – Actual current use of the property regardless of zoning. This consists of four alphanumeric characters. The 1st character denotes the general classification (e.g., 0 = residential, 1 = commercial, etc.). The 2nd character further defines the type of property within the major classification. The 3rd and 4th characters indicate additional characteristics, and the presence of specific features.

Last Sale Amount – Dollar amount of the last sale price of the parcel.

Last Sale Date – Date of the last sale.

*Number of Rental Un*its – The total number of rental units on the property.

For each parcel, the characteristics of up to five buildings are reported. The information available in each of these fields is summarized in Table 6.

Table 6: Building variable in the Los Angeles County Assessor's data

Field	Description
Design Type	A 4-character code describing the original purpose for which the improvement was intended, providing the building has not been extensively remodeled.
Quality, Class, Shape	A 5-character code identifying the class of construction, quality of construction, and shape of the perimeter. See below for additional information.
Year Built	Original year the structure was built.

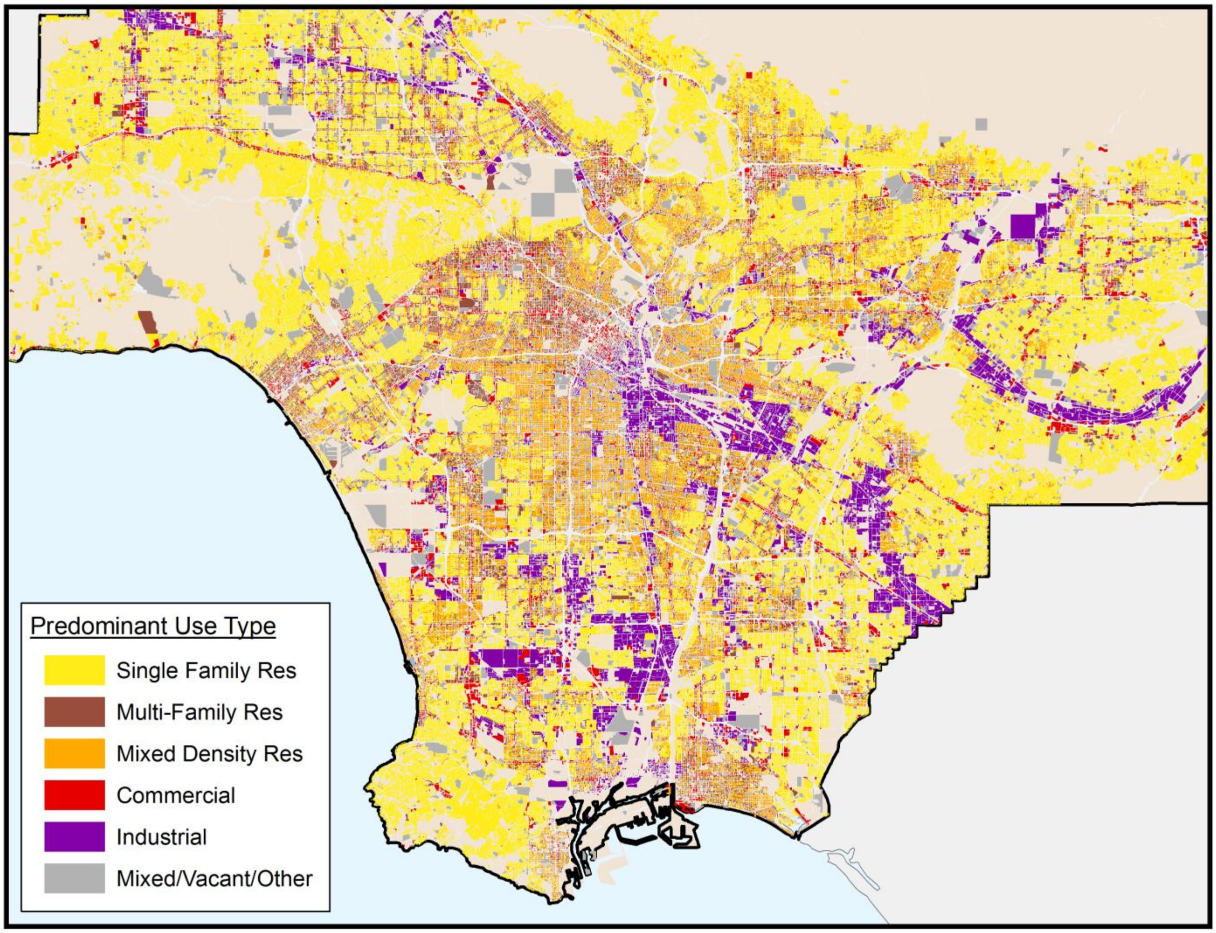
Number of Units	The number of stores, residential units, etc., contained in a multiple unit type structure. This code follows a similar scheme to the “UseCode”.
Number of Bedrooms	Number of bedrooms present in a single residence or the total bedrooms in the apartment
Number of Bathrooms	Number of bathrooms present in a single residence or the total bathrooms in the apartment
Square Footage	The total area in square feet of the main structure
Unit Cost Main	The dollar cost per square foot for the main structure
RCN Main	The cost of replacing the main structure (square feet x unit cost)

Quality, Class, Shape (QCS) code – Additional details:

> A three part (5-character) code designating Quality, Class, and Shape of the improvement. The first character denotes the building class. For example, “A” represents a building having a fireproofed structural steel frames carrying all wall, floor, and roof loads. Wall, floor, and roof structures are built of noncombustible materials. The next 1, 2 or 3 characters represent the quality of the construction. The quality class ranges from 1 to 12.5 and gives a relative assessment of the construction quality of the structure. The last character denotes the shape of the perimeter. Shape classifies the structure by how regular or irregular the shape of the structure is, to determine the aspect ratio of the front/back and sides and whether the structure is a simple box or has more complex geometry.

To make the information in the Assessor’s database comparable to other datasets, including the energy consumption data, it was aggregated into EAZs. During this process, additional transformations were applied to some of the variables. For example, the original 4-character use code, which had over a thousand unique values, was cross-walked into a set of 21 new use categories (see Chapter 3, Table 3), which were derived from PECAS floorspace categories. All of the relevant variables from the Assessor’s dataset – for both parcels and buildings – were then aggregated to EAZs, using both the original use code and the new use categories. The reason for doing this was that later steps utilize both the original 4-character use codes, and the new categories.

Figure 16: Predominant Land Use Types (100% minimum threshold)



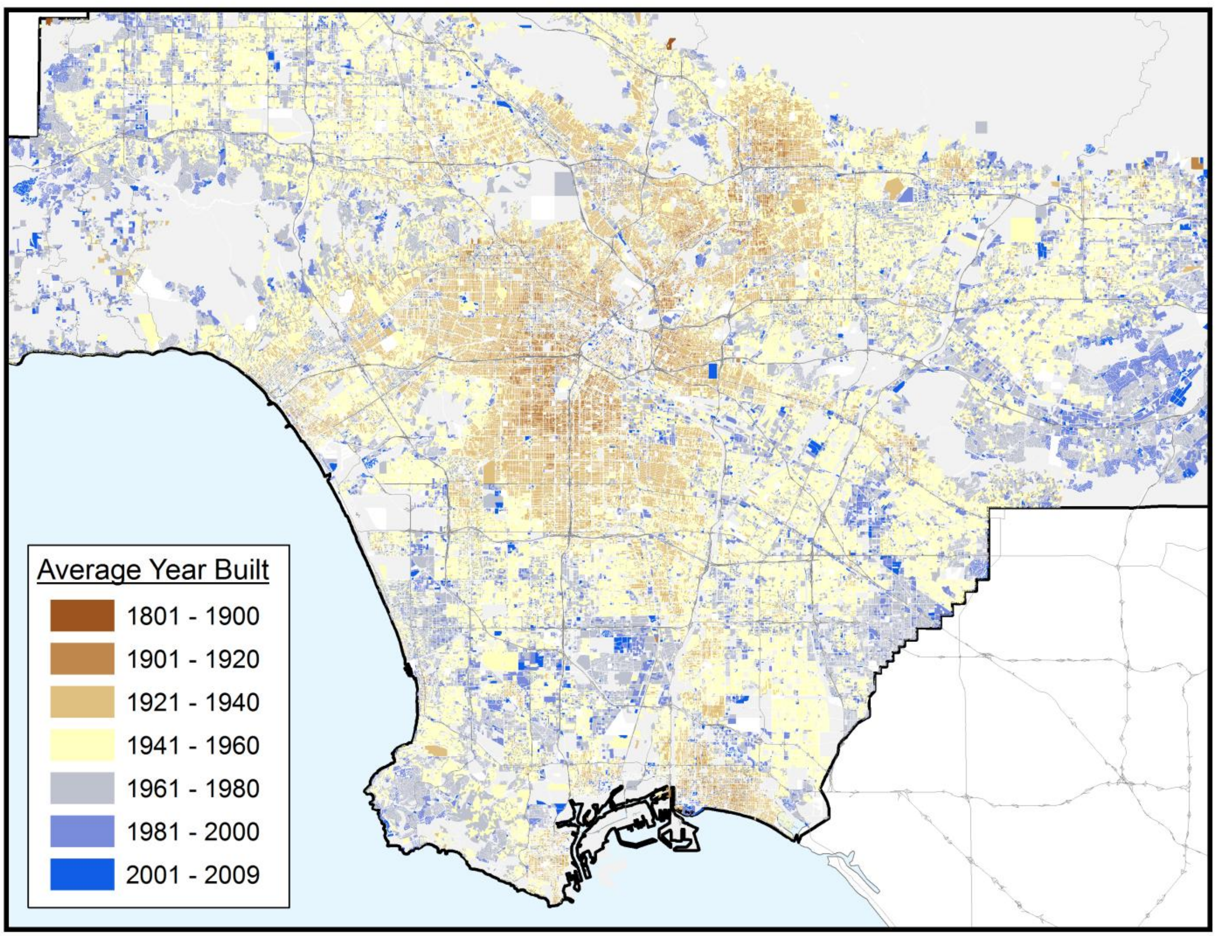
To identify EAZs of a predominant use type, we queried the table with the values summarized by original 4-character use code to identify the use codes that are associated with the majority of the developed square feet in the parcel. The use codes are passed to a custom T-SQL function, which classifies each EAZ into one of several general categories, based on the proportion of square footage in the zone:

- *Single Family Residential*
- *Multi-Family Residential*
- *Commercial*
- *Industrial*
- *Vacant*
- *Other*

The total square footage in each of these categories is calculated, and then passed to another function, which calculates if any of the six general types comprises a proportion of the total that is above a specified minimum threshold. If so, the EAZ is identified as being predominantly composed of that type (Figure 16). If no single type falls above the minimum threshold, then the EAZ is classified as mixed. A Mixed Residential category is also used to identify EAZs for which neither the Single Family Residential nor the Multi-Family Residential uses reach the threshold used in the computation but the sum of the two categories of floorspace types reach the threshold. Minimum threshold values were tested at 70, 80, 90, 95, 99, and 100%. This was done to examine how changing the minimum threshold value would affect the classification of EAZs into predominant use type categories. In addition to identifying predominant types, the total square footage falling into each of these categories was calculated for each EAZ, and added to the final dataset.

Average building age was calculated for each EAZ, using the "Year Built" attribute (Figure 17). The average building age was added to the final dataset as a potential explanatory variable for energy consumption.

Figure 17: Average Building Age



Several additional attributes were summarized using the information available in the use type, design type, and QCS codes. For example, the 4-character use type codes indicated residential buildings that had a pool on the property. For each EAZ, the number of parcels (total and proportion) and the building square footage (total and proportion), was calculated.

The design type code is another 4-character code, which provides information on the original purpose for which the building was intended. It contains information on technological features, such as the type of heating and cooling present in the building. The information in this field was summarized into a few general categories. For each of these categories, the number of buildings, total square footage, and proportion of square footage was calculated at the EAZ level.

A similar process was done to summarize the construction class information contained in the QCS codes. As with the design type codes, the number of buildings, total square footage, and proportion of square footage was calculated at the EAZ level, for each class code (Table 7).

Table 7: Construction Class Codes

Class	Description
A	Buildings have fireproofed structural steel frames carrying all wall, floor, and roof loads. Wall, floor, and roof structures are built of noncombustible materials.
B	Buildings having fireproofed reinforced concrete frames carrying all wall, floor, and roof loads. Wall, floor, and roof structures are built of noncombustible materials.
C	Buildings having exterior walls built of a noncombustible material such as brick, concrete block, or poured-in-place concrete. Interior partitions and roof structures are built of combustible materials. Floor may be concrete or wood frame.
D	Buildings having wood or wood and steel frame.
S	Those specialized buildings that do not fit in any of the above categories.

American Community Survey (ACS)

The United States Census Bureau conducts the American Community Survey on an ongoing basis to provide current information on demographic, social, economic, and housing characteristics. The 5-year estimate (2006 – 2010), which is centered on the target year of 2008, was selected to provide data on these characteristics within LA County. ACS data is also made available in 1-year and 3-year estimates, but the geographic resolution of these datasets is much coarser than the 5-year estimate and, because they are based on smaller samples, the data are less reliable.

For this project, a subset of data was extracted, covering the spatial extent of LA County. A PostgreSQL database was built using raw census files, and queried to produce the subset of variables that were of interest for this project. The data can be accessed at several levels of census geography, but block groups were used because they provide the most useful set of attributes at the highest spatial resolution. There are 6,425 block groups in LA County.

A full description of the 2006-2010 ACS 5 Year Summary File, including all available attributes, can be found on the website of the US Census Bureau.[26] The attributes selected for use in this project are as follows:

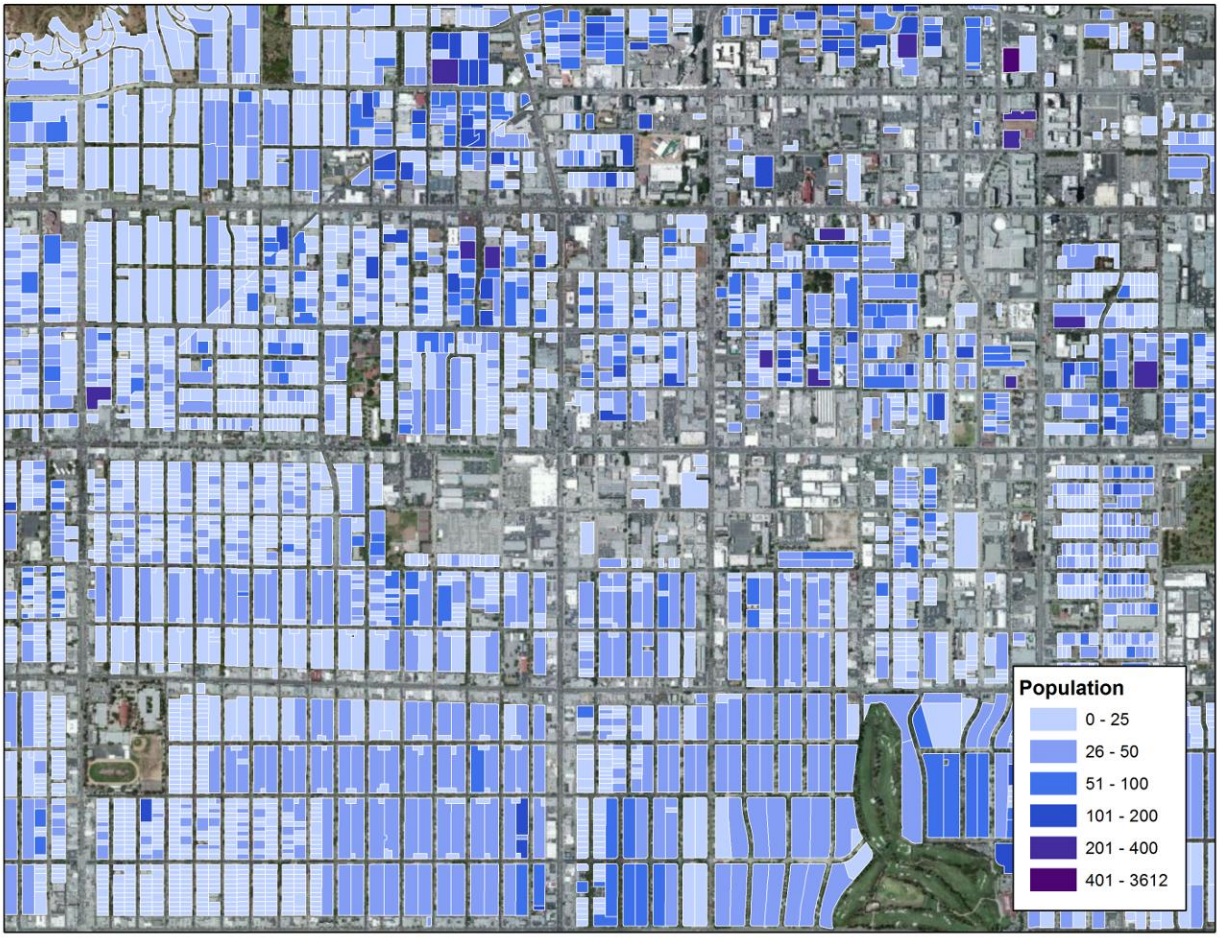
- Total population
- Median age
- Population by age category
- Household occupancy status
- Median household income
- Median household income by race
- Hispanic or Latino origin by race

After extracting the ACS data, each of the variables was allocated to the EAZ polygons. The ACS data were originally available at the level of census block groups, which are generally larger than the EAZs (on average, about 70 EAZs per block group). As a result, the ACS data had to be allocated using an appropriate method for each variable.

Population was disaggregated to EAZs using the amount of residential square footage in each EAZ. Residential square footage was obtained from the LA County Assessor's Parcel Database, which crosswalks directly to EAZs. A Python script was developed to control the disaggregation process. It begins by converting the EAZ polygons to points, and uses a spatial join to attach the ACS attributes to the EAZ points. The total residential square footage is first calculated for a block group, and then the proportion that each EAZ contributes to that total is calculated. This proportion is then multiplied by the total population of the block group, to allocate it to EAZs in the same ratio as the amount of residential square footage. In rare cases (~0.4% of all persons), some amount of population is estimated to occur in a block group where the LA County Assessor's database does not record any non-vacant residential square footage. In this case, the population is allocated to EAZs using the proportion of total area in the EAZ polygons. Figure 18 illustrates the result of the population disaggregation process, where population has been allocated to the EAZs that contain residential buildings.

[26] See http://www2.census.gov/acs2010_5yr/summaryfile/ACS_2006-2010_SF_Tech_Doc.pdf

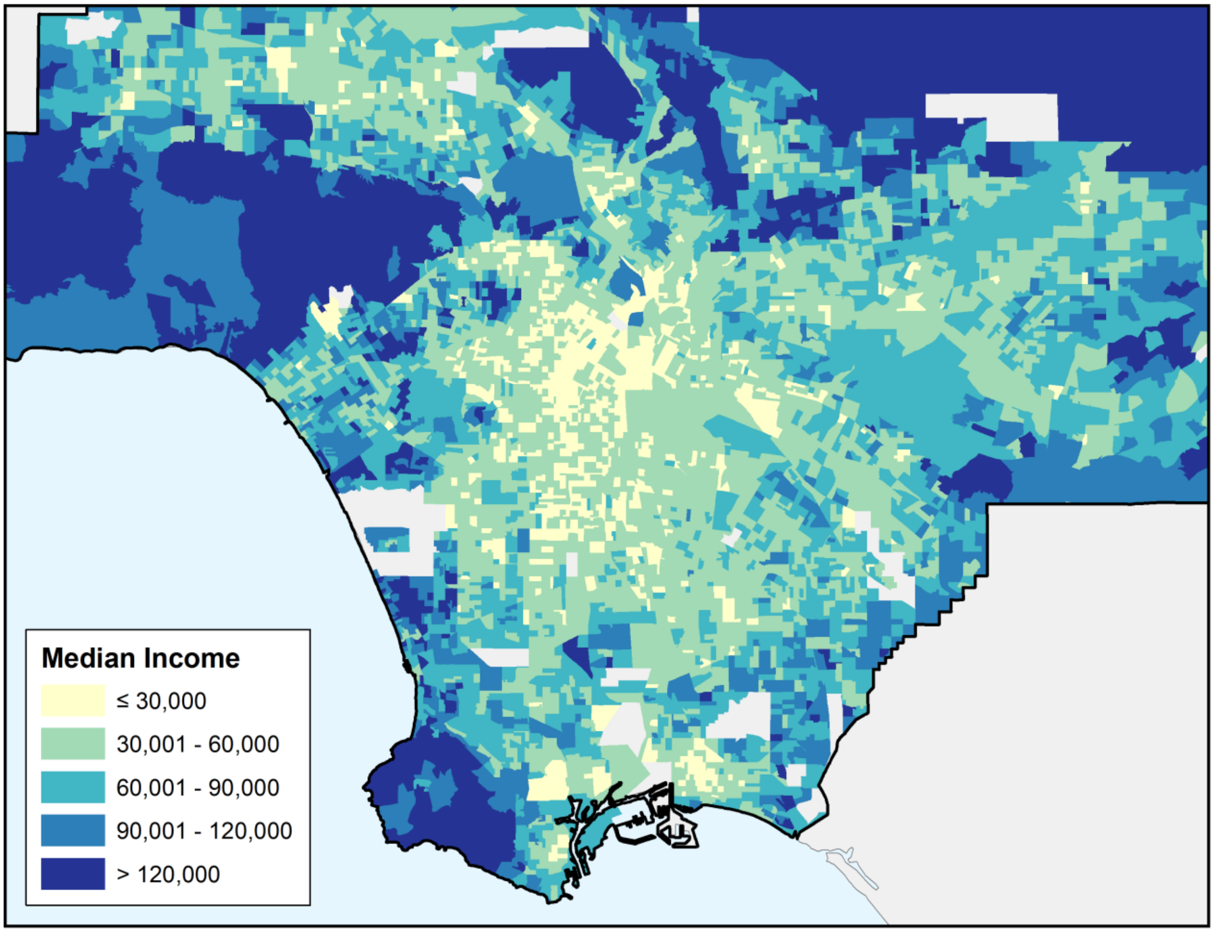
Figure 18: Population values after being allocated to EAZs.



The median age attribute for each block group was allocated to the EAZs using the same EAZ-to-block group spatial relationship developed for population. It is assumed that the median age in the block group is a relatively good representation of the median age in the corresponding EAZs. The ratio of total population in several age categories (under 18, 18 to 29, 30 to 64, and 65 plus) was also calculated for each block group and assigned to the EAZs. The same was done for the ratio of occupied vs. unoccupied housing units, and Hispanic or Latino origin by race. Median household income and median household income by race were also assigned to the EAZs from the ACS block group data.

In this way, the ACS variables were processed in order to provide estimates of the demographic and economic characteristics of the population in each EAZ. They are used as explanatory variables in the energy use model.

Figure 19: Median Household Income.



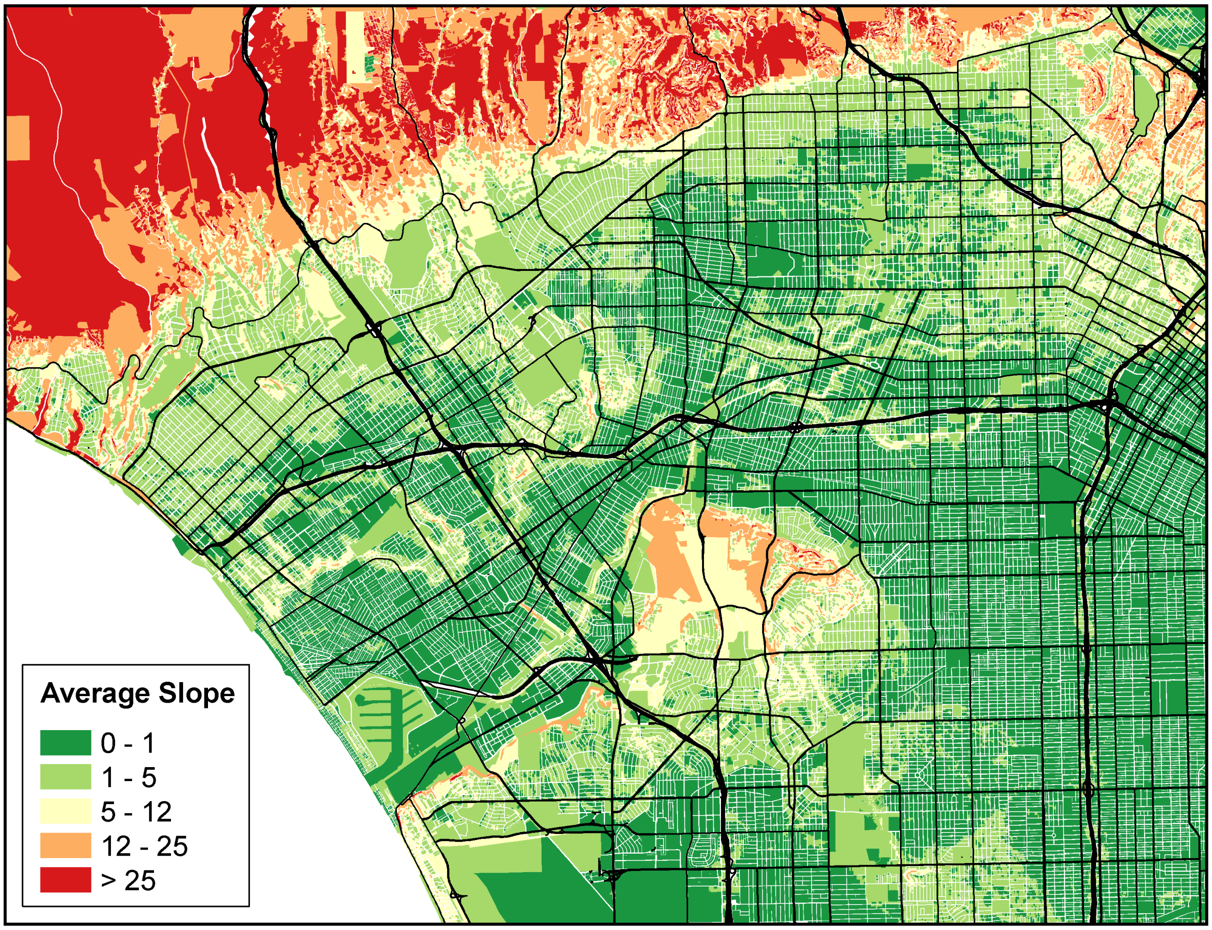
Geomorphological data

Mean slope and aspect, by parcel, were calculated for the purpose of representing solar exposure, which can impact levels of energy consumption. This information was attached to the parcels to aid in the understanding of its variation across the study area.

Geographically, the mean slope and aspect are determined for units in the Los Angeles County parcel dataset. The slope and aspect values are derived from the U.S.G.S. 10 meter Digital Elevation Model. This resolution of data is sufficient to get a reasonably accurate value for mean slope and aspect by parcel. The mean slope and aspect are two fields in the parcel database that respectively measure the average slope of a parcel (in percent slope) and average aspect of a parcel (in degrees from North).

The Los Angeles County GIS Parcel data was rasterized using a U.S.G.S. 10 meter Digital Elevation Model as a template. The slope and aspect of the DEM were then derived using standard surface analysis tools available in GIS software. Finally, the zonal statistics tool was used to calculate the mean slope and aspect of each parcel. These mean values were then joined back to the parcel database to be used as explanatory variables in the energy consumption model.

Figure 20: Average slope (degrees)



Climate data

Average values of climate variables, namely Maximum Temperature, Minimum Temperature, Precipitation, Potential Evapotranspiration, and Actual Evapotranspiration, are calculated at the level of Energy Analysis Zones (EAZs). Information on climate variables was obtained from a refined version of the Parameter-elevation Regressions on Independent Slopes Model (PRISM).

The original PRISM datasets were processed and downscaled by researchers at the Information Center for the Environment (ICE) of the University of California, Davis and the USGS for the California Energy Commission, Public Interest Energy Research (PIER) Program 2010 Vulnerability and Adaptation (V&A) Study. As part of the V&A project, several additional climate variables were derived from the downscaled PRISM temperature and precipitation data, using a regional water balance model, the Basin Characterization Model (BCM). The result is a total of 14 climatic and hydrologic variables, which are available at a resolution of 270 meters:

- Maximum Temperature
- Minimum Temperature
- Precipitation
- Potential Evapotranspiration
- Runoff, Recharge
- Climate Water Deficit
- Actual Evapotranspiration
- Sublimation
- Soil Water Storage
- Snowfall
- Snowpack
- Snowmelt
- Excess Water

The finer resolution of climate data (downscaled from original 4-KM PRISM data to 270-meter) enables us to associate climate information with County Assessor's parcel geography and EAZ geography. Out of the 14 variables, 5 were selected for use in this study. They are summarized in Table 8 below.

Table 8: Climate Variables obtained from the PRISM and BCM Models

Variable	Code	Units	Description
Maximum Temperature	*tmax*	Celsius	Maximum monthly temperature
Minimum Temperature	*tmin*	Celsius	Minimum monthly temperature
Precipitation	*ppt*	mm	Total monthly precipitation (rain or snow)

Potential Evapotranspiration*	*pet*	mm	Potential amount of water that can evaporate from the ground surface or be transpired by plants if available water is not limiting
Actual Evapotranspiration†	*aet*	mm	Actual amount of water that evaporates from the surface or is transpired by plants

* Modeled on an hourly basis from solar radiation (which is modeled using topographic shading), corrected for cloudiness, and partitioned on the basis of vegetation cover to represent bare-soil evaporation, and evapotranspiration due to vegetation

† Calculated to be the same as pet, while soil water content remains above the wilting point.

The climate data were available in raster files at a resolution of 270 meters. The five climate variables selected for use in this study were allocated to parcels by overlaying them with the County Assessor's parcel GIS dataset. The results are then summarized at the EAZ level.

Building Climate Zones

The California Energy Commission has established 16 zones in California, which are used in conjunction with California's Title 24 *Building Energy Efficiency Standards*[27] to dictate the minimum efficiency standards that are required for new construction in an area. Each zone has distinct climatic conditions which determine the types of energy efficiency features that are the most appropriate. The climate zones are based on energy use, temperature, weather and other factors, and are essentially geographic areas with similar climatic conditions. They were defined using weather station data from across the state, and are based primarily on summer and winter mean temperatures (California Energy Commission, 1995). Additionally, for ease of enforcement, they are kept fairly consistent with jurisdictional boundaries. The five climate zones found within LA County are summarized in Table 9.

[27] Also referred to as "Standard Climate Zones", climate zones are used by the CEC to dictate building energy standards. These climate zones are different from the Forecasting Climate Zones. A potential cause of confusion is that there are also 16 forecasting climate zones. The standard climate zones, used in this study, are based on climatic conditions and population centers, independent of utility service area, whereas the forecasting climate zones are based on utility electric service area boundaries and climate.

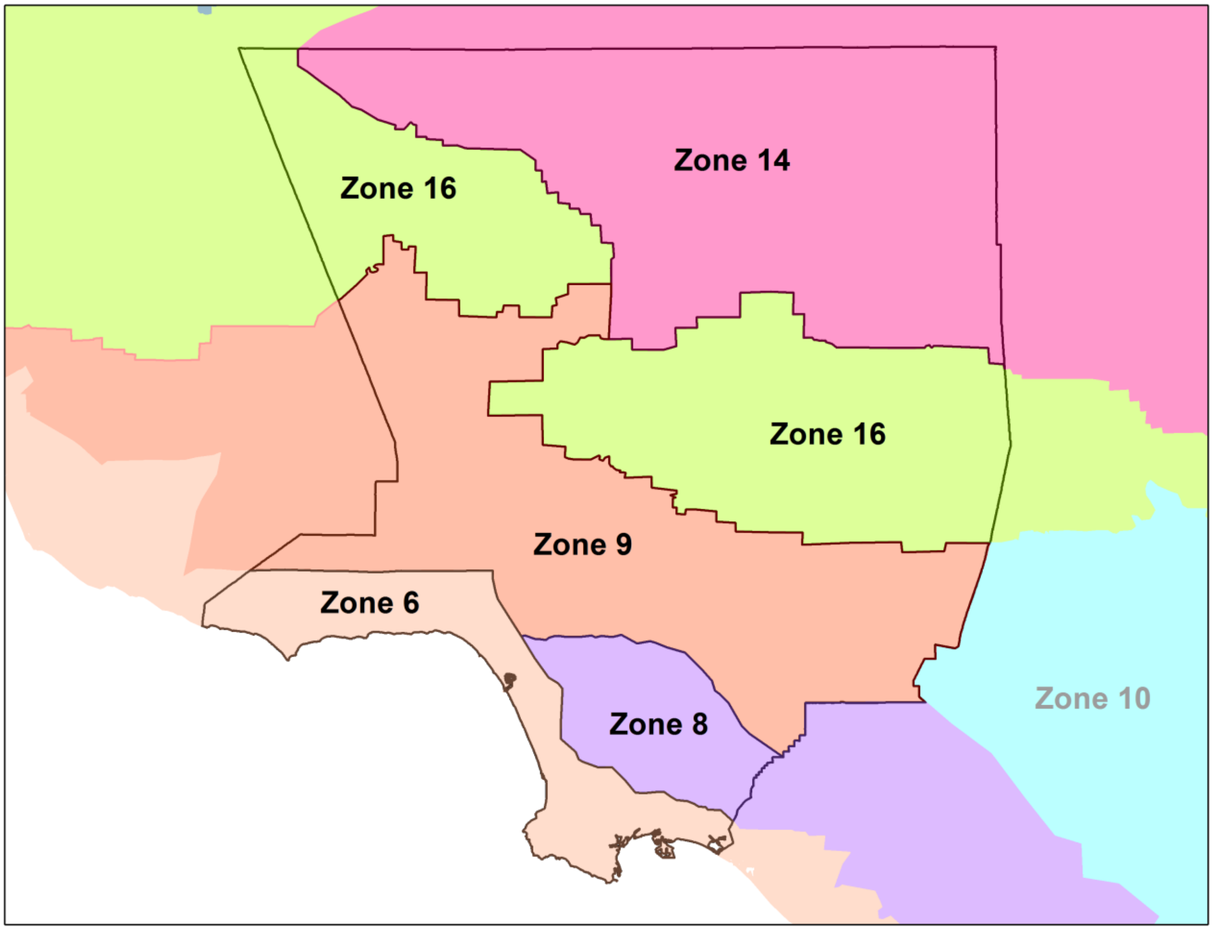
Table 9: Building Climate Zones in LA County

Zone #	Representative City	Description*
6	Los Angeles	Includes the beaches at the foot of the southern California hills, as well as several miles of inland area where hills are low or nonexistent. The Pacific Ocean is relatively warm in these latitudes and keeps the climate very mild.
8	El Toro	Inland from the coast, but still influenced by marine air. Since this zone is not directly on the coast the temperatures in the summer are warmer, and in the winter, cooler.
9	Pasadena	Both coastal and interior weather influences the Southern Californian inland valley climate zone. The inland winds bring hot and dry air, and marine air brings cool and moist air. Compared to the coast, summers are warmer and winters are cooler.
14	China Lake	Medium to high desert, the continental mass influences this interior climate more than the ocean. This zone is characterized by wide swings in temperature, both between summer and winter and between day and night
16	Mount Shasta	High, mountainous and semiarid region above 5,000 feet in elevation. The climate is mostly cold, but seasonal changes are well defined and summer temperatures can be mild.

* Climate Zone descriptions from *The Pacific Energy Center's Guide to: California Climate Zones and Bioclimatic Design (2008)*

For each of the 16 zones, the California Energy Commission has established typical weather data, prescriptive packages, and energy budgets. An energy budget is the maximum amount of energy that a building, or portion of a building, can be designed to consume per year (California Energy Commission, 2008).

Figure 21: California Building Climate Zones in Los Angeles County



CHAPTER 7: Energy Consumption Patterns for Building Operations

This section presents the results of the analysis on energy consumption for building operations in Los Angeles County that the researchers developed using the energy database built for this project. In the original research plan for the project, the researchers initially planned to develop jointly estimated models, or structural equation models, to analyze the consumption of both electricity and natural gas in buildings in Los Angeles County simultaneously. The very comprehensive energy database that was built as part of this project, developed at a very detailed level of spatial analysis (it includes almost 450,000 Energy Analysis Zones), has a great potential to investigate energy consumption in buildings in connection with the characteristics of the building stock, geographical location, climate and geomorphological variables and sociodemographic traits. However, the development of this plan was limited by the reduced availability of energy consumption data for these two energy sources. In particular, the lack of spatial overlap between the areas of service of the Los Angeles Department of Water and Power, which provided data on electricity consumption, and the Long Beach Gas and Oil, which provided data on natural gas consumption, hampered the ability to estimate models for the consumption of these two energy sources simultaneously. Still, the rich energy database built as part of this project allows a wide variety of meaningful analyses on the relationships between energy use and other variables of interest in the area of study. The remainder of this section of the report describes the analyses that were carried out in the study, through the presentation of summary descriptive statistics, first, and through the estimation of econometric models for the consumption of energy use in different categories of buildings.

It is important to note that the estimation of jointly estimated models, and of structural equation models, would have allowed the estimation of energy consumption models for each one of the two energy sources, electricity and natural gas, while accounting for the contemporary consumption of the other form of energy in each area. Natural Gas and Electricity are substitutes for some end use purposes. In particular, they are common substitutes for heating purposes in residential buildings (as well as for some purposes in non-residential buildings). According to the U.S. Department of Energy, about 71% of the California homes heat using natural gas during the cold season[28], while 22% of California homes use electricity instead. These percentages provide a clear example of the importance of controlling for the consumption of all other energy sources when studying the distribution, and relationship with other variables, of the consumption of one of these energy sources in a building.[29] Otherwise, for example, the estimation of a model to explain electricity consumption in buildings might incur

[28] Percentages for the use of natural gas for heating purposes in residential homes are lower in other parts of the country. At national level, 51% of homes are heated using natural gas, and 30% using electricity.

[29] Only 7% of California homes use other energy sources for heating purposes (mainly propane). This percentage is usually lower in highly urbanized areas, as in Los Angeles County and the use of these other sources of energy is not explicitly treated in this study.

in an "unobserved variables bias". If the unobserved variable (natural gas consumption, in this example) is highly correlated with the dependent variable (e.g. electricity consumption), the estimated coefficients from the regression model will be biased, and will significantly differ from their true values.[30]

Energy data for Los Angeles County

The limited availability of energy consumption data, which were provided only by the two utilities LADWP and LBGO, did not allow the researchers to populate the entire 450,000 EAZ energy database with energy consumption data for these two energy sources. In particular, at the time of writing this report, the lack of spatial overlap in the utility data does not allow the joint estimation of energy consumption models for electricity and natural gas in this project, as originally planned. For the reason, the researchers have developed an alternative approach that still allows exploring the relationships between energy consumption in buildings and the many variables of interest in the database.

The following subsections of the report present the results of the analysis of energy consumption in buildings that was performed where energy consumption data were available. The results are then compared to independent sources, and in particular to simpler energy consumption models that were developed using data from the RASS and CEUS energy consumption studies developed by the California Energy Commission respectively for residential and commercial buildings in California.

The authors also want to stress how, while they worked on the development of this alternative plan to investigate energy consumption in buildings, and took all necessary steps to develop alternative approaches that could reduce the disruptions caused by the limited availability of utility data on the quality of the research, they also kept working in close cooperation with the funding agency and the colleagues at UCLA on trying to obtain additional energy consumption data from the remaining utility companies in Los Angeles County. Additional efforts have been made, as previously described in this report, for this purpose. Talks are currently underway with the major IOUs in the area of study, and the energy consumption data from these utilities might become available in the near future. At the time these data will be available, it will be possible to update the current analyses, and use the full potential of the large energy database that has been created as part of this research. The data management and analytical tools that were developed as part of this project are of general validity and could be applied to the complete energy database, as soon as the new data become available, disclosing the full potential and depth of information contained in the almost 450,000 records (EAZs) database, with more than 1,000 explanatory variables, that has been created for the project.

---

[30] Accordingly, the estimated coefficients in an econometric model that does not control for an unobserved variable that is correlated, in a statistically significant way, with the explanatory and the dependent variables, are biased, and they will tend to underestimate, or overestimate, the effect of the explanatory variables on the dependent variable, depending on the sign of the correlation and the nature of the interaction among the variables.

Out of the 448,380 EAZs that compose the energy database, energy consumption data for electricity are available for 149,812 EAZs included in the LADWP area of service and natural gas data are available for 20,426 EAZs included in the LBGO area of service, for a total of 170,238 EAZs with available energy consumption data. Electricity consumption data provided by LADWP are available for all years from 2005 to 2010. Consumption data for year 2008 were selected for the analyses of this project, as most of the other variables (including the information for the building stock from the Assessor's data) are for this year. Annual electricity consumption by EAZ for all other years was also loaded into the energy database, as they provide additional information on the energy use in the EAZs.[31] The researchers used consumption data for natural gas for 2011, given the high proportion of missing records in the LBGO data for previous years, as described in Chapter 5. The EAZs with available energy data cover roughly 40% of the total number of EAZs, with about one third of the total EAZs included in the LADWP area of service, and less than 5% of the EAZs included in the LBGO subsample.

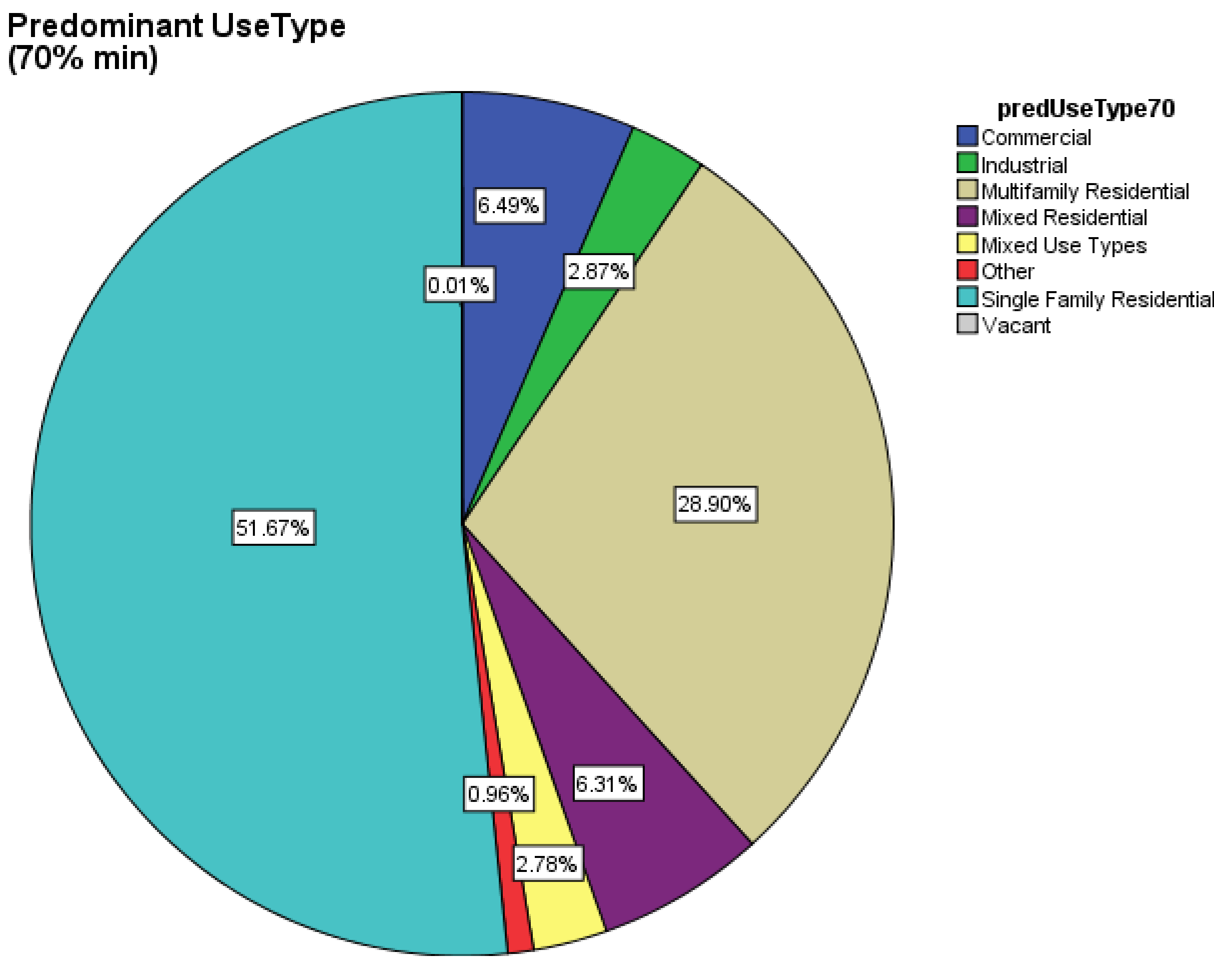
Table 10: Energy Analysis Zones by CEC climate zone in the total sample

		Frequency	Percent	Valid Percent	Cumulative Percent
Climate Zone	6	33,452	19.7	19.7	19.7
	8	28,735	16.9	16.9	36.5
	9	107,729	63.3	63.3	99.8
	16	322	.2	.2	100.0
	Total	170,238	100.0	100.0	

Table 10 reports the distribution of the number of EAZs in the total sample by CEC building climate zone. CEC Climate Zones 6, 8, 9 and 16 are represented in the sample with the available energy consumption data. Climate Zone 16 is not well represented in the sample (very small sample size, only 322 cases across the entire sample). Only climate zones 6 and 8 are present in the LBGO subsample. The additional climate zone 14, which is present, in the Northern part of the Angeles County, is not covered in the areas of service of the two utility companies and therefore it is not included in the sample.

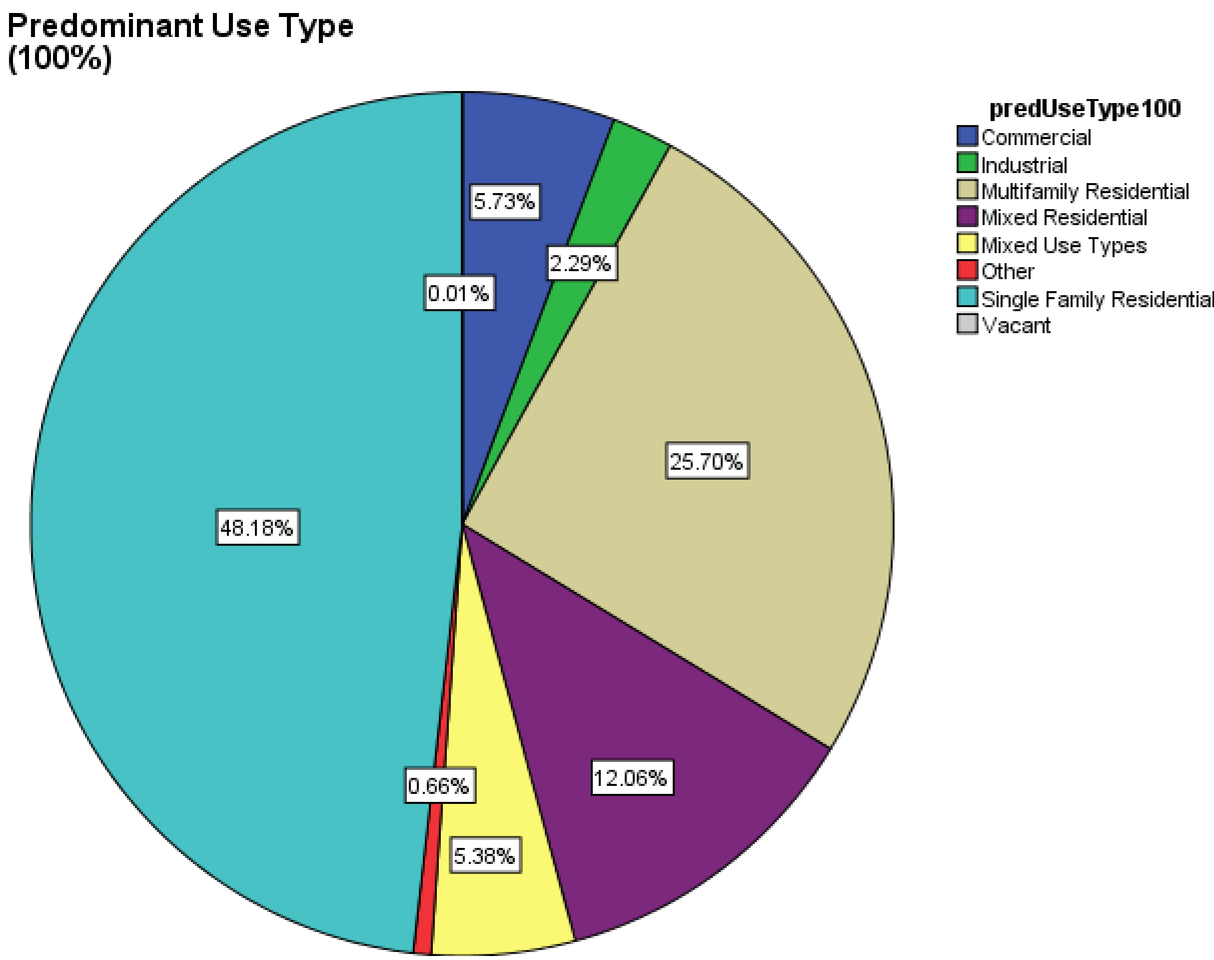
[31] Future extensions of this project could focus on the time series analysis of electricity consumption by Energy Analysis Zones, using the information contained in the energy database.

Figure 22: Predominant Use Types in Energy Analysis Zones (70% threshold, N=170,238)



The information on the building stock in Los Angeles County contained in the energy database was used to code the predominant land use observed in each Energy Analysis Zone. We controlled the predominant land use that was observed in each zone, using different threshold levels to code an EAZ as belonging to a specific predominant land use if the proportion of developed space belonging to that land use exceeded the threshold. The "Other" category includes the remaining land use types that did not fit into the residential, commercial, industrial, or vacant categories: these include institutional, government, recreational, farm and miscellaneous uses. We coded the zone in the "Mixed Use Types" category, which aggregate mixed land use areas, if no dominant use type reached the threshold level. Figure 22 shows the predominant land use categories obtained with a 70% threshold, separating zones into Single Family Residential, Multifamily Residential, Mixed (Density) Residential, Commercial, Industrial, Other (including government, institutional, etc.), Mixed Use Types (when no use type reached the 70% threshold), and Vacant (when more than 70% of the floorspace contained in the EAZ is vacant).

Figure 23: Predominant Use Types in Energy Analysis Zones (100% threshold, N=170,238)



As building types tend to mix in a city, we also controlled for other thresholds of the predominant land use, and in particular measured at 70%, 80%, 90%, 95%, 99% and 100%. The results from the different coding assumptions showed little differences between the various threshold levels. Figure 23 shows the map of predominant land use types at 100% level (all developed floorspace in an Energy Analysis Zone needs to share the same use type to be assigned to that category). In the rest of the analyses presented in this chapter, we will always refer to the 100% Predominant Use Type category, in order to identify zones in a sharper way, and reduce disturbances associated with the total amount of energy consumed for different end uses and building types.[32]

[32] At least for the LADWP dataset, energy consumption records do not distinguish between energy consumed by residential customers or by other customers in the same geographic area.

Table 11: Energy Analysis Zones by climate zone and predominant use type 100% (N=170,238)

		Climate Zone				Total
		6	8	9	16	
Pred Use Type (100%)	COM	1,676	1,578	6,245	1	9,500
	IND	568	892	2,286	0	3,746
	MFR	10,290	7,993	25,225	2	43,510
	MXR	4,098	6,824	9,814	15	20,751
	MIX	1,568	2,621	4,728	6	8,923
	OTH	187	254	657	0	1,098
	SFR	15,063	8,572	58,767	298	82,700
	VAC	2	1	7	0	10
Total		33,452	28,735	107,729	322	170,238

Table 11 shows the crosstabulation of EAZs by predominant use type (rows) and climate zone (columns) for the records with available energy consumption data. Tables 12 and 13 provide the breakdown of the numbers from Table 11 in the two different areas of service for LADWP and LBGO.

Table 12: EAZs by climate zone and predominant use type 100% (LADWP subsample, N=149,812)

		Climate Zone				Total
		6	8	9	16	
Pred Use Type (100%)	COM	1,005	1,266	6,245	1	8,517
	IND	355	827	2,286	0	3,468
	MFR	5,371	6,864	25,225	2	37,462
	MXR	1,833	5,794	9,814	15	17,456
	MIX	857	2,355	4,728	6	7,946
	OTH	111	206	657	0	974
	SFR	10,349	4,565	58,767	298	73,979
	VAC	2	1	7	0	10
Total		19,883	21,878	107,729	322	149,812

Table 13: EAZs by climate zone and predominant use type 100% (LBGO subsample, N=20,426)

		Climate Zone		Total
		6	8	
Pred Use Type (100%)	COM	671	312	983
	IND	213	65	278
	MFR	4,919	1,129	6,048
	MXR	2,265	1,030	3,295
	MIX	711	266	977
	OTH	76	48	124
	SFR	4,714	4,007	8,721
Total		13,569	6,857	20,426

The energy database contains a number of variables that can be useful to investigate energy consumption patterns in buildings. For instance, Table 14, below, shows the distribution of the age of the buildings by climate zone for EAZs with available energy consumption data.

Table 14: EAZs by building age and climate zone (N=170,238)

		Climate Zone				Total
		6	8	9	16	
Age Category	*Missing*	253	521	1,249	0	2,023
	1920 or Older	1,429	2,944	4,970	0	9,343
	1921 to 1940	5,868	9,067	19,250	19	34,204
	1941 to 1960	13,134	11,186	40,867	139	65,326
	1961 to 1980	7,474	3,569	27,697	122	38,862
	1981 to 1990	3,643	878	9,152	6	13,679
	1991 to 2000	818	365	2,397	32	3,612
	2001 to 2007	804	198	2,048	4	3,054
	2008 or Newer	29	7	99	0	135
Total		33,452	28,735	107,729	322	170,238

As the numbers from Table 12-14 demonstrate, information on buildings that were built in recent years is quite limited in the database (in particular, considering that the Assessor's data are referred to the year 2008). For this reason, for all further analyses on energy consumption, the two most "recent" categories of building age were merged into one category "Built in 2001, or Newer", which groups all buildings built in the new century. In addition, buildings built in the 1980s and 1990s were merged in a unique category. This category is of particular interest to study the impact of energy efficiency standards, which were first introduced in California with the Title 24 building standards, whose effects started to be measureable for buildings built after 1980.[33] In addition, buildings that are located in climate zone 16 are not sufficiently represented in the sample.

Residential Sector

This subsection of the report presents the results from the analysis of energy consumption for the residential sector. Energy consumption records for either electricity or natural gas are available for a total of 170,238 Energy Analysis Zones in Los Angeles County. After filtering out the records that do not contain information on residential areas and removing incomplete records and outliers from the sample, the sample that contains information on energy consumption (either for electricity or natural gas) for the residential sectors contains 132,514 EAZs. In the creation of this dataset, the researchers decided also to remove the records in which mixed land uses where observed, as the utility data, at least for the larger utility that provided data (LADWP), did not allow them to separate energy use by purpose. Also considering the small sample of EAZs with mixed land uses, and to avoid difficulties in the allocation of energy consumption to different types of customers (e.g. commercial vs. residential) in the same Energy Analysis Zone, it was decided to focus the rest of the analysis on the more homogenous data for purely residential areas. The rest of the analysis reported in this subsection of the report are therefore referred to the EAZs with predominant land use coded in the categories SFR (Single Family Residential), MFR (Multi-Family Residential) or MXR (Mixed Residential), which can include various typologies of building types belonging to either the single family or multifamily categories in the same EAZ.

Table 15: Residential EAZs by use type, building age and climate zone (N=132,514)

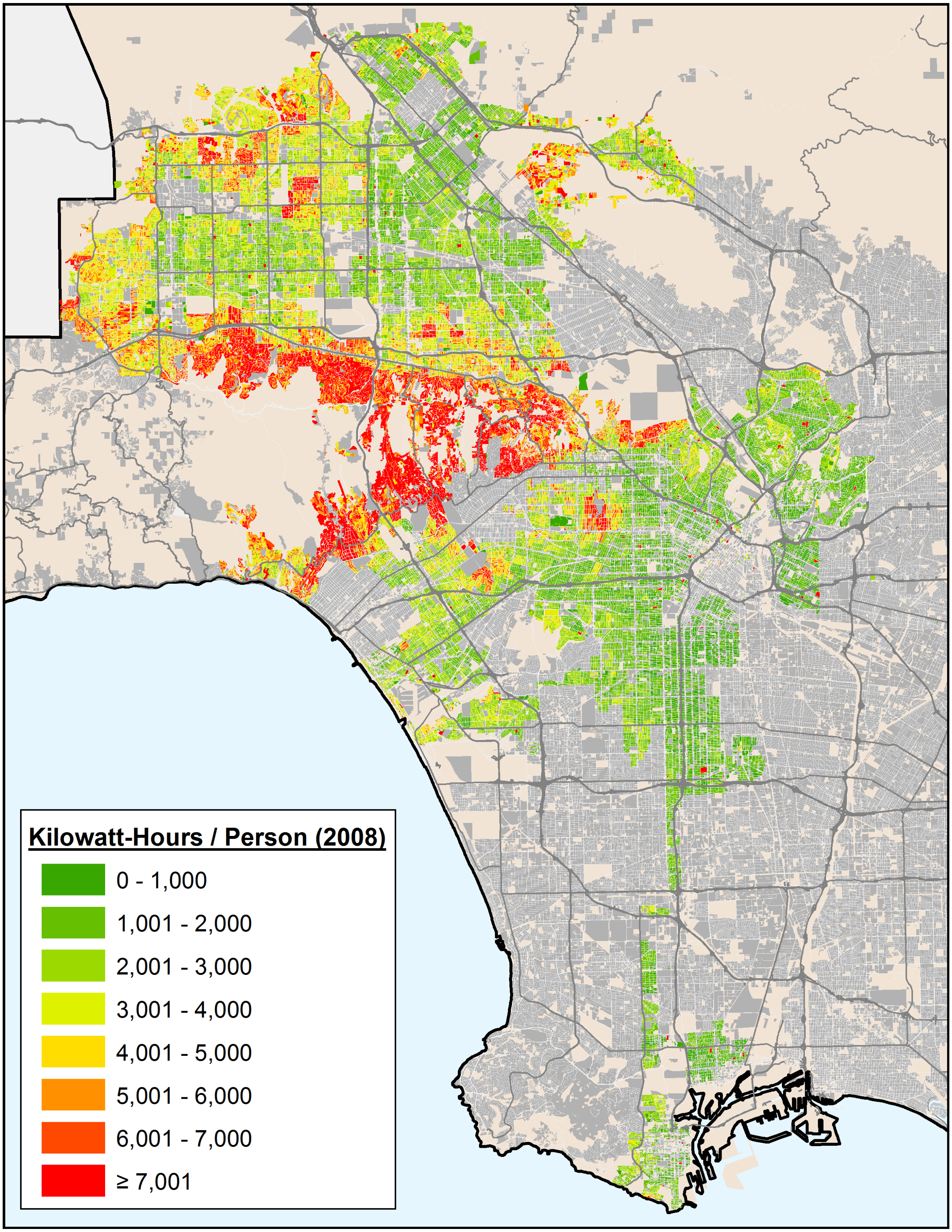
Climate Zone			predUseType100			Total
			MFR	*MXR*	*SFR*	
6	*Age Category*	*1920 or Older*	415	402	299	1,116
		1921 to 1940	1,242	1,820	1,369	4,431
		1941 to 1960	2,609	1,438	6,295	10,342
		1961 to 1980	2,683	221	2,897	5,801
		1981 to 1990	1,138	24	1,827	2,989
		1991 to 2000	87	10	490	587
		2001 or Newer	64	0	591	655
	Total		8,238	3,915	13,768	25,921
8	*Age Category*	*1920 or Older*	396	1,639	422	2,457
		1921 to 1940	1,314	3,577	2,338	7,229
		1941 to 1960	2,719	1,309	4,794	8,822

33 The first building standards were approved in California with the "Title 24" regulation approved in 1978. Several additional updates and new regulations have been introduced in the following years in the State. More information is available from http://www.energy.ca.gov/title24/standards_archive/

		1961 to 1980	1,929	101	579	2,609
		1981 to 1990	439	12	173	624
		1991 to 2000	136	4	78	218
		2001 or Newer	94	4	39	137
	Total		7,027	6,646	8,423	22,096
9	***Age Category***	*1920 or Older*	1,027	1,986	844	3,857
		1921 to 1940	4,870	4,785	5,384	15,039
		1941 to 1960	6,405	2,179	24,942	33,526
		1961 to 1980	6,083	256	15,179	21,518
		1981 to 1990	2,414	14	4,446	6,874
		1991 to 2000	418	3	1,273	1,694
		2001 or Newer	312	1	1,383	1,696
	Total		21,529	9,224	53,451	84,204
16	***Age Category***	*1920 or Older*	0	0	0	0
		1921 to 1940	1	2	15	18
		1941 to 1960	1	7	115	123
		1961 to 1980	0	1	114	115
		1981 to 1990	0	0	3	3
		1991 to 2000	0	0	30	30
		2001 or Newer	0	0	4	4
	Total		2	10	281	293
Total	***Age Category***	*1920 or Older*	1,838	4,027	1,565	7,430
		1921 to 1940	7,427	10,184	9,106	26,717
		1941 to 1960	11,734	4,933	36,146	52,813
		1961 to 1980	10,695	579	18,769	30,043
		1981 to 1990	3991	50	6449	10490
		1991 to 2000	641	17	1871	2529
		2001 or Newer	470	5	2017	2492
	Total		36796	19795	75923	132514

The aggregation of the floorspace types reported in Chapter 3 of the report is used for the aggregation of these floorspace types. We further distinguish, in the development of the energy analyses, between single-family residential units "without a pool (simply regarded as "single family housing" in the report) and single family units "with a pool". The distinction between these two categories of residential units correspond to observed trends in the energy data, as homes with a pool are usually found to consume more energy (in particularly for electricity), all else equal, than other single family homes. Table 15 reports the crosstabulations of the EAZs included in the residential sample by climate zone, age category of the building (rows) and predominant use type (columns).

Figure 24: Residential Electricity use per capita in the LADWP area of service (in 2008)



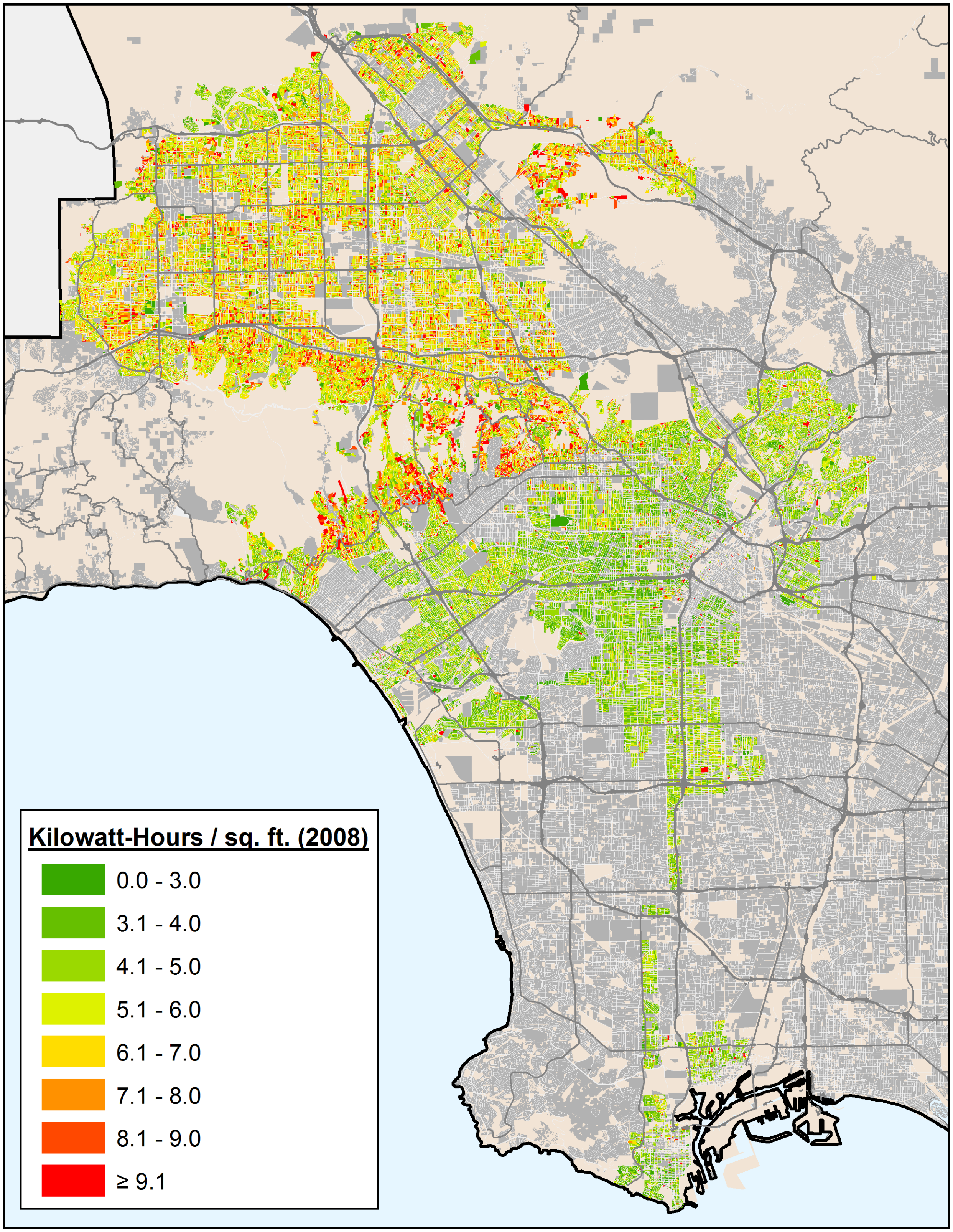
As mentioned in Chapter 4 of this report, EAZs were created as a logical aggregation of parcels and Zip+4 areas. As such, they vary in size and population (larger EAZs are associated with more complex relationships that link many Zip+4s and parcels). For this reason, the total energy consumption per EAZ is not a very interesting metric to represent spatially, as it is largely influenced by the size of each EAZ, and by the total population and amount of built floorspace in each zone.

The availability of demographic data and information on the building stock in Los Angeles County allowed computing derived measures of energy intensity in the area of study. Measures of electricity consumption as "electricity (respectively, natural gas) use per capita", or "electricity (respectively, natural gas) use per residential unit" are of larger interest, at least for the purposes of mapping energy use on a more homogenous and easily readable scale. Figure 24 shows the variation of electricity use per capita (by EAZ) in the LADWP area of service in 2008. Hilly areas of the city predominantly associated with lower density housing and larger single family homes are the areas where the highest rates of electricity consumption per capita are observed. Analogously, Figure 25 shows the variation of electricity use per sq. ft. of residential space. The different demographic distribution, average household size and size of the residential units in the various areas in the map contribute to mitigate the differences in electricity use per sq. ft. between areas. However, higher energy consumption per sq. ft. of residential space is associated with housing units located in the northern part of the map, which are located further away from the ocean and have higher thermal variation. The integration of several different sources of data in the energy database allows studying several additional relationships between electricity consumption and the possible explanatory variables contained in the dataset.

We estimated econometric models for electricity consumption based with the data obtained from the energy database. As we could not follow our original plan to develop jointly estimated models (or structural equation models) to model electricity consumption, while at the same time accounting for the consumption of natural gas (and vice versa), we opted for the estimation of cross-sectional multiple linear regression models that predict the consumption of each one of these energy sources, in separate analyses[34].

[34] As previously mentioned, the inability to account for the consumption of the other energy source during the estimation of the consumption models might be a source of unobserved variable biases in the estimated coefficients. We will later discuss this topic in more details when comparing the results from the estimated energy consumption models with other sources of data. Any eventual bias in the data could be removed through the access to more complete energy consumption data for both electricity and natural gas data in the area of study, and re-estimating models of energy consumption with the energy database and the analytical tools that were developed for this project.

Figure 25: Electricity use per square foot of residential space in the LADWP area of service (in 2008)



Several different functional forms where tested to create robust, and meaningful, specifications that can explain electricity consumption in the LADWP area of service. We first estimated models to predict electricity use per capita. Thanks to the large amount of variables in the dataset, we were able to test several different model specifications with or without a constant term and combinations of explanatory variables. We tested various specifications for this electricity consumption model, with or without a constant term. Table 16 summarizes the results from the model of electricity use per capita with a constant term, which was selected as best fitting to describe the electricity consumption per capita.

Table 16: Linear regression model for residential electricity use per capita (LADWP area)

Variable	Unstandardized Coefficients	Standardized Coefficients	p-value
Constant	2,490.731		<.001
Population per unit	-446.581	*-.368*	<.001
Median Income	.017	*.350*	<.001
Avg. Slope	8.737	*.024*	<.001
Avg. SF unit size (sq. ft.)	.227	*.125*	<.001
Avg. SF unit size, with pool (sq. ft.)	.450	*.273*	<.001
Avg. MF unit size (sq. ft.)	-.087	*-.025*	<.001
Year Built (1981 to 2000)	179.714	*.031*	<.001
Year Built (2001 or newer)	109.408	*.009*	<.001
Climate Zone 6	-589.561	*-.116*	<.001
Climate Zone 8	-337.045	*-.068*	<.001
Aspect - South	-179.678	*-.051*	<.001
Aspect - North	42.502	*.005*	.008
Sample Size (N)	*115,987*		
R Square	*.559*		

Dependant Variable: *Annual Electricity Consumption per capita* (KWh)

The model for electricity user per capita has a rather good goodness of fit (R-square = 0.559) and it includes several variables of interest. Electricity use per capita tends to increase with income[35],

[35] Higher income individuals tend to leave in larger, more comfortable houses, have smaller household sizes, and tend to make higher use of appliances and other energy-intensive devices. Even after

while it decreases with the number of people that live in the household: as expected, all else equal, individuals that live in larger groups in the same housing unit tend to consume less electricity per capita. Electricity use per capita tends to increase also with the slope of the area where the EAZ is built: this variable is probably a proxy also for the geographical location of the household. In the area of study, larger, more luxury houses are generally located on the steeper areas of the hills surrounding the city.

Electricity consumption is lower for individuals that live in the Climate Zones 6 and 8 (comparing to Climate Zones 9 and 16 as reference in the dataset). These areas, located closer to the ocean, benefit from the proximity to this vast body of water, and consequentially usually require lower amounts of energy for heating during winter and for cooling during the summer season. A similar effect, although smaller in terms of absolute impact on energy consumption, is associated with the aspect of the EAZ where the individual lives. Areas that face the North tend to have slightly higher electricity consumption per capita, while the consumption is lower for areas facing the South (compared to areas facing East or West, used as reference).[36]

The age of the building is a significant predictor of electricity use in the building: in particular, individuals that live in newer buildings (built after 1980) tend to consume more energy than those that live in older buildings, probably as an effect of larger use of modern appliances and increased use of A/C and other facilities that overcompensate for the increased efficiency of the building. Several different specifications were tried to model the influence of the age of the building on electricity use. Among the main findings, a reduction in the electricity use per capita is registered for individuals that live in recently built (or renovated, after 2000) homes. Overall, electricity consumption per capita is higher in homes built between 1981 and 2000 than in older buildings. It then declines for individuals that live in newer buildings (built or renovated after 2000) if compared to the buildings built in the previous era (1981-2000). This effect can be probably explained by the effects of the improvements in the standards for the energy efficiency in buildings and by the increased proportion of retrofits in the existing building stock.[37]

As expected, the impact of the size of the housing units is an important predictor of the electricity consumption per capita. Individuals that live in larger homes tend to use more electricity, and the effect is amplified for individuals that live in a house with a pool: the estimated coefficient for housing unit size "with pool" is always higher than the one for single family home without a pool in all models that were estimated. This effect sums with another

---

accounting separately for the household size and the size and type of housing unit, the income variable is still statistically significant and has a positive estimated coefficient.

[36] Please note that apart from the different exposure to the sunlight, the different aspect of an area is also a proxy for the location of the building in the county, as in particular in the LADWP area of service, a limited number of residential lots are built in areas that predominantly face the North direction and most of them are predominantly located in specific part of the area of study.

[37] The variable for age of the building is an average across the entire Energy Analysis Zone and it includes newly built buildings but also major renovation of previously existing buildings.

relationship observed in the dataset, which is that the average size of a home with pool is usually larger than the average size of a house without a pool in the area of study. This effect increases even more the difference in the electricity consumption between individuals that tend to live in a house with a pool and those that do not.

The negative coefficient for the average size of a multifamily home might look rather counterintuitive, at first, and it greatly differs from the alternative model specifications estimated without a constant term. However, the negative coefficient for the size of the multifamily family housing unit might be explained by the need to compensate the excessive value of the constant term for individuals that live in multifamily houses. For this reason, a modified model was estimated, including the possibility for the constant term to assume a different value for the individuals that live in Energy Analysis Zones dominated by multifamily housing units. The results of the modified model are summarized in Table 17. The significance and signs of all the variables in the model resemble the values from the previous model. However, the presence of the additional constant modifier for the Multi-Family units reduces the size of the constant to 2,177.31 Kwh, with a positive, and statistically different from zero, coefficient for the size of the multifamily home in which the individual lives.

Table 17: Linear regression model for residential electricity use per capita (LADWP area, with modified constant for Multi-Family housing units)

Variable	Unstandardized Coefficients	Standardized Coefficients	p-value
Constant	3,104.219		<.001
MF Constant (modifier)	-926.914	*-.269*	<.001
Population per unit	-498.963	*-.411*	<.001
Median Income	.015	*.319*	<.001
Avg. Slope	10.143	*.028*	<.001
Avg. SF unit size (sq. ft.)	.061	*.033*	<.001
Avg. SF unit size, with pool (sq. ft.)	.431	*.262*	<.001
Avg. MF unit size (sq. ft.)	.343	*.100*	<.001
Year Built (1981 to 2000)	256.540	*.044*	<.001
Year Built (2001 or newer)	147.441	*.012*	<.001
Climate Zone 6	-566.260	*-.112*	<.001
Climate Zone 8	-334.472	*-.067*	<.001
Aspect - South	-153.161	*-.043*	<.001

Aspect - North	45.316	*.006*	.004
Sample Size (N)	*115,987*		
R Square	*.578*		

Dependant Variable: *Annual Electricity Consumption per capita* (KWh)

One of the purposes of this study is to develop models of energy consumption that can be easily applied in a modeling approach that simulates the development of buildings in a complex area such as Los Angeles County. For this reason, we developed some simplified models to model energy consumption in Los Angeles County. We will present them starting from the simplest (and more parsimonious) model. In this part of the analysis of electricity consumption, we focused on both models that include a constant term and models that do not. These models are built for the purpose of developing estimates for energy consumption of individual buildings and are based on a number of inputs. This type of model can generate energy estimates that can be applied to a building inventory to project energy use in our area of study.

Table 18 summarizes the results of the estimation of a simple model of energy consumption that simply estimates the consumption of electricity depending on the size of the residential unit. The model differentiates the weight that a square foot of each of the three residential floorspace types can have on residential electricity consumption through the adoption of different slopes in the model for the three residential floorspace types. The model is estimated without a constant term, in order to simplify its application to an energy assessment of the building stock through the estimation of a unique coefficient for each of the three floorspace types that can be easily applied to the building inventory developed for Los Angeles County.

Table 18: Regression model for residential electricity consumption in residential EAZs (LADWP area)

Variable	Unstandardized Coefficients	p-value
Total sq. ft. SF housing	4.395	<.001
Total sq. ft. SF housing, with pool	5.792	<.001
Total sq. ft. MF housing	4.379	<.001
Sample Size (N)	*115,987*	
R Square	*.791*	

Dependant Variable: *Total Annual Electricity Consumption in EAZ* (KWh/EAZ). Regression model through the origin.

The interpretation of the unstandardized coefficients from this model is very simple, as each estimated coefficient represents the estimated electricity consumption of one square foot of that floorspace type (in KWh/sq. ft.). We want to call attention to the goodness of fit measure for this model (the R-squared), in this case, refers to a model through the origin. This measure of the goodness of fit cannot be directly compared to the R-square measure of the goodness of fit for a model with an intercept (as the models that presented later in this section).

We developed an alternative model using a stratified sample, created through the separation of the different residential densities in the sample. The purpose of this model is to isolate the effects of the different types of residential floorspace types on electricity consumption in buildings reducing the perturbation introduced by the presence of multiple types of floorspace types in the same EAZ.[38] Table 19 shows the results of the estimation of a linear regression model estimated using the stratified sample composed of more homogenous residential areas. In this sample, all records from EAZs with mixed residential building types were filtered out.

Table 19: Regression model for residential electricity consumption in residential EAZs (LADWP area, separated residential land uses)

Variable	Unstandardized Coefficients	p-value
Total sq. ft. SF housing	4.261	<.001
Total sq. ft. SF housing, with pool	5.244	<.001
Total sq. ft. MF housing	4.361	<.001
Sample Size (N)	*71,371*	
R Square	*.738*	

Dependant Variable: *Total Annual Electricity Consumption in EAZ* (KWh/EAZ). Regression model through the origin.

Table 20 presents a modified version of the previous regression model, still based on the same stratified sample, which accounts for the location of EAZs in different climate zones. This model estimates the annual electricity consumption (in KWh/residential unit) using the average square footage of the units located in each EAZ as explanatory variables. Given the distribution of the EAZs by climate zone in the sample, it is not possible to estimate a separate coefficient for the climate zone 16 (for which the sample size is particularly small). Similarly, climate zones 6 and 8, which showed a similar behavior in terms of energy consumption in all modeling analyses, are grouped in a unique climate area for the purposes of this analysis (this ensures larger sample sizes for all subsamples). For this reason, the model shown in Table 21 is based on the estimation of different coefficients for the three floorspace types for residential units respectively located in the CEC Title 24 climate zones 6 or 8, and for those located in the climate zones 9 or 16.

---

[38] The process of aggregation of the Energy Analysis Zones, imposed by the need to treat the information provided by the utility companies at the Zip+4 level imposes some averaging of the variables for the building stock across an EAZ. This might reduce the explanatory power of the variables in the model. The separation of the land uses in this model tried to isolate the effects of the different building types on energy consumption without these confounding factors.

Table 20: Regression model for electricity consumption in residential areas (LADWP area, separated residential land uses, with climate zones)

Variable	Unstandardized Coefficients	p-value
Sq. ft. SF housing (CZ6 or CZ8)	3.886	<.001
Sq. ft. SF housing, with pool (CZ6 or CZ8)	3.969	<.001
Sq. ft. MF housing (CZ6 or CZ8)	3.950	<.001
Sq. ft. SF housing (CZ9 or CZ16)	4.574	<.001
Sq. ft. SF housing, with pool (CZ9 or CZ16)	5.069	<.001
Sq. ft. MF housing (CZ9 or CZ16)	4.522	<.001
Sample Size (N)	*71,371*	
R Square	*.870*	

Dependant Variable: *Annual Electricity Consumption per residential unit* (KWh/ residential unit). Regression model through the origin.

The models that have been so far presented in this section are estimated using the data contained in the energy database developed as part of this project with a regression model through the origin (without intercept, to simplify the computation of the energy consumption per square foot of residential unit). However, one problem with this kind of models is that forcing the intercept to a value of zero might reduce the ability of the estimated coefficients to correctly explain the variance in the dependant variable.

For this reason, Table 21 presents the results of a modified version of the previous model that includes an intercept and that explains the energy consumption in buildings depending on the amount of square feet of the various typologies of residential buildings and their location in the different climate zones. Similarly to what was done with the regression model that explains electricity consumption per capita, this model also allows the constant term to vary for single family vs. multifamily housing units, using a MF constant modifier in the regression.

This model has a high goodness of fit, and it is able to estimate the electricity consumption per household using a constant term (which measures the common consumption to all households, independently from the square footage) and an additional term proportional to the size of the residential unit. The coefficients (the "slope" in the regression model) are allowed to vary for the various residential types and for the impact of the climate zones on electricity consumption. In addition, the model makes a rather realistic representation of electricity consumption in buildings, with a different constant term for residential units in multifamily buildings, which have different electricity consumption profiles than single family homes.

Table 21: Regression model for residential electricity consumption in residential EAZs (LADWP area, separated residential land uses, with climate zones)

Variable	Unstandardized Coefficients	p-value
Constant	4,163.054	<.001
MF Constant (modifier)	-3,865.776	<.001
Sq. ft. SF housing (CZ6 or CZ8)	1.468	<.001
Sq. ft. SF housing, with pool (CZ6 or CZ8)	2.675	<.001
Sq. ft. MF housing (CZ6 or CZ8)	3.647	<.001
Sq. ft. SF housing (CZ9 or CZ16)	2.266	<.001
Sq. ft. SF housing, with pool (CZ9 or CZ16)	3.612	<.001
Sq. ft. MF housing (CZ9 or CZ16)	4.233	<.001
Sample Size (N)	*71,371*	
R Square	*.611*	

Dependant Variable: *Annual Electricity Consumption per residential unit* (KWh/residential unit).

The dataset that is used in these analyses has one severe limitation due to the limited information released by the utility companies in Los Angeles County. Thus, it cannot account for the consumption of natural gas in the estimation of the electricity models presented above. For this reason, the researchers also accessed alternative datasets that can provide information on the relationships between electricity and natural gas consumption and they compared the results from this present analysis to other independent sources, in order to make a validity check of the estimated models of energy consumption. Table 22 shows a simple model of electricity consumption that was developed using the California Energy Commission RASS (Residential Appliance Saturation Survey) dataset for the five largest utilities in the State. The model is developed using a similar approach to the previous models, and it contains a constant term (that is also in this case allowed to differ for multifamily housing units).

For the way the RASS survey data are structured, we cannot differentiate between houses with a pool and without. The simple regression model that is estimated with the RASS data has a lower explanatory power, and different estimated coefficients than the model that is presented in the previous table. Besides, it has the ability to control for the natural gas consumption, as the model is built using a dataset that contains information on both electricity and natural gas. We will return to the joint estimation of the electricity and natural gas consumption towards the end of this section.[39]

Table 22: Regression model for residential electricity consumption from RASS data

Variable	Unstandardized Coefficients	p-value
Constant	2,536.964	<.001
MF Constant (modifier)	-615.481	<.001
Sq. ft. SF housing unit	2.229	<.001
Sq. ft. MF housing unit	1.529	<.001
Sample Size (N)	*13,826*	
R Square	*.299*	

Dependant Variable: *Total Annual Electricity Consumption per household* (Kwh/residential unit).

In order to compare the results from the model estimated with the energy database built with this project and the simple model developed using the RASS data, we used the information related to the average characteristics of residential units in the area of study to build an electricity consumption function for the "average" household in the study. Table 23 contains the results of this comparison. The estimated coefficients from the model are multiplied by the average household sizes for the various categories of residential units respectively in the LADWP and in the RASS dataset.

The estimated electricity consumption predicted by the model for the "average" housing unit in the LADWP dataset is slightly higher than for the correspondent housing unit in the RASS dataset. However, this is probably reasonable for the Los Angeles area, if compared to the rest of the state (RASS data are collected for the entire State of California). Overall, the estimations of electricity consumption that was estimated in the study seem reasonable. The model is also able to account for the variation of energy consumption by climate zone, with higher levels of electricity consumption forecasted for housing units located in climate zones 9 and 16, which are located further away from the ocean. This appears to be reasonable, in consideration of the higher use of energy, for cooling and heating purposes respectively during the summer and winter seasons.

[39] Linear regression models cannot account for an additional unobserved variable that is not included in the regression. This topic will be discussed later through the estimation of a simple structural equation model for the estimation of electricity and natural gas consumption.

Table 23: Estimated residential electricity consumption per housing unit[40]

Energy database	Constant	MF Constant	KWh/sq. ft.	Avg. Sq. ft./Unit	Total KWh/Unit
SF housing (CZ6 or CZ8)	4,163.054		1.468	1,638.74	**6,568.724**
SF housing, with pool (CZ6 or CZ8)	4,163.054		2.675	2,095.46	**9,768.426**
MF housing (CZ6 or CZ8)	4,163.054	-3,865.776	3.647	878.07	**3,499.614**
SF housing (CZ9 or CZ16)	4,163.054		2.266	1,638.74	**7,876.429**
SF housing, with pool (CZ9 or CZ16)	4,163.054		3.612	2,095.46	**11,731.877**
MF housing (CZ9 or CZ16)	4,163.054	-3,865.776	4.233	878.07	**4,014.165**
RASS data	**Constant**	**MF Constant**	**KWh/Sq. ft.**	**Avg. Sq. ft./Unit**	**Total KWh/Unit**
SF housing	2,536.964		2.229	1,746.84	**6430.670**
MF housing	2,536.964	-615.481	1.529	951.77	**3376.739**

Note: Comparison based on estimated coefficients from Table 21 (LADWP area) and Table 22 (RASS data).

In a similar way to what was developed for the estimation of electricity use in residential buildings, we also estimated similar linear regression models to predict the natural gas consumption in residential buildings in Los Angeles County. Table 24 shows the estimated coefficients for a linear regression model for natural gas use per capita. The R-square (measure of goodness of fit) for this model is 0.153, lower than in the model for the electricity

[40] Electricity consumption estimates from the model seem to be consistent with data from other sources, too, and in particular with the average annual amount of electricity consumed per household in the LADWP area, which is estimated at the level of about 6,500 KWh (in 2008) in the adopted forecasts for the California Energy Demand 2010-2020 developed by the California Energy Commission (Kavalec and Gorin, 2009).

consumption per capita. This is in line with the lower explanatory power of all the models estimated with the natural gas dataset in this study, which has smaller sample size, and is referred to a smaller geographic region.[41]

Table 24: Linear regression model for residential natural gas use per capita (LBGO area, with modified constant for Multi-Family housing units)

Variable	Unstandardized Coefficients	Standardized Coefficients	p-value
Constant	181.965		<.001
MF Constant (modifier)	-20.655	*-.085*	<.001
Population per unit	-30.092	*-.318*	<.001
Median Income	.000	*.090*	<.001
Sq. Ft. SF unit (with/without pool) size	.028	*.194*	<.001
Sq. Ft. MF unit size	.009	*.034*	.001
Year Built (1981 to 2000)	-21.266	*-.054*	<.001
Year Built (2001 or newer)	-36.504	*-.029*	<.001
Sample Size (N)	*16,527*		
R Square	*.153*		

Dependant Variable: *Annual Natural Gas Consumption per capita* (Therms)

In the preferred (best) model for natural gas consumption per capita, the difference in the impact of the square footage of single family homes "with a pool" and "without a pool" on natural gas consumption per capita is not significant. For this reason, given also the limited sample size for the LBGO area of service (the only area where natural gas consumption data are available for this study), the researchers chose a more parsimonious specification for this model. A unique term estimates the impact of the size (in square feet) of the residential units (with or without the pool) on the consumption of natural gas per capita in this model.[42] Natural gas

[41] Limited variance is observed for many variables in this smaller dataset, e.g. for the size of multi-family residential units and the location of EAZs in different climate zones. In addition, complete records for the annual consumption of natural gas were provided by the utility company (see Chapter 5) only for 2011. Later in the report, we will discuss the possible effects of this temporal mismatch (with the other variables in the database) on the estimated coefficients, and on the goodness of fit of the models of natural gas consumption.

[42] Even if the estimated coefficients for the impact of a square foot of residential unit on natural gas consumption per capita are the same for houses with and without a pool, in practical applications, higher levels of consumption of natural gas per capita are observed in houses with a pool, as an effect of the larger average size of these housing units (which affects the distribution of this explanatory variable).

consumption per capita is usually lower in multifamily homes, and it decreases with the household size (i.e. "population per unit" in the model). As expected, average natural gas consumption increases with an increase in income (as confirmed by other studies, as an effect of the different characteristics of the buildings, lifestyles and energy use).

Finally, an important role is associated with the age of the building: average natural gas consumption per capita is lower for individuals that live in newer buildings. This confirms the effects of the energy efficiency standards: Title 24, approved in 1978, is responsible for the reduction in the consumption of natural gas per capita for individuals that live in buildings built in 1980s and 1990s (compared to older buildings). The reduction becomes even larger for individuals that live in even more recent buildings, built or largely renovated after 2000, when efficiency standards become even more stringent.[43] The impact of climate zones on the consumption of natural gas per capita is not found to be statistically significant in this model. But it is important to note that the data from LBGO cover only two (and rather similar) climate zones, respectively number 6 and 8, and therefore the possibility to study the impact of climate zones on energy consumption with the available data is limited.

Following the same approach used for the electricity consumption, the researchers also estimated some more simplified models to predict the natural gas consumption in a residential unit in the area of study, which can be of interest for the assessment of energy consumption for the building stock in the area of study. Table 25 shows the results of the estimation of the final model that estimates natural gas consumption for a residential unit in the LBGO area of service. The model is estimated with an intercept (which can differ for residential units located in single family or multifamily buildings). This simplified model of natural gas consumption predicts the natural gas consumption in a residential unit as a linear function (including a constant term) of the residential unit size. The model predicts slightly different natural gas consumption for single family homes located respectively in climate zone 6 and climate zone 8. The total natural gas consumption of single family homes in climate zone 8 tends to increase at a slightly higher rate with size, than for houses located in climate zone 6. As expected, natural gas consumption in multifamily housing units is found to be lower than in single family homes. Interestingly, the estimated coefficients for the size of the residential units located in multifamily buildings are not statistically different from zero. Natural gas consumption does not vary in statistically significant way with the size of the unit in this area of service (however, not a very large variation in the square footage for this type of residential units is observed in the sample, making the estimation of this coefficient more difficult).

[43] Additional policies promoted by the utility companies and the increased awareness on energy efficiency are also probably co-responsible for the increased efficiency of buildings built in more recent years.

Table 25: Regression model for residential natural gas consumption in residential areas (LBGO area, separated residential land uses, with climate zones)

Variable	Unstandardized Coefficients	p-value
Constant	278.229	<.001
MF Constant (modifier)	-58.756	<.001
Sq. Ft. SF unit (with/without pool) SF housing (CZ6)	.109	<.001
Sq. Ft. SF unit (with/without pool) SF housing (CZ8)	.115	<.001
Sample Size (N)	*16,527*	
R Square	*.204*	

Dependent Variable: *Annual Natural Gas Consumption per residential unit* (Therms/residential unit).

Also for the natural gas consumption, the researchers developed a simple model of energy consumption using the data from the CEC RASS survey, which can be used for comparison of the results from the model estimated with the energy database for this project. Table 26 summarizes the regression model for natural gas consumption estimated using the RASS data.

Table 26: Regression model for residential natural gas consumption from RASS data

Variable	Unstandardized Coefficients	p-value
Constant	283.924	<.001
MF Constant (modifier)	-143.301	<.001
sq. ft. SF housing unit	.128	<.001
sq. ft. MF housing unit	.093	<.001
Sample Size (N)	*13,826*	
R Square	*.266*	

Dependent Variable: *Total Annual Natural Gas Consumption per household* (Kwh/residential unit).

The RASS model provides a useful comparison to validate the results from the analysis of this project. Table 27 provides some comparison for the natural gas consumption predicted by the model from Table 25 and the RASS model from Table 26 for some categories of housing units with "average" characteristics in the respective samples. Total natural gas consumption is not very different for the two models, and is consistent with expectations: consumption of larger homes with a pool is higher than for single family homes without a pool, and residential units in multifamily buildings tend to consume less natural gas than single family homes in the same climate zone.

Table 27: Estimated natural gas consumption per housing unit

Energy database	Constant	MF Constant	Therms/ Sq. ft.	Avg. Sq. ft./Unit	Total KWh/Unit
SF housing (CZ6)	278.229		.109	1477.50	**439.277**
SF housing, with pool (CZ6)	278.229		.109	1883.56	**483.538**
MF housing (CZ6)	278.229	-58.756		812.437	**219.473**
SF housing (CZ8)	278.229			1477.50	**448.142**
SF housing, with pool (CZ8)	278.229		.115	1883.56	**494.839**
MF housing (CZ8)	278.229	-58.756	.115	812.437	**219.473**
RASS data	**Constant**	**MF Constant**	**Therms/ Sq. ft.**	**Avg. Sq. ft./Unit**	**Total KWh/Unit**
SF housing	283.924		.128	1,746.84	**507.519**
MF housing	283.924	-143.301	.093	951.77	**229.138**

Note: Comparison based on estimated coefficients from Table 25 (LBGO area) and Table 26 (RASS data).

Overall natural gas consumption in the LBGO housing units tend to be slightly lower than in the housing units predicted by the RASS model, but this result is somewhat expected if considering the geographical areas the two datasets refer to. RASS data contain households in the entire State of California, and it is expected that natural gas consumption in residential units in Los Angeles County (and in general in Southern California) tends to be lower than in the rest of the State, as one of the primary end uses for natural gas is heating residential housing units.

As the RASS data allow studying both electricity and natural gas consumption simultaneously, we also developed a simple model of jointly estimated regression equations that analyzes the relationships between square footage of the two housing types (SF vs. MF) and energy consumption. Figure 26 shows the modeled relationships among the studied variables in this model: two equations respectively predict electricity and natural gas consumption of each household. The explanatory variables, as in the previous examples using the RASS data, are the square footage of the two types of housing units, single-family residential and multifamily residential. The errors of the two dependent variables, electricity consumption and natural gas consumption, are allowed to be correlated.

Figure 26: Jointly estimated models of household energy consumption from the RASS data

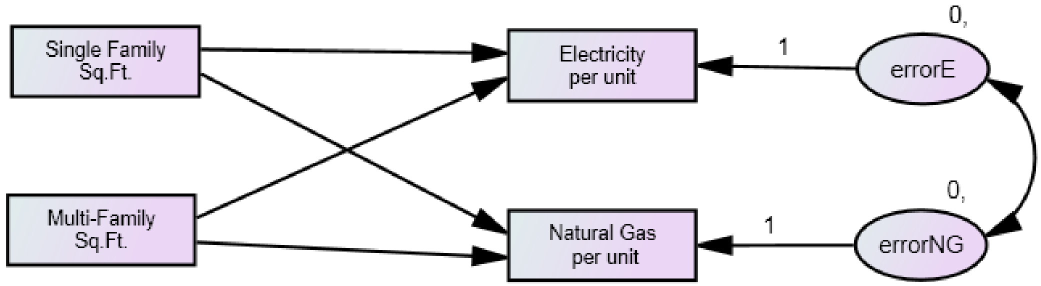
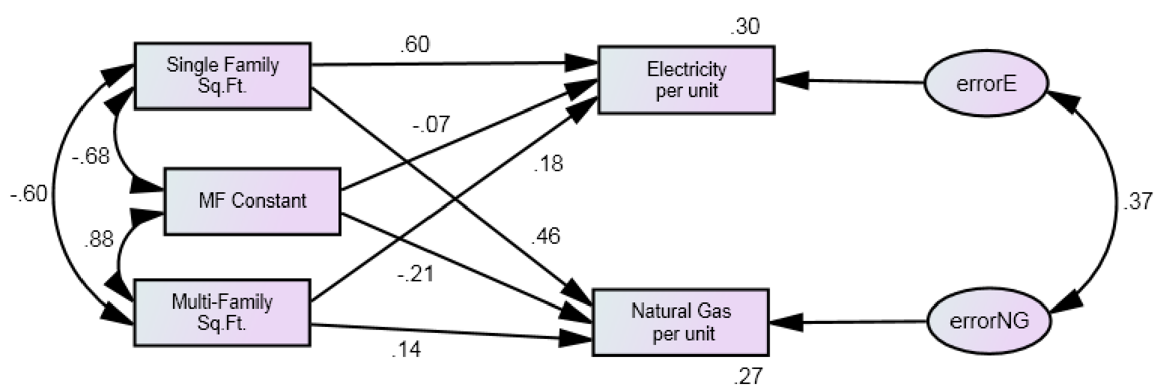


Figure 27 shows the final specification for this model. Standardized coefficients from the estimation of the model are reported in the figure. The inclusion of the MF constant allows the constant term to differ for Multifamily housing, in a similar way to the linear regression models presented earlier in this section. The estimation of this simple model produces coefficients that are equal in sign and magnitude to those from the estimation of the linear regression models in Table 22 and Table 26 and used for the computations in Table 28. The standard errors of the coefficients, however, are smaller, as an effect of the jointly estimated model that is able to deal with some of the violation of the assumptions of the multiple linear regression model and that leads to the estimation of more efficient estimators.

Figure 27: Standardized estimated coefficients for the jointly estimated models of household energy consumption from the RASS data



The model with the RASS data accounts for the correlation among electricity and natural gas consumption levels. The correlation is statistically significant and positive (0.37). The result, which might be counterintuitive, is actually explained by the relationships among the consumption patterns of these two types of energy sources. Even if natural gas and electricity can be important substitutes for some energy end uses (e.g. heating in residential units), the total consumption of the two energy sources will tend to be a function of the size of the housing units and the characteristics of the household. Larger households are associated with larger consumption of both electricity and natural gas. Therefore, the correlation among the total consumption levels of the two energy sources tends to be positive, as well as the regression weights for the total amount of square footage for both types of residential units.

The comparison of the energy models estimated with the energy database built for this project with the simpler models built from the RASS data is useful to provide some benchmark to check the reasonability of the estimated consumption patterns analyzed in this study. The results from our study were also compared to other data sources and studies in the literature, for instance on the annual consumption of electricity per capita in the residential sector in the State of California[44] (Kandel *et al*., 2008). Also this comparison did not highlight any significant issues or concerns on the validity of the results that were estimated as part of the current research and its application for the estimation of energy use in the residential sector in Los Angeles County.

Non-Residential Sector

Energy consumption in buildings other than residential dwellings constitutes an important fraction of the total energy use in urban areas. The service sector includes all commercial and public buildings: under this definition, many different types of buildings are grouped, including retail and stores, offices, schools, restaurants, hotels, hospitals, museums, etc. The non-residential service sector accounts for a wide variety of uses and energy services. Among the most relevant in terms of the energy consumption are heating, ventilation and air conditioning [HVAC], domestic hot water, lighting, refrigeration, food preparation, etc.[45] In non-domestic buildings, the type of use and activities make a huge impact on the quality and quantity of energy services needed. Office and retail are among the most energy intensive building types, typically accounting for over 50% of the total energy consumption for non-residential buildings. Hotels and restaurants, hospitals and schools are other building types that consume a large share of energy in the service sector.

In this subsection of the report, we will discuss the analysis of energy consumption in non-residential buildings that was developed in the project. The analysis of energy consumption in non-residential buildings focused on the floorspace types presented in Chapter 3. Differently

[44] In the cited study, electricity consumption per capita in the residential sector in California was estimated to be 2,369 KWh/person, on average. The study provides information on the variation of electricity consumption per capita depending on income, household size, number of cooling/heating days, and electricity price (this last variable is not analyzed in our research, as it is a constant in a cross-sectional study) through a time series analysis of energy consumption in California. The results from the study are consistent with the results from our research.

[45] Developed countries have witnessed a sharp increase in energy consumption in the service sector during the last few decades. According to estimates from the EIA, in the USA, energy consumption in the service sector has expanded from about 11% to 18% of the total energy budget from the 1950s until the beginning of the 21th century.

from the analysis of energy consumption in residential buildings presented in the previous section, the investigation of energy consumption in the service sector had an additional difficulty: a relevant number of energy analysis zones used in the project had heterogeneous land use characteristics. This is primarily due to the need, in the definition of the EAZs, to aggregate Zip+4 areas (the level of aggregation at which the utility companies provides energy consumption data for the area of study), which in more densely populated areas often tend to include contiguous parcels with different land uses. The separation of the land uses is important for a more accurate analysis of energy consumption in the study. However, the energy consumption data (at least in the LADWP dataset) do not disaggregate energy consumption by end use or sector. For these reasons, EAZs were classified using the predominant land use type classification with a 100% threshold. EAZs belonging to the non-residential sector were further subdivided using the floorspace type categories presented in chapter 3.

After filtering out records that contained information on residential EAZs (either in predominantly residential EAZs or as a minor portion of the total floorspace type in the EAZs), the non-residential dataset includes 13,442 EAZs. Of this total number of EAZs, 12,183 EAZs have information on electricity consumption data (in the LADWP area of service) and 1,259 EAZs have information on natural gas consumption (in the LBGO area of service).

The researchers further subdivided the sample for non-residential buildings, in order to identify energy trends in zones containing different building types. In particular, 1,768 EAZs contain office buildings, while 2,406 EAZs contain commercial/retail buildings (including a smaller subset of 91 EAZs that contain malls and big box retail space) and 3,338 EAZs contain warehouses and distribution facilities or light industrial buildings[46]. In the development of the energy consumption analysis, due to the difficulty of separating different energy purposes in the total consumption data (electricity consumption data from LADWP did not include information on the final energy end use), we focused the analysis of energy consumption on EAZs that had only limited variation in the contained floorspace types.

Information on government buildings and secondary education buildings (colleges and universities) was not used in the computation of energy consumption by square foot, because of the large amount of missing entries in the Assessor's data for these categories.[47] This could cause large departures from current energy consumption when trying to develop estimates

---

[46] This study did not attempt to study energy consumption in heavy industrial buildings (heavy industries and factories). Energy consumption in the industrial sector is in fact rather difficult to study, and it highly depends on the activities that are performed in the building. This is in particular true for heavy industrial buildings, where the type of manufacturing/production that is hosted in the building is directly responsible for the highest proportion of the energy consumed in the building. For this reason, the energy consumption and environmental impacts of heavy industries are better studied through dedicated analysis that focus on the assessment of the energy consumption of the specific facilities and/or the stationary emissions associated with each plant/factory.

[47] The computation of the total energy consumption for these buildings (Chapter 8) uses proxies for the unitary energy consumption per square foot of these buildings from similar floorspace categories.

while compensating for the missing information (up to 96% of government buildings have missing information on their size in the Assessor's data) that would need to be interpolated in the dataset. An additional group of 1,873 EAZs was removed from the dataset and treated separately, as these EAZs included land use types that are coded as "mixed land uses" in the Assessor's data. This category contains a variety of combinations of different land use codes, many of which are rather rare and difficult to study in statistical terms. These mixed land use zones, overall, contain various possible combinations of residential and commercial buildings. Mixed use EAZs were treated separately in this analysis, and their energy consumption was computed through a combination of the residential and commercial energy forecasting tools, as explained in the following Chapter 8 of this report. The remaining EAZs identified smaller groups of building types (which respectively include hospitals and educational/religious buildings) or were associated with the presence of heterogeneous building types belonging to different categories.[48]

In this part of the research, we estimated the average energy consumption in non-residential buildings using a different approach from the energy models developed for residential units. For the non-residential sector, the building "units" do not represent a valuable metrics to evaluate energy consumption. Rather, a valuable measure of energy consumption (and of energy efficiency) that is usually adopted in most energy studies expresses the building energy consumption in terms of unitary electricity and natural gas consumption per square foot of developed floorspace (by building type). We identified different clusters of buildings, and estimated the average energy consumption for square foot of developed floor space of each group of building types, using the data contained in the energy database. We also attempted to estimate regression models for energy consumption in non-residential buildings. However, the energy consumption models estimated at the EAZ level, using the limited sample size for the non-residential sector, did not have a satisfactory goodness of fit. For this reason, we did not use regression models to estimate energy consumption in non-residential buildings. Instead, we used the data contained in the energy database to estimate average values for electricity and natural gas consumption per square foot of developed floorspace in non-residential buildings. This approach is more appropriate, given the limited energy consumption data provided by the utility companies (in particular, sample sizes for some non-residential building categories are very small in the natural gas datasets), and the rather high variance often observed in the energy consumption variables as an effect of the many categories of non-residential buildings included in each zone.[49]

---

[48] The analysis of energy use in zones that included many different types of buildings was further complicated by the lack of information on the end use and purpose of energy consumption in the utility data. The investigation of energy consumption in more complex combinations of building types will be possible when more disaggregated data will be available, both at the spatial level and by purpose/energy end use.

[49]The estimation of energy consumption patterns in non-residential buildings could significantly improve with the access to more spatially detailed information on energy consumption data from the utility companies, which include also information by final energy end use, and the information. This, in addition

Table 28 reports the annual average consumption of electricity and natural gas per square foot of office buildings (including both high rise and low rise office space) computed using the data contained in the energy database. It is important to note that the energy consumption per square foot of built floorspace tend to have a rather large dispersion from the average values (as indicated by the standard deviation in parentheses), as an effect of the small sample and the variety of different types of building that are aggregated in the same category. Besides, climate zone do not significantly affect the consumption of energy for this category of buildings (at least, in the limited sample that is available; a similar effect is observed for the other floorspace types described later in this section).

In order to verify the average values for the consumption of electricity and natural gas per square foot of built office space from this study, we compared the results from the analysis to available independent studies. In particular, the results are rather consistent with the forecasts for energy consumed in office buildings estimated with the CEUS survey from the California energy Commission (CEC, 2006). The annual electricity consumption per square foot of office building computed in this study (14.147 KWh/sq. ft.) is contained between the values predicted by the CEUS study respectively for a small office (13.10 KWh/sq. ft.) and a large office (17.70 KWh) in the State of California, as well as the CEUS estimates for a small office (13.25 KWh/) and a large office (17.91 KWh) located in Southern California.[50]

Table 28: Average annual electricity and natural gas consumption in office buildings

	Electricity Consumption (KWh/Sq. ft.)	**Natural Gas Consumption (Therms/Sq. ft.)**
Office Space	14.1469	.2319
sample size	*1612*	*156*
(Std. Deviation)	*(8.7534)*	*(.2113)*

The average consumption of natural gas from this study is higher than in the CEUS data (0.105 to 0.219 Therms/Sq. ft. in California) and in particular it is much higher than the estimated consumption in the SCE area (0.08 to 0.13 Therms/Sq. ft., as reported in CEC, 2006). This

to the access to the complete dataset for all utilities in the Los Angeles County would contribute to map energy consumption patterns with more certainty, and to increase the goodness of fit of the estimated models.

[50] Please note that the cited study refers to estimates for energy consumption in the Southern California Edison (SCE) area of service. Please note that SCE has not (yet) provided utility consumption data for this research project. SCE serves an area of service that is predominantly located further away from the coastal climate than LADWP, and this might explain the reason for which our estimates is closer to the lower boundary of the interval. Besides, the majority of office square feet in our dataset come from small office buildings.

discrepancy might be due to the small sample size for natural gas in this study. Also considering the high variance observed in the annual natural gas consumption in our sample, we recommend reducing the coefficient for the consumption of natural gas for the office floorspace type to a value included in the range from the CEUS study (we recommend 0.12 Therms/Sq. ft.; this value will be used in the computations in Chapter 8).

Table 29 reports the average electricity and natural gas consumption per square foot of commercial and retail floorspace. These average energy consumptions are computed respectively with data from 2,237 EAZs commercial EAZs in the LADWP areas of service and 169 EAZs in the LBGO areas of service.

Table 29: Average annual electricity and natural gas consumption in general commercial and mall/big box retail space

	Electricity Consumption (KWh/Sq. ft.)	Natural Gas Consumption (Therms/Sq. ft.)
General Commercial*	21.5085	.2683
sample size	*2237*	*169*
(Std. Deviation)	*(14.4696)*	*(.2529)*

*It includes energy consumption for malls and big box retail.

The data include both neighborhood commercial stores and shops, larger store facilities and malls and big box retail. The decision to include malls and big box retail in the same category with other commercial facilities was due to the rather small sample size for these types of floorspace types (in particular for natural gas consumption). Average consumption of electricity and natural gas per square foot of developed, non-vacant space are in the range of the statewide estimates for electricity consumption (from 13 KWh/Sq. ft. for general commercial to 40.99 KWh/Sq. ft. for food stores; please note that this category may contain several categories of heterogeneous commercial floorspace types), and natural gas consumption (.260 to .276 Therms/Sq. ft.). Also in this case, the average natural gas consumption per square foot from this study tends to be higher than the estimates built on CEUS data for the SCE area.

Table 30: Average annual electricity and natural gas consumption in primary K12 educational and religious space

	Electricity Consumption (KWh/Sq. ft.)	Natural Gas Consumption (Therms/Sq. ft.)
Educational and Religious Space*	7.7074	.2539
sample size	*433*	*61*
(Std. Deviation)	*(8.0254)*	*(.1554)*

*It includes primary K12 educational and religious floorspace types.

We also computed average annual energy consumption for K-12 schools and religious space (we grouped these two "educational/religious" floorspace types as the sample sizes for each individual category were rather small). Table 30 summarizes the average consumption of electricity and natural gas for one square foot of these floorspace types. Please note that the smaller sample size (in particular for the natural gas) and observed variance in the sample contribute to the higher uncertainty associated with the average consumption of energy in these floorspace types.

Electricity consumption per square foot of educational/religious space is very similar to the range of values for schools from the Statewide (7.46 KWh/Sq. ft.) and SCE specific results (8.22 KWh/Sq. ft.). Natural gas consumption values tend, however, to be much higher than the values from the CEUS dataset. Also in consideration of the small sample size for the natural gas dataset and the incomplete information for school and religious buildings in the Assessor's data (which might be responsible for inaccurate estimation of the total square footage of developed floorspace in the dataset), we recommend adopting a more conservative value for the average consumption of natural gas, closer to the statewide averages, in the area of .15 Therms/Sq. ft.

Table 31: Average annual electricity and natural gas consumption in hospitals and health facilities

	Electricity Consumption (KWh/Sq. ft.)	**Natural Gas Consumption (Therms/Sq. ft.)**
Hospitals	20.658	1.0394
sample size	*44*	*10*
(Std. Deviation)	*(11.002)*	*(.5797)*

We also attempted to compute average consumption of electricity and natural gas for hospitals and health facilities (Table 31). Results for hospitals are based on a small number of EAZs, in particular for the consumption of natural gas. Average annual electricity consumption per square foot of hospital space is in line with the estimates for California and the SCE area. Annual average natural gas consumption per square foot is about 25% higher than in the estimates from CEUS.[51]

Finally, we computed annual electricity and gas consumption for light industrial space and warehouses. These two floorspace types proved to have rather similar results, especially for the estimation of the annual average electricity consumption.

[51] Please note that the sample size for the natural gas data for hospitals includes only 10 EAZs and is probably not statistically representative of the wide variety of health facilities and hospitals that can be found in Los Angeles County.

Table 32: Average annual electricity and natural gas consumption in light industrial space and warehouses

	Electricity Consumption (KWh/Sq. ft.)	Natural Gas Consumption (Therms/Sq. ft.)
Light Industrial and Warehouses	9.7004	.1289
sample size	*3119*	*219*
(Std. Deviation)	*(11.6942)*	*(.1759)*

The consumption of electricity for this category of floorspace types tend to be slightly higher than the estimates for warehouses (the only category for which an estimate is available in the CEUS data). Current estimates for electricity consumption in warehouses range between 4.45 KWh/Sq. ft. and 20.02 KWh/Sq. ft. in the case of refrigerated warehouses (only a subset of the this floorspace category). Estimates for energy consumption by square foot for light industrial space are not common, as electricity and natural gas consumption in industrial areas tend to vary depending on the activities that are performed in the industrial site.[52] In this study, we merged the light industrial sector and warehouses for the rather similar patterns that are found in energy consumption trends for these more limited categories.

We do not include heavy industries in our energy consumption estimations, as the forecasts for this subcategory of industrial activities would be too difficult to predict at the unit of floorspace, given the dramatic differences in the use of energy that is registered in different industrial fields. The annual average consumption of natural gas in light industrial space and warehouses differ more significantly in our dataset than the electricity consumption. In particular, natural gas consumption, which for the two categories of floorspace types lies way above the estimates from other sources for warehouses, was recomputed separately respectively for the light industrial space and the warehouses. Final results showed an average of .1045 Therms/Sq. ft. in warehouses and .1534 Therms/Sq. ft. for light industrial space.

Still, we call the attention of the reader on the large variation in the energy use of natural gas for different purposes in these building types, and limited amount of information[53] that is available

[52] Documents as the *California Energy Demand 2010 - 2020, Adopted Forecast* and its 2012 revised version (Kavalec *et al*., 2012) provide estimates for total electricity consumption in the area of study for the entire industrial sector, and by unit of production (in $ dollars). For the reasons mentioned before, it is difficult to estimate electricity consumption per unit of square foot in the industrial sector.

[53] Moreover, natural gas consumption data available from LBGO are referred to 2011, while the information from the Assessor's data are for 2008. This temporal mismatch might cause distortions in the process used to match energy data with the information on the building stock, as the number of square feet in EAZs for specific floorspace categories might be underestimated, for instance in the case of new building construction or renovation with upgrades (or change of use) in some parcels between 2008 and 2011.

through the analysis of the available energy consumption data for natural gas in this project. When more complete data for natural gas consumption are available from all utility companies in the area of study, it will be possible to revise the estimation of energy consumption for the various categories of floorspace. This will result in more robust and accurate estimates of energy consumption in buildings for this study, using the large amount of data developed as part of the project and contained in the energy database and the methodologies developed (and the others that were originally planned but could not be applied yet) in this research.

CHAPTER 8:
Pilot Energy Baseline and Building Energy Consumption in Los Angeles County

The purpose of this project is to develop an analytical methodology to study energy consumption for building operations in Los Angeles County. It provides an opportunity to study energy consumption patterns in buildings for many research purposes, i.e. through the possibility to assess the impact of several variables, such as income, climate zones and the characteristics of the housing units, on energy consumption per capita or by unit of developed floorspace. Additionally, the study aims at informing modeling applications on the energy consumption that is associated with the most common building types and land uses in the area of study. In this way, the results of the study can inform land use models and studies on the assessment of the energy consumption and greenhouse gas emissions associated with the current land use patterns. It also provides useful information to study energy trends and possible impacts on the energy demand associated with future land development and modifications in the location of residences and economic activities in Los Angeles County.

The results from the study provide useful insights on the distribution of energy consumption by building types/sectors, and their geographical variation with climate zones and other variables of interest. Moreover, the results from the analyses presented so far are useful to provide the needed inputs to estimate the proportion of GHG emissions that are associated with the various sectors of the building stock. In the following sections, applications of the results from the study are discussed. Some of these results will be of immediate application in an urban metabolism study that focuses on the analysis of environmental impacts of different sectors (economic activities, buildings, transportation) in Los Angeles County.

Building electricity consumption

This section applies the results from the previous analyses of energy consumption patterns in buildings in Los Angeles County to build an assessment of energy use for building operations in the area of study. We estimate the energy consumption for the various categories of buildings in the area of study using the estimated coefficients from the energy consumption models presented in the Chapter 7 of this report.

Table 33: Electricity consumption by categories of building types in the LADWP area of service

Floorspace type	*No. of Units*	*Sum of Sq. Ft.*	*Total Energy Consumption (MWh)*
a) Residential Sector			
SF residential[1]	381,385	619,160,875	3,197,437
SF residential with pool	99,773	259,130,732	1,487,315
MF residential[2]	1,046,667	808,409,437	4,044,014
Total Residential Sector	***1,527,825***	***1,686,701,045***	***8,728,765***

Floorspace type	*Sum of Sq. Ft.*	*Total Energy Consumption (MWh)*
b) Non-Residential Sector		
Developed amusement park space	747,662	18,972
General commercial	153,640,527	3,898,549
Government operations space	7,564,050	126,242
Office space[3]	168,855,095	2,818,140
Hospital space	20,594,888	501,920
Mall and big box retail space	34,669,370	879,717
Mixed use space	43,114,895	805,728
Primary K-12 education space	76,623,166	696,715
Secondary education space	3,770,901	34,288
Religious space	22,589,725	205,403
Warehouse & distribution space	94,140,152	1,077,337
Industrial space[4]	196,424,142	2,793,103
Total Non Residential Sector	***822,734,573***	***13,856,115***
Total LADWP area	**2,509,435,618**	**22,584,880**

Note: [1]includes urban mobile homes; [2]includes apartments, joined and GQ residential; [3]high and low density office space; [4]light and heavy industrial space.

Table 33 summarizes the results for the electricity consumption in the LADWP area by categories of floorspace/building type.[54] Estimates for the energy consumption from building categories are built using the number of units and the amount of developed floorspace (in square feet) by each floorspace category that has been determined as part of the building inventory for Los Angeles County. The assessment of the developed floorspace by category in the area of study is primarily based on the data from Assessor's parcel records, with the inclusion of the adjustments described in Chapter 3 of the report, to account for missing information on specific floorspace categories in the Assessor's data.

In the definition of the electricity consumption assessment summarized in the table we primarily used the estimates for unitary electricity consumption from Chapter 7. Energy consumption values for general commercial were used also for the developed amusement park space, as direct observations for this floorspace types did not allow us to estimate specific energy use estimates. Similarly, energy estimates for office space were used for government buildings, and secondary education space shares the same energy consumption patterns of the educational/religious space (estimated for primary K-12 and religious space). In the development of the energy assessment, we do not consider parking space (which was previously included in the building assessment from Chapter 3).

[54] Results were scaled to match total electricity consumption in the LADWP area of service. Original results produced a slight underestimation of the total amount of annual electricity consumption in the LADWP area of service, probably due to an underestimation of the building stock in the area of service.

The mixed use space includes EAZs that are coded with various combinations of land uses according to the Assessor's data. From the distribution of floorspace in this category, we could categorize this floorspace in two main sub-categories: EAZs that include various combinations of office and commercial space (46.05% of the square feet that fit in this category) and EAZs that include various combinations of commercial and residential space (53.95% of the mixed use square feet). Accordingly, we could estimate energy consumption profiles for this category of space using the appropriate combinations of floorspace types that resembled these sub categories of mixed use developed floorspace.

Finally, as mentioned before, industrial areas are very difficult to predict through the estimation of consumption patterns, as energy intensity highly depends on the specific activities that are carried out in each plant/establishment. The electricity consumption for the industrial sector reported in table 33 was matched to the reported quantity of electricity consumed by the industrial sector in the LADWP area in 2008. For this specific economic/building category, only the computation of electricity consumption in light industrial buildings is computed using the method from Chapter 7: the remaining amount of energy that is consumed in the industrial sector is therefore assigned to the heavy industries that are present in the area.

We initially produced estimates for electricity consumption that (from the comparison with LADWP total consumption data) slightly underestimate total annual electricity consumption in the LADWP area for all sectors (estimates are about 9% lower than observed consumption, on average). Results were scaled to match the total consumption of energy in the LADWP area of service. The results from Table 33 match the total consumption of energy in the region, and they provide an interesting breakdown of the electricity consumption by sector/building category.

The estimation of the energy consumption from buildings is based on the energy consumption patterns that have been identified and on the amount of developed floorspace types that is present in each Energy Analysis Zone in Los Angeles County. This allows the creation of the energy summaries at all levels of geography. The results presented in Table 33 are summarized for the entire LADWP area.

Energy consumption in buildings can be summarized also at different geographical levels of spatial aggregation, for example at the city level, and can provide information for specific comparisons among regions in the County. Using the results that were computed for the areas where energy consumption data were available, we can compute estimates of electricity consumption for building operations in the entire area of Los Angeles County. The results are scaled to match the total consumption of electricity in the region (65,163 GWh in 2008; original electricity forecasts from the application of the model were 10% lower than the total electricity consumption in the County). Table 34 shows the summary of energy consumption by categories of building types in the entire Los Angeles County. The same assumptions on the aggregation of buildings from the computation for the LADWP area are used also in this computation. In addition, assumptions for the effects of the climate zones for the climate zone 14 (not included in the sample that was used to estimate the electricity consumption model) were derived from the literature, based on the average number of cooling degree days and heating degree days in this climate zone.[55]

Table 34: Electricity consumption for building operations in Los Angeles County

Floorspace type	*No. of Units*	*Sum of Sq. Ft.*	*Total Energy Consumption (MWh)*
a) Residential Sector			
SF residential[1]	1,288,923	2,042,801,703	10,437,144
SF residential with pool	246,609	603,038,633	3,414,202
MF residential[2]	1,847,195	1,639,365,847	7,785,426
Total Residential Sector	***3,382,727***	***4,285,206,183***	***21,636,772***

Floorspace type	*Sum of Sq. Ft.*	*Total Energy Consumption (MWh)*
b) Non-Residential Sector		
Developed amusement park space	7,723,730	176,252
General commercial	462,507,541	10,554,181
Government operations space	35,523,283	533,175
Office space[3]	439,230,653	6,592,491
Hospital space	55,574,937	1,218,044
Mall and big box retail space	144,934,874	3,307,338
Mixed use space	87,648,920	1,500,426
Primary K-12 education space	183,716,885	1,502,286
Secondary education space	39,241,821	320,887
Religious space	73,282,749	599,246
Warehouse & distribution space	584,560,624	6,016,097
Industrial space[4]	689,292,379	11,206,050
Total Non Residential Sector	***2,803,238,395***	***43,526,473***
Total in Los Angeles County	**7,088,444,578**	**65,163,245**

Note: [1]includes urban mobile homes; [2]includes apartments, joined and GQ residential; [3]high and low density office space; [4]light and heavy industrial space.

The proposed assessment of energy consumption in Los Angeles County is based on the results of the analysis of energy consumption patterns that were developed as part of this project. As such, it is an attempt to depict energy consumption phenomena for the entire county. Indeed, large variation in energy consumption might be observed in specific areas, with local results that might differ significantly from the trend that has been estimated for the entire area of

[55] In the final estimates of energy consumption, we assume an increase in energy consumption from climate zone 9 to climate zone 14 that is similar to the increase that was estimated respectively between electricity consumption in climate zones 6 and 8 and in climate zones 9 and 16.

study. As highlighted in the previous sections of this report, the results from the study could be crosschecked and better verified when utility data will become available from all utility companies that operate in the area of study. With these pieces of additional information[56], it will be possible to improve the model specification and the estimation of the energy consumption pattern used in the study to improve its goodness of fit and the correspondence to local patterns and specific energy use profiles in the region.

Building natural gas consumption

Similarly to what was done for the annual electricity consumption in the LADWP area, we also computed the annual natural gas consumption by building type in the LBGO area of service. Table 35 summarizes the distribution of natural gas consumption by building type in the LBGO area.

Table 35: Natural gas consumption for building operations in the LBGO area of service

Floorspace type	*No. of Units*	*Sum of Sq. Ft.*	*Total Energy Consumption (thousands of Therms)*
a) Residential Sector			
SF residential[1]	58,797	85,714,229	25,643
SF residential with pool	5,905	11,931,720	2,949
MF residential[2]	113,624	83,032,856	24,618
Total Residential Sector	***178,326***	***180,678,805***	***53,209***

Floorspace type	*Sum of Sq. Ft.*	*Total Energy Consumption ((thousands of Therms)*
b) Non-Residential Sector		
Developed amusement park space	-	-
General commercial	17,719,684	8,453
Government operations space	281,239	134
Office space[3]	14,683,911	6,055
Hospital space	2,670,069	3,561
Mall and big box retail space	3,046,412	1,453
Mixed use space	4,230,176	2,420
Primary K-12 education space	3,304,993	1,175
Secondary education space	809,309	288
Religious space	3,078,276	1,095
Warehouse & distribution space	7,332,787	1,362

[56] To date, energy consumption data for electricity and natural gas associated with the majority of the population and building stock in Los Angeles County have not been provided yet by the utility companies, and could not be included in the analyses for this project.

Industrial space[4]	12,771,676	14,062
Total Non Residential Sector	***69,928,532***	***40,058***
Total LBGO area	**250,607,337**	**93,267**

Note: [1]includes urban mobile homes; [2]includes apartments, joined and GQ residential; [3]high and low density office space; [4]light and heavy industrial space.

Estimates from the natural gas consumption models from this study are able to estimate natural gas consumption in residential dwellings quite accurately (total energy consumption in the residential sector was initially slightly above, about 1%, the actual amount of energy consumption in the residential sector in the LBGO). Natural gas consumption in the non-residential sector was forecasted with lower accuracy, and needed more adjustments to match the total consumption in the area. The difference in the estimated and observed consumption of natural gas for non-residential buildings might be due to the small sample size that was used to estimate natural gas consumption and the rather large dispersion observed in the data. [57]

Table 36: Natural gas consumption for building operations in Los Angeles County

Floorspace type	*No. of Units*	*Sum of Sq. Ft.*	*Total Energy Consumption (thousands of Therms)*
a) Residential Sector			
SF residential[1]	1,288,923	2,042,801,703	687,049
SF residential with pool	246,609	603,038,633	160,282
MF residential[2]	1,847,195	1,639,365,847	463,439
Total Residential Sector	***3,382,727***	***4,285,206,183***	***1,310,770***
Floorspace type		*Sum of Sq. Ft.*	*Total Energy Consumption (thousands of Therms)*
b) Non-Residential Sector			
Developed amusement park space		7,723,730	2,934
General commercial		462,507,541	176,914
Government operations space		35,523,283	13,401
Office space[3]		439,230,653	143,854
Hospital space		55,574,937	59,069
Mall and big box retail space		144,934,874	55,346
Mixed use space		87,648,920	41,504
Primary K-12 education space		183,716,885	52,082

[57] The accuracy of the estimation of natural gas for the various categories of non-residential buildings would certainly benefit from the availability of more complete and spatially disaggregated utility data, and it could be improved in future extensions of the project when the data become available.

Secondary education space	39,241,821	11,046
Religious space	73,282,749	20,801
Warehouse & distribution space	584,560,624	86,048
Industrial space[4]	689,292,379	1,059,701
Total Non Residential Sector	***2,803,238,395***	***1,722,700***
Total in Los Angeles County	**7,088,444,578**	**3,033,469**

Note: [1]includes urban mobile homes; [2]includes apartments, joined and GQ residential; [3]high and low density office space; [4]light and heavy industrial space.

Table 36 computes the natural gas consumption for building operations in Los Angeles County. The results are based on the estimated consumption patterns for each category of buildings, and are scaled to the totals for annual natural gas consumption in the County in 2008. As visible from the data in the table, the total volume of natural gas consumed in the LBGO area of service is only a very limited fraction of the total amount of natural gas consumed in Los Angeles County. LBGO is a small MOU that provides natural gas in Los Angeles County. The largest share of the total supply of natural gas consumed in Los Angeles County[58] is provided by Southern California Gas Company.

GHG emissions from building operations

Using the results from this study, it is possible to estimate the proportion of greenhouse gas (GHG) emissions associated with the energy consumption in the various sectors of the building stock, by the various zones in Los Angeles County. Table 37 uses the results from the assessment of the electricity consumption that was presented earlier to compute an estimate of GHG emissions associated with electricity consumption for building operations in the LADWP area of service.[59]

For simplicity of exposition in this report we have reported only the total CO_2 equivalent emissions associated with the use of electricity in the LADWP area. A more complete assessment of GHG emissions could be created using disaggregate emission factors that differentiate the environmental impact of electricity production among the emission components for various greenhouse gases (or for any other pollutant emissions[60]). The results reported in Table 36 are based on the use of LADWP specific mix of energy sources for the

58 Unfortunately, to date, SCG has not provided utility data that could be included in our energy database and used for the estimation of energy consumption models for this study.

59 Please note that the amount of GHG emissions associated with the industrial sector, even if labeled as "industrial space" actually refers to the total amount of GHG emissions (and energy consumption) associated with the consumption of energy in the industrial sector, as it is not easy to separate energy consumed for building operation from the energy consumed for other activities in industrial facilities.

60 Alternative estimates could be created for specific pollutant emissions if, for instance, of interest in a study on the impact of human activities on health.

production of electricity, and on the use of the GHG emission estimates reported by the U.S. Environmental Protection Agency (EPA) Emissions & Generation Resource Integrated Database (eGRID)[61] for this specific utility.

Table 37: GHG emissions associated with electricity consumption for building operations, including grid losses, in the LADWP area of service

Floorspace type	*Total Adjusted Emissions (lb CO2 Equivalent)*	*Total Adjusted Emissions (Metric Tons CO2 Equivalent)*
a) Residential Sector		
SF residential[1]	3,724,159,956	1,689,251
SF residential with pool	1,732,324,406	785,769
MF residential[2]	4,710,195,737	2,136,509
Total Residential Sector	***10,166,680,099***	***4,611,529***
Floorspace type	*Total Adjusted Emissions (lb CO2 Equivalent)*	*Total Adjusted Emissions (Metric Tons CO2 Equivalent)*
b) Non-Residential Sector		
Developed amusement park space	22,096,772	10,023
General commercial	4,540,768,848	2,059,658
Government operations space	147,037,859	66,695
Office space[3]	3,282,380,582	1,488,863
Hospital space	584,603,212	265,172
Mall and big box retail space	1,024,635,884	464,767
Mixed use space	938,458,181	425,677
Primary K-12 education space	811,486,827	368,084
Secondary education space	39,936,182	18,115
Religious space	239,239,196	108,517
Warehouse & distribution space	1,254,810,141	569,172
Industrial space[4]	3,253,219,276	1,475,635
Total Non Residential Sector	***16,138,672,959***	***7,320,379***
Total LADWP	**26,305,353,058**	**11,931,907**

Note: [1]includes urban mobile homes; [2]includes apartments, joined and GQ residential; [3]high and low density office space; [4]light and heavy industrial space.

The results presented in Table 37 are based on energy consumption patterns that were estimated as part of this study and on the specific emission factors for LADWP. They also include an adjustment factor that accounts for power grid losses between the points of consumption and the points of generation (that are not already factored in the eGRID output

[61] More information on the Emissions & Generation Resource Integrated Database program from the U.S. EPA can be found on http://www.epa.gov/cleanenergy/energy-resources/egrid/index.html.

emissions rates). For these reasons, the researchers adjusted the output emission rates for electricity consumption to account for transmission and distribution line losses, which account on average for 8.21% of electricity production in the Western region of the United States.[62]

Table 38: GHG emissions associated with electricity consumption for building operations, including grid losses, in Los Angeles County

Floorspace type	*Total Adjusted Emissions (lb CO2 Equivalent)*	*Total Adjusted Emissions (Metric Tons CO2 Equivalent)*
a) Residential Sector		
SF residential[1]	12,156,485,721	5,514,089
SF residential with pool	3,976,633,237	1,803,770
MF residential[2]	9,067,942,312	4,113,149
Total Residential Sector	***25,201,061,270***	***11,431,009***
Floorspace type	*Total Adjusted Emissions (lb CO2 Equivalent)*	*Total Adjusted Emissions (Metric Tons CO2 Equivalent)*
b) Non-Residential Sector		
Developed amusement park space	205,285,928	93,116
General commercial	12,292,802,781	5,575,922
Government operations space	621,006,794	281,684
Office space[3]	7,678,491,305	3,482,905
Hospital space	1,418,695,608	643,510
Mall and big box retail space	3,852,165,996	1,747,313
Mixed use space	1,747,595,055	792,696
Primary K-12 education space	1,749,761,848	793,679
Secondary education space	373,748,124	169,529
Religious space	697,961,750	316,590
Warehouse & distribution space	7,007,146,106	3,178,388
Industrial space[4]	13,052,055,155	5,920,313
Total Non Residential Sector	***50,696,716,450***	***22,995,644***
Total in Los Angeles County	**75,897,777,720**	**34,426,653**

Note: [1]includes urban mobile homes; [2]includes apartments, joined and GQ residential; [3]high and low density office space; [4]light and heavy industrial space.

Table 38 summarizes the total greenhouse gas (GHG) emissions associated with electricity consumption in Los Angeles County, computed using the average GHG emission factors for the generation of electricity associated with the utilities that operate in Los Angeles County. Also

[62] The actual amount of electricity that needs to be produced to satisfy electricity demand is therefore obtained by the consumption divided by (one minus the grid gross loss as a decimal).

for this table, results are expressed in total annual CO2 equivalent emissions, and they do contain the same adjustment factors to account for power grid losses on the electric grid.[63]

Similarly to what has been done for the consumption of electricity, Table 39 contains the GHG emissions associated with the consumption of natural gas in Los Angeles County.

Table 39: GHG emissions associated with natural gas consumption for building operations in Los Angeles County

Floorspace type	*Total Adjusted Emissions (lb CO2 Equivalent)*	*Total Adjusted Emissions (Metric Tons CO2 Equivalent)*
a) Residential Sector		
SF residential[1]	9,238,058,320	4,190,313
SF residential with pool	2,155,149,629	977,559
MF residential[2]	6,231,399,650	2,826,515
Total Residential Sector	***17,624,607,598***	***7,994,388***
Floorspace type	*Total Adjusted Emissions (lb CO2 Equivalent)*	*Total Adjusted Emissions (Metric Tons CO2 Equivalent)*
b) Non-Residential Sector		
Developed amusement park space	39,452,976	17,896
General commercial	2,378,791,161	1,079,002
Government operations space	180,183,727	81,730
Office space[3]	1,934,258,431	877,365
Hospital space	794,235,291	360,259
Mall and big box retail space	744,187,976	337,558
Mixed use space	558,067,450	253,135
Primary K-12 education space	700,292,331	317,647
Secondary education space	148,522,107	67,368
Religious space	279,691,818	126,866
Warehouse & distribution space	1,157,002,142	524,807
Industrial space[4]	14,248,737,015	6,463,118
Total Non Residential Sector	***23,163,422,425***	***10,506,752***
Total in Los Angeles County	**40,788,030,023**	**18,501,139**

Note: [1]includes urban mobile homes; [2]includes apartments, joined and GQ residential; [3]high and low density office space; [4]light and heavy industrial space.

GHG emissions for the consumption of natural gas are expressed in terms of total CO_2 equivalent associated with the consumption of natural gas in buildings in Los Angeles County.

[63] CO_2 equivalent emissions are computed using a weighted factor to account for the presence of different utility companies, with different energy mixes, operating in the area.

Estimates of CO_2 equivalent for natural gas have been computed using an average value of 13.446 lb CO_2 equivalent/Therm, according to the specifications suggested by the CPUC.[64]

The total GHG emissions associated with the consumption of electricity and natural gas compose the total emissions associated with building operations. Table 40 summarizes the total distribution of GHG emissions by the building types that are responsible for the consumption of energy.

[64] The emission factor for the consumption of natural gas is based on the ClimateSmart computation, which includes both the emissions from the customers' use of natural gas and an estimate of the emissions associated with gas delivery.

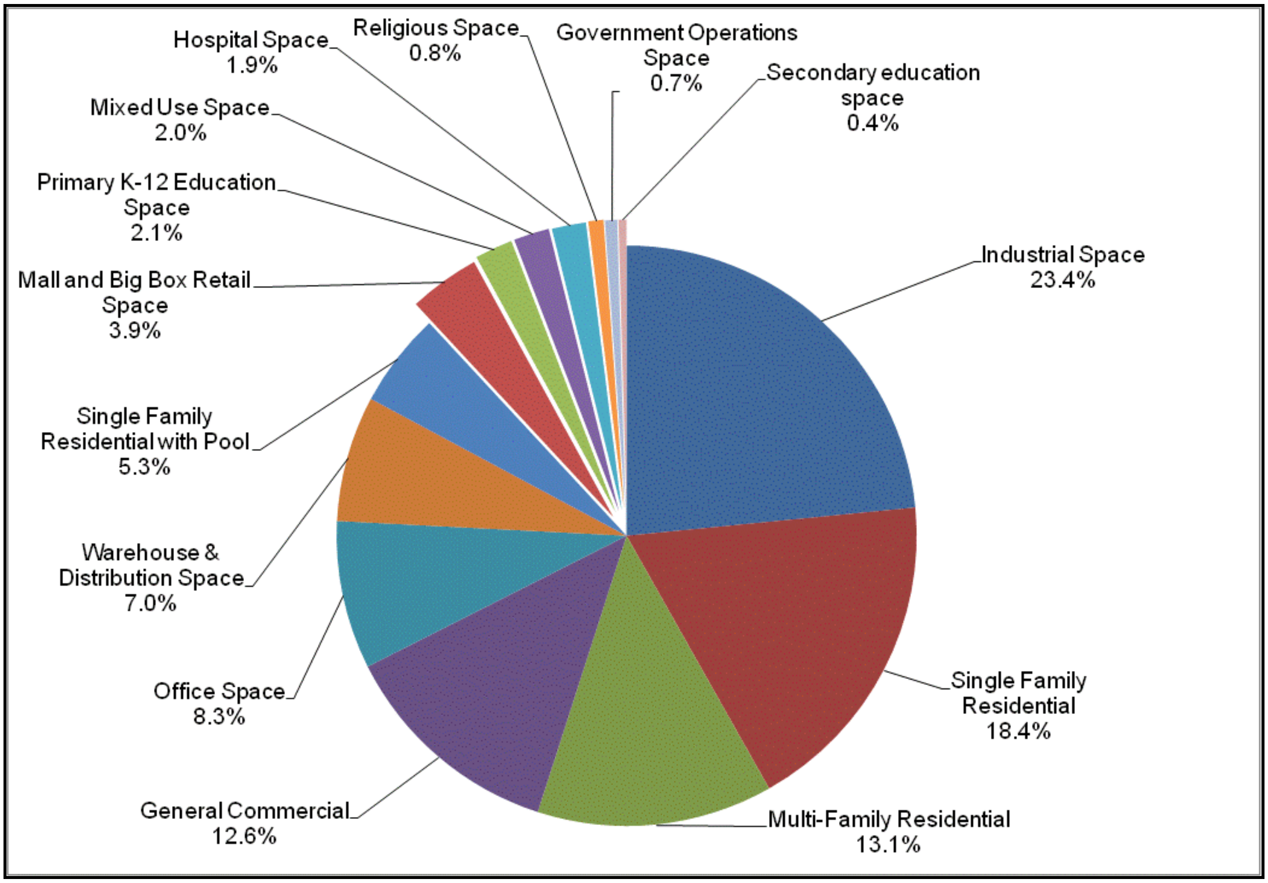
Table 40: GHG emissions associated with electricity and natural gas consumption for building operations in Los Angeles County

Floorspace type	*Total Emissions from Electricity Cons. (MetricTons CO2 Equivalent)*	*Total Emissions from Natural Gas Cons. (MetricTons CO2 Equivalent)*	*Total Emissions from Building Operations (MetricTons CO2 Equivalent)*
a) Residential Sector			
SF residential[1]	5,514,089	4,190,313	9,704,402
SF residential with pool	1,803,770	977,559	2,781,330
MF residential[2]	4,113,149	2,826,515	6,939,665
Total Residential Sector	***11,431,009***	***7,994,388***	***19,425,397***
Floorspace type	*Total Emissions from Electricity Cons. (MetricTons CO2 Equivalent)*	*Total Emissions from Natural Gas Cons. (MetricTons CO2 Equivalent)*	*Total Emissions from Building Operations (MetricTons CO2 Equivalent)*
b) Non-Residential Sector			
Developed amusement park space	93,116	17,896	111,012
General commercial	5,575,922	1,079,002	6,654,923
Government operations space	281,684	81,730	363,414
Office space[3]	3,482,905	877,365	4,360,270
Hospital space	643,510	360,259	1,003,769
Mall and big box retail space	1,747,313	337,558	2,084,871
Mixed use space	792,696	253,135	1,045,831
Primary K-12 education space	793,679	317,647	1,111,326
Secondary education space	169,529	67,368	236,898
Religious space	316,590	126,866	443,456
Warehouse & distribution space	3,178,388	524,807	3,703,195
Industrial space[4]	5,920,313	6,463,118	12,383,431
Total Non Residential Sector	***22,995,644***	***10,506,752***	***33,502,395***
Total in Los Angeles County	**34,426,653**	**18,501,139**	**52,927,792**

Note: [1]includes urban mobile homes; [2]includes apartments, joined and GQ residential; [3]high and low density office space; [4]light and heavy industrial space and processes.

Figure 28 shows the percentage distribution of GHG emissions associated with the consumption of energy for building operations in Los Angeles County. Please note that in this pie chart "industrial" emissions include GHG emissions associated with industrial building operations and also those associated with the energy consumed for other activities, as in industrial processes and manufacturing, as it was not possible to separate the different energy end uses for industries. The residential sector (including the various types of residential units) accounts for the largest share of total GHG emissions from the building stock.

Figure 28: GHG emissions from different building types in Los Angeles County[65]



The results contained in this estimate of greenhouse gas emissions for the area of study are not supposed to be exact accounts of the actual GHG emissions, but they provide some good metrics to compute the order of magnitude and proportion of the GHG emissions associated with energy consumption in buildings in the region.

Several sources of error and uncertainty might affect these results: first, the results are estimated through assumptions on the electricity and natural gas consumption patterns in buildings located in Los Angeles County, but they are estimated only on data provided by only two utility providers in the region (LADWP, for the electricity data, and LBGO for the natural gas data).

---

[65] Values for industrial space include not only building operations but also the energy consumed (and GHG emissions) associated with energy consumption for other activities in the building facilities.

Second, utility data that are used for the estimation of the energy consumption model were provided by these utility companies at the Zip+4 level of spatial aggregation and do not discriminate for energy purpose or end use. Therefore, it is possible that some of the relationships between energy consumption and the explanatory variables in the energy database might have been "watered down" due the necessary process of aggregation of parcels and Zip+4 data in the Energy Analysis Zones. This might contribute to the generation of some confounding factors, for instance for areas with mixed residential and commercial land uses, where the portions of energy consumed for the different purposes cannot be separated. Third, not all parcels in the database had a valid mailing address that could be matched to a Zip+4, and therefore matched to information on energy use for that area. Luckily, the amount of parcels that could not be matched to valid Zip+4 codes is rather small, but still they account for a number of buildings that could not be included in the energy database. The results of this study are based on the assumption that unmatched parcels were uniformly distributed in the database, and therefore that they do not generate distortions in the spatial patterns of buildings in the EAZs and in the resulting definition and interpretation of the energy consumption profiles.[66] Finally, an additional possible source of error in the study is due to the lack of accurate information on the "non-property tax" buildings in Los Angeles County. This problem, which is common to many studies that attempt to assess, and model, land use patterns in a region, reduces the validity of the energy estimates associated with these buildings.

Even after accounting for all the issues above, this research represents an important milestone in the development of a methodology for the assessment of energy consumption in buildings, based on the integration of data from a number of different sources, and developed at a high level of spatial details (about 450,000 Energy Analysis Zones). As part of the study, the researchers developed an important set of analytical tools for the analysis of energy consumption in buildings, which are of general validity. The accuracy of the results and of the projections/forecasts can be increased when more detailed utility consumption data become available, and in particular if detailed consumption data are provided by the utility companies for both natural gas and electricity for the same areas and times (ideally, for the entire county and for multiple years), and at a good level of temporal and spatial aggregation (ideally, monthly data, at the parcel level).

---

[66] The proportion of unmatched parcels might be responsible for an underestimation of the number of sq. ft. in the EAZs, and therefore for some resulting errors in the estimation of energy consumption per residential unit and square foot of developed floorspace. However, the results for total energy consumption in the region were scaled to match energy consumption totals in 2008 in Los Angeles County. This reduces the risk of overestimation (or underestimation) of CO_2 emissions in the assessment of the GHG emissions from energy consumption of the building stock.

An analytical tool to forecast future energy use in Los Angeles County

This study provides information on the energy consumption (for electricity and natural gas) in buildings in Los Angeles County. The results from the study are useful to create estimates for energy consumption by sectors (e.g. residential vs. commercial) and for specific building types, which are of valuable use to explore the relationships between energy consumption patterns and the characteristics of the building stock, of the natural environment and individuals' sociodemographic features. Moreover, they inform the PECAS land use modeling system on the energy consumption component of land use and can be used as part of studies that attempt to quantify the impact of human settlements and communities on the consumption of resources and on the generation of GHG emissions and other environmental externalities.

The estimates provided by this study can be also used to provide forecasts on the modifications in energy demand that would result from changes in the land use and the (re)location of residences and economic activities. Especially if in combination with a land use model, the results from the study can inform on the expected modifications in the energy demand and in the resulting GHG emissions, by sector, under different scenarios of development of land use.

The study has been designed, on purpose, to use categories of floorspace types that are as consistent as possible with the floorspace types adopted in the California Statewide PECAS modeling framework. Additionally, the results from the study, and any future updates and extensions, could be easily applied also to inform other land use models on building energy consumption, or to inform simplified sketch models for a specific region[67].

The results from the study and the application of the land use modeling framework, updated with the building energy consumption component, may provide useful information, for instance, on the modification of the resulting energy demand and GHG emissions associated with a change in land use that favors an increase in density of residential units. This could be the case for policies that support the development of more compact developments (e.g. compared to a different scenario in which more conventional suburban developments are built). The results from the study can inform on the expected magnitude of the change in energy consumption that would derive from such modifications, as well as can be useful to estimate expected differences in energy demand forecasts under different development scenarios. Another use is to estimate (all else equal, i.e. not varying the total amount of developed space) the effects on energy consumption of the implementation of policies to further increase energy efficiency in buildings, e.g. reducing the energy consumption for some specific categories of

[67] The use of improved sketch models, like Urban Footprint, is quickly spreading in metropolitan planning organizations (MPOs) and research institution, as a quick way to represent the interaction between the land use and the transportation and energy systems, usually with more simplified approaches and faster times of development than full land use models. The accuracy and comprehensiveness of this new generation of models is, however, significantly increasing, providing a valuable alternative to more complex, and expensive to build and maintain, models.

buildings in the current inventory, or simulating a target rate of building retrofits that would reduce energy consumption and GHG emissions for specific building categories.

It is important to note how the current study is based on the analysis of energy consumption data from 2008, and that it therefore does not consider the impact of additional technological development introduced since then, or still to develop. Policies that account for these additional gains in energy efficiency could be, however, accounted for as external inputs that further modify the energy consumption estimates created through the simulation of the previously scenarios of land use development. Finally, the impact of specific factors, like the spread of photovoltaic (PV) panels or other solutions that contribute to modify energy consumption through increased energy efficiency and/or cogeneration (up to the level of zero energy balance in buildings) is not explicitly studied in this research. If not necessarily associated with a reduction in the total energy consumption per se, PVs are responsible for a change in the energy demand that is requested from the grid and in the overall energy mix associated with the energy consumed in a study area. These factors could be included in future extension of the project, using the same research methods and analytical tools that were developed for this project, and they would lead to further developments in the application of land use models (providing additional and more detailed information on the characteristics of the buildings and their energy sources and technology).[68]

Future extensions of this research will focus on refining the analysis of energy consumption patterns for different building types, using the more detailed energy consumption data that will be provided by the utility companies in the region. Additional research is needed to study the impact of specific technological solutions, building structures, vintage and shape of the buildings on energy consumption. Moreover, the Assessor's data provide limited information on the adoption of new technologies in buildings (PVs, green roofs, higher efficiency standards, cogeneration and zero energy buildings, etc.). However, data from additional sources, like satellite imagery and surveys on the adoption rate of PVs and cogeneration facilities can be useful to provide additional sources of data to estimate the contribution to energy conservation and GHG reduction strategies of these solutions.

The use of the estimated patterns of building energy consumption in land use modeling solutions is particularly useful for understanding the impact of modifications of land use on energy consumption for building operations. Future extensions of this project will focus on the integration of the analysis of the building stock and the energy consumption for building operations developed in this study to develop forecasts for future trends in energy consumption in the region, under different assumptions of land use development, demographic and economic trends and modifications in the building efficiency standards.

[68] This eventual extension of the project would also need the access to reliable sources of data to complement the energy database for the region of study, as information on the adoption of these solutions is still rather fragmented (especially in terms of the spatial distribution in a region), and in continuous evolution.

REFERENCES

Armstrong, M. M., Swinton, M. C., Ribberink, H., Beausoleil-Morrison, I., & Millette, J. (2009). Synthetically derived profiles for representing occupant-driven electric loads in Canadian housing. *Journal of Building Performance Simulation*, *2*(1), 15–30.

Beccali, M., Cellura, M., Lo Brano, V., & Marvuglia, A. (2008). Short-term prediction of household electricity consumption: Assessing weather sensitivity in a Mediterranean area. *Renewable and Sustainable Energy Reviews*, *12*(8), 2040–2065.

Bell, M. (2004). Energy efficiency in existing buildings: the role of building regulations. *COBRA 2004 Proceedings of the RICS Foundation Construction and Building Research Conference*. Retrieved from http://www.leedsmet.ac.uk/as/cebe/projects/cobra04-1.pdf

California Energy Commission (1978) Building Energy Efficiency Standards: New Residential and New Nonresidential Buildings. CEC-400-1978-001.

California Energy Commission. (1995) California climate zone descriptions for new buildings: Directory. CEC-400-95-041.

California Energy Commission. (2006) California Commercial End-Use Survey. Consultant Report prepared by Itron, Inc. for the California Energy Commission. CEC-400-2006-005.

California Energy Commission. (2008) 2008 building energy efficiency standards for residential and nonresidential buildings: Regulations/standards. CEC-400-2008-001-CMF.

California Energy Commission (2011) Integrated Energy Policy Report, Final commission report, CEC-100-2011-001-CMF.

Castleton, H. F., Stovin, V., Beck, S. B. M., & Davison, J. B. (2010). Green roofs; building energy savings and the potential for retrofit. *Energy and Buildings*, *42*(10), 1582–1591.

Chingcuanco F. and E. J. Miller (2012) A microsimulation model of urban energy use: Modelling residential space heating demand in ILUTE, *Computers, Environment and Urban Systems*, 36(2), 186-194.

Darley, J.M. (1978) Energy conservation techniques as innovations and their diffusion, *Energy and Buildings*, 1, 339–343.

Department for Communities and Local Government. (2006). *Building a Greener Future: Towards Zero Carbon Development*. UK: Department for Communities and Local Government.

Department for Communities and Local Government: London. (2007). *Building Regulations: Energy efficiency requirements for new dwellings*. London.

Department of Trade and Industry. (2007). *Meeting the Energy Challenge: A White Paper on Energy* (GOV. UK/NATIONAL_STATISTICS No. CM 7124). UK: Department of Trade and Industry. Retrieved from http://www.berr.gov.uk/files/file39387.pdf

European Parliament (2012) Directive 2002/91/CE on the energy performance of buildings, December 16, 2002.

Gobakis, K., Kolokotsa, D., Synnefa, A., Saliari, M., Giannopoulou, K., & Santamouris, M. (2011). Development of a model for urban heat island prediction using neural network techniques. *Sustainable Cities and Society*, *1*(2), 104–115.

Griffith, B., Long, N., Torcellini, P., Judkoff, R., Crawley, D., & Ryan, J. (2007). *Assessment of the technical potential for achieving net zero-energy buildings in the commercial sector* (Technical Report No. NREL/TP-550-41957). Golden, CO: National Renewable Energy Laboratory. Retrieved from http://pmcgroup.biz/downloads_files

Howard, B., Parshall, L., Thompson, J., Hammer, S., Dickinson, J., & Modi, V. (2012). Spatial distribution of urban building energy consumption by end use. *Energy and Buildings*, *45*, 141–151.

Høyer, K. G., & Holden, E. (2003). Household consumption and ecological footprints in Norway–does urban form matter? *Journal of Consumer Policy*, *26*(3), 327–349.

Jebaraj, S., & Iniyan, S. (2006). A review of energy models. *Renewable and Sustainable Energy Reviews*, *10*(4), 281–311. doi:10.1016/j.rser.2004.09.004

Johnston, D., Lowe, R., & Bell, M. (2005). An exploration of the technical feasibility of achieving CO2 emission reductions in excess of 60% within the UK housing stock by the year 2050. *Energy Policy*, *33*(13), 1643–1659.

Johnston, David. (2003) *A physically-based energy and carbon dioxide emission model of the UK housing stock*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.194.7917

Kalogirou, S. A., & Bojic, M. (2000) Artificial neural networks for the prediction of the energy consumption of a passive solar building. *Energy*, *25*(5), 479–491.

Kandel, A., M. Sheridan and P. Mcauliffe (2008) A Comparison of Per Capita Electricity

Consumption in the United States and California. California Energy Commission, Staff Paper. Publication Number: CEC-200-2009-015.

Kavalec, C. and T. Gorin (2009) *California Energy Demand 2010 - 2020, Adopted Forecast*. California Energy Commission. Publication Number: CEC-200-2009-012-CMF.

Kavalec, C., N. Fugate, T. Gorin, B. Alcorn, M. Ciminelli, A. Gautam, K. Sullivan and G. Sharp (2012). *Revised California Energy Demand Forecast 2012* - 2022. California Energy Commission, Electricity Supply Analysis Division. Publication Number: CEC-200-2012-001-SD-V2.

Keirstead J. (2006) Evaluating the applicability of integrated domestic energy consumption frameworks in the UK, *Energy Policy*, 34(17), 3065-3077.

Keirstead J. and A. Sivakumar (2012) "Using Activity-Based Modeling to Simulate Urban Resource Demands at High Spatial and Temporal Resolutions", *Journal of Industrial Ecology* (in press).

Kennedy, C., J. Cuddihy, and J. Engel-Yan (2007) "The changing metabolism of cities", *Journal of Industrial Ecology*, 11(2), 43–59.

Kolokotroni, M., Davies, M., Croxford, B., Bhuiyan, S., & Mavrogianni, A. (2010). A validated methodology for the prediction of heating and cooling energy demand for buildings within the Urban Heat Island: Case-study of London. *Solar Energy*, *84*(12), 2246–2255.

Lee, W. L., Yik, F. W. H., & Burnett, J. (2007). Assessing energy performance in the latest versions of Hong Kong Building Environmental Assessment Method (HK-BEAM). *Energy and Buildings*, *39*(3), 343–354.

Lopes, M.A.R., C.H. Antunes and Martins (2012) Energy behaviours as promoters of energy efficiency: A 21st century review, *Renewable and Sustainable Energy Reviews*, 16(6), 4095-4104.

Marszal, A. J., Heiselberg, P., Bourrelle, J. S., Musall, E., Voss, K., Sartori, I., & Napolitano, A. (2011). Zero Energy Building – A review of definitions and calculation methodologies. *Energy and Buildings*, *43*(4), 971–979.

Mihalakakou, G., Santamouris, M., & Tsangrassoulis, A. (2002). On the energy consumption in residential buildings. *Energy and buildings*, *34*(7), 727–736.

Miller, N. (2011) Comparing six neighborhood scale urban patterns using building energy simulation. Paper presented at the Computer in Urban Planning and Urban Management (CUPUM) Conference, Lake Louise, Canada, July 2011.

Pacific Energy Center (2008). *The Pacific Energy Center's Guide to: California Climate Zones and Bioclimatic Design.* Retrieved from: http://www.pge.com/includes/docs/pdfs/about/edusafety/training/pec/toolbox/arch/climate/california_climate_zones_01-16.pdf

Pérez-Lombard, L., Ortiz, J., & Pout, C. (2008). A review on buildings energy consumption information. *Energy and Buildings*, *40*(3), 394–398.

Pérez-Lombard, L., J. Ortiz, R. González and I. R. Maestre (2009) A review of benchmarking, rating and labelling concepts within the framework of building energy certification schemes, *Energy and Buildings*, 41(3), 272-278.

Pincetl, S., S. Murphy, P. Bunje, P. Burns, M. Chester, G. Circella, J. Ferrell, D. Flaming and M. McCoy (UCLA Institute of the Environment and Sustainability). 2012. *Methodology to Establish Regional Energy Baselines.* California Energy Commission. CEC-500-10-033.

Scott Kelly (2011) Do homes that are more energy efficient consume less energy? A structural equation model of the English residential sector, *Energy*, 36(9), 2011, 5610-5620.

Shorrock, L. D., & Dunster, J. E. (1997). The physically-based model BREHOMES and its use in deriving scenarios for the energy use and carbon dioxide emissions of the UK housing stock. *Energy Policy*, *25*(12), 1027–1037.

Steemers, K. (2003). Energy and the city: density, buildings and transport. *Energy and Buildings*, 35(1), 3–14.

Tso, G. K. F., & Yau, K. K. W. (2007). Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. Energy, 32(9), 1761–1768.

ULTRANS and HBA Specto (2011) California Statewide Travel Demand Model. Final System Documentation: Technical Note (18 Volumes). Institute of Transportation Studies, University of California, Davis, 2011.

Van Raaij, W. F. and T. M. M. Verhallen (1983a) A behavioral model of residential energy use. *Journal of Economic Psychology*, 3(1), 39-63.

Van Raaij, W. F. and T, M. M. Verhallen (1983b), Patterns of residential energy behavior, *Journal of Economic Psychology*, 4(1–2), 85-106.

Yu, Z., F. Haghighat, B. Fung, and H. Yoshino. (2010) A decision tree method for building energy demand modeling, *Energy and Buildings*, 42, 1637-1646.

GLOSSARY

AIN	Assessor Identification Number
ANN	Artificial Neural Network
CCAR	California Climate Action Registry
CBECS	Commercial Buildings Energy Consumption Survey
CEC	California Energy Commission
CED	California Energy Demand
CEUS	Commercial End-Use Survey
CPUC	California Public Utilities Commission
CSV	Comma Separated Value
DGS	Department of General Services
eGRID	Emissions & Generation Resource Integrated Database
EIA	Energy Information Administration
EPA	Environmental Protection Agency
EPSG	European Petroleum Survey Group
GHG	Greenhouse Gas
GIS	Geographic Information System
GQ	Group-Quarter
GRIP	Greenhouse Gas Regional Inventory Protocol
GSA	General Service Administration
GW	Gigawatt
GWh	Gigawatt hours
HVAC	Heating, Ventilation, Air Conditioning
IEPR	Integrated Energy Policy Report
ICLEI	International Council for Local Environmental Initiatives
IOU	Investor-owned utility
KW	Kilowatt
KWh	Kilowatt hours
LADWP	Los Angeles Department of Water and Power
LBGO	Long Beach Gas and Oil
MH	Mobile Home
MW	Megawatt
MWh	Megawatt hours
PIER	Public Interest Energy Research
PV	Photovoltaic
RASS	Residential Appliance Saturation Study
RECS	Residential Energy Consumption Survey

SCE	Southern California Edison Company
SEM	Structural Equation Model
SCG	Southern California Gas Company
SFD	Single Family Detached
WRI	World Resources Institute
WCSB	World Business Council for Sustainable Development

APPENDIX A

Energy analysis Zones nest very well in the county and city boundaries for communities located in Los Angeles County. The following table reports the distribution of EAZs by city inside Los Angeles County.

Table A-1: Distribution of Energy Analysis Zones by city in Los Angeles County

City Name	*Number of EAZs*
Agoura Hills	1071
Alhambra	4935
Arcadia	3760
Artesia	740
Avalon	0
Azusa	1802
Baldwin Park	2500
Bell	1271
Bell Gardens	935
Bellflower	2666
Beverly Hills	2004
Bradbury	111
Burbank	5667
Calabasas	1220
Carson	4161
Cerritos	2558
Claremont	2020
Commerce	905
Compton	4381
Covina	2666
Cudahy	423
Culver City	2529
Diamond Bar	2904
Downey	4419
Duarte	1071
El Monte	3407

El Segundo	1286
Gardena	3480
Glendale	10813
Glendora	2999
Hawaiian Gardens	539
Hawthorne	3710
Hermosa Beach	1593
Hidden Hills	186
Huntington Park	2141
Industry	644
Inglewood	6267
Irwindale	229
La Canada Flintridge	1402
La Habra Heights	633
La Mirada	2518
La Puente	1281
La Verne	1799
Lakewood	3198
Lancaster	7804
Lawndale	1073
Lomita	1033
Long Beach	22888
Los Angeles	154022
Lynwood	2235
Malibu	1217
Manhattan Beach	2921
Maywood	971
Monrovia	2346
Montebello	2653
Monterey Park	3481
Norwalk	3850
Palmdale	7242
Palos Verdes Estates	955

Paramount	1754
Pasadena	8943
Pico Rivera	2348
Pomona	6223
Rancho Palos Verdes	2454
Redondo Beach	3835
Rolling Hills	112
Rolling Hills Estates	393
Rosemead	1810
San Dimas	1758
San Fernando	1124
San Gabriel	2272
San Marino	1029
Santa Clarita	8528
Santa Fe Springs	1466
Santa Monica	6952
Sierra Madre	985
Signal Hill	923
South El Monte	937
South Gate	3118
South Pasadena	1613
Temple City	1691
Torrance	8163
Vernon	558
Walnut	1517
West Covina	4904
West Hollywood	2420
Westlake Village	617
Whittier	4734
LA County (Total)	448,380

The following Figures A-1 and A-2 and A-3 show some examples (for the same geographic area) of the spatial overlap of the Energy Analysis Zones with the U.S. Census Block Group and the Traffic Analysis Zones (TAZs) developed at the University of California, Davis, for the analysis of transportation demand in the California Statewide Travel Demand Model (CSTDM) (ULTRANS, 2011). Figure A-3 shows the overlap of the Energy Analysis Zones with the larger Land Use Zones (LUZs), also developed at the University of California, Davis, for the analysis of land use activities in the PECAS model (ULTRANS, 2011).

Figure A-1: Spatial overlap of Energy Analysis Zones and Block Groups

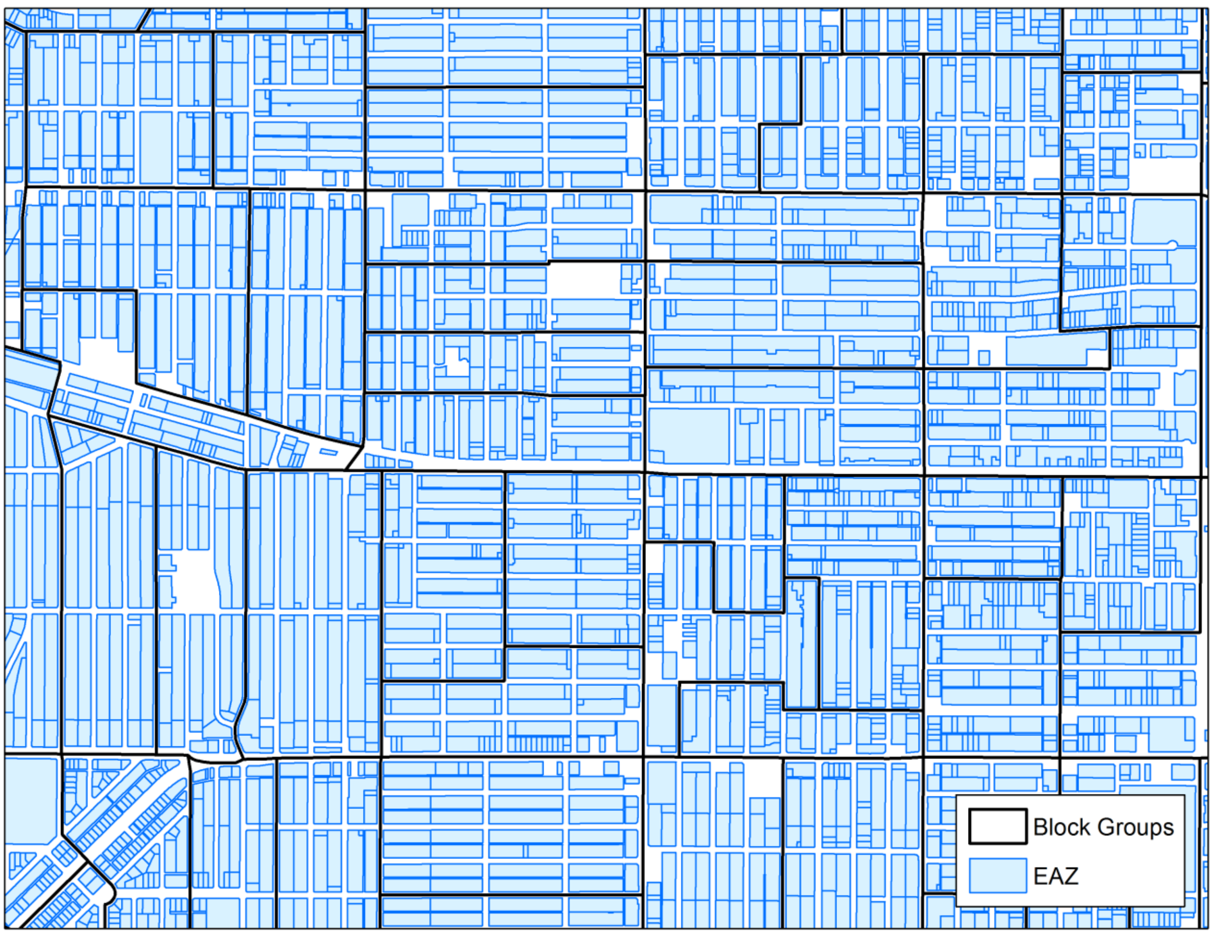


Figure A-2: Spatial overlap of Energy Analysis Zones and Traffic Use Zones

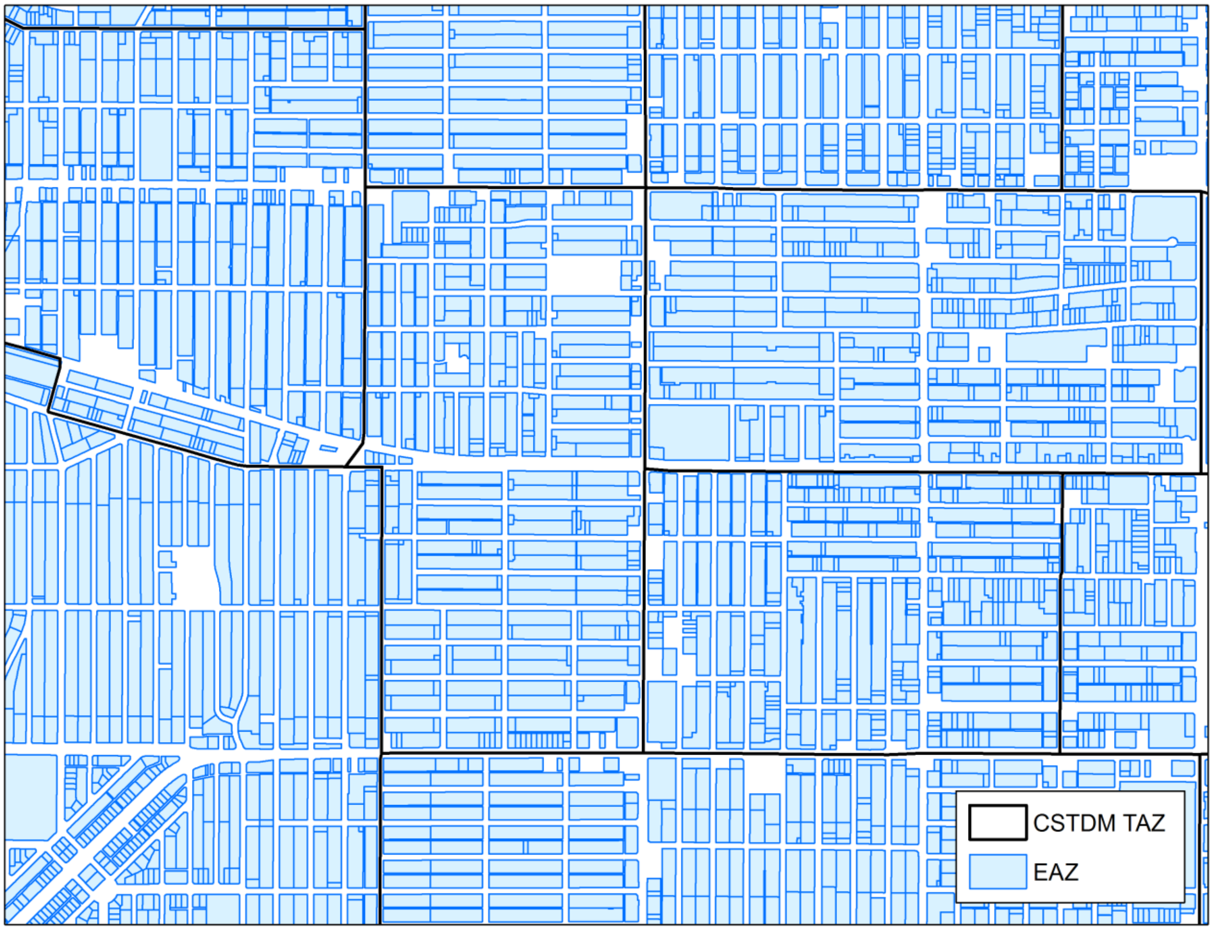


Figure A-3: Spatial overlap of Energy Analysis Zones and Land Use Zones