A DATA-DRIVEN APPROACH FOR PREDICTING SOLAR ENERGY POTENTIAL OF BUILDINGS IN URBAN FABRIC

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

AYÇA DURAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF ARCHITECTURE
IN
ARCHITECTURE

JULY 2022

Approval of the thesis:

A DATA-DRIVEN APPROACH FOR PREDICTING SOLAR ENERGY POTENTIAL OF BUILDINGS IN URBAN FABRIC

submitted by **AYÇA DURAN** in partial fulfillment of the requirements for the degree of **Master of Architecture** i**n Architecture, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Cânâ Bilsel
Head of the Department, **Architecture** _____

Assoc. Prof. Dr. İpek Gürsel Dino
Supervisor, **Architecture, METU** _____

Assist. Prof. Dr. Çağla Meral Akgül
Co-Supervisor, **Civil Engineering, METU** _____

Examining Committee Members:

Assoc. Prof. Dr. Funda Baş Bütüner
Architecture, METU _____

Assoc. Prof. Dr. İpek Gürsel Dino
Architecture, METU _____

Assist. Prof. Dr. Çağla Meral Akgül
Civil Engineering, METU _____

Assoc. Prof. Dr. Murat Göl
Electrical and Electronics Engineering, METU _____

Assoc. Prof. Dr. Yasemin Afacan
Interior Architecture and Environmental Design, Bilkent Uni. _____

Date: 05.07.2022

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name Last name : Ayça Duran

Signature :

ABSTRACT

A DATA-DRIVEN APPROACH FOR PREDICTING SOLAR ENERGY POTENTIAL OF BUILDINGS IN URBAN FABRIC

Duran, Ayça
Master of Architecture, Architecture
Supervisor: Assoc. Prof. Dr. İpek Gürsel Dino
Co-Supervisor: Assist. Prof. Dr. Çağla Meral Akgül

July 2022, 120 pages

Energy-efficient buildings that use clean and sustainable energy sources are urgently needed to reduce the environmental impact of buildings and mitigate climate change in cities. Buildings have great potential in harvesting solar energy by their solar exposure capacity. Developments in PV technologies also encourage the integration of PV systems into architectural applications. However, urban contexts can limit solar energy generation capacity of buildings by shading building envelopes and reducing available space for building-integrated PV (BIPV) deployment. Urban scale analyses are needed to understand the impact of urban form on solar energy generation potential. Nevertheless, urban scale analyses require tedious modeling and expertise, which may not be available during building design development or post-occupancy. This thesis presents a data-driven approach for predicting BIPV generation per façade and roof surface area utilizing machine learning (ML) techniques. As a first step, a simulation-based data generation method is applied to four neighborhoods in Ankara, representing different street network patterns. Three ML models are utilized and trained with simulation outputs. Two groups of predictors, simple and complex, are formed to allow predictions in various contexts. Overall, the best prediction model for facades has an R^2 value of 0.98 and nRMSE

of 10.83%. For roofs, R^2 and nRMSE could reach up to 0.94 and 3.13%, respectively. Complex predictors offer better prediction performances, whereas basic predictors enable the prediction of BIPV potential independent of a 3D model. Varieties observed in prediction performances for different neighborhoods also imply the need for urban morphology inclusive approaches in predicting BIPV potential.

Keywords: Solar energy, Urban fabric, BIPV, Machine learning, Building energy performance.

ÖZ

KENT DOKUSU İÇİNDEKİ BİNALARIN GÜNEŞ ENERJİSİ POTANSİYELİNİ TAHMİN ETMEYE YÖNELİK VERİYE DAYALI BİR YAKLAŞIM

Duran, Ayça
Yüksek Lisans, Mimarlık
Tez Yöneticisi: Doç. Dr. İpek Gürsel Dino
Ortak Tez Yöneticisi: Dr. Öğr. Üyesi Çağla Meral Akgül

Temmuz 2022, 120 sayfa

Binaların çevresel etkilerini azaltmak ve kentlerde iktim değişikliğini önlemek için temiz ve sürdürülebilir enerji kaynakları kullanan verimli binalara ihtiyaç duyulmaktadır. Binalar, güneşe maruz kalma kapasiteleri ile güneş enerjisini hasat etmede büyük potansiyele sahiptir. PV teknolojilerindeki gelişmeler de bu sistemlerin mimari uygulamalara entegrasyonunu teşvik etmektedir. Ancak, kentsel bağlamlar yapı kabuğunu gölgeleyerek ve binaya entegre PV (BIPV) sistemleri için kullanılabilir alanı azaltarak binaların güneş enerjisi üretim kapasitesini sınırlayabilir. Kent formunun binaların güneş enerjisi potansiyeline etkisini anlamak için kentsel ölçekte analizlere ihtiyaç vardır. Fakat, kentsel ölçekte analizler, detaylı modelleme ve ihtisas gerektirir. Bu gereklilikler bina tasarım sürecinde veya sonrasıda sırasında mevcut olamayabilir. Bu tez, makine öğrenmesi (ML) tekniklerini kullanarak cephe ve çatı yüzey alanı başına BIPV enerji üretimini tahmin etmek için veriye dayalı bir yaklaşım sunar. İlk adım olarak, Ankara ilindeki farklı sokak ağı örüntüsüne sahip dört mahalleye simülasyon tabanlı veri üretme yöntemi uygulanmıştır. Üç ML modeli seçilmiş ve simülasyon verileriyle eğitilmiştir. Basit

ve karmaşık olmak üzere iki öngörücü grubu oluşturularak, çeşitli bağlamlarda tahminlemeye olanak sağlanmıştır. Genel olarak, cepheler için en iyi tahmin modeli 0,98 R^2 ve %10,83 nRMSE değerlerine sahiptir. Çatılar için en iyi R^2 ve nRMSE sırasıyla %0,94 ve %3,13 değerlerine ulaşmıştır. Karmaşık öngörücüler daha iyi performans tahmini sağlarken basit öngörücüler üç boyutlu modelden bağımsız olarak BIPV potansiyelinin tahmin edilmesini sağlamaktadır. Farklı mahalleler için tahmin performanslarında gözlemlenen çeşitlilikler, BIPV potansiyelini tahmin etmede kentsel morfolojiyi gözeten yaklaşımlara olan ihtiyacı işaret etmektedir.

Anahtar Kelimeler: Güneş enerjisi, Kent dokusu, Fotovoltaik, Makine öğrenmesi, Bina enerji performansı.

To my family

Şenay, Vedat, and Ayberk

ACKNOWLEDGMENTS

First and foremost, I owe my deepest gratitude to my thesis supervisors Assoc. Prof. Dr. İpek Gürsel Dino, Assist. Prof. Dr. Cagla Meral Akgül and Assoc. Prof. Dr. Sinan Kalkan. Since the first days of my undergraduate studies, Prof. Gursel Dino has contributed to my professional and personal development in many ways. I consider myself extremely lucky to have the opportunity to work with her. Having the chance to work with Prof. Dr. Meral Akgül was a great privilege. Our stimulating conversations have motivated me and contributed to my academic development significantly. I am incredibly thankful to Assoc. Prof. Dr. Sinan Kalkan, for accepting being the *ghost supervisor* of my thesis. Our discussions deepened my interest in interdisciplinary studies.

I would like to thank the examining committee members Assoc. Prof. Dr. Funda Baş Bütüner, Assoc. Prof. Dr. Murat Göl, and Assoc. Prof. Dr. Yasemin Afacan for their invaluable feedback.

My dearest friends, Elif Ecem Pala, Defne Çakır, Meltem Kumru, and Abdullah Kırmacı, and Burcu Ün have been a source of motivation during this thesis. We have shared almost everything from tears to laughter for many years. Sezgin Baloğlu deserves to be mentioned explicitly. He was always there to help me whenever I needed and motivated me greatly in my personal and academic pursuits during these years.

My colleagues, Neris Parlak Temizel, Bengisu Derebaşı, Bilge Karakaş, Sezin Sarıca, Ömer Faruk Ağırsoy, and Çağrı Burak Başkol, made my research assistantship at METU a stimulating and cheerful experience.

Finally, my deepest gratitude is due to my family, Şenay, Vedat, and Ayberk. Their presence, unlimited support, love, and friendship are what I will always appreciate the most in my life. I dedicate this thesis to them.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

ABBREVIATIONS

PV	Photovoltaics
BIPV	Building Integrated Photovoltaics
BAPV	Building Attached Photovoltaics
ML	Machine Learning
ZEB	Zero Energy Buildings
DSR	Design Science Research
AC	Alternative Current
DC	Direct Current
POA	Plane of Array
STC	Standard Test Conditions
CAD	Computer Aided Design
3D	Three Dimensional
SVF	Sky View Factor
SEF	Sky Exposure Factor
MSE	Mean Square Error
RMSE	Root Mean Square Error
nRMSE	Mean-normalized Root Mean Square Error
MAE	Mean Absolute Error

CHAPTER 1

INTRODUCTION

Climate change, depletion of natural resources, and global warming are the inevitable outcomes of continuous population growth, rapid urbanization, and economic developments in recent years (Shafique et al., 2020). Currently, slightly more than half of the population worldwide lives in cities, and it is expected to reach 68% by 2050 (United Nations, 2019). An increase in the urban population implies a growing energy demand due to varying needs brought by modernization and change in comfort conditions. United Nations (2020) declared that cities consume close to 70% of the global energy and are responsible for about 75% of the greenhouse gas emissions.

Non-renewable energy sources constitute approximately 64% of the energy demands in the building sector (Gutierrez & Durrant, 2018) Onsite energy generation from renewables and electrification initiatives in the building sector is one of the important strategies to eliminate the dependency on fossil fuels and reducing building carbon footprints. The share of electricity in total building energy demands was around 31% in 2015. According to 2050 target of the International Renewable Energy Agency (IRENA), electricity will constitute 56% of energy use, and the share of non-renewables will be reduced to only 23% of building energy use by 2050 (Gutierrez & Durrant, 2018). The energy transition, a continual process of leaving fossil fuels for clean and low-carbon energy supplies, will change the energy supply and consumption systems with the increasing electrification in the built environment.

Solar, wind, geothermal, and biomass energies are increasingly considered for building applications in this transition process. Among the other alternatives for electricity generation, solar energy has become a key energy resource worldwide since it is readily available and abundant (R. Zhu et al., 2020). PV power generation

has exponentially expanded over the past decades (Ahmed et al., 2020). This expansion can be attributed to the decrease in the price of PV systems, improvement in system efficiency, modularity in installation, economical cost of maintenance and operation, the extended service life, the reduction in carbon emissions, and the environmental friendliness of the technology.

Buildings and Solar Energy

Passive solar design strategies and active solar energy technologies have been developed and involved in design development. Initially, buildings were designed to attain the direct use of solar energy by adjusting their location, geometry, building components, and materials. Deliberate decisions on the selection of site, placement, form, and orientation with a thoughtful arrangement of windows and choice of building materials lead to absorption and storage of solar heat and improvement in daylighting benefits (Hegger, 2003). In recent decades, not only passive design strategies but also active solar technologies have been involved in architectural design strategies for the efficient utilization of solar energy. From being the major shareholder of global energy use, buildings have the potential to generate the energy they demand due to the high potential for harvesting solar energy from their envelope.

Photovoltaics (PV) generate electricity from solar energy absorbed by the solar panels without fuel consumption or pollutant emission (IEA SHC, 1996). From small residential units to large commercial buildings, PV systems are being considered during the early design processes (REN21, 2020). Since the integration of solar technologies through PV into buildings represent an essential branch of approaches for reducing dependency on fossil fuels, these systems are recommended to be applied or even become an obligation in some cases where new buildings are designed in European Union countries (COST Action TU1205 (BISTS), 2015). Similarly, their utilization is required by the zero-energy building (ZEB) target of the International Energy Agency (IEA, 2015).

Urban Contexts and Solar Energy

The increasing rates of urbanization and growing energy demands in cities make buildings the primary target for integrating PV systems. However, urban contexts can alter the energy yield of building-integrated PV (BIPV) systems. Building spatial layout and heights may impact the mutual shading and limit solar-harvesting potential. Studies aimed to establish a link between urban form and solar potential. The relationship between urban block typology (Shi et al., 2021; Zhang et al., 2019), street canyons (Mohajeri et al., 2019), and urban density (Chokhachian et al., 2020) are widely studied in the literature.

Although many aspects of urban morphology have been studied in the literature, urban layouts were rather simplified and standardized in the majority of studies. Recent research often represented urban areas with simplified grid plans characterized by streets crossing each other perpendicularly (Aghamolaei et al., 2019; Chokhachian et al., 2020; Natanian et al., 2019; Poon et al., 2020). Therefore, the accuracy of the representation of real urban environments is questionable. This simplified representation might result in biased results preventing generalizable conclusions about the impact of urban contexts on solar energy potential. Although many aspects of urban morphology have been thoroughly examined for solar energy potential assessment, the impact of street network patterns on solar energy potential remains unaddressed and should be explored in detail.

The Need for Data-driven Approaches in Performance-based Design Tools

Accurate and fast urban scale analyses are needed to assess the solar potential of building envelopes during the design development phases. Immediate feedback from urban scale analyses is essential to aid in design decision-making. Design problems are accepted as ill-defined problems that often refer to unclear problem definitions. For this reason, designing is not a regular problem-solving activity but an interplay of *finding* problems and *solving* them (Cross, 2001). Schön identifies this behavior of design activity as *the reflective practice* in detail (Schön, 1983). In comparison with the other disciplines, a higher number of new *problem structures, or*

formulations, are generated, and alternative solutions are explored by architects (Akin, 2001). It is reported that architects continue to propose alternative solutions even when they have already settled on a satisfactory one. Therefore, there is a need for rapid response from the urban scale solar potential analysis to assess a high number of design alternatives and intervene in design development.

Urban scale analyses require tedious modeling and expertise. As a result, the increase in the number of design alternatives increases the effort required for modeling and computation. Although the energy output of a BIPV module can be rapidly simulated when shadings are not considered, the inclusion of shading analysis significantly increases computation cost and requires detailed 3D modeling of an urban environment. In district-scale assessments, larger city models should be provided, and longer simulation times are needed to assess each city block accurately.

In addition to supporting design development, predicting the variability of renewable resources is critical since they are affected by environmental conditions; hence do not secure constant power (Apt & Jaramillo, 2014). The difficulty with BIPV is that irradiance on an established system is often inhomogeneous due to non-optimal locations or shading of BIPV modules, even when mounted on the same plane. This indicates that a single irradiance measurement or irradiation value for the whole system is not always suitable and the standard definition of performance ratio is difficult to apply (Martín-Chivelet et al., 2022). This variability results in different adverse effects on the electric grid system, such as the stability, reliability, and planning of the operation (Das et al., 2018). Accurate forecasting of BIPV power is critical for stabilizing and securing grid operation, especially in large-scale integration of BIPV systems.

Predictive Models for Solar Energy Potential

A large number of studies focused on predicting PV generation in different time scales with various input and methods, such as numerical weather prediction (NWP), image-based, statistical methods, and artificial neural networks. In terms of the time horizon, short-term projections are needed for power smoothing operations, real-

time electricity dispatch monitoring, PV storage management, and energy marketing or pricing (Raza et al., 2016). Besides the energy bidding and securing operation, long-term PV power forecasting is necessary for planning the electricity production, transmission, and distribution organization (Das et al., 2018). Especially for large-scale BIPV applications, long-term predictions enable the assessment of total energy generation potential and comparative analysis of site selection for BIPV deployments.

PV energy yield can be predicted directly or indirectly. Direct models use historical data, whereas indirect models forecast solar irradiation and then estimate power generation through PV simulation software (Das et al., 2018). Although many studies have utilized NWP, image-based, and hybrid machine learning (ML) techniques for indirect PV forecasting, predicting PV generation indirectly with architectural and urban design parameters has not been thoroughly explored. A prediction method based on parameters of buildings and urban contexts has the potential to inform urban designers, architects, and legislative bodies about the PV potential of buildings in cities in the energy transition process.

Nevertheless, the majority of studies in the literature focused on predicting rooftop PV generation through historical data. Energy generation from facades received considerably less research interest. On the other hand, facades occupy greater surface area in most cities, implying a larger BIPV deployment potential. With the recent trends of decreasing the cost of PV and increasing efficiencies and aesthetical properties, such as a wide range of shape, color, size, and transparency options, PV installments on facades are encouraged. However, facades are more vulnerable to shading problems than rooftops in cities (Martín-Chivelet et al., 2022). PV generation from vertical surfaces in urban areas fluctuates more than from rooftops (Brito et al., 2017). In addition to rooftop PV potential, energy generation from facades should be accurately predicted with architectural and urban design parameters in cities.

1.1 Problem Statement

The energy transition from fossil fuels to clean energy sources encourages electrification in the built environment. The deployment of BIPV systems on the building envelope is one of the most viable solutions for onsite electricity generation. However, urban contexts can limit BIPV electricity generation. Therefore, rapid feedback from urban scale analyses is needed for performance-driven design decisions. Building upon the identified gaps in the literature, this thesis focuses on the following research problems,

- Data-driven methods that facilitate performance-based design choices are needed to assess BIPV applications in urban environments. Indirect prediction models trained with quantifiable design parameters have the potential to be actively used by the architects, urban planners, legislative bodies, and building occupants in the design and post-occupancy period of a building.
- Facades usually have larger surfaces in cities, and developments in PV technology and lower prices encourage BIPV application on vertical surfaces. However, majority of research focuses on rooftop BIPV potential. Prediction models should also have the capacity to evaluate the BIPV potential of façade surfaces. In addition to rooftop BIPV potential, the assessment of façades for energy harvesting is needed.
- Assessment of urban environments based on street network patterns is an unexplored field of study in the literature on the interaction between solar energy and urban form. Although street networks have been partially considered in solar potential studies, to the best of the author's knowledge, street network patterns and their impact on solar energy potential have not been studied thoroughly yet.

1.2 Aims and Objectives

This thesis mainly aims to explore the potential of machine learning methods that can guide performance-driven architectural design approaches. In order to achieve this purpose, the quantitative analysis of urban form indicators is performed and followed by the development of predictive ML models. Several ML techniques were explored to select the best-performing technique for the specified problem. This exploration has potential to result in an extendable and customizable tool available to reuse by architects, urban designers, building occupants, and policymakers.

During the development of the proposed methods, the impact of urban morphology on BIPV generation capacity in cities is also explored. With a specific focus on the street network patterns, four urban areas representing different street network patterns are analyzed comparatively in terms of BIPV energy generation. The aim of this exploration is to identify urban morphologies with higher solar energy generation potential to inform urban development decisions.

The proposed method is developed by analyzing four districts in Ankara, Turkey, with different street network patterns. The solar energy potential of the study areas from façade and roof surfaces and the prediction performance of the ML models are reported. One of the districts, namely the Bahcelievler district, is studied under the scope of a project entitled “Tool development for urban building energy modeling (UBEM) against climate change supported by machine learning and building simulations.” The project is funded by the Scientific and Technological Research Council of Turkey under grant number TUBİTAK 120M997.

In summary, the aims and objectives of this thesis are:

- Identification and analysis of building and context-related factors affecting solar energy generation capacity of buildings in cities.
- Comparative analysis of different street network patterns based on district-scale energy generation capacity through BIPV.

- Development of insight for studied urban areas to support future urban development and district-scale retrofit initiatives that target the use of renewable technologies.
- Development and comparative analysis of ML models that can evaluate design decisions for BIPV energy generation capacity in urban contexts.

1.3 Research Questions

Considering the research gaps and aims of this study introduced in the previous sections, the main research question addressed by this research is:

To what extent can data-driven methods accurately predict solar energy potential of buildings in urban contexts?

To answer this overarching question, four sub-questions posed and addressed are:

What factors affect the solar energy potential of building envelopes in urban environments?

What is the impact of urban morphology, specifically the street network pattern, on solar energy potential?

What would be the impact of district-scale BIPV deployments on meeting building electricity demands?

What is the accuracy with which the solar energy potential is predicted using machine learning methods?

1.4 Scope of the Thesis

This thesis aims to incorporate the broad research areas of urban morphology and PV forecasting with the aims and objectives described above. After shortly introducing these fields, a specific focus on each subject is sustained throughout the

study to ensure clarity and quality of the research. The research is limited to the following aspects of urban morphology and solar potential prediction,

- Street network pattern constitutes the specific focus on urban morphology in the comparative analyses. The selection of cases is primarily related to the identified street network patterns, gridiron, organic, radial, and no pattern.
- Solar potential predictions are limited to predicting the annual power output per BIPV application area (kWh/m^2) for buildings in Ankara. A simulation-based data generation method is used to obtain the data for training the ML models. A single BIPV technology is applied to all buildings to maintain the focus on design parameters. Similarly, the same typical meteorological year weather file for Ankara is utilized in simulations.

1.5 Thesis Methodology and Structure

This thesis conducts design science research (DSR) to address specified research gaps in the literature. DSR is a form of scientific knowledge production to solve problems by innovative constructions (Dresch & Valle Jr., 2015). The problem is formulated as the lack of design support that can inform decision-making by assessing the solar energy potential of buildings in urban contexts. A data-driven method that can aid in performance-based design decisions is proposed to solve this research problem as an innovative construct. The capacity of the solar potential assessment method based on machine learning to address the identified issues is explored. Overall, the thesis methodology can be summarized in four steps; problem identification, intervention, evaluation, and reflection and learning (Cole et al., 2005). These steps are followed throughout this research (Figure 1.1).

Chapters 1 and 2: In the *problem identification* phase, research motivations are introduced. The need for the conducted study is discussed with related works in literature. Building upon the identified knowledge gaps, objectives of the study are determined.

Chapter 3: The *intervention* step is characterized by constructing an artifact to tackle identified problems in the previous step. A method based on ML techniques for evaluating the BIPV potential of building envelopes in cities is developed to address the need for data-driven tools that can aid in performance-based design decisions. In this process, urban areas are modeled and analyzed through quantitative simulations. Simulations outputs are used for training the predictive models based on three ML techniques.

Chapters 4 and 5: In the *evaluation* step, the proposed model/method is applied to the case studies as a proof of concept. The accuracy of predicting the BIPV potential of buildings in four neighborhoods of Ankara is reported at this step. The comparative analysis of different ML models in predicting BIPV potential is presented according to the selected performance metrics.

Chapter 6: Potentials and contributions of the research are discussed in the *reflection and learning* phase by revisiting the research questions and aims of the study. The applicability and viability of the proposed method are articulated.

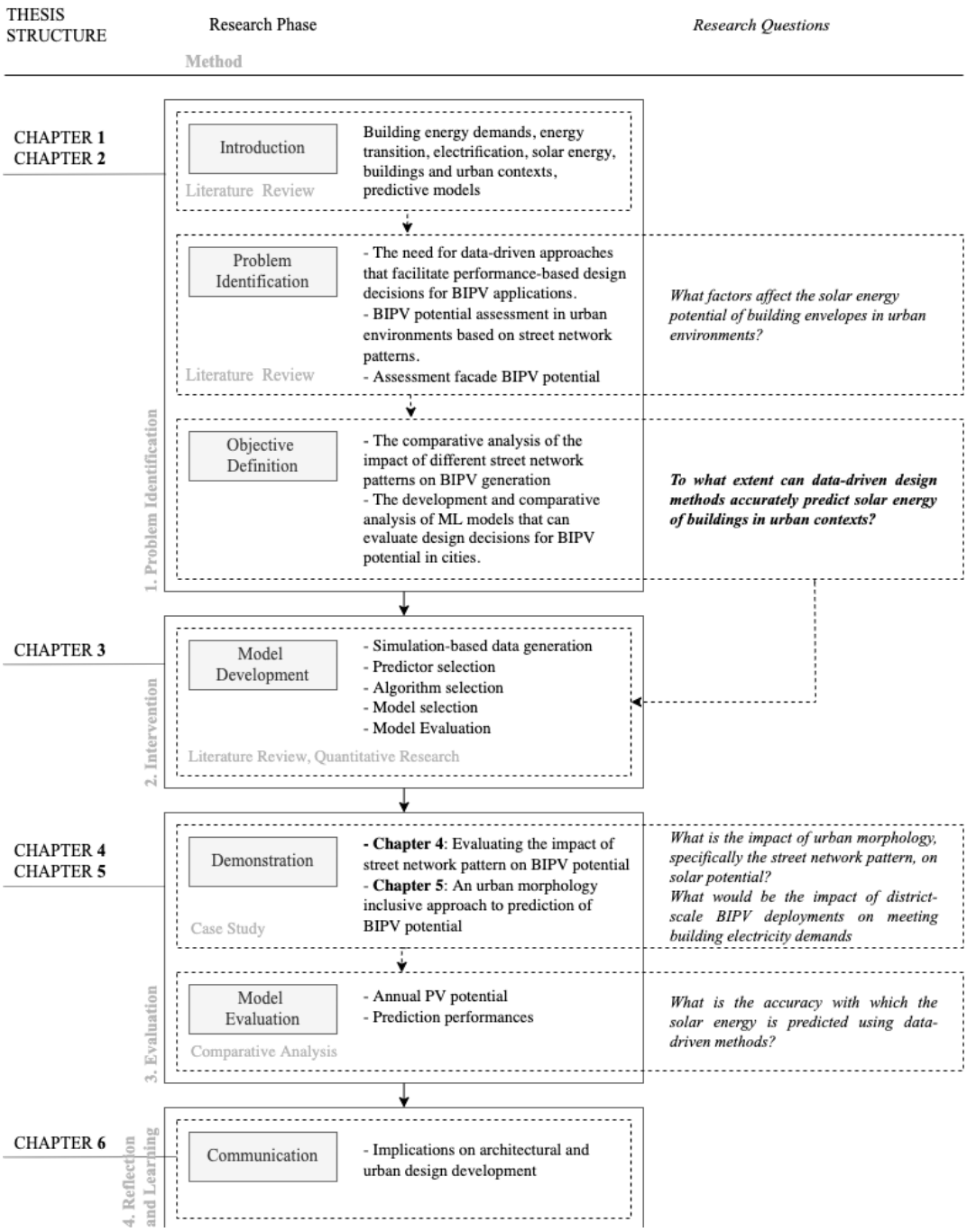


Figure 1.1 Research phases, related research questions, methods and chapters.

CHAPTER 2

LITERATURE REVIEW

In this chapter, solar energy and buildings, the application of active solar systems to buildings, factors affecting BIPV potential, and methods for estimating energy generation from PV systems are examined with the previous research conducted on these subjects. Reviewed concepts and their relationships are illustrated as a diagram in Figure 2.1.

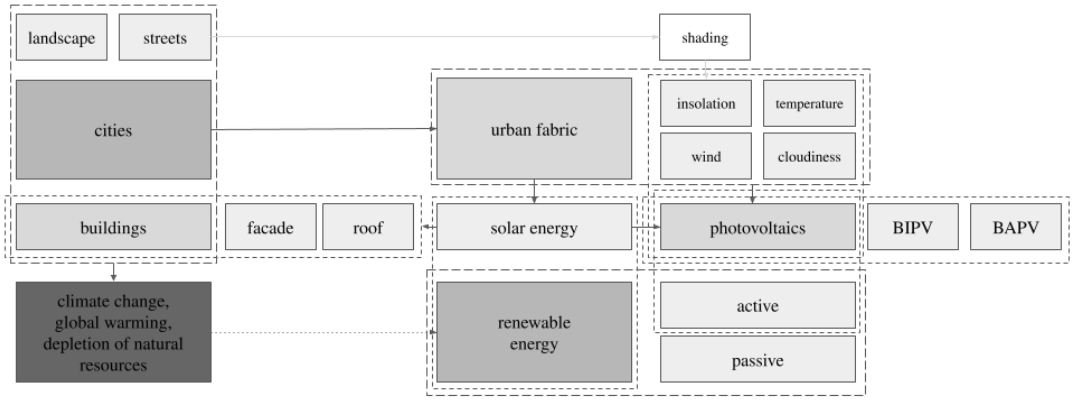


Figure 2.1 Diagram of reviewed concepts and their relations.

2.1 Solar Energy and Architecture

The sun, on which we depend for light and heat, provides a tremendous amount of energy, equal to approximately 1.8×10^{11} MW, over the atmosphere (Wengenmayr & Bührke, 2011). The energy received from radiation components (Figure 2.2) by the landmasses of the earth is 3000 times more than the amount of energy demand worldwide (Schittich, 2012). However, the energy demand is mainly met by fossil fuels which are finite and pollute the environment.

With the depletion of these limited resources and the sociopolitical conditions of recent years, control over energy has become more prominent. Solar energy appears to be not only a sensible environmental solution but also a contributor to world peace (Schittich, 2012). Ways to benefit from this abundant and free energy source, in addition to its environmental and economic benefit, are searched by many different parties, including investors, governments, engineers, city planners, and architects.

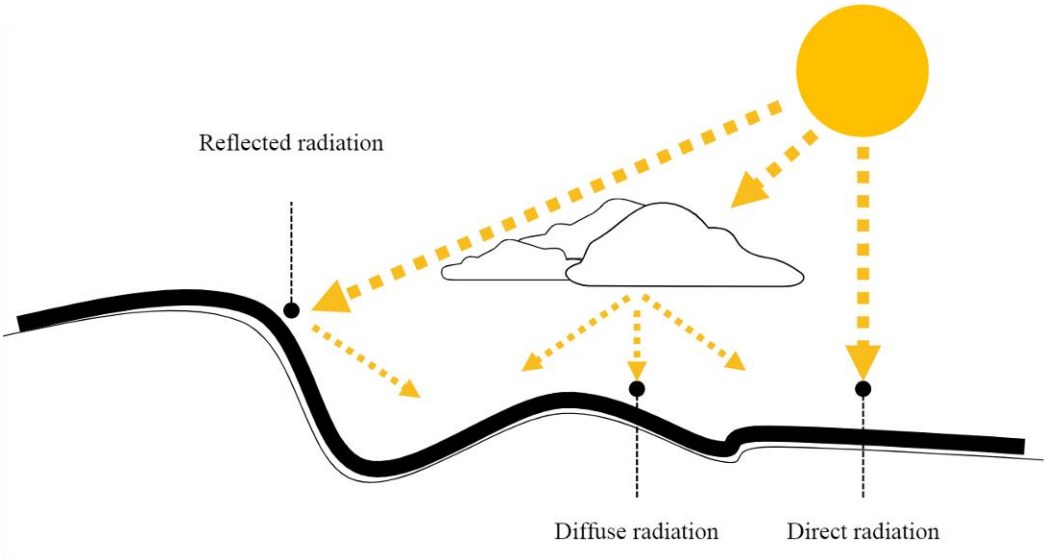


Figure 2.2 Diagram of solar radiation components (Hachem-Vermette, 2020).

Buildings are one of the primary energy consumers worldwide. Reducing energy consumption and even moving from being energy consumers to energy generators is necessary to mitigate climate change and lower carbon emissions in the building sector through architectural design considerations. *Solar architecture* evolved as a response to these problems, originating from the design approaches of Ancient Greek to the integration of solar energy technologies today (Butti & Perlin, 1980).

Design approaches that benefit from solar energy can be divided into two main groups, passive and active, complementing each other (Schittich, 2012). Passive design strategies include adjusting building location, geometry, building components, and materials to utilize the solar energy directly by absorption and storage of solar heat and improvement in daylighting benefits (Hegger, 2003). Active

solar design approaches feature technologies that convert solar energy into a more useable form of energy, such as hot water and electricity. The term *solar building* is proposed for a building that employs various methods, such as windows and sunspaces or solar water collectors and photovoltaics, to harness solar energy (Hestnes, 1999). It should also be highlighted that solar architecture is not restricted to interventions in a particular building operation. In addition to architectural design decisions, city planning and urban design approaches and material choices have a major influence on solar potential. (Schittich, 2012).

Zero Energy Building (ZEB) concept, with various definitions, has been introduced to signify mainly buildings with zero energy consumption (Marszal et al., 2011). Achieving *zero* in buildings is suggested to be possible with appropriate architectural design and an airtight building envelope. Reducing energy use is followed by utilizing renewable energy concepts (IEA, 2015). In 2008, IEA Task 40 / Annex 52 "Towards Net Zero Energy Solar Buildings" was initiated for five years by IEA to apprise and spread the NetZEB concept in the building sector. Meanwhile, the European Commission and Parliament stated that the buildings after 2019 should be Nearly Zero Energy Buildings" in 2010 (EU, 2010). The European Union countries are obliged to develop national plans to adopt this directive. Consequently, PV electricity generation has been the common property of leading NetZEB examples (Musall et al., 2010). These developments and initiatives strongly encourage the design of solar buildings that integrate active solar energy technologies and passive solar design approaches.

Integration of solar energy systems into architecture requires interaction between various disciplines, using new technologies, and recognizing architectural and aesthetic considerations (Wall et al., 2012). To tackle these issues, IEA commenced the IEA SHC Task41 "Solar Energy and Architecture" in 2009. The task comprises three subtasks: architectural quality, tools and methods for architects, case studies, and communication guidelines based on the barriers detected in a*chieving solar energy in architecture* (Wall et al., 2012). It is claimed to be both a challenge and an opportunity for architecture to employ these technical and functional systems while

achieving an architecturally satisfying comprehensive concept (Schittich, 2012). In this regard, this thesis aims to contribute to the development of solar energy-generating extensive design concepts by proposing a method that architects and urban designers can use in design development phases.

2.1.1 Electrification in the Built Environment

Electrification in the built environment also encourages the deployment of solar panels increasingly. Fossil fuels (oil, coal, and natural gas) have been the primary source of electricity generation over the centuries, and they continue to play an important role today. Nonetheless, consumption of these finite resources generated environmental concerns over carbon emissions and climate change. When fossil fuels are burned, carbon dioxide and other hazardous chemicals are released into the atmosphere, significantly contributing to climate change and environmental pollution.

Electrification is the transition process from using solid, liquid, or gaseous fossil fuels (such as natural gas or fuel oil) to using electricity to power various end uses (Deason et al., 2018). Fossil fuels are mainly used in buildings for heating and cooking (Gutierrez & Durrant, 2018). Gas-fueled furnaces and household water heaters or boilers may be replaced with heat pumps driven by electricity in buildings. Electricity-driven modern cookstoves are potential alternatives to cooking with traditional biomass.

Solar energy appears to be one of the key solutions for decarbonizing and electrifying the built environment. Electricity prices from utility-scale solar photovoltaics decreased by 85% between 2010 and 2020 (Irena, 2021), making the technology more cost-competitive. Currently, solar and wind energy are cheaper energy sources than fossil fuels. How to accommodate large proportions of variable renewable energy into power grids, the role of electrification, solutions for decarbonizing heating and transport demand, and more integrated long-term planning of energy

systems become increasingly important with the energy transition (Gutierrez & Durrant, 2018).

2.1.2 Solar Energy Potential for Buildings

Assessment of solar energy potential is needed for performance-driven design decisions to improve building environmental performance and support planning of future energy systems in the energy transition process. Researchers have identified four key metrics for solar energy potential assessment in studies considering urban form and buildings: PV generation, solar thermal yield, solar irradiation, and load match index (Natanian et al., 2019). PV generation capacity, also referred to as PV potential, can be described as the overall electricity energy output of a PV system (kWh). The solar thermal yield is the amount of energy harvested from the sun to heat water (kWh). Solar irradiation is the summation of energy received by a surface per square meter area (kWh/m^2), which can be considered as the most rudimentary metric since all the others depend on solar irradiation. Lastly, the load match index (LMI) indicates the temporal coverage of energy generated to the energy demanded on-site (%) for selected time resolution, such as hourly, daily, monthly, or yearly.

This thesis focuses on electrical energy generation from urban BIPV systems. Therefore, the first and the last metrics, BIPV generation and LMI, are further examined. Total BIPV generation is simply the summation of hourly energy generation from a BIPV system. It can also be normalized by dividing to the total deployment area. LMI provides an extensive understanding of how much of the energy generated on-site can meet the building energy demands (Shi et al., 2021) and how much energy is needed from the grid. Total BIPV generation E_i and LMI (Sartori et al., 2012) can be calculated with the following equations,

$$E_i = \sum_{year} G_i(t) \tag{1}$$

$$LMI_i = \frac{1}{n} \times \sum_{year} \min\left[1, \frac{G_i(t)}{L_i(t)}\right] \quad (2)$$

where *G* and *L* are energy generation and load, respectively; *i* is the energy carrier, and *t* is the time interval, such as hour, day, or month. *N* represents the number of data samples. For instance, 12 for the monthly or 8760 for the hourly time interval.

PV system generation and grid operation can also be quantified in many ways. Salom et al. (2011) systematically reviewed the load matching and grid interaction indicators for NetZEB buildings. Researchers identified four metrics which focus on the onsite energy load and generation under the group of load match indicators: LMI, solar fraction, load cover factor (LCF) and self-consumption factor. These metrics aim to indicate how much energy load can be met by onsite generation and how much energy is needed from the grid. Similar to LMI, LCF is calculated with the following equation,

$$LCF_i = \frac{\sum_i^{i+N} \min[G_i(t),\ L_i(t)]}{\sum_i^{i+N} L_i(t)} \quad (3)$$

2.2 Active Solar Technologies

Active solar systems transform solar energy into usable thermal and electrical energy sources. PV technologies, solar thermal collectors, and hybrid PV/thermal collectors constitute widely utilized active solar technologies (Hachem-Vermette, 2020). Solar thermal collectors generate thermal energy from solar radiation, while PV systems generate electrical energy. This research concentrates on the electric energy generation potential of buildings; therefore, building-integrated PV systems are introduced in detail.

2.2.1 BIPV Applications for Buildings

PV applications in buildings are divided into two main categories: building-integrated PV (BIPV) and building-adapted PV (BAPV). Traditional building materials replace PV panels in BIPV applications, whereas modules are added later onto the building envelope in the BAPV applications (Figure 2.3). While attached systems can be a practical consideration for retrofit scenarios, integrated systems can become a part of building skin with a design perspective leading to a comprehensive concept in the early design phase (Schittich, 2012). The following section presents BIPV applications, having considered PV systems integral to a design concept.

Figure 2.3 PV integration gradient, redrawn from Peronato (2019)

Roof Integrated PV Applications: Roofs generally receive a tremendous amount of solar radiation. The interest in rooftop BIPV installations in residential or low-rise buildings is increasing with the developments in PV technology and the drop in prices (Hachem-Vermette, 2020). Although rooftops contain space-occupying elements such as chimneys, mechanical ventilation, or air-conditioning units that reduce the usable area for PV installation, and increasing variety of PV unit sizes overcome this problem. Demonstrative applications can be found in Figure 2.4.

Facade Integrated PV Applications: Whole building façade can be designed with a BIPV system where BIPV cells are used as cladding materials with various shapes, opacities, colors, and energy efficiencies. In Figure 2.5, three BIPV applications on facades are illustrated. It should be noted that façade-integrated PV systems produce less energy than roof-integrated PV systems because of the tilt angle of BIPV panels

(Hachem-Vermette, 2020). The factors affecting BIPV energy output are discussed more in detail in the following section.

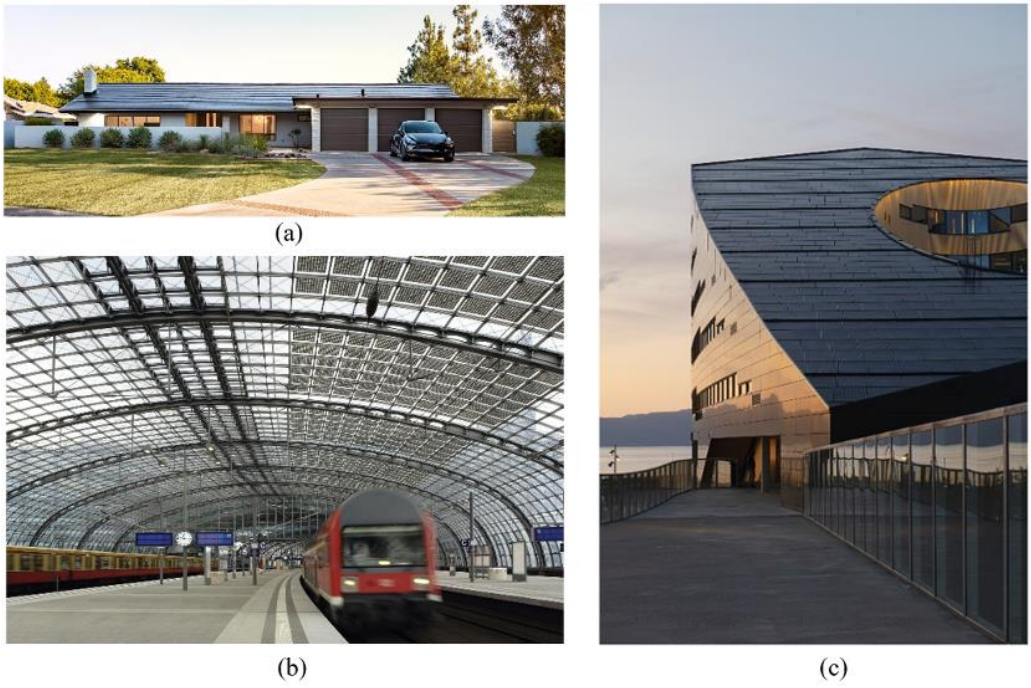


Figure 2.4 (a) Tesla solar roof (Tesla, 2020), (c) Powerhouse Brattørkaia, Trondheim, Norway (Snøhetta, 2019), (b) Berlin Central Station, Berlin (Meinhard von Gerkan and Jürgen Hillmer, 2006)

Other Integrated PV Applications: Solar panels can substitute traditional materials used for shading devices, spandrels, balcony parapets, railings, and also windows and skylights, thanks to advances in PV technology enabling colored and transparent, and semi-transparent solar panels (Ghosh, 2020).

2.2.2 PV Technologies

PV systems convert solar radiation into direct current (DC) electricity with semiconducting materials in PV cells through the photovoltaic effect (IEA, 2012). Direct, diffuse, and reflected radiations contribute to the generation of power output

from solar cells. The DC output is converted into alternating current (AC) by inverters to be used by buildings appliances (Hachem-Vermette, 2020). The proportion of the energy of solar incident on the cell surface converted into electrical power defines the efficiency, i.e., performance ratio, and it is calculated under specific lab design conditions, called standard test conditions (STC) (Shubbak, 2019).

Figure 2.5 (a) Nursery +E in Marburg, Germany, (Opus Architekten, 2014) , (b) EFH_125 in Bregenz, Austria (k_m architektur, 2001), (c) SwissTech Convention Center at EPFL Campus in Ecublens, Switzerland (Richter Dahl Rocha & Associès, 2014).

Common types of PV cells based on the technology used can be found in Table 2.1. Crystalline cells are the dominating technology, with a 93% share among the other types of technologies (Shubbak, 2019). They are mainly used for civil applications and are the most common type of technology used in buildings. On the other hand, new technologies, such as transparent PV cells and biophotovoltaic panels, emerge and enrich the variety of architectural applications (Hachem-Vermette, 2020).

Table 2.1 Typical electrical efficiency values at STC (η_{STC}) for standard opaque PV technologies (Martín-Chivelet et al., 2022).

PV Technology	η_{STC}	
	Typical	Range
Mono-crystalline silicon (m-Si)	0.157	0.084–0.221
Poly-crystalline silicon (p-Si)	0.149	0.062–0.204
Silicon heterostructures (HIT)	0.173	0.126–0.197
Amorphous silicon (a-Si)	0.064	0.053–0.088
Micromorphous silicon (a-Si/μc-Si)	0.082	0.063–0.104
Cadmium telluride (CdTe)	0.139	0.094–0.170
Copper indium (gallium) selenide (CI(G)s)	0.117	0.055–0.167

Moreover, PV systems can be stand-alone or grid-tied based on their relation to the utility grid. While stand-alone systems are not connected to the grid and store excess energy, grid-tied systems supply the utility grid with the energy generated (Hachem-Vermette, 2020). Stand-alone systems can meet the electricity demand in rural areas where connection to utility power is not provided. On the other hand, grid-tied systems can help reduce electricity demand for fossil fuels in cities by feeding excess energy back onto the utility grid. Therefore, both methods have advantages based on the context and should be considered during the design process.

2.3 Factors Affecting BIPV Potential

A BIPV system is electrically very similar to a "normal" PV system (Martín-Chivelet et al., 2022). However, due to building-specific boundary circumstances, the electrical performance of a BIPV system may vary from that of a PV plant. BIPV systems are more susceptible to partial shading, have inferior back ventilation, and their irradiance is more impacted by reflections from neighboring surfaces. In addition, non-optimal tilts may have an effect on angle-dependent and/or soiling losses. Martín-Chivelet et al. (2022) discussed the loss types in three categories. The first group consists of system-related factors, such as solar spectral changes, optical reflections, surface soiling, wiring, and losses in electronics. Researchers identified some losses as site-dependent, which are related to meteorology and latitude. Factors other than these two groups are considered as design-dependent and associated with suitable design, mounting, and maintenance variables.

Using a similar approach, but with a stronger emphasis on design-dependent elements, the factors influencing BIPV performance are categorized under three themes, climate and location, PV system, and buildings and the urban setting. (Figure 2.6). It is important to note that the boundaries of these groups are flexible due to interactions between different themes.

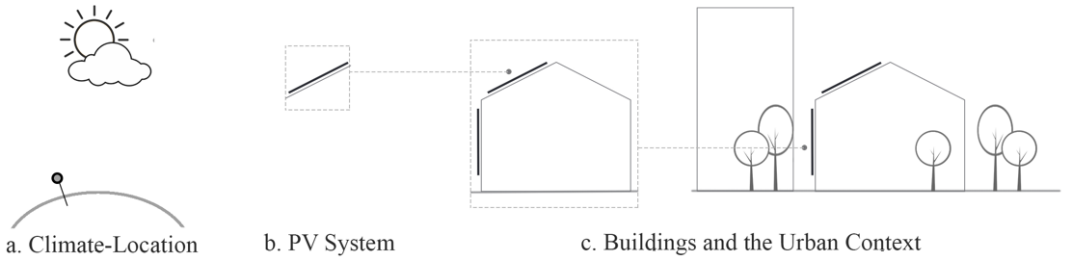


Figure 2.6 Factors influencing PV potential.

2.3.1 Climate and Location

Daily and seasonal variations in the solar angles (Figure 2.7 b), relative humidity, cloudiness, and air quality can be the factors in this group (Hachem-Vermette, 2020). The most crucial climate-dependent factors are concerned with incoming solar radiation since only a small proportion of the solar radiation energy incident on a PV system's surface is converted into electricity. Therefore, among the other climate-dependent parameters, such as humidity, dry bulb temperature or wind, solar radiation is reported to be the most significant factor influencing the power output of a PV system (Das et al., 2018).

The location of the PV system on earth can partially impact the solar incident on a surface. An identical system design in a different location can yield different energy outputs. This result can be explained by the angle of latitude (φ), which designates the latitude of a place. While the Equator has a φ of 0°, poles have ± 90°. A suggestion for optimal tilt angle simply states that the tilt angle should be equal to latitude during spring and fall, whereas it should be -15° to latitude in summer and +15° in winter seasons (Duffie & Beckman, 1980) to maximize energy yield.

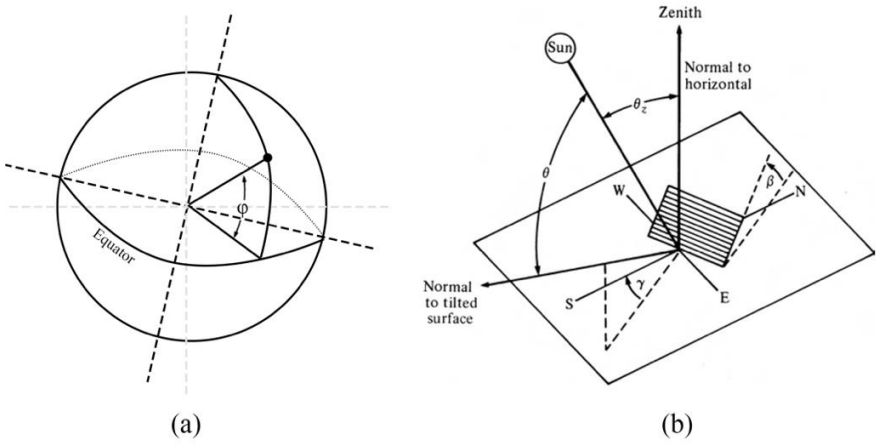


Figure 2.7 (a) Angle of latitude of a point (φ), (b) Solar angles, including angle of incidence (θ_z), zenith (θ_z), tilt (β), and azimuth (γ) for a tilted surface (Twidell & Weir, 2006).

The zenith angle describes this relationship with the sun and a given location. It represents the angle between sun rays and the vertical direction. Therefore, depends on the time of the day. Midday offers the highest potential for PV generation in most cases. The azimuth angle is a climate-based factor determining the direction from which the sunlight is coming to a surface; in other words, it represents the orientation of a surface. Although PV generation is possible with diffuse and reflected radiation, direct radiation is most desired. Therefore, in the northern hemisphere, West, South, and East facing surfaces are more convenient for PV application.

In addition, different physical aspects of climate, such as temperature, humidity, rain, wind, fog, snow, cloudiness, solar irradiation, and air quality determined by pollution, directly or indirectly, affect the solar energy generation potential of buildings in urban contexts. For instance, the heating of solar cells affects electrical performance and power output adversely (Hachem-Vermette, 2020). Therefore, high temperatures are undesirable for higher solar energy potential. Urban Heat Islands (UHI) effect observed in dense metropolitan areas poses a risk of decreased power output by increasing cell temperatures (Berardi & Graham, 2020). Wang et al. (2006) reported that due to the UHI effect and attenuation of solar radiation brought by air pollution, the daily PV power output in urban areas could be declined by 13% compared to rural areas. On the other hand, local wind speeds areas might help the cooling of PV cells.

2.3.2 PV System

PV system output is primarily affected by PV technology (see Sect. 2.2.2). Due to the design of the system, only around one-fifth of solar radiation incident is converted into electricity, while the rest is converted into heat. The heating of solar cells affects both the electrical performance and power output adversely (Dubey et al., 2013). BIPV cell temperatures are more significant than the air when coupled with the building absorptivity of building materials.

Previously mentioned specific lab design conditions, i.e., STC, are applied to test and quantify the efficiency of PV devices. STC is characterized by the insolation intensity of 1 KW/m^2 perpendicular to the PV surface with a cell temperature of 25 °C and spectral irradiance representing sunlight when it is at the highest position in the sky (IEC, 2016). Nevertheless, cell temperatures are generally above 25 °C. The temperature coefficient of solar cells depends on the technology, and crystalline silicon is the most sensitive technology to temperature increases (Martín-Chivelet et al., 2022). For amorphous and crystalline silicon PV, a one °C increase in cell temperature results in a 0.2% to 0.45% drop in efficiency (Kalogirou, 2014).

The majority of the factors affecting PV performance also affect cell temperature. The incoming solar radiation, which is essential for energy generation, might increase cell temperature and lower the power output. In addition, the PV efficiency drops due to degradation caused by high temperatures over the years. NREL of the US has reported that the performance of PV cells drops 0.5% every year on average (Jordan & Kurtz, 2012), and the lifetime of a BIPV installment is guaranteed for 25 years of power production (Peng et al., 2011).

Partial shading caused by nearby urban elements affect design of the BIPV systems. Under particular situations, shaded cells might be reverse-biased, therefore dissipating the energy created by the other cells in the string (Martín-Chivelet et al., 2022). To prevent such hot spots, all modules are equipped with bypass diodes. As partial shading happens more often in several BIPV systems, these bypass diodes must be carefully selected to provide a long lifespan even under frequent partial shading situations. Connecting PV modules with varying irradiances may lower the string current to the module current corresponding to the lowest irradiance. In general, all parallel strings should have the same voltage level under all working conditions, and in some circumstances, fuses or string diodes are required to prevent reverse bias of parallel strings.

On-site energy generation can match temporal energy loads. In grid-tied and off-grid systems, the surplus energy can be stored within batteries with the help of an inverter.

System losses during DC to AC conversion can also be included in this group. An inverter aims to convert DC power output to AC with the highest possible efficiency while keeping the voltage on the AC side constant (Kalogirou, 2014). Therefore, inverter selection also influences the amount of usable power output for daily use.

2.3.3 Buildings and Urban Contexts

BIPV systems are more susceptible to partial shading, have poor back ventilation, and their irradiance is more influenced by reflections from neighboring objects (Martín-Chivelet et al., 2022). Factors related to buildings and urban contexts that directly affect the shading of BIPV systems are presented in a more in-depth literature review. Several recent publications on solar energy and urban design are chosen for this evaluation. Articles that have mentioned at least one keyword in both categories (Table 2.2) in their abstract are examined. Overall, 22 articles published between 2012 to 2022 are reviewed (Aghamolaei et al., 2019; Ahmadian et al., 2021; Boccalatte et al., 2022; Chatzipoulka et al., 2016; Chokhachian et al., 2020; Lau et al., 2017; Lobaccaro et al., 2019; Mahaya et al., 2022; Martins et al., 2016a; Mohajeri et al., 2016; Morganti et al., 2017; Nakazato et al., 2021; Natanian et al., 2019; Poon et al., 2020; Ramkumar et al., 2019; Sarralde et al., 2015; Shi et al., 2017, 2021; Tian & Xu, 2021; Vulkan et al., 2018; W. Wang et al., 2021; Zhang et al., 2019).

Table 2.2 Categories and keywords for the reviewed publications.

Category	Keywords
1	Urban form, urban morphology, urban morphology indicators, morphological parameters, urban fabric.
2	Solar energy, solar access, solar energy potential, solar potential, PV potential, PV generation, PV energy/power generation, Energy load match, solar performance

Figure 2.8 Frequency of parameters studied in articles on solar energy and urban form (Only parameters with frequency > 1, vv: viceversa, l: length, h: height, w: width, N: number).

The frequency of studied parameters in relation to categories of buildings or urban contexts can be found in Figure 2.8. During this investigation, it is noticed that, although there are several well-established parameters, such as window to wall ratio (WWR), floor area ratio (FAR), and plot ratio, researchers have also experimented with new metrics derived from simple measures, such as façade to roof area ratio or nearest neighbor ratio. However, if a parameter is observed in only one study, it is not included in the summary figure. In addition, although the calculation methods were the same for some metrics, parameters were named differently. More frequent naming is preferred in such cases, and the figure is presented with the dominant naming convention.

Overall, the examined parameters are grouped into the building and urban context categories. Identifying the difference between building-related and environment-related factors was a difficult classification problem. If a parameter can only be measured with building-related basic measures, such as building height, width, or area, it is counted in the building-related parameters. However, if a parameter can only be quantified with the consideration of the urban land or neighboring buildings or elements of urban landscape, then it is included in the parameters related to the urban context.

2.3.3.1 Buildings

Building-dependent factors can be related to architectural design decisions. Parameters relating to building geometry, such as building typology, orientation, and dimensions of the footprint geometry, have been investigated most frequently in the building category based on the examined studies (Figure 2.8). Consideration of BIPV in the conceptual design phase is critical. The decision on application areas, such as roof, wall, or window, alters many other factors that consequently constrain the available PV technology and the PV performance.

Building typology was one of the most frequently studied metrics related to buildings. Natian et al. (2019) studied five building typologies; courtyard, scatter, slab in two orientations, and high-rise. They have found that courtyard typology provided the highest average LM index values. Shi et al. (2021) created a pool of blocks with 178 different and irregular blocks, classified in the inventory of 18 block typologies. Block typology identified with shop houses resulted in the highest solar energy penetration, an equivalent metric to LMI.

Building typology can also affect the usable area for PV applications. A recent study compared six roof types: flat & shed, gable, hip, gambrel &mensard, cross/corner hip & gable, and complex in terms of the functional rooftop area (Mohajeri et al., 2018). For instance, the usable roof area to building footprint area is the highest on gable roofs compared to other roof types. The roof design can influence the placement, tilt, and orientation of solar panels, which in turn affects the overall solar energy potential. The available area dedicated to PV installment can also be affected by decisions related to the façade design. For exterior wall applications on façades, the window-to-wall ratio (WWR), arrangement of the windows on façade, and complexity of the façade geometry can be the major factors affecting the suitable area for BIPV application. For roof applications, roof pitch and the position of the chimney and air-conditioning units can be the major factors influencing solar energy generation.

In terms of overall capacity, the shape factor and the ratio of the total building volume to the surface area may indicate the greatest power production potential based on the maximum accessible surface area for BIPV application. Moreover, decisions on the building form can narrow down the suitable PV technology. For instance, while flat BIPV panels are appropriate for large flat surfaces, small BIPV tiles would be better integrated into curved forms.

2.3.3.2 Urban Contexts

Parameters quantifying the features of urban context constitute the largest portion of studied metrics (Figure 2.8). In an urban setting, distance to neighboring buildings or trees, urban block form, site coverage, neighbor ratio, and albedo of neighboring surfaces influence the solar potential of buildings (Boccalatte et al., 2020; Mohajeri et al., 2016; Natanian & Wortmann, 2021; R. Zhu et al., 2020). Those also impact cell temperature and efficiency, exhibiting complex relationships depending on the context in which the building and PV are located (Natanian & Wortmann, 2021).

Urban form, revealing the impact of context and neighboring buildings, has been an important variable affecting solar energy potential. From dispersed to compact neighborhoods, researchers have found that energy production capacity decreases from 94% to 79% for roofs and 20% to 3% for facades (Mohajeri et al., 2016). A recent study examined ten urban morphology parameters by conducting a correlation analysis with roof and facade irradiance (Poon et al., 2020). Researchers have found that SVF, a measurement of "the fraction of the overlying hemisphere occupied by sky" (Oke, 1981), correlates significantly with solar irradiation. SVF and SEF are the two metrics that focus on the relationship between the urban contexts and the sky. SEF was determined as the ratio between the solid angle of the sky patch visible from a certain position and the solid angle of the hemisphere centered at the same place (Zhang et al., 2012). Poon et al. (2020) have reported that both metrics have similar correlations with façade irradiance.

Researchers have also identified the urban morphology indicators that lower solar energy potential. A recent study has shown that a simple parameter, building height, decreases the solar energy potential of the surrounding buildings by 15% (Lobaccaro et al., 2019). Similarly, Natanian et al. (2019) reported that a higher floor area ratio and lower distance between buildings negatively affect PV production.

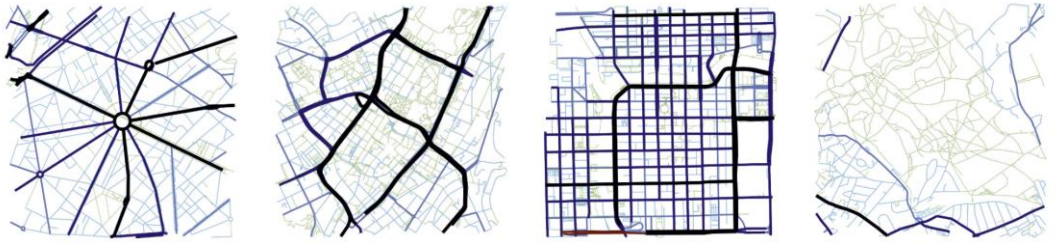
Street geometry has also attracted research interest in urban morphology and solar energy potential studies. Mohajeri et al. (2019) studied 1600 street canyons in terms

of orientation, width, length, SVF, and asymmetric aspect ratio. The received yearly solar radiation by street surfaces and facades is strongly influenced by street orientation. Researchers have found that for surfaces, the most radiation is received from roadways aligned with WNE-ESE, whereas the highest radiation is obtained from facades facing SSW. Another study has conducted sensitivity analysis between nine parameters; shape factor, floor area ratio, plot ratio (the ratio of floor area to plot area), aspect ratio (the ratio of building height to the width of the distance between buildings), verticality, the distance between buildings, and building width, height, depth (Martins et al., 2016b). The results revealed that the aspect ratio and the distance between buildings have a substantial effect on solar energy potential, implying the importance of street networks. In addition, street networks were also indirectly considered for the solar potential. For instance, researchers have considered street network density when selecting the size of the studied urban area for solar energy potential assessment (D. Zhu et al., 2020).

Street network pattern of an urban area is also connected to the street width, building density, plot ratio, and building interval parameters which affect solar potential. Street network pattern has distinguished itself from other urban morphology parameters focusing on various details of roads such as width, aspect ratio, intersection density, or orientation since it involves the visual and intuitive perspective of humans different from traditional metrics (W. Chen et al., 2021). Four street network patterns have been frequently observed: gridiron, organic, radial, and no pattern (Snellen et al., 2002). Biljecki et al. proposed a method based on deep learning for visually classifying road networks into these four distinct categories (Figure 2.9).

The gridiron pattern can be characterized by the streets crossing each other perpendicularly, while in the radial pattern, roads direct movement to a central point. Urban areas with organic patterns can be identified with streets in the form of curved polylines, designating the natural growth of the paths and roadways in the city. On the other hand, no pattern regions are urban areas that cannot be represented using the previously established patterns. Researchers have studied nine cities worldwide

to uncover underlying patterns in their road networks through clustering with a classification accuracy of 87.5% and aimed to explore its relation with urban viability. Similarly, Hachem-Vermette (2020) considered the impact of site layout in relation to road network patterns on solar radiation and urban resilience. Three hypothetical layout options are generated with rectilinear, radial, and hexagonal street organizations while keeping the building density fixed.



(a) Radial (Paris). (b) Organic (Stuttgart). (c) Gridiron (Chicago). (d) No pattern (London).

Figure 2.9 Four road networks are classified by the tool developed by W. Chen et al. (2021)

In solar potential studies, most of the recent research focusing on urban morphology has considered urban areas with gridiron street network patterns (Aghamolaei et al., 2019; Chokhachian et al., 2020; Natanian et al., 2019; Poon et al., 2020; Vermeulen et al., 2018). However, this might result in biased results preventing generalizable conclusions about the impact of urban contexts. Although street networks have been partly considered in solar energy potential studies, such as road ratio, street width, or building interval, to the authors' knowledge, street network pattern and its impact on solar potential, there is a knowledge gap in the assessment of different configurations.

2.4 Estimating BIPV Energy Output

Variability of the energy generated from BIPV systems requires accurate forecasts at different spatial and temporal scales for cost-efficient integration of these systems into the grid. In this section, firstly, major electrical performance models for BIPV

are introduced. Many BIPV software tools utilize the described electrical performance models. In the following subsection, ML techniques for BIPV power forecasting will be elaborated further. Different forecasting approaches and related work will be discussed.

2.4.1 Electrical Performance Models

The theoretical efficiency and energy output of solar cells can be calculated with several methods. The general working principles of the accurate and straightforward models available for building applications are investigated. Those potential calculation models are mainly based on the plane of array (POA) irradiance, cell, and module temperature or circuit design (M. Wang et al., 2021).

First, the efficiency of a system can be derived from the conversion efficiency, which is the ratio of the solar energy received by the panels to the system's power output. The conversion efficiency is calculated at the previously mentioned STC, considering the area of POA and the POA irradiance, as shown in the following equation.

$$E = \eta \cdot A \cdot I_{POA}$$

where E is the DC power output, η is the conversion efficiency of the PV module, A is the area of the PV array, I_{POA} is the irradiance on the plane of the PV array.

As module conversion efficiency is affected by the cell temperature and the STC is not valid on the site application, a correction is integrated into the previous calculation to make more accurate estimations in the second model. Conversion efficiency is adjusted based on the temperature difference between the operational conditions and STC using temperature coefficient γ (Dubey et al., 2013).

$$\eta = \eta_{STC} \cdot [1 + \gamma(T_{cell} - T_{STC})]$$

The Sandia Array Performance Model commonly estimates cell and module temperature with two separate equations. The equation for module temperature

T_{module} takes ambient air temperature T_a, solar irradiance incident on module and wind speed WS at 10-meter height into consideration together with empirically determined coefficients a and b (King et al., 2003). Later, cell temperature T_{cell} is estimated with the module temperature T_{module}, reference solar irradiance E on module surface and temperature difference ΔT between the cell and module back surface at the reference solar irradiance on the module E_0 of 1000W/m^2.

$$T_{module} = E \cdot \{e^{a+b \cdot WS}\} + T_a$$

$$T_{cell} = T_{module} + \frac{E}{E_0} \cdot \Delta T$$

The third model is characterized by an electrical equivalent circuit representation of a PV cell, and the most common circuit-based model is the one-diode model (Bader et al., 2019). A diode is a specialized electrical component that allows current flow in only one direction. However, electrical properties affecting PV performance are beyond the focus of this study.

Furthermore, several computer software tools blend different BIPV performance models and use open-source libraries. PVWatts is one of these software offering an online calculation tool developed by the National Renewable Energy Laboratory (NREL) of the US Department of Energy (NREL, 2014). It takes two types of inputs; field and advanced. While field inputs include system size (DC output in kW), module type (PV technology), system losses (PV efficiency in %), array type (e.g., fixed roof mount, fixed open track, etc.), tilt angle (degrees), and azimuth angle (degrees); advanced inputs include DC/AC ratio, inverter efficiency, and ground coverage ratio as user-defined inputs (Dobos, 2014).

Researchers from Sandia National Laboratories conducted research to compare the measured and calculated PV-system performances of several models, including the described four models. The described models are reported to have around 10% error or less (Cameron et al., 2008). For easy and fast evaluation of each design alternative,

PVWatts integrated into environmental design software Ladybug Tools will be used in the further stages of this thesis study.

2.4.2 Machine Learning Models

The definition of ML includes computational approaches for accurate forecasts or performance enhancement (Mohri et al., 2018). Several learning types are associated with various tasks, such as classification, regression, ranking, clustering, and dimensionality reduction. Classification is used to assign an item to a category, regression predicts the actual value of an item, and ranking is used to order items (Mohri et al., 2018). Depending on the learning tasks, learning scenarios can also vary. Supervised, unsupervised, semi-supervised, and transductive inference are the main types of learning scenarios. In supervised learning, a predictive model is built based on the known input and output data to predict unknown outputs, and it can handle classification, regression, and ranking tasks (Mohri et al., 2018). BIPV power prediction can be achieved with supervised learning techniques for regression tasks.

Forecasting PV electricity has advanced considerably in recent years (Das et al., 2018). Forecasting methods are grouped under several categories in the literature. The first category focuses on the data source and predicted output by dividing models into categories of direct and indirect. In indirect forecasting models, solar irradiance on various time scales has been predicted using various techniques such as NWP, statistical methods, and artificial neural networks. Predicted solar irradiance is then supplied to a PV simulation software to forecast PV energy output. On the other hand, direct forecasting models rely on historical data recorded from existing PV systems.

In addition, forecasting methodologies can also be categorized according to time horizons (Das et al., 2018). The forecast horizon of a few minutes to several hours is categorized as very short-term or nowcasting. Most commonly, nowcasting is performed with all-sky cameras that forecast cloud motion and solar irradiance.

Short-term forecasting spans 1 to 3 days, while long-term forecasting can range from a week to a year. NWP models are widely used in short-term, medium-term, and long-term forecasting. Short-term forecasts are needed for power smoothing operations, real-time electricity dispatch monitoring, PV storage management, and energy marketing or pricing (Raza et al., 2016). On the other hand, long-term PV power forecasting is required for planning the electricity generation, transmission, and distribution organization, in addition to energy bidding and securing operations (Das et al., 2018). Long-term projections allow the evaluation of overall potential and comparative analysis of the BIPV deployment sites, particularly for large-scale BIPV applications.

In the case of BIPV, both modeling and forecasting are more complicated than in traditional PV due to the prevalence of non-optimal module installations and the addition of shading from the surrounding topography and objects, which makes modeling much more challenging (Jakica et al., 2019). Beyond the difficulty in estimating solar irradiance, it is difficult to predict the BIPV potential of facades owing to partial shading from surrounding buildings, trees, and other architectural aspects (Saretta et al., 2020). The literature is dominated by the studies focusing on rooftop BIPV modeling and direct forecasting with ML (Martín-Chivelet et al., 2022). This thesis aims to contribute to the knowledge in indirect and long-term BIPV forecasting with architectural design parameters and ML methods.

ML algorithms utilized in the PV power prediction models include but are not limited to multivariate linear regression (MLR), kernel ridge (KR), decision trees (DT), multi-later perceptron (MLP), support vector machines (SVM), Gaussian process regression (GPR) and *k*-nearest neighbor (*k*NN). Markovics and Mayer (2022) recently investigated 24 ML models for short-term PV power forecasting directly. Researchers have reported that KR provided the most accurate results, whereas it required high memory usage. Instead, MLP, a feed-forward artificial neural network (ANN) structure, also provided similar accuracy with lower training time. Similarly, Das et al. (2018) suggested that ANN-based models have a good prediction performance. On the other hand, a study focusing on the empirical comparison of

supervised learning methods has reported that boosted trees and random forests (RF) outperform among ten different methods, including ANN, SVM, and several other methods mentioned above. (Caruana & Niculescu-Mizil, 2006). In a recent study, RF also outperformed in predicting annual rooftop irradiation (Swierc et al., 2020). From the reviewed literature, it can be deduced that there is room for exploration with different ML methods.

The success of a learning algorithm is related to the data utilized in the training process (Mohri et al., 2018). The capacity of ML models to establish a link between inputs and outputs enables their application in prediction tasks. (Voyant et al., 2017). Most of the PV power output prediction studies focused on the electrical properties of the system and climatic factors (Akhter et al., 2019; Das et al., 2018). Instead, this thesis aims to forecast BIPV generation with building and context-dependent factors, which are generally overlooked in solar energy forecasting studies. A few studies in literature considered factors related to buildings and their context as input parameters for forecasting solar energy generation (Table 2.3). Most of these studies consider shading as an input with sky-related metrics, such as SVF, sun coverage factor, or shading factor (SF). Another common feature of these studies is predicting the long-term solar potential of rooftops with annual or monthly outputs. Prediction performances are evaluated with a variety of metrics, most commonly with RMSE and nRMSE.

Table 2.3 Reference studies that consider buildings and urban contexts in prediction of solar energy potential with ML.

Prediction model	Resolution	Location	Input features	Output feature	Performance Evaluation	Reference	Accuracy
SVM	City (commune)	Rooftop	Roof area, building SF, tree SF, monthly global tilted solar radiation (predicted with roof tilt, azimuth, latitude, longitude and altitude).	Monthly diffuse horizontal, global horizontal, extra-terrestrial horizontal radiation.	RMSE, nRMSE (6-fold CV)	(Assouline et al., 2017)	nRMSE = 4.68%, 5.83% and 0.04% (for outputs, respectively)
LR, DT, **SVM**, GPR	Surface	Rooftop and facade	SVF, sun coverage factor, ground view factor, building view factor	Annual insolation	nRMSE (5-fold CV)	(Bredemeier et al., 2021)	nRMSE = 8.8% and 3.2% (rooftop and façade)
Polynomial regression	City	Any surface	SVF, sun coverage factor	Annual solar irradiation	R, relative deviation	(Calcabrini et al., 2019)	R = 0.97, relative deviation < 10%
LR, *k*NN, SVM, **RF,** ELM-E	City	Rooftop	Horizontal irradiation, roof shading, roof tilt, roof aspect, sky visibility, horizon maps	Annual solar irradiation (kWh/m^2)	RMSE, MAE, MBE, R^2, training and prediction time	(Swierc et al., 2020)	R^2 = 0.84, RMSE = 94.69, MAE (%) = 7.66

Bold: in the case of multiple models, bold indicated the best performing model and accuracy is reported according to it.

CHAPTER 3

A METHOD FOR BIPV POTENTIAL ESTIMATION IN URBAN CONTEXTS

The method proposed to estimate building BIPV potential consists of three steps. (Figure 3.1). Firstly, geometric information about the buildings is acquired from open data sources. In the second step, BIPV potential is calculated based on simulations that can assess and automate the process of estimating solar energy potential. Lastly, ML models are utilized to predict BIPV generation with the basic and complex design parameters.

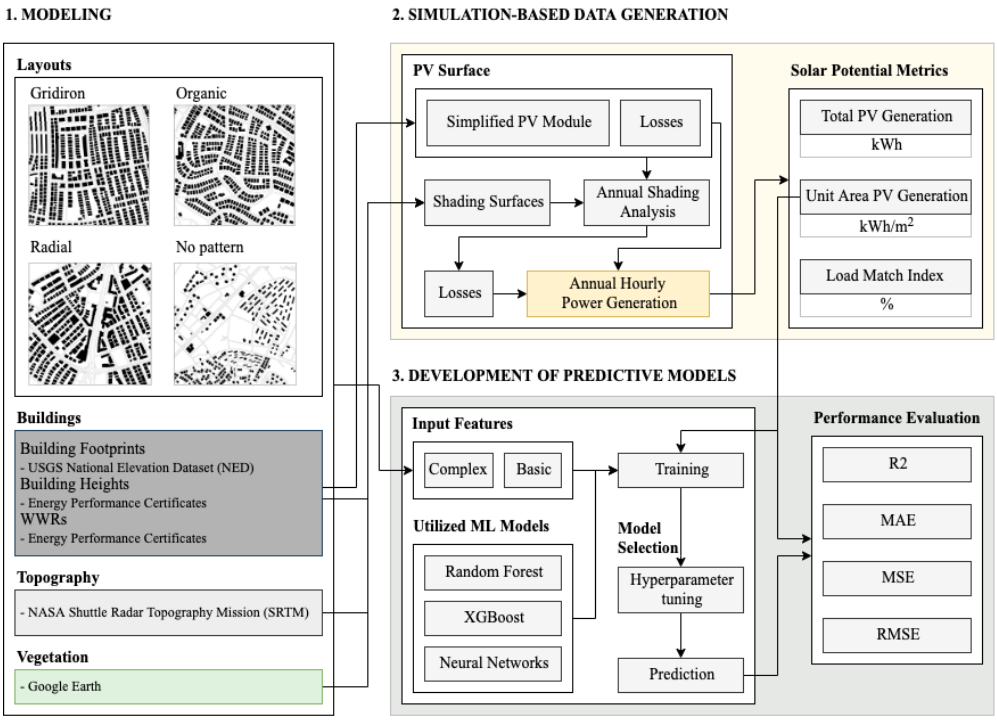


Figure 3.1 Proposed method for estimating BIPV generation capacity of buildings in urban contexts.

The described methodology is partially and fully applied to two different case studies to answer several research questions of the thesis (Table 3.1). Chapter 4 explores the

relationship between the street network pattern and BIPV generation. Chapter 5 the main research question of the research is addressed. Development and comparative analysis of ML models that can predict the BIPV potential of buildings in cities are presented.

Table 3.1 Summary of the case studies with related chapters, research questions(s) and the coverage of the proposed methodology.

Chapter	Research Question(s)	Steps
Chapter **4**: Evaluating the impact of street network pattern on BIPV potential	What is the impact of urban morphology, specifically the street network pattern, on solar potential?	1, 2.
Chapter **5**: An urban morphology inclusive approach to prediction of BIPV potential	What is the accuracy with which the solar energy potential is predicted using machine learning methods? What would be the impact of district-scale BIPV deployment on meeting building electricity demands?	1, 2, 3.

3.1 3D Modeling

3D modeling of built environments is critical in the assessment of BIPV potential since BIPV surface insolation is widely affected by surface orientation, tilt, and shading. Urban areas are modeled using Rhino and Grasshopper parametric modeling interface (Rutten, n.d.). A more accurate representation of existing neighborhoods can be achieved by retrieving several types of geometric information about the study areas. Four types of geometric information, namely urban layouts, buildings, topography, and vegetation, are considered in the modeling process.

Urban layouts, building footprints, and topography are retrieved from the Cadmapper online web tool, which relies on one of the crowd-generated data

initiatives, namely OpenStreetMap (OSM) (CADMAPPER LLC, 2013; Haklay & Weber, 2008). Cadmapper converts geometric data, such as building footprints, 3D buildings, parks and roads, contours, and topography, from open data sources of OSM, NASA, and USGC into structured CAD files. Although OSM offers extensive data on building footprints in most cities worldwide, only a small number of buildings are available with building height information.

Building height and WWR information is retrieved from the energy performance certificates (EPC) provided by the Turkish Ministry of Environment, Urbanization, and Climate Change. EPC includes many different items that directly and indirectly affect building performance. Building construction materials, floor area, number of floors, WWR, energy consumption, and renewable energy generation are among many items that can be found in the EPCs. In addition to building heights, WWR information is obtained from the EPCs in the modeling phase. WWRs are assumed to be equal on different façades, and windows are placed on façade surfaces homogenously.

On the other hand, EPCs are not available for every building. For missing building height information, mod of building heights within each parcel is used for missing data imputation. This assumption relies on the generalization that most buildings aim at the highest permitted building height limitation, which is in most cases the same for all buildings in a parcel. Building footprints are extruded with the height information to form building masses.

Missing WWR information is generated by a density estimation fitted to the obtained limited data. In statistics, the probability density function is an essential concept that describes the distribution of a random quantity X with a probability density function $f(x)$ (4). Density estimation is explained as the process of constructing an estimate of the density function from observed data (Silverman, 1961). Density estimations can be both parametric and nonparametric. For parametric estimations, the density function underlying the data may then be approximated by extracting variance and mean estimates from the data and putting them into the normal density calculation.

$$P(a < X < b) = \int_a^b f(x)dx \qquad \text{for all } a < b \tag{4}$$

For estimating the distribution of WWR values, a nonparametric density estimation approach, namely kernel density estimation (KDE), is preferred. KDE (5) relies on more flexible assumptions derived from the observed data have a high capacity to capture distributions of non-uniform and non-symmetrical datasets. In KDE, the data will be permitted to speak for themselves in deciding the estimate of f more than if f were confined to fall inside a specified parametric family. The kernel estimator with s kernel K is defined by

$$f(x) = \frac{1}{Nh}\sum_{i=1}^{N} K\left(\frac{x - x_i}{h}\right) \tag{5}$$

where h is the bandwidth. In other words, the kernel estimator is a sum of the distributions placed at the observations with the kernel function K for the shape of the distributions and the bandwidth h for their width. Gaussian distribution is selected as the kernel function, and Silverman's rule of thumb is applied in the choice of bandwidth. Each data point is weighted uniformly in the distributions implying equal importance of observations. Based on the fitted KDE, missing WWR values are generated.

Vegetation is also considered in the geometric models with the data obtained from the satellite images of Google Earth (Google, n.d.). Green tones are filtered out of the color scheme, and trees are placed at random data points. A basic tree geometry is represented with a 2m high trunk, 3m high volume of branches, and leaves with a 2m radius. Trees are divided into two categories based on their leaf abscission characteristics, coniferous and deciduous. The ratio of the number of randomly selected coniferous trees to the total number of trees is set to 58% based on the data for the forests of Turkey in 2020 provided by the Ministry of Agriculture and Forestry (T.C. Tarım ve Orman Bakanlığı, 2020).

3.2 Simulation-based Data Generation

3D models and typical meteorological year (TMY) weather files are used as inputs for the simulation-based data generation method. PV generation from rooftop and façade surfaces for the modelled urban areas is estimated with the PVWatts (Dobos, 2014), implemented inside the Ladybug Tools (Roudsari & Pak, 2013) add-on for Grasshopper.

3.2.1 Hourly PV Generation Estimation

Several computer software blends different PV performance models and uses open-source libraries. PVWatts is one of these software offering an online calculation tool developed by the National Renewable Energy Laboratory (NREL) of the US Department of Energy (Dobos, 2014) as introduced in Section 2.4.1. It takes two types of inputs; field and advanced. Field inputs include system size (DC output in kW), module type (PV technology), system losses (PV efficiency in %), array type (e.g., fixed roof mount, fixed open track, etc.), tilt angle (degrees), and azimuth angle (degrees). Advanced inputs are DC/AC ratio, inverter efficiency, and ground coverage ratio as user-defined inputs (Dobos, 2014). For the empirical BIPV performance estimation, first, weather files specific to a location are utilized. The parameters obtained from the weather files for PV generation calculations are beam and diffuse irradiance, ambient temperature, and wind speed at 10 m height. PVWatts is also implemented in Ladybug Tools, an environmental design software connecting computer-aided-design interfaces, and is available for use during the design development phase (Roudsari & Pak, 2013).

3.2.2 PV Surface and Module Properties

Detailing each BIPV unit is computationally expensive in large-scale BIPV simulations with shading analysis. Therefore, the application area, which is

limited to façade and roof surfaces in this research, is assumed to be covered with BIPV modules. Façade and roof surfaces with more than 10 m^2 area are considered available for PV application. 70% of the roof and 90% of opaque façade surfaces are assumed to be available for PV deployment. For façade surfaces, the tilt angle is assumed to be parallel to vertical wall surfaces. Similarly, roof surfaces are covered with solar panels aligned with the tilt angle of roof surfaces. Module parameters are set based on a commercial monocrystalline solar module provider (Table 3.2) (TW SOLAR, 2022).

Table 3.2 Solar module parameters.

Parameters	Value
Module efficiency, η_{PV}	20.08%
Temperature coefficient	-0.40%/°C
DC to AC derate factor	0.85

3.2.3 Shading from Urban Fabric

BIPV modules are often subject to inhomogeneous irradiance due to placement and shading from various objects in urban contexts. Therefore, a single irradiance measurement or irradiance value for the whole system is not always suitable, and applying the usual performance ratio definition is not possible (Martín-Chivelet et al., 2022). Researchers recommended using time-resolved shading analysis in the case of partial shading instead of commonly used simple SVF methods for heterogenous irradiance.

This study considers shading from neighboring buildings during the simulations for BIPV generation with the calculation of the annual hourly shading factor (SF). Hourly SF values are obtained with the "Sunpath Shading" component of Ladybug tools. Although BIPV potential without shading can be quickly estimated, SF calculations increase the computation cost. Computational cost is highly correlated

with the number of surfaces that shade the analyzed surface. Therefore, the description of the analysis region and the detail of the BIPV surface is critical.

Neighbor buildings and trees nearby and topography are involved in SF calculations as context geometries. Buildings inside a circle (r = 5*h*), centering the analyzed BIPV surface of a building with building height *h*, are included in SF calculations (Figure 3.2). SF calculations are conducted for a number of points on the analyzed BIPV surface. SF of each point is averaged, and a single value is obtained for a BIPV. An annual value for the analysis surface is utilized in the simulations for the selected roof or façade surface. Tree leafless period and diffuse radiation are also taken into consideration in SF analysis.

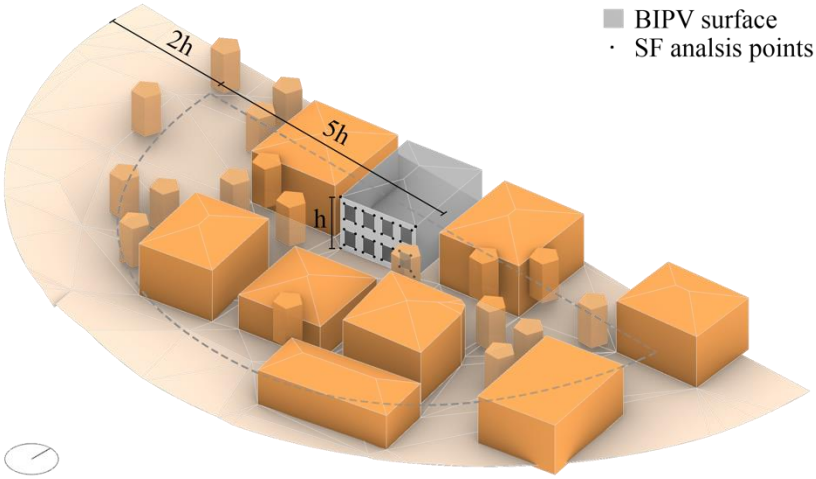


Figure 3.2 Example urban setting in shading analysis.

3.3 Development of Predictive Models

Three supervised ML models are trained with the subset of data generated from simulations. Firstly, a number of parameters are recorded from the 3D model and simulations as input and output features for predicting solar energy production. In the model development process, ML models are fine-tuned to lower prediction errors. The last step involves the evaluation of models with the selected performance evaluation metrics.

3.3.1 Data Description

Machine learning models for regression tasks take a set of input parameters with numerical values and process them to predict numeric output variable(s). Since this thesis aims to predict PV generation (kWh/m^2) with parameters relating to urban form, quantitative measures of urban form are investigated. After reviewing the urban morphology indicators studied in recent literature, two groups of input parameters, basic and complex, are proposed for the fast and accurate prediction of PV generation per façade (Table 3.3) and roof (Table 3.4) surface area.

Table 3.3 Summary of input variables and descriptions for façade prediction.

Predictor Group	Info Group	Description	Variable Name	Unit
Basic	Building	Facade height	fHeight	m
		Window to wall ratio	WWR	%
		Facade azimuth angle	fAzimuth	°
		Facade elevation	fElevation	m
	Context	Neighbor building surface height	nbHeight	m
		Neighbor building surface width	nbWidth	m
		Neighbor building surface azimuth angle	nbAzimuth	°
		Neighbor building surface distance	nbDistance	m
		Neighbor building surface elevation	nbElevation	m
		Neighbor building surface relative orientation to the studied surface	nbLocation	°
Complex	Building	Facade height	fHeight	m^2
		Window to wall ratio	WWR	%
		Facade azimuth angle	fAzimuth	°
	Context	Sky exposure factor	SEF	%

Basic parameters may be easily estimated by hand, using a master plan and on-site observations, while complex parameters require a 3D model of the target building and its immediate surroundings. Most studies in the literature require a 3D model or aerial images in a particular resolution which may not always be available (Jakubiec & Reinhart, 2013; Romero Rodríguez et al., 2017). In this respect, basic predictors allow designers, occupants, or anyone interested in assessing a building surface can easily estimate annual energy generation potential from BIPV.

Table 3.4 Summary of input variables and descriptions for roof prediction.

Predictor Group	Info Group	Description	Variable Name	Unit
Basic	Building	Roof tilt	rTilt	°
		Roof azimuth angle	rAzimuth	°
		Roof elevation	rElevation	m
	Context	Neighbor building surface height	nbHeight	m
		Neighbor building surface width	nbWidth	m
		Neighbor building surface azimuth angle	nbAzimuth	°
		Neighbor building surface distance	nbDistance	m
		Neighbor building surface elevation	nbElevation	m
		Neighbor building surface relative orientation to the studied surface	nbLocation	°
Complex	Building	Roof tilt	rTilt	°
		Roof azimuth angle	rAzimuth	°
	Context	Sky exposure factor	SEF	%

Reliable and high-quality datasets are essential for ML applications. Outliers and unbalanced distributions of input features might degrade the prediction performance. Data preprocessing plays a critical role in this step to enhance prediction performance. After removal of outliers, Simple yet powerful methods of descriptive statistics and data visualizations will be presented to assess data quality. A

correlation matrix will be generated to capture variable dependency. The Spearman correlation coefficient (ρ) will be calculated to evaluate the monotonic relationship between variables.

ML models are trained with a subset of data, and their performance is evaluated with unseen observations. Therefore, datasets are split into training and test datasets. A balanced sampling of four urban areas with different street network patterns was necessary for this process. For each network pattern, randomly selected %80 of the data points are used for training, while the remaining 20% are used for prediction of baseline models.

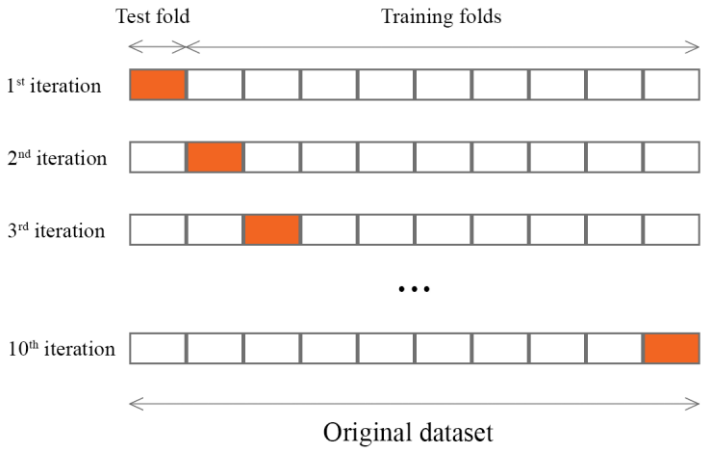


Figure 3.3 Diagram of 10-fold cross-validation method.

However, this randomness in the selection might also create a bias. For instance, selected data points for the test dataset can be dominated by poorly learned observations by the ML. A widely trusted 10-fold cross-validation method is applied to overcome such problems and provide more reliable results (Kohavi, 1995). In 10-fold cross-validation, the dataset is randomly split into ten mutually exclusive datasets of approximately equal size (Figure 3.3). A rotating test fold is predicted with the data different than the test fold. Overall prediction performance is evaluated by averaging the prediction performance of each iteration. The method aims to reduce the uncertainty resulting from the selection of train and test splits.

3.3.2 Utilized ML Models

In the tool development process, three ML models, RF, XGBoost, and MLP, which offered accurate predictions of various building performance metrics, were utilized.

RF (Breiman, 2001)*:* RF is an ensemble model that combines the predictions of a number of different regression trees (CART). The predictions made by RF are based not on a single tree, as is the case with CART, but rather on a collection of trees. Consequently, it produces more reliable and steady predictions than CART. For multiple tree growth, bootstrap samples and randomization are used in. A selection of variables picked at random and have the same size as the training data is passed to an algorithm comparable to CART. A forest composed of several trees, or an ensemble of trees, is created. The forecast of each tree is averaged for regression tasks.

XGBoost (T. Chen & Guestrin, 2016)*:* The scalable and innovative tree-boosting technique known as XGBoost offers cutting-edge performance. In RF, separate trees are grown independent of one another without knowing the performance of the previously grown tree. On the other hand, while using the boosting approach, the growth of trees is determined by the outcomes of previously constructed trees. Conceptually, the goal is to increase the performance of each new tree by teaching it from the errors of its predecessors. Individual split points are exploited for random subsampling while building trees. XGBoost offers a scalable version of the boosting algorithm with effective parallelization.

ANN (Anil et al., 1996)*:* ANN is a data processing system that discovers the input-output connection from the acquired data. The most prevalent and simple ANN design is the feed-forward multilayer perceptron (MLP). It consists of one input layer, one or more hidden neuron layers, and one output layer. A neural network processes inputs by multiplying them by weights and biases. After passing the results to activation functions, outputs are then obtained. An activation function determines

the output of a node given a single input or many inputs. The final output is the prediction made by the ANN for the input or inputs provided.

3.3.3 Model Selection

ML models were initialized with parameters and hyperparameters, which affect prediction performance (Agrawal, 2021). During the training phase, parameters are adjusted to characterize the relationship between the inputs and outputs. On the other hand, hyperparameters are the higher-level parameters that interact with the data characteristics and learning capacity of the algorithm. While the model parameters are tuned in reference to the datasets without an intervention, hyper-parameters are set manually at the initialization of the model.

In order to maximize prediction performance, the bias-variance tradeoff should be optimized through hyperparameter tuning. If not appropriately fine-tuned, underfitting or overfitting problems can be observed. Hyperparameters of different ML techniques vary from each other based on the principles guiding the prediction processes.A short description of the utilized ML models and hyperparameter grids can be found in Appendix A.

Utilized ML models have different hyper-parameters. For instance, *max_depth* implies the depth of the tree and is highly associated with the overfitting problem for RF. The *min_sample_split* and *min_samples_leaf* are the parameters controlling the number of instances required to form a split and a leaf node, respectively. The *n_estimators* control the number of trees in the forest.

MLP hyperparameters vary from the other tested decision three-based methods. The *hidden_layer_sizes* sets the number of hidden layers and neurons in each layer. The *learnin_rate* controls the adaptation of the model to the problem. From an oversimplified perspective, a small learning rate might require a longer learning process. In contrast, a shorter training might be enough for a high learning rate yet might be subject to converging a suboptimal solution. Learning rate interacts with

other parameters, including the *batch_size*. The *batch_size* determines the number of samples sent through the network. *Max_iter* defines the total number of iterations. For the tested solvers, *max_iter* defines the number of epochs, i.e., how many times the network will see each data point. The *alpha* hyperparameter is the regularization term that aims to prevent overfitting by limiting the size of weights. Increasing *alpha* may solve the overfitting problem by decreasing variance. Lastly, the *solver* identifies the weight optimization algorithm for nodes. The *momentum* parameter should also be optimized if the *solver* is stochastic gradient descent (sgd). It is an additional term that accelerates convergence by adjusting the gradient vectors in the right direction.

Most machine learning libraries have default values for the hyperparameters for fast and easy implementation of the models with a decent prediction performance. However, tuning hyperparameters is critical since it exploits the full potential of ML models. The tuning can be achieved by training and testing models for various hyperparameter combinations and selecting the combination that gives the lowest errors. Grid and random search are the two common fine-tuning methods (Agrawal, 2021). Grid-search explores all potential combinations of pre-defined hyperparameter values. It might be exhaustive and computationally costly. Instead, random search assesses only a pre-defined number of combinations rather than testing all possible combinations. Although it does not guarantee the best combination, it is a more time-efficient method.

In this study, the random search method is preferred and coupled with the cross-validation method described in Section 3.3.1. The number of iterations is set to 50, and a 5-fold CV is applied. This application is also called “nested cross-validation.” In the process, each set of hyperparameters is evaluated using 5-fold cross-validation. Unlike the previously discussed method, the provided training dataset is split into five folds rather than the original dataset. This method prevents retrieving biased results specific to a particular training dataset split.

3.3.4 Model Evaluation

In ML applications, the goodness of a prediction can be measured by several metrics. These metrics aim to assess prediction error and accuracy. Without a generally established consensus in the literature, many different metrics are utilized for model evaluation. R-Squared (R^2), Mean square error (MSE), root mean square error (RMSE), mean-normalized root mean square error (nRMSE), and mean absolute error (MAE) are the selected performance metrics, which are calculated using the equations below,

$$R^2 = 1 - \frac{\frac{1}{n}\sum_{i=1}^{n} (y_{predicted} - y_{true})^2}{\frac{1}{n}\sum_{i=1}^{n} (y_{predicted} - y_{true})^2}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_{predicted} - y_{true}\right)^2$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_{predicted} - y_{true}\right)^2}$$

$$\text{nRMSE} = \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_{predicted} - y_{true}\right)^2}}{\frac{1}{N}\sum_{i=1}^{N} W_{true}}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|y_{predicted} - y_{true}\right|$$

where, $y_{predicted}$ and y_{true} represent the predicted PV generation and simulated PV generation, respectively. *N* is the length of the dataset.

Although several different metrics can be used in evaluating model performance, consistency is advised in fitting or optimizing models (Yang et al., 2020).

Researchers have also highlighted RMSE as a metric that should be reported in solar forecasting studies.

CHAPTER 4

EVALUATING THE IMPACT OF STREET NETWORK PATTERN ON BIPV POTENTIAL

This chapter explores the impact of urban morphology on the PV generation potential. Four representative street network pattern categories, gridiron, radial, organic, and no pattern, are parametrically modeled (Table 4.1) in reference to a recent study investigating the classification of road network patterns (W. Chen et al., 2021). This chapter is presented at Sustainable Built Environment Conference 2022: Innovations for the Urban Energy Transition at TU Delft with the title "Impact of Urban Street Network on PV Generation Capacity of Buildings."

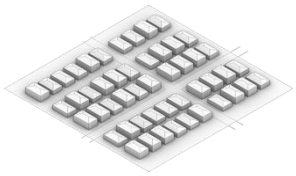
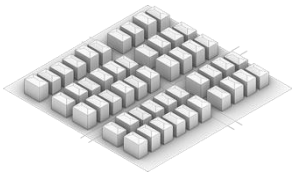
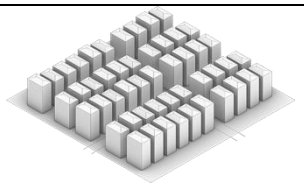
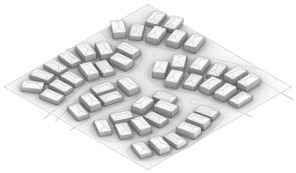
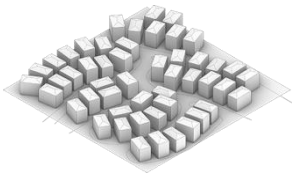
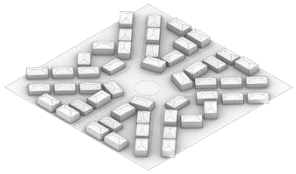
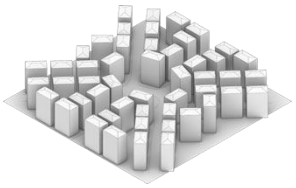
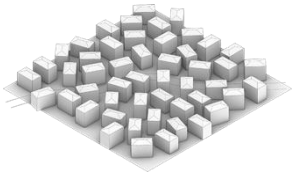
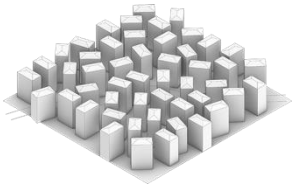
4.1 Model and Data Description

For synthetic neighborhoods, only the urban layout and building data are considered in the modeling phase. Identical buildings with a footprint area of 15×40 m^2 are located in each urban area. Simplified footprint extrusions in various heights represent building masses. Facade detailings such as balconies or claddings are not considered. The WWR is set to 0.25 identically for each building homogenously distributed to the façade surface. Only opaque facade surfaces are considered for solar PV generation. For the roofs, pitched roof geometry is generated over the building masses.

Each urban area is sampled with a 250×250 m^2 flat land area. Forty-five identical buildings with a footprint area of 15×40 m^2 are located on flat terrain. The building density is kept constant for each street network pattern to explore the impact of building configurations, in other words, urban form, with varying street network patterns. In addition to different street network patterns, three building height

classifications: low-rise, middle-rise, and high-rise, are also examined. The floor height for each building in different scenarios was set to 3 m.

Table 4.1 Studied street network patterns and building height options.

	Low-rise (2-3 floors)	Middle-rise (6-7 floors)	High-rise (12-13 floors)
			
			
			
			

In the shading analysis, for low-rise configuration, buildings inside a circle centering the analyzed PV surface and having a 48 m diameter are included in shading factor calculations as context geometries (Figure 4.1). The same method is used for mid-rise and high-rise buildings with diameters of 64 and 96 meters, respectively.

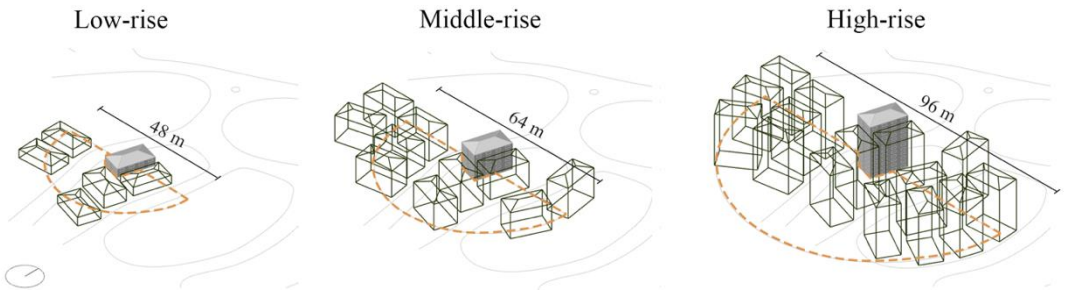


Figure 4.1 Context geometries included in shading analysis.

4.2 Results

Four street network patterns, gridiron, organic, radial, and no pattern are modelled with three uniform building height configurations. The energy generation capacity of modelled urban areas is calculated in the city of Ankara with a warm-summer Mediterranean climate. Unit area BIPV generation (kWh/m^2) for each network pattern was selected as the comparison metric. Total BIPV generation from façade surfaces was divided by the total façade area, including all four orientations of building geometry. Total generation from the surfaces of a pitched roof is divided by the total area of roof surfaces. Overall, facade BIPV generation capacity per roof area (Mean=143.698, Std Dev=0.743) is greater than the roof BIPV generation capacity per façade area (Mean=43.650, Std Dev=9.606).

4.2.1 Façade BIPV Generation

An increase in the building height for densely built homogenous urban environments resulted in a decline in unit area BIPV generation (Figure 4.2). This result can be related to the impact of shading, which increases with building height. The highest unit area generation capacity is observed in the low-rise urban areas for all four street network patterns where shading from neighbor buildings is minimal. In addition, urban areas with no pattern street network category outperform all evaluated building height options. Irregular building intervals and a wide range of facade orientations observed in no pattern areas might result in an extended insolation period.

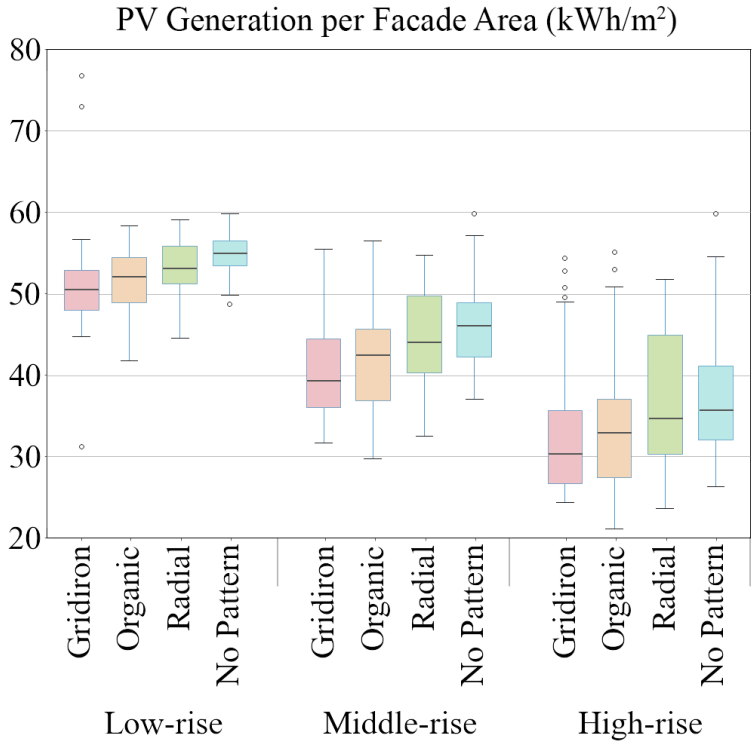


Figure 4.2 Facade PV generation (kWh/m^2).

4.2.2 Rooftop BIPV Generation

Contrary to the façade generation, BIPV generation per roof area was largely unaffected by the change in building heights (Figure 4.3). The BIPV generation capacities are observed to be very similar in evaluated building height options. The surfaces with no shading define the upper threshold of the roof BIPV generation. Therefore, the maximum generation is determined by the same upper limit in all tested floor height options and street network patterns. Since the tested urban areas have homogenous building heights, the shading effect is very limited. Due to the negligible shading effect, a narrow range of BIPV generation is observed for roof surfaces.

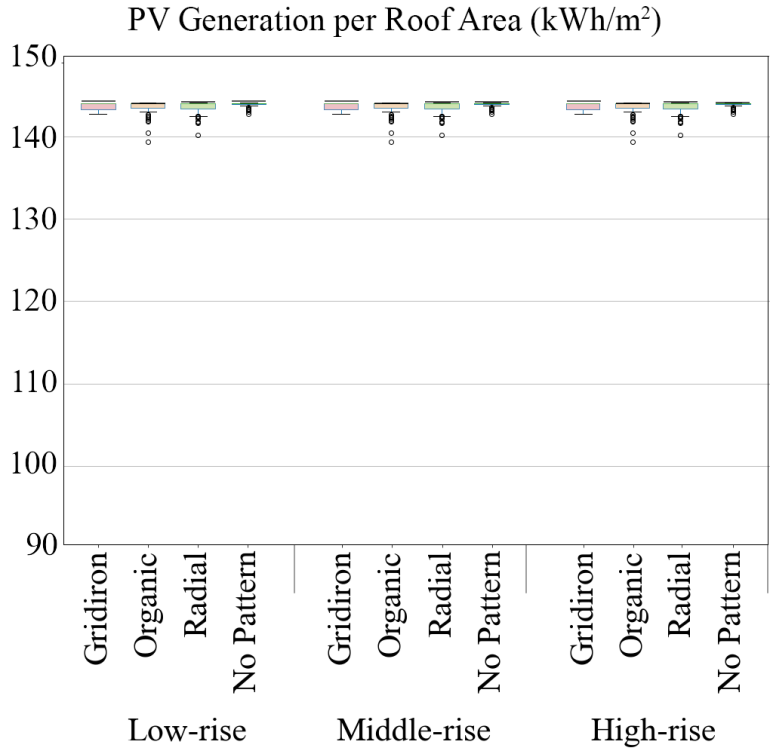


Figure 4.3 Roof PV generation (kWh/m^2).

4.3 Discussion

Chapter 4 aims to compare the BIPV generation capacity of four street network patterns with three building height configurations in a city with a warm-summer Mediterranean climate. Total energy generation from the façade and pitched roof surfaces were calculated and normalized with the gross application area for the comparative analysis. BIPV generation from roof surfaces was almost three times greater than the facades, which is a common finding in the literature since most rooftops receive greater solar irradiation than facades. Street networks with no pattern have offered higher BIPV generation capacity per PV deployment area for facades. However, the impact of street network patterns and building heights has been negligible for BIPV generation from roof surfaces. Since this study examined uniform urban areas with similar building heights, the influence of shading on roofs was minimal.

The results of this study can be validated further by testing the described street network patterns in different orientations within the city fabric. The observed high energy generation potential of the no pattern street networks for facades could be challenged with a study design considering multiple orientations of streets. The impact of urban morphology on rooftop BIPV potential might also be observed with mixed building height options. The impact of street network patterns on solar energy potential in existing built environments can be explored in future work.

CHAPTER 5

AN URBAN MORPHOLOGY INCLUSIVE APPROACH TO PREDICTION OF BIPV POTENTIAL

Chapter 5 aims to explore how accurately machine learning algorithms can forecast BIPV generating potential with parameters related to buildings and urban contexts. Firstly, four urban areas exhibiting different street network patterns in Ankara, Turkey (Figure 5.1) are selected and modeled with the method described in Section 3.1. Several urban morphology indicators are recorded from 3D models and utilized in the prediction of BIPV potential. A simulation-based method is applied for data generation. Three ML algorithms are trained, tested, and evaluated with simulation outputs.

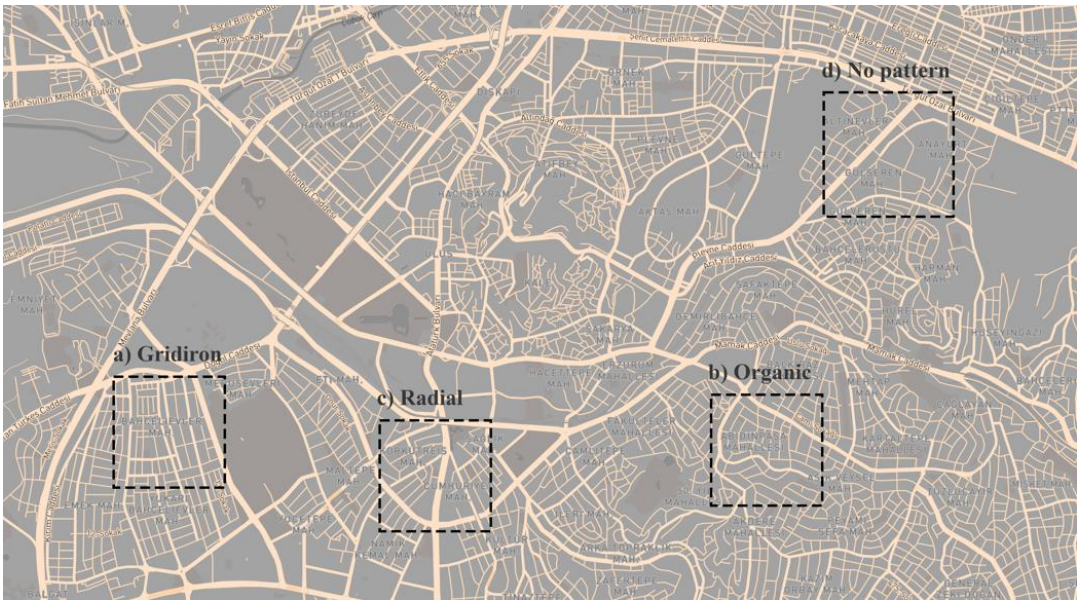


Figure 5.1 Studied neighborhoods in Ankara.

5.1 Model and Data Description

Four urban areas are sampled from the urban fabric of Ankara. Bahçelievler District is chosen for its likelihood of the gridiron pattern (a) with roads crossing perpendicularly and forming rectangular parcels (Figure 5.2). It should be noted that the perfect gridiron pattern in Ankara was hard to observe since the city has grown in an unplanned manner in the past decades. Organized industrial zones located in the suburbs were among the most representative of the gridiron pattern; however, they were out of the interest of this thesis focusing on urban districts.

Abidin Paşa District is selected for the organic pattern (b). It exhibits a more organically shaped street network with curvilinear paths crossing at various angles (Figure 5.3). Residential buildings mostly occupy the district. In addition to Abidin Paşa, the majority of residential districts built in Ankara in the 1960s are representative of the organic street network pattern.

A combination of Kızılay, Cumhuriyet, Sağlık, and Korkutreis Districts offers the most accurate representation of radial network patterns (c) (Figure 5.4). These old neighborhoods form a focal point on the Sıhhıye Square. However, only a partial radial pattern could be achieved. Districts in the North have more organic characteristics; therefore, they were not included in the selected urban areas for the radial pattern. The area is occupied primarily by commercial buildings, offices, and state-owned buildings.

Lastly, Gülseren, Altınevler, and Anayurt Districts exhibit less organized urban areas and are selected for no pattern category (d) (Figure 5.5). The area is dominated by a gated community, namely, Gülseren TOKI, designed by the Housing Development Administration of the Republic of Turkey (TOKI).

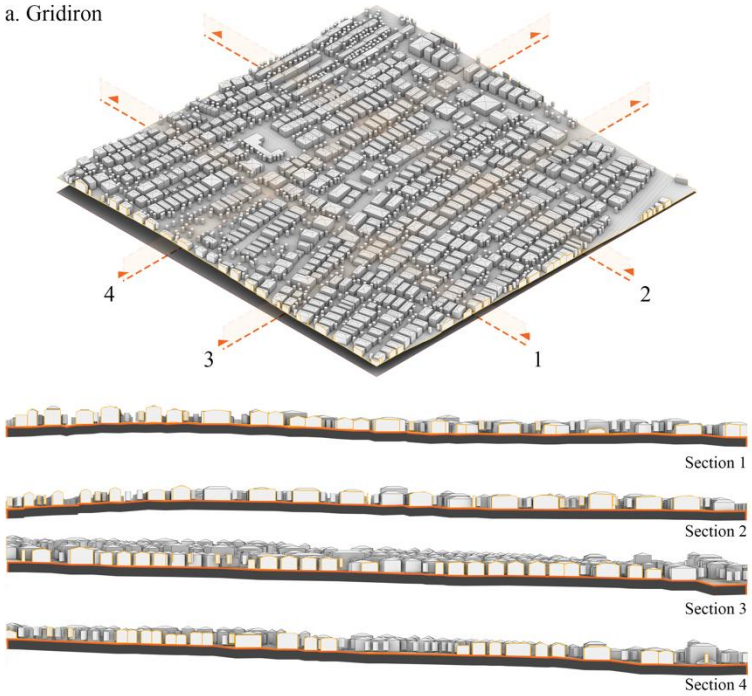


Figure 5.2 Selected urban area representing gridiron street network pattern.

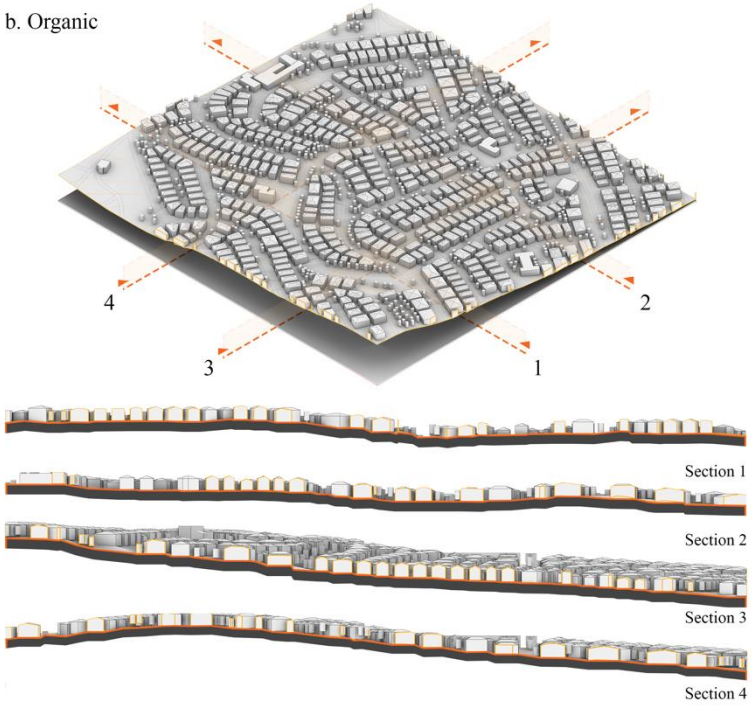


Figure 5.3 Selected urban area representing organic street network pattern.

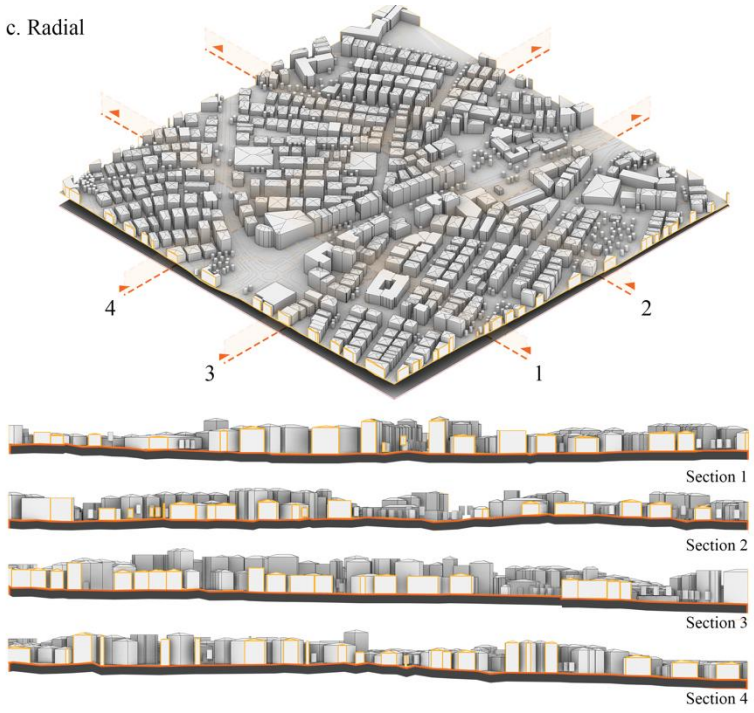


Figure 5.4 Selected urban area representing radial street network pattern.

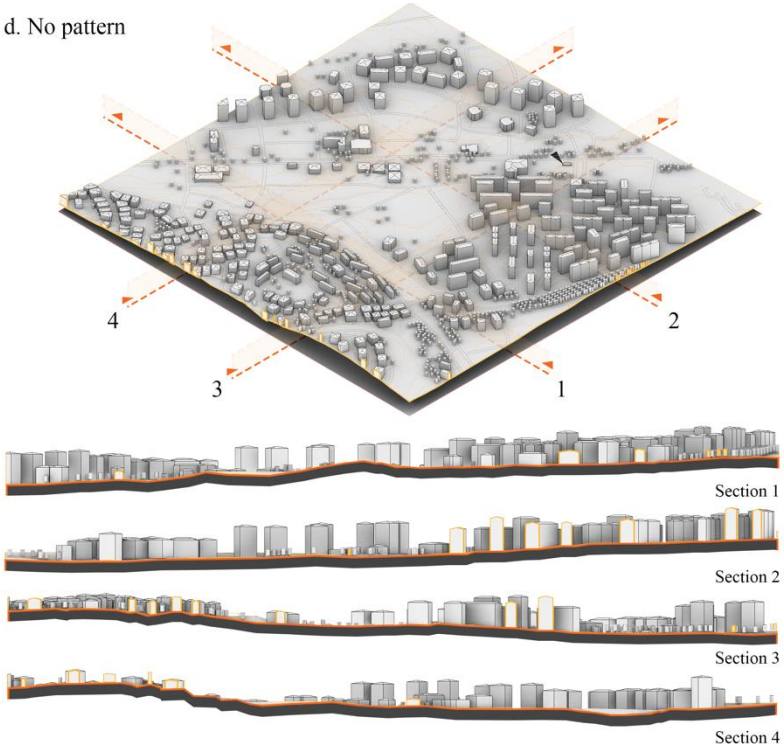


Figure 5.5 Selected urban area representing no pattern street network.

Four urban areas representing different street network patterns are modeled to generate the dataset for ML models. Parameters quantifying the urban form are calculated from the 3D models based on geometric information. Overall, the dataset comprises 1942 data points, denoting the total number of buildings examined (Figure 5.6). Buildings in the gridiron pattern have the largest share in the dataset with 646 buildings, whereas no pattern has the smallest share due to the lower building density.

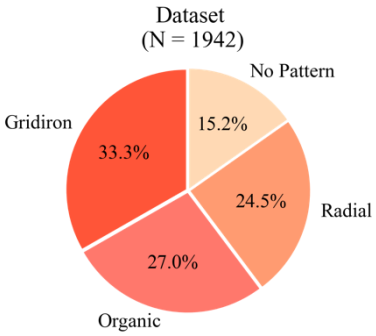


Figure 5.6 Distributions of the data points based on street network patterns.

Districts also vary from each other in terms of the building functions. Non-residential buildings dominate the urban area representing the radial pattern since it is located in one of the major central business districts in Ankara. The remaining three neighborhoods are mostly residential, with the share of non-residential buildings below 8% (Table 5.1). Building functions can be associated with the occupied hours and building energy system operations.

Table 5.1 Proportion of residential and non-residential buildings.

	Gridiron	Organic	Radial	No pattern
Non-residential Residential	5.1% 94.9%	1.9% 98.1%	5.0% 95.0%	7.4% 92.6%

Building function can be associated with building occupancy and energy consumption patterns in operation. Non-residential buildings, such as schools, libraries, and commercial centers, are generally occupied during the daytime. On the other hand, residential buildings might have a variety of occupancy patterns depending on different occupant profiles and behavior. In addition, the electricity demand for non-residential buildings can be higher than that of residential buildings. Comfort conditions are aimed to be sustained for a larger number of occupants, and many types of equipment, such as computers, kitchenware, and air-conditioning, are regularly run in offices and commercial buildings.

5.1.1 Energy Performance Certificates

Energy performance certificates (EPCs) can be retrieved from the Ministry of Environment Urbanization and Climate Change of Turkey by entering the door numbers of buildings to the Building Energy Performance (BEP) application (*BepTR*, n.d.). Individual door numbers can be retrieved from the address inquiry system of the Ministry of Internal Affairs of Turkey, which is also open to public inquiry (*Adres Sorgulama*, n.d.).

Table 5.2 Availability of EPCs.

	Gridiron	Organic	Radial	No pattern
Number of buildings	642	425	476	297
Number of EPCs collected	193	146	43	40

EPCs contain information about building geometry, construction, and environmental performance, which is recorded during the inspection of government officers. EPCs of buildings located in the studied four neighborhoods in Ankara were acquired. Unfortunately, EPCs were not available for all buildings in the study areas during the data acquisition phase of this thesis (Table 5.2). All information reported in this section is based on the available EPCs in February 2022. Geometric data, date of

construction (Figure 5.7 left), construction materials, energy consumption, energy performance categories, and energy generation from PV are offered data items in EPCs. Unfortunately, the energy generation section was empty on obtained EPCs since no PV installations were observed.

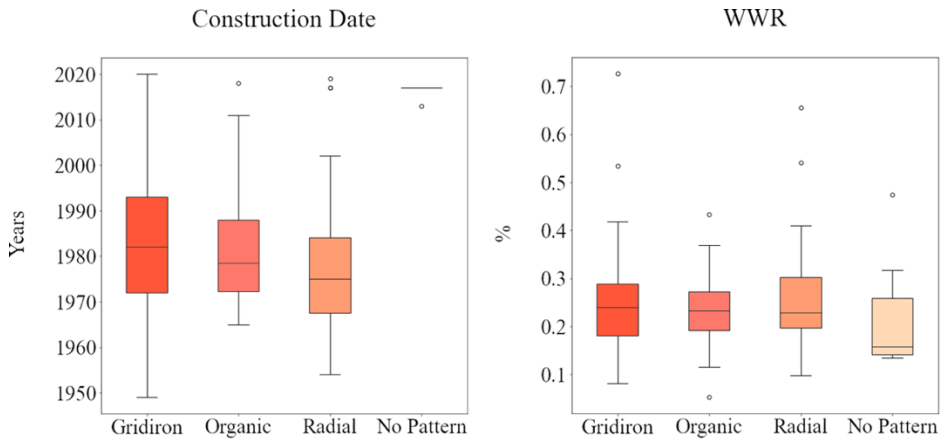


Figure 5.7 Distribution of the building construction date (left) and WWR (right) based on EPCs.

WWR is one of the recorded parameters for geometric data available on EPCs (Figure 5.7 right) affecting the available BIPV deployment area on facades. Gridiron, organic, and radial patterned urban areas have very similar mean WWR values of around 0.25. No pattern study area has an average WWR of around 0.16 only, although the available EPCs were from building built between 2010 and 2020. The majority of the buildings in the radial study area were built in the 90s. A few buildings built in the 2000s appear as outliers. The gridiron study area has similar characteristics to the radial in terms of building age, except with an increased number the new buildings built in the past decades.

5.1.2 Urban Morphology Indicators

Studied urban areas were analyzed in terms of their formal qualities. Among the many indicators of urban morphology, discussed in Section 2.3.1.4, several metrics are calculated (Figure 5.8) and reported in Table 5.3. Except for the no pattern study

area, land areas were set to 750 × 750 m^2. Due to the smaller number of buildings in the urban area representing no pattern street networks, a larger land area, 1125 × 1125 m^2, was studied.

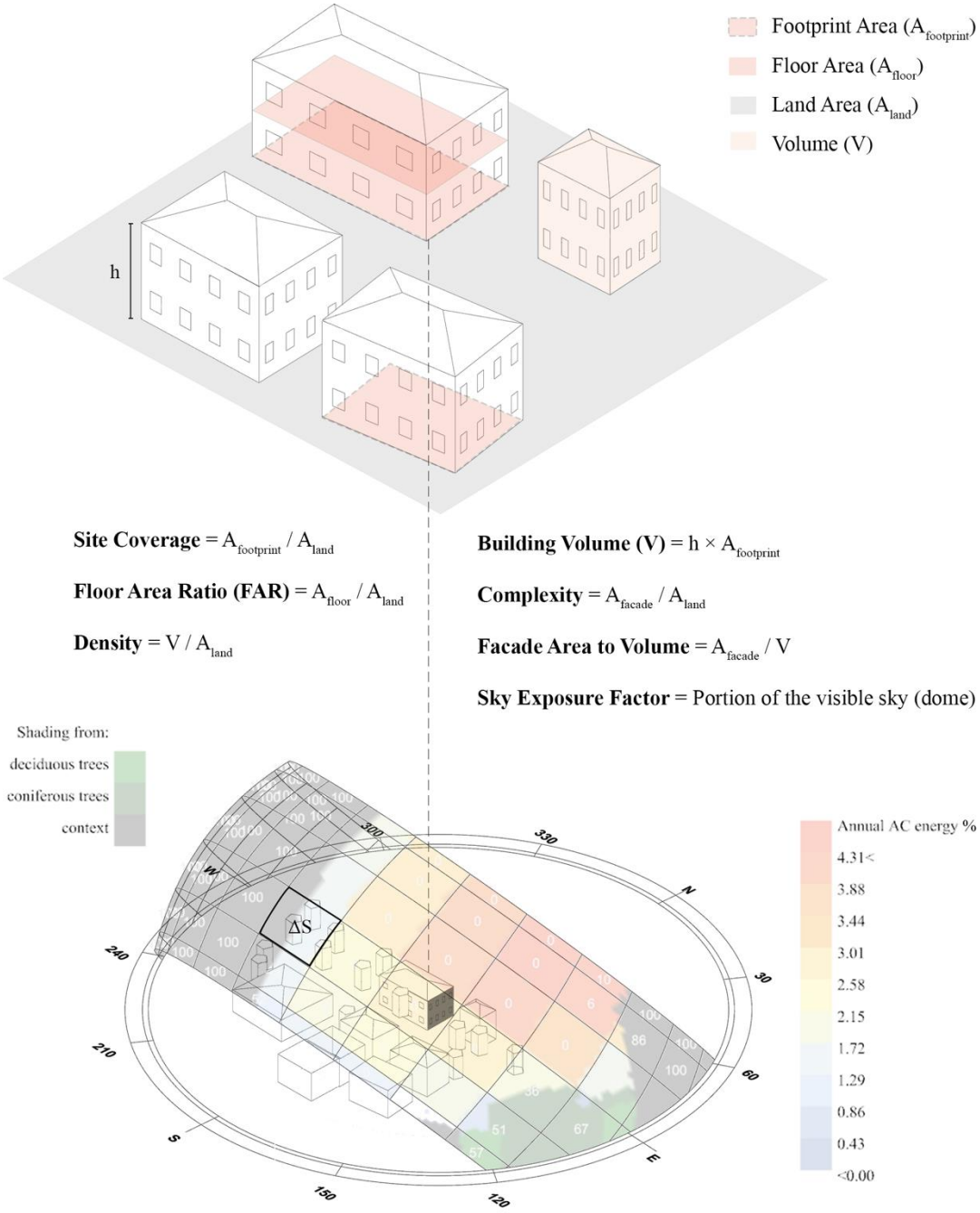


Figure 5.8 Studied urban morphology indicators.

The gridiron study area has the highest number of buildings and the largest building footprint area. In contrast, the number of buildings and the footprinwt area occupied by buildings is the smallest in no pattern study area. However, it is the leader in average building height. Therefore, although it has the smallest footprint area, it has the second-largest total floor area. However, the standard deviation (std) of building heights in no pattern area is the largest, implying that the building heights are non-uniform. Std of the radial urban area is also considerably larger with a value of around two floors (6.75 m) compared to the gridiron and organic study areas.

Table 5.3 Analysis of studied urban areas based on urban morphology indicators.

Parameters	Unit	Gridiron	Organic	Radial	No pattern
Number of buildings	-	642	525	476	297
Mean building height	m	11.56	11.80	18.79	22.72
Total footprint area	km^2	0.203	0.164	0.180	0.129
Total floor area	m^2	0.786	0.643	1.120	1.004
Building height std.	m	2.31	1.41	6.75	11.89
Site coverage	-	0.36	0.29	0.32	0.10
Floor area ratio	-	1.40	1.14	1.99	0.79
Density	m	4.19	3.42	5.97	2.38
Façade area /volume	m^{-1}	0.23	0.23	0.21	0.20
Complexity	m^{-1}	0.96	0.79	1.23	0.47
Average SEF Facade	%	47.00	48.96	41.23	52.44
Average SEF Roof	%	92.19	94.20	89.45	94.53

The radial study area is the most densely built urban area and has the highest floor area ratio. Again, no pattern study area offers the lowest values in these urban density metrics, sustaining its state as the least dense study area. The radial study area also comes to the forefront in terms of complexity. As expected, no pattern region has the smallest complexity value since it has the largest land area. However, having the

largest complexity value, the radial study indicates that the total building façade area is the largest in this study area.

Calculations for the SEF are conducted with the buildings, trees and topography within the close proximity area as described in Section 3.2.3. SEF has a very similar approach to calculation SF. The same points chosen on the BIPV surface for SF analysis are also used in SEF calculations. Results for the multiple points are averaged, and a single SEF value is obtained for a single BIPV surface. Results of SEF calculations show that the radial study has the smallest exposure to the sky from both façade and roof surfaces. This result can be linked with the previously reported urban density metrics for the radial pattern. The radial urban area is the densest study area with the lowest exposure to the sky. Following the same interpretation, it can be said that no pattern study area has the largest exposure to the sky and offers a less compact urban area.

Average SEF values for facades have an 11.21% range, whereas the gap between the maximum and minimum values is only 5.08% for roofs. As expected, roofs receive less shading and are more open to the sky than facades. The selection of the study areas is also responsible for this result. Similar building heights prevent abrupt changes in building heights and rooftop shading.

5.2 Results

The proposed method for predicting BIPV potential in urban contexts is demonstrated in the urban buildings of Ankara located in four regions. The results are organized into two sections. The first section reports the solar potential of building modeled in urban areas in Ankara with different morphologies. The second section presents the BIPV potential prediction performances with ML models.

5.2.1 Annual BIPV Potential

Solar potential is evaluated with three metrics, total and unit area PV generation (1), LMI (2), and LCF (3). Considering application areas, three solar potential metrics are evaluated for the only facade, only roof, and combined roof and facade applications. After a surface-based simulation approach for more accurate estimations, results are aggregated for each building, and building-based results are presented in the following subsections.

5.2.1.1 BIPV Generation

Annual BIPV generation from roof and façade surfaces per building can reach around 3 MWh considering the extremities (Figure 5.9). However, the majority of the building BIPV generation, either roof or façade surfaces, lay between 0 to 1.5 MWh, with an average of around 450 kWh in total for a year. Total BIPV generation can be positively linked to the amount of application area when the shading from the surrounding environment is not considered.

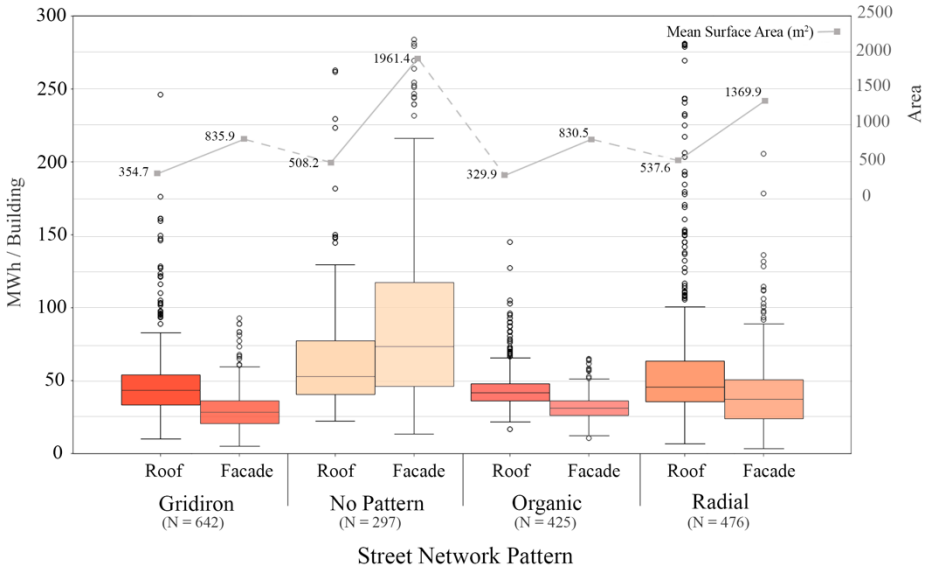


Figure 5.9 Annual BIPV Generation per building.

However, if shading is considered, façade surfaces are more prone to less insolation than roofs. Many studies in the literature report that facades receive only around one-third of rooftop irradiance, confirming the results of the simulation-based total BIPV generation. The statistical description of the data can be found in Appendix A.

For three of the studied network patterns, the BIPV generation of roof surfaces was greater than the façades. However, this relationship was not observed in the no pattern region. No pattern study area offers a less dense urban area, which can reduce the shading caused by the buildings nearby on facades. In addition, the ratio of mean BIPV-facade area to mean BIPV-roof area is the most significant in the no pattern case. Although roofs receive more irradiation than facades, excessive façade area may surpass the roof PV generation.

5.2.1.2 Unit Area BIPV Generation

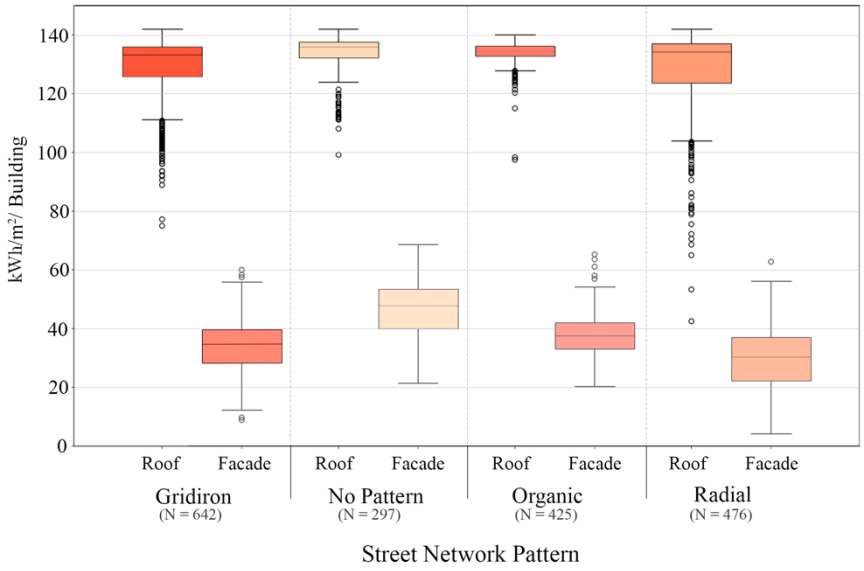


Figure 5.10 Annual BIPV generation per building facade and roof area.

PV generation per m^2 application area gives more comparable results independent of the largeness of the PV deployment area. The unshaded roofs set the upper limit for roof potential, while the lower limit varies between the studied four urban areas. The

statistical description of the data can be found in Appendix A. The effect of shade and different insolation levels may be recognized most clearly with the comparative analysis of energy generation per application area. The average generation per façade area is 27.44% of the average generation per roof area. No pattern study area overperformed in both application areas.

5.2.1.3 Load Matching

BIPV generation can be directly used onsite or stored. Due to the inclusion of batteries in the process, the storage of photovoltaic (PV) energy is both ecologically and economically expensive. Consequently, the direct use of energy produced onsite is advantageous. For grid integration and possible energy savings from onsite energy production, it is crucial to understand the temporal energy match and surplus. In this study, residential and nonresidential lighting (Q_L) and equipment (Q_{Eq}) energy consumption are considered for load matching. Hourly energy demand is calculated using continuous EnergyPlus standard schedules and loads for midrise apartments and open offices (Table 5.4). Hourly plots of the annual schedules can be found in Appendix C.

Table 5.4 Energy loads for residential and non-residential buildings.

Building function	Load type	Load (W/m^2)
Residential	Q_L	11.8404
	Q_{Eq}	3.875
Non-residential	Q_L	11.8404
	Q_{Eq}	7.6424

Temporal load match and surplus energy over a year for four neighborhoods are shown in Figure 5.11. While interpreting the LMIs, it is essential to consider the ratio of residential buildings to non-residential buildings and the FAR. The radial pattern study area is dominated by the offices and commercial buildings, which result in the

highest energy demand of around 83 GWh. On the other hand, the remaining three study areas are mainly occupied by residential buildings, and their energy demand range between 30 to 60 GWh.

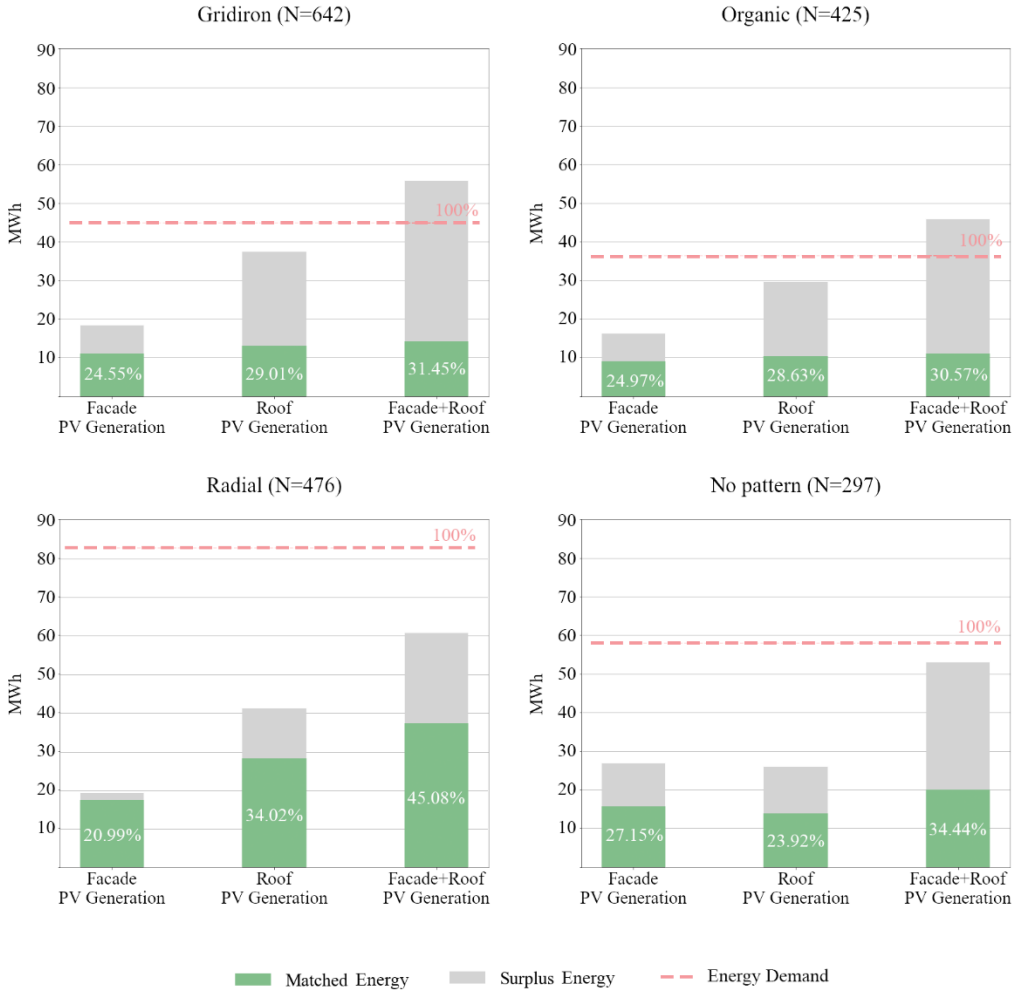


Figure 5.11 Matched and surplus energy for a year.

Although the radial design has the largest energy consumption, it also offers the highest temporal load match potential. This result can be related to the simultaneity of the occupied hours of buildings and energy generated hours by BIPV. The decrease in surplus energy also confirms that energy generation is increasingly matched in the radial pattern. For the other study areas, the proportion of surplus

energy to PV generation is smaller than that of the radial pattern. If stored, BIPV generations can meet the electricity demand of gridiron and organic study areas.

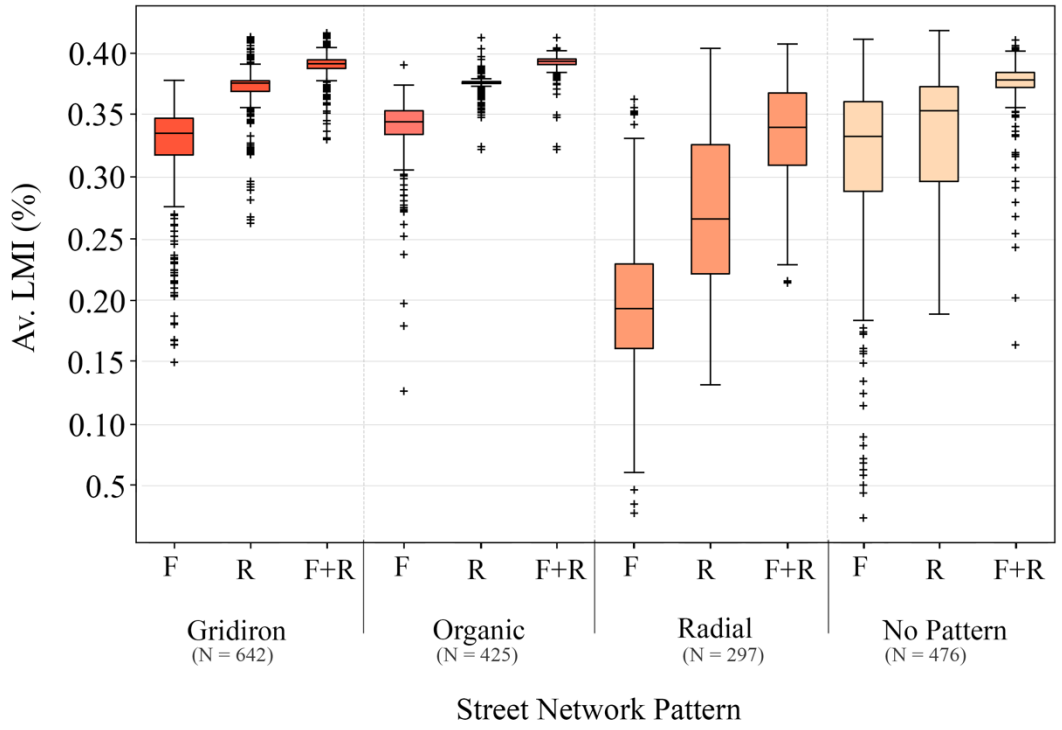


Figure 5.12 Building LMI for three BIPV applications, roof (R), façade (F) and roof and façade (R+F).

Hourly LMI, Eq (2), is averaged for the buildings in study areas. The distribution of building LMI in different morphologies can be found in Figure 5.12. According to Av. LMI, the organic study area, offers the greatest load match potential for all three application types. LMI prioritizes the ratio of energy generated to energy used. In the radial pattern, the energy used is considerably greater. Therefore, the radial pattern has the lowest LMI values.

However, Av. LCF results (Figure 5.13) show that buildings in the radial pattern have the greatest load covering potential with BIPV, particularly if applied to roofs or facades and roofs. Gridiron and organic patterned study areas offer very similar Av. LCF distributions. LCF, Eq (3), prioritizes the amount of energy matched

proportionally to the need. Therefore, LMI and LCF metrics give different points of view regarding BIPV generation and load matching.

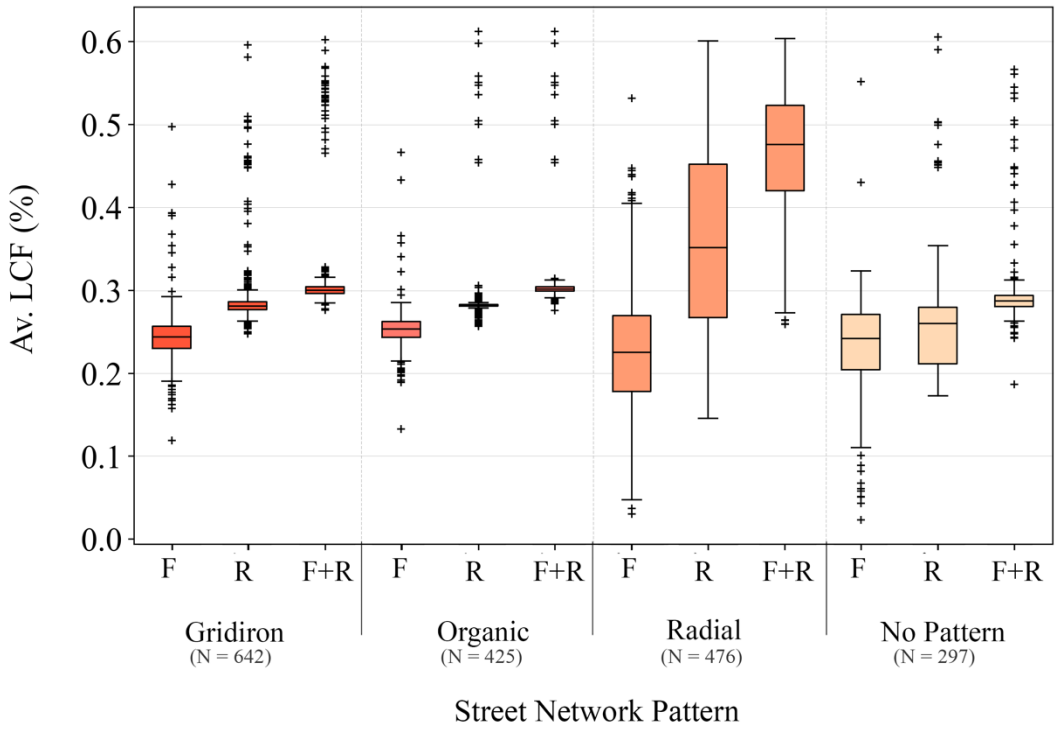


Figure 5.13 Building LCF for three BIPV applications, roof (R), façade (F) and roof and façade (R+F).

5.2.2 Predicting Annual BIPV Generation with ML

Data recorded from four neighborhood models representing different street network patterns and having different morphological characteristics are merged and shuffled to generate data necessary for training and testing prediction models. Energy generation per BIPV applied surface area (kWh/m^2) constitutes the output variable of façade and roof BIPV generation prediction models. Energy generation per m^2 of opaque surface area on vertical façade surfaces constitutes the output variable for façade predictions. For roofs, BIPV generation is divided by the inclined surface area that constitutes a closed roof geometry. Unit area BIPV generation from roof and

facade surfaces is predicted with two groups of design parameters; basic and complex, as described in Section 3.3.1.

These two sets of predictors are utilized for training three prediction algorithms for two application areas, roof and façade. RF, XGB, and MLP are the selected machine learning algorithms that have provided successful results in the literature for various research problems. After data preprocessing, ML models (2 × 3 × 2) are trained, model hyper-parameters are finetuned, and prediction performances are presented with the results of 10-fold CV. Prediction performances of ML models are compared with performance metrics of R2, MAE, MSE, RMSE, and nRMSE.

Façade and roof prediction performances are presented in two sections with two subsections. In the first subsection, the overall performance of built ML models is reported, whereas the second section presents the accuracy with which various morphologies were predicted.

5.2.2.1 Façade

The monotonic relationship between input and output parameters gives an initial understanding of which input features influence the target feature. Correlations between the basic and complex input features and the output feature of façade PV generation (kWh/m^2) are presented in Figure 5.14 and Figure 5.15, respectively.

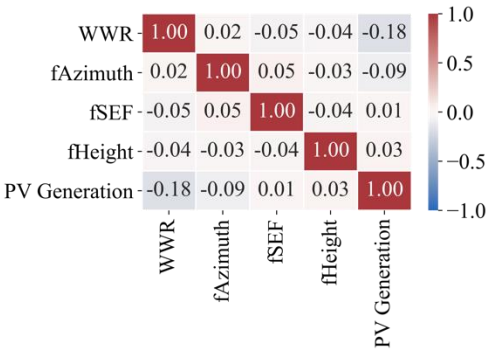


Figure 5.14 Correlations between the complex and output features for facades.

When the heat map for complex features and the output feature is examined, it can be seen that there is no correlation between the complex features and the output feature. In addition, the predictors do not have any correlations in themselves. Therefore, no linear dependency is observed in the complex predictors. This result can be associated with probable more complex relationships with the output feature. For instance, the increase in fHeight does not necessarily bring about an increase in PV generation. Shading might also increase and result in a reduction in PV generation.

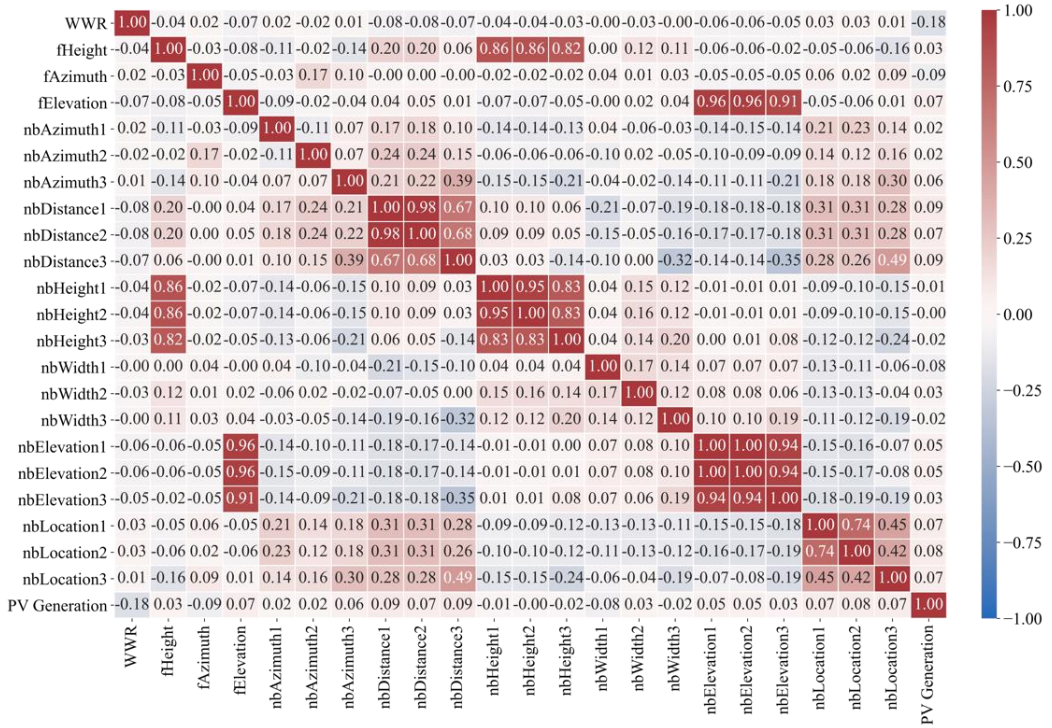


Figure 5.15 Correlations between the basic and output features for facades.

For the basic predictors, WWR has a weak negative linear relationship with the output feature (ρ = -0.18). On the other hand, strong positive correlations are observed for input features belonging to the same category. For instance, the height of the first neighbor surface (nbHeight1) is strongly correlated with the second and third surfaces (nbHeight2 and nbHeight3). The same relationship can be observed in also the distance and elevation features. The reason for these correlations can be explained by the similarity of buildings located in close proximity. Neighbor

buildings tend to change together with the studied buildings in this research due to the more homogenous character of studied urban areas. Correlations observed between the context and building parameters belonging to the same category, such as nbHeight and fHeight, also support this argument. Similarly, façade elevation (fElevation) is linked to the elevation of neighbor buildings (nbElevation).

5.2.2.1.1 Prediction Performance

ML model hyper-parameters are fine-tuned with the random search method explained in Section 3.3.3. Fine-tuned model hyper-parameters can be found in Appendix D. The success of ML models in predicting PV generation per façade area is tested with the 10-fold CV method. Error metrics are presented with the mean and std values in Table 5.5. XGB model outperforms in predicting PV generation with both sets of predictors. Complex predictors exhibit greater predictive capacity in comparison to basic predictors.

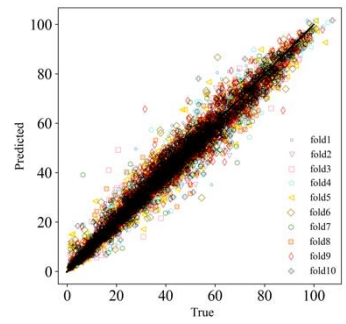
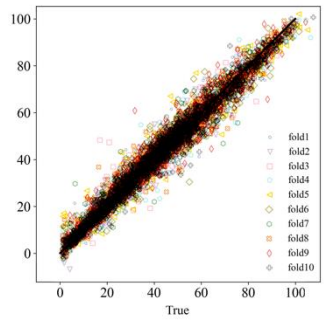
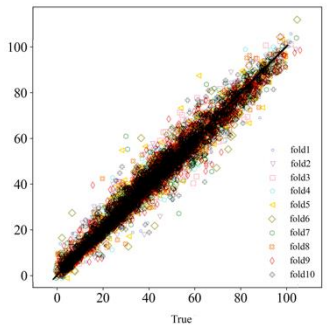
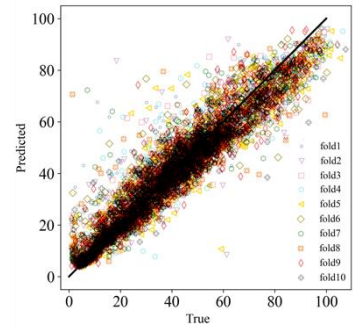
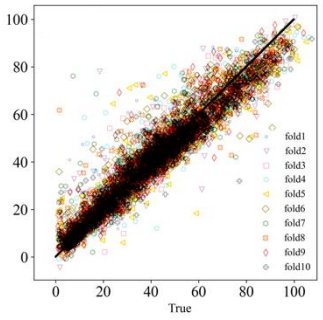
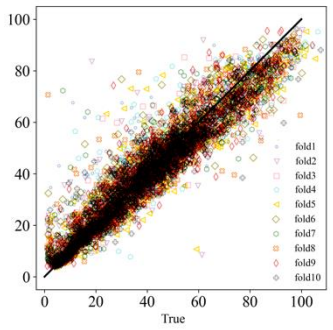
Table 5.5 10-fold CV results for facade prediction models (mean ± std).

		RF	XGB	MLP
Complex	R^2	**0.97** ± 0.00	**0.97** ± 0.00*	**0.97** ± 0.00
	MAE	2.92 ± 0.10	**2.85** ± 0.08*	2.87 ± 0.17
	MSE	19.16 ± 1.55	**17.04** ± 1.18*	17.93 ± 2.16
	RMSE	4.37 ± 0.18	**4.13** ± 0.15*	4.23 ± 0.25
	nRMSE	11.48 ± 0.53%	**10.83** ± 0.44%*	11.12 ± 0.62%
Basic	R^2	0.90 ± 0.01*	**0.92** ± 0.01	0.91 ± 0.01
	MAE	5.18 ± 0.21*	**4.81** ± 0.22	5.11 ± 0.20
	MSE	58.70 ± 5.81*	**48.85** ± 6.76	55.58 ± 5.15
	RMSE	7.65 ± 0.38*	**6.97** ± 0.48	7.45 ± 0.36
	nRMSE	20.09 ± 1.13%*	**18.17** ± 1.24%	%19.60 ± 1.00%

Bold: best mean value, *: lowest std.

With complex predictors, the XGB model provides the most accurate predictions. R^2 values are very close to each other in the utilized models. However, the error metrics show a slight difference in favor of the XGB model. On average, predictions are 10.83% off from the actual values according to nRMSE with complex features. RF and MLP result in comparable errors. XGB also outperforms the models trained with basic predictors. However, errors increased significantly. Predictions of the best-performing increase to 18.17% nRMSE. Results of the 10-fold CV are visualized with the scatter plots of predicted and true values for different folds in Table 5.6.

Table 5.6 Predicted and true values according to 10-fold CV results for predicting facade PV generation (kWh/m^2).

	RF	XGB	MLP
Complex			
Basic			

5.2.2.1.2 Predicting Different Urban Fabrics

The effectiveness of models in predicting a variety of urban topologies may be leveraged to improve the accuracy of prediction for each urban morphology. For

façade predictions, the success of ML models in predicting studied urban morphologies is presented with 10-fold CV results (Table 5.7). ML model that performs the best in each predictor category is reported. Due to the variety in the length of data samples for different categories, the number of test data (N) varies. Overall, three categories, gridiron, organic, and no pattern, were interchangeably predicted the best with respect to different performance metrics, whereas the prediction performance for the radial pattern was always the poorest.

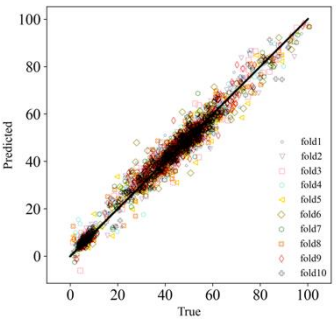
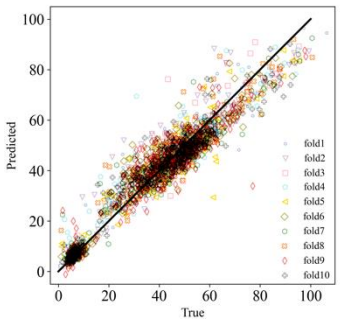
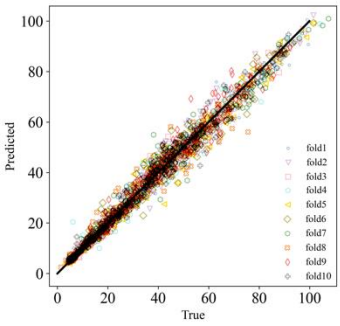
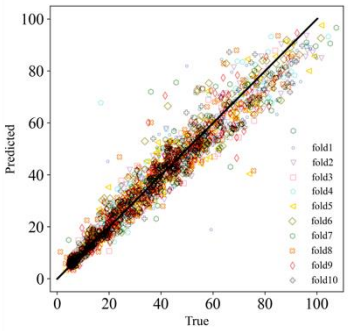
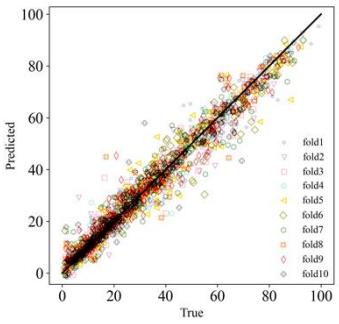
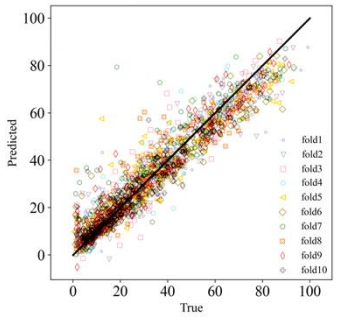
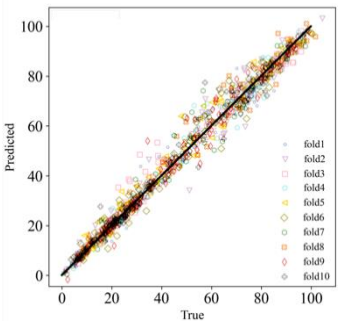
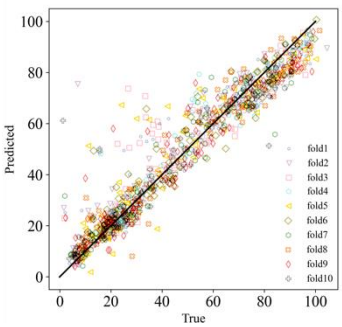
Table 5.7 Prediction performances for different urban fabrics (mean ± std).

		Gridiron (N = 223)	Organic (N = 177)	Radial (N = 159)	No pattern (N = 95)
Complex (XGB)	R^2	0.97 ± 0.00	**0.98** ± 0.00	0.95 ± 0.01	**0.98** ± 0.00*
	MAE	**2.55** ± 0.15*	**2.55** ± 0.21	3.46 ± 0.30	2.95 ± 0.19
	MSE	**13.24** ± 1.57*	13.37 ± 2.29	25.66 ± 4.36	17.81 ± 1.81
	RMSE	**3.63** ± 0.21*	3.64 ± 0.33	5.05 ± 0.46	4.22 ± 0.21*
	nRMSE	9.92 ± 0.51%*	9.43 ± 0.84%	14.92 ± 1.34%	**8.88** ± 0.63%
Basic (XGB)	R^2	**0.93** ± 0.01*	**0.93** ± 0.02	0.89 ± 0.03	0.90 ± 0.04
	MAE	4.34 ± 0.32*	**4.28** ± 0.50	5.44 ± 0.55	5.80 ± 0.74
	MSE	**37.79** ± 5.84*	39.35 ± 10.00	58.22 ± 14.55	76.91 ± 25.18
	RMSE	**6.13** ± 0.47*	6.22 ± 0.86	7.58 ± 0.92	8.67 ± 1.41
	nRMSE	16.76 ± 1.51%*	**16.10** ± 2.13%	22.35 ± 2.06%	18.25 ± 3.05%

Bold: best mean value, *: lowest std

The prediction performance for radial pattern deteriorates the overall prediction performance. This interpretation can be drawn from the decrease in prediction performance of the XGB model, as shown in Table 5.5. Gridiron, organic, and no pattern morphologies are predicted with nRMSE ≤ 9.92% on average, whereas, with the inclusion of data samples from radial pattern to the test dataset, the results recede to nRMSE = 10.83% on average.

Table 5.8 Predicted and true values according to 10-fold CV method for predicting facade PV generation (kWh/m^2).

	Complex (XGB)	Basic (XGB)
Gridiron (N = 254) (RF)		
Organic (N = 208)		
Radial (N = 188)		
No pattern (N = 117)		

On the other hand, it can be argued that patterns other than the radial category do not dominate the prediction performance as the observed prediction performances are alternately best in various categories depending on performance metrics. True and predicted values for different morphologies based on the best-performing ML models are shown in Table 5.8.

5.2.2.2 Rooftop

The monotonic relationships between the basic and complex input features and the output feature of roof PV generation (kWh/m^2) are presented with ρ values in Figure 5.16 and Figure 5.17, respectively.

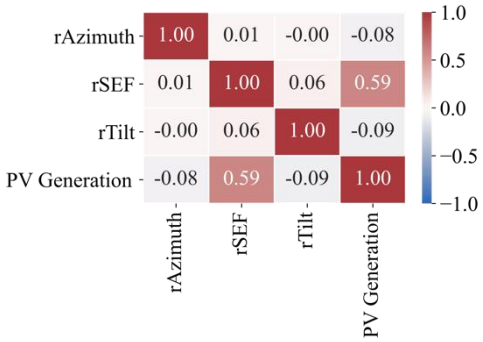


Figure 5.16 Correlations between the complex and output features for rooftops.

Rooftop PV generation per unit area is moderately correlated (ρ = 0.59) with the roof SEF (rSEF) in complex parameters. On the other hand, roof tilt and azimuth angle do not have a monotonic relationship with the output or other input features in the complex set of predictors for roof PV potential prediction.

Correlations observed between the output and input features for the basic predictors were less than 0.1, indicating that the output feature has almost no relationship with the input features. Similar to the context-related features of façade prediction, features describing the characteristics of neighboring buildings are correlated with each other. In addition, a weak positive relationship between neighbor surface location and distance is observed ($0.28 < \rho < 0.50$). However, this relationship is

presumably random since the location of a neighboring building is not related to the distance from the analysis surface.

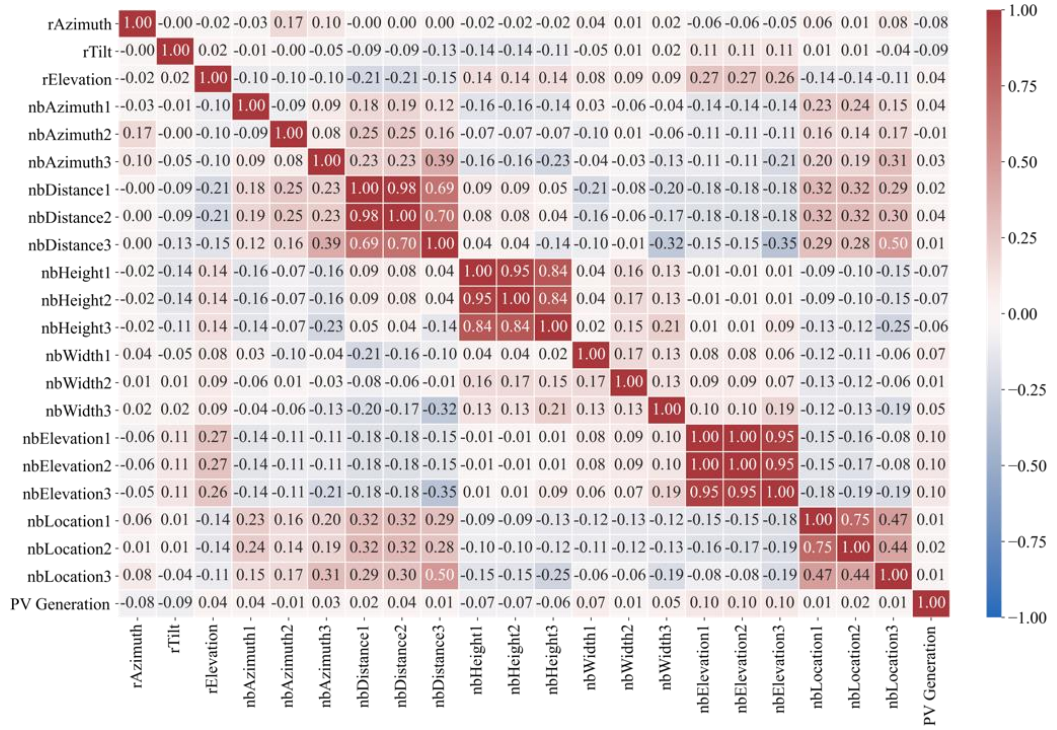


Figure 5.17 Correlations between the basic and output features for rooftops.

5.2.2.2.1 Prediction Performance

The random search method is applied to fine-tune model parameters for each predictor set. Fine-tuned model parameters are presented in Appendix D. The success of fine-tuned prediction models in prediction rooftop PV generation (kWh/m^2) is explored with a 10-fold CV. The results are presented with mean values and standard deviations in Table 5.9. RF model outperforms when predicting with complex predictors, whereas XGB offers the most accurate predictions with basic predictors. Predictions with complex features result in smaller errors compared to basic features, likewise the façade predictions.

The RF and XGB models give the highest mean R^2 value of 0.94 with a standard deviation of 0.01 when making predictions with complex predictors. RF model

considerably outperforms when all performance metrics are involved. It gives the best mean values with the lowest std, which shows that the predictions are both more accurate and precise when compared to other models.

Table 5.9 10-fold CV results for best performing roof PV generation prediction models (mean ± std).

		RF	**XGB**	MLP
Complex	R^2	**0.94** ± 0.01*	**0.94** ± 0.01*	0.92 ± 0.01*
	MAE	**1.73** ± 0.12	2.06 ± 0.10*	3.25 ± 0.24
	MSE	**16.65** ± 1.52*	17.66 ± 1.67	22.65 ± 3.08
	RMSE	**4.08** ± 0.18*	4.20 ± 0.20	4.75 ± 0.32
	nRMSE	**3.13** ± 0.13%*	3.32 ± 0.15%	3.64 ± 0.25%
Basic	R^2	0.66 ± 0.04	**0.74** ± 0.03*	0.54 ± 0.03*
	MAE	4.94 ± 0.29	**4.53** ± 0.19*	6.61 ± 0.30
	MSE	92.00 ± 15.47	**72.40** ± 10.75*	125.90 ± 14.26
	RMSE	9.56 ± 0.80	**8.49** ± 0.63*	11.20 ± 0.63*
	nRMSE	7.34 ± 0.06%*	**6.51** ± 0.51%	8.59 ± 0.49%

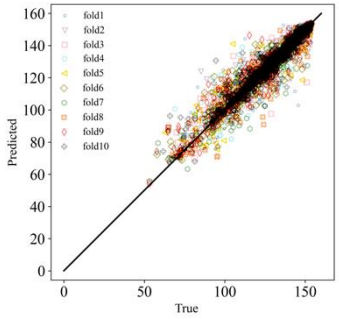
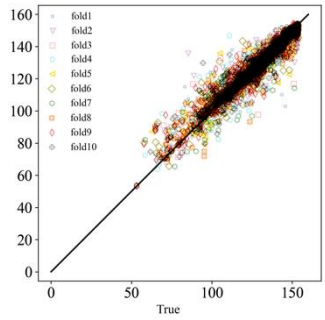
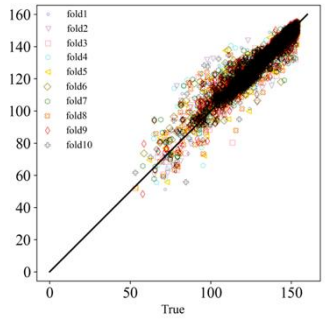
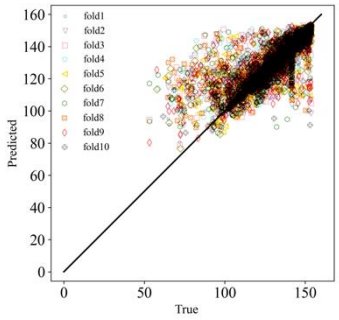
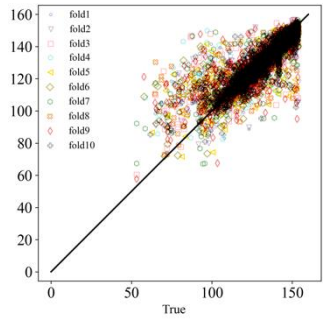
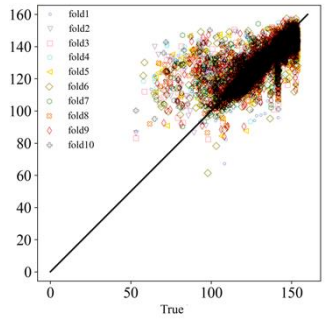
Bold: best mean value, *: lowest std

Basic predictors have the lowest prediction error in the XGB model with R^2 values of 0.74. R^2 values are considerably lower in predictions with basic predictors than in models trained with complex predictors. Similarly, nRMSE reaches 6.51% with basic predictors, whereas it was only 3.13% with complex predictors.

When compared to façade predictions, although R^2 values of roof predictions are lower than the façade predictions, nRMSE is lower than those of façade. This result can be explained by the variety in the range of predicted values. 75% of the values fall in the range of 120.9 to 153.6 kWh/m^2 for roofs, whereas the same portion of data is ranged between 22 to 68 kWh/m^2 for facades (Appendix A-İ). To illustrate, the 10 kWh/m^2 error is larger in façade predictions proportional to its mean value.

True and predicted values for different morphologies based on the best-performing ML models are shown in Table 5.10.

Table 5.10 Predicted and true values according to 10-fold CV results for predicting roof PV generation (kWh/m^2).

	RF	XGB	MLP
Complex			
Basic			

5.2.2.2.2 Predicting Different Urban Morphologies

Prediction accuracy of different urban morphologies has the potential to indicate the weaknesses and strengths of prediction models with respect to different urban morphologies based on street network patterns. 10-fold CV results are presented with mean and std values for predicting rooftop BIPV generation in Table 5.7. Roof surfaces of buildings in urban areas with organic morphology significantly predicted

the best with basic and complex predictor sets. On the other hand, the radial pattern predictions were the poorest in both predictor groups.

Table 5.11 Prediction performances for different urban fabrics (mean ± std).

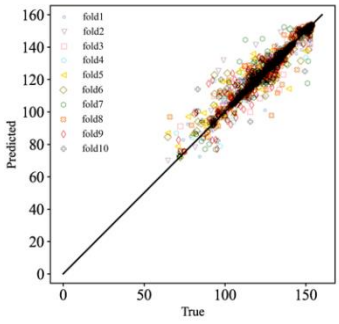
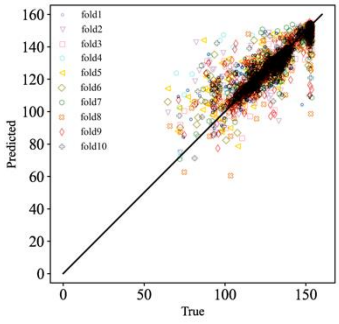
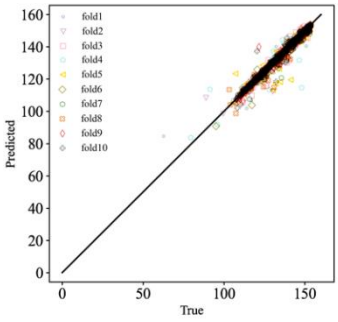
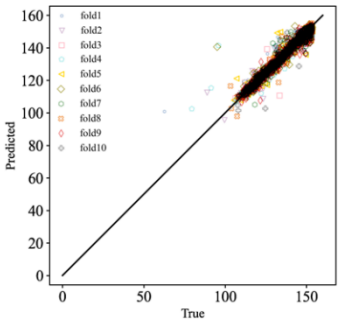
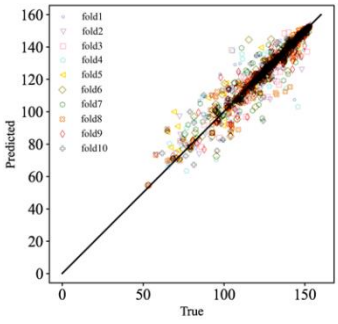
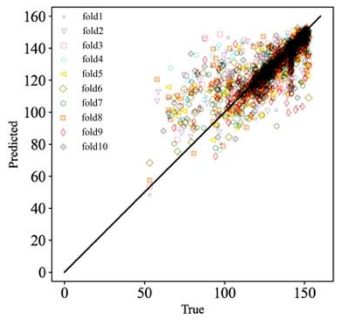
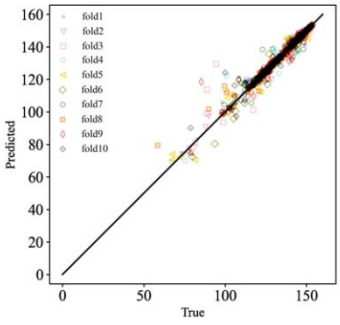
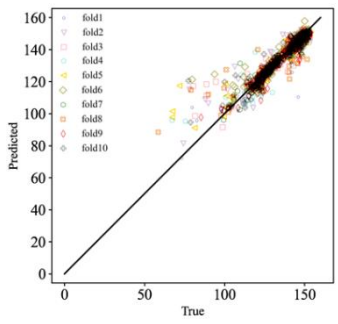
		Gridiron (N = 254)	Organic (N = 208)	Radial (N = 188)	No pattern (N = 117)
Complex (RF)	R^2	0.93 ± 0.02*	**0.97** ± 0.02*	0.92 ± 0.02*	0.96 ± 0.02*
	MAE	1.98 ± 0.24	**0.910** ± 0.15	2.63 ± 0.30	1.19 ± 0.14*
	MSE	19.59 ± 4.57	**5.22** ± 2.91*	29.82 ± 7.49	9.30 ± 4.89
	RMSE	4.40 ± 0.52*	**2.22** ± 0.58	5.42 ± 0.67	2.96 ± 0.76
	nRMSE	3.43 ± 0.41%*	**1.66** ± 0.44%	4.25 ± 0.55%	2.23 ± 0.57%
Basic (XGB)	R^2	0.71 ± 0.06	**0.93** ± 0.03*	0.59 ± 0.06	0.85 ± 0.05
	MAE	5.13 ± 0.35	**1.97** ± 0.15*	7.45 ± 0.56	3.14 ± 0.41
	MSE	81.37 ± 16.67	**11.69** ± 5.17*	149.44 ± 29.63	34.35 ± 12.74
	RMSE	8.98 ± 0.91	**3.37** ± 0.68*	12.17 ± 1.21	5.76 ± 1.16
	nRMSE	7.00 ± 0.75%	**2.52** ± 0.51%*	9.52 ± 0.98%	4.33 ± 0.989%

Bold: best mean value, *: lowest std.

Although BIPV generation in urban areas with organic patterns can be predicted with nRMSE as low as 2.52% with basic predictors, overall predictions were made with 6.51% nRMSE (Table 5.9). This drop in the overall prediction performance with basic predictors can be associated with the poor prediction performance observed in gridiron and radial patterns. However, organic and no pattern morphologies were predicted more successfully when compared prediction performance of the overall dataset and gridiron and radial morphologies.

The prediction performances based on different morphologies is more homogenous when compared to predictions with basic features. The gap between the highest and lowest nRMSE is 2.59 % for complex predictors, whereas it is 7% in basic predictors. True and predicted values for different morphologies based on the best-performing ML models are shown in Table 5.12.

Table 5.12 Predicted and true values according to 10-fold CV method for predicting roof BIPV generation (kWh/m^2).

	Complex (RF)	Basic (XGB)
Gridiron (N = 254) (RF)		
Organic (N = 208)		
Radial (N = 188)		
No pattern (N = 117)		

5.3 Discussion

A data-driven method consisting of three main steps is proposed and applied to four existing urban areas in Ankara. The results show that the ML models have the potential to be used during the design development processes for buildings to be built in a city fabric. The described methodology can be adapted to estimate the BIPV potential of buildings in many cities of Turkey since the utilized data sources are publicly available and can be accessed easily. The application of the method to cities worldwide is also possible if alternative data sources to EPCs are available. The described method is viable and performs well with an R^2 value as high as 0.97 and an nRMSE as low as 3.13% for predicting BIPV potential. However, the research can be improved and challenged in several ways. The following paragraphs discuss possible limitations and improvements based on the three main steps of the methodology.

The **3D Modeling** step is performed using several open-data sources. Open-data sources increase research opportunities and contribute to knowledge production in literature. However, it should be noted that crowd-sourced data may have flaws and be out-of-date. For instance, new constructions can be missing in OSM databases. Therefore, the timeliness of the city model and the dataset depends on the timeliness of the open-source data, which cannot be assured.

Simulation-based data generation approach allows the estimation of the BIPV potential for every building in a city. Simulations have the power to test many scenarios when large-scale real-world applications are not feasible and not possible. In a sense, simulations allow researchers to foresee the possible outcomes of real-world applications. In this study, 1942 buildings from four districts are assessed as if they had applied BIPV modules into their envelopes. In reality, the application of this scenario comes with many hardships related to grid operation, manufacturing, economics, and legislation. The power of the simulations is once again pronounced against real-world problems. However, simulation outputs may deviate slightly from

the actual results of on-site installments. Therefore, validations with historical data are advised to achieve more accurate results.

The **predictive model development** phase focuses on training and testing three state-of-the-art ML techniques for the proposed tool in this thesis. Supervised ML techniques applied in this study allow forecasting unknown outcomes of an activity/problem by learning from the known outcomes of similar activities/problems. BIPV generation per building facade and roof area constitute the problem we aimed to predict in this thesis. Although promising results were achieved with the utilized models, it should be noted that there is always room for improvement by updating the datasets, tuning model hyperparameters, and utilizing different ML models. In the following paragraphs results of the predictive models are discussed further with specific concentrations:

Impact of input features: Two groups of input features, basic and complex, are used to predict the output features. Basic predictors can be calculated easily with a few measurements, either from estimations from images of a building with nearby urban elements or simply by looking at an urban setting. In contrast, complex predictors require 3D models and annual shading analysis. These two groups of predictors can be a viable method for prediction in different contexts based on model dependency. Complex features provided better prediction performance compared to predictions with basic features. For façades, complex predictors resulted in up to 0.97 R^2, whereas the best results with basic predictors were only 0.92. The difference between complex and basic predictors was even more apparent for roof predictions. The best model for roof predictions with complex features resulted in an R^2 value of 0.94, whereas it was only 0.74 with basic predictors.

Understanding what the input features signify may provide insight into the fact that complex features offer better prediction performance. Complex features include a key parameter, fSEF, and rSEF, that informs the model about the shading factor. In addition to that, the orientation, tilt, and core geometric information about the analysis surfaces are included in the complex predictors. Both predictor sets include

the core information about the analysis geometry. However, parameters that describe the shading conditions on the site, which are indicated under the context category in Table 3.3 and Table 3.4 for facades and roofs, respectively, vary significantly. The prediction performance difference in complex and basic features could be associated with the success of input features in representing the shading conditions. Since the calculation of SEF involves more sophisticated shading analysis, complex features outperform in the prediction of BIPV generation.

ML models: Three ML models, RF, XGB, and MLP, are utilized during the exploration of prediction performances for the proposed method. BIPV generation from façade and roof surfaces was predicted with two sets of predictors. Hyper-parameters of each model was fine-tuned with the random search method, and the results are presented in Appendix D. Four (Complex-Façade, Basic-Façade, Complex-Roof, Basic-Roof) of the twelve ML models developed are selected for the proposed tool based on 10-fold CV results (Table 5.13). Three of the selected ML models used the XGB method, whereas one was RF. XGB method exhibited relatively better prediction performance addressing the research problem of this thesis.

Table 5.13 Selected ML models.

BIPV Application	Predictor type	Selected Model
Facade	Complex	XGB
	Basic	XGB
Roof	Complex	RF
	Basic	XGB

In addition to the utilized ML techniques, different methods, such as support vector machines, kernel ridge regression, and ensemble models, can be explored. RF and XGB are also ensemble models that provided relatively good results in this study. Other ensemble models, which combine several different ML techniques, such as adaptive boosting, and light gradient boosting, can be utilized in future studies.

Furthermore, selected models are fine-tuned with the random search method. The search grid or the number of iterations can be extended to scan a more expansive search space to obtain better prediction performance. Fine-tuned parameters could be improved by applying various other hyper-parameter tuning approaches, such as grid search, Bayesian optimization, or evolutionary algorithms.

Prediction performance: The proposed method is aimed to aid in design development by allowing real-time evaluation of design alternatives. Therefore, accurate predictions are desired to inform design decisions. A mean relative error of around 10% was found acceptable for the early design development phase in many studies (Martín-Chivelet et al., 2022). The developed tool in this thesis predicts BIPV generation for rooftop surfaces with 3.13% and 6.51% nRMSE with complex and basic predictors. For facades, nRMSE reaches up to 10.83% and 18.17% with complex and basic predictors, respectively.

Basic predictors offer lower prediction performance than façade predictions, as discussed above. However, during the early design phase, it is important to assess initial design ideas even before drafting or massing studies. Although errors are higher, the effort required to estimate BIPV is significantly lower with basic predictors. Basic predictors can potentially democratize the assessment of BIPV since the input features can be easily estimated even with eye. On the other hand, complex predictors perform well with metrics directly associated with solar energy generation in the literature. However, simple shoebox models are needed for the calculation of complex predictors, more specifically, the SEF.

Predicting different urban morphologies: Numeric data recorded from four urban areas representing different urban morphologies constitute the datasets necessary for training and testing ML models. During the selection of the ML methods, data from four neighborhoods were aggregated. Models trained with the combination of four neighborhoods were also tested to explore which urban morphologies were predicted with which accuracy by the proposed tool. This exploration can indicate the strengths and weaknesses of the developed tool according to street network patterns.

Results have shown that generalizing the best-predicted morphology is not possible for facades and roofs. In contrast, the radial morphology was considerably poorly predicted with all developed ML models. This result can be associated with the morphological differences in the data recorded from radial urban areas. Several urban morphology indicators are presented in Table 5.3. The radial pattern is the densest urban area studied in this thesis, with 5.97 buildings per land area. Similarly, FAR is the highest among the other urban areas with 1.99. These two metrics imply that the urban area representing the radial pattern is more densely built than other urban areas studied.

In addition, the complexity is observed to be the highest in the radial study area. Complexity is one of the metrics that show how undulating an urban form is. Researchers have associated the highest complexity with lower solar availability on facades and ground (Chatzipoulka et al., 2016). Conforming the finding of previous studies, unit area PV generation was the lowest in the radial pattern. Higher complexity observed in the radial pattern can be associated with decreased solar potential and might negatively impact the prediction performance.

The ratio of visible sky dome, SEF, from roof and façade surfaces was also the smallest in the radial study area compared to other morphologies. This indicates that buildings in the radial study area were more shaded. SEF is positively correlated with the solar potential in the literature (Poon et al., 2020). Similar to the implications of higher complexity, lower SEF can be associated with a drop in BIPV potential and prediction performance. All in all, significant differences in prediction performances for different street network patterns imply the need for urban morphology inclusive approaches in predicting BIPV potential.

CHAPTER 6

CONCLUSION

The energy transition from fossil fuels to renewable energy sources encourages electrification in the built environment, and BIPV is one of the most viable solutions for onsite power production. Buildings should and will increasingly play an active role in energy generation since they are one of the greatest energy consumers. In this regard, the main objective of the thesis was to develop a method that can evaluate energy generation potential from building envelopes in cities. A number of research questions were posed and addressed in the process. The findings and contributions of the research are summarized in the following subsections.

6.1 Revisiting Research Questions

This research tackles the question of to what extent data-driven methods can predict solar energy potential. Assessment of solar energy generation from building envelope is gaining more importance since limited energy resources of the world are running low day by day, and the use of fossil fuels contributes significantly to the effects of climate change. Buildings in cities are one of the largest energy consumers that should become more energy-efficient and actively involved in energy generation from renewables. Therefore, a method is developed as a response to the need for data-driven design tools that can guide performance-driven design decisions. Prediction models relying on ML techniques have the potential to become viable design tools that can aid in decision-making for designing performance-driven buildings. In order to address the main research question, several sub-questions were posed and answered during this research,

What are the drivers of the solar energy potential of buildings in urban environments? The research question aims to identify and analyze factors affecting the solar energy generation capacity of buildings in cities. Three categories were proposed in the identification process: climate and location, PV system, and building and context. Architectural and urban design interventions are further examined for the development of the proposed method. A thorough literature review shows that parameters describing building geometry, such as building pattern/typology, building/façade orientation, and building w/l/w to l and plot ratio, are the metrics frequently studied in relation to solar energy. These metrics are reflected in the inputs of the proposed methods as surface height, orientation (azimuth angle), tilt, and elevation. On the other hand, the relationship between a building and its context is critical since shading since BIPV systems are more commonly affected by partial shading, the irradiance affected by reflections from nearby objects, and extreme tilts, which may affect PV system losses. (Martín-Chivelet et al., 2022). Site coverage, av/std/min-max building height, and building density are the metrics more frequently examined for describing the urban contexts. These metrics are associated with input parameters of the developed tool under two categories, complex and basic. SEF, belonging to the complex input category, was found very powerful in determining the solar energy generation potential. In the basic category, neighbor surface distance, location (in angles), height, width, and elevation are used to inform the ML models about the impact of the context.

What is the impact of urban morphology, specifically the street network pattern, on solar potential? Different street network patterns based on district-scale energy generation capacity through BIPV are comparatively analyzed to answer this research question in CHAPTER 4. The same number of uniform buildings on the same land area with four street network patterns and three building height configurations compared. Urban areas without an identifiable street network offered greater solar energy generation potential for facades. The relationship between the four street network patterns remained the same with different building height options for facades as follows, $E_{\text{No pattern}} > E_{\text{Radial}} > E_{\text{Organic}} > E_{\text{Gridiron}}$. On the other hand, the

difference between four morphologies for rooftop BIPV generation was negligible. This result might be associated with the fact that the influence of shading on roofs was minimal since the examined urban areas had uniform building heights.

What would be the impact of district-scale BIPV deployment on meeting building electricity demands? The development of insights for urban areas to support future urban development and district-scale retrofit initiatives that target the use of renewable technologies is important. Four neighborhoods in Ankara are studied in this research. BIPV systems can be grid-connected or stand-alone. Grid-connected systems rarely have batteries for storage since they are neither environmentally friendly nor economical. Therefore, self-consumption is most favorable. To evaluate temporal load match, hourly energy generations and electricity demand (for lighting and equipment) are calculated. The highest energy match is observed for the radial urban area. For a typical year, if building facades and roofs are deployed with BIPV, %45 of the energy generated on-site can match the energy demand. Sole roof and façade BIPV have the potential to cover 34.02% and 20.99% of the electricity demand in the radial district, respectively. On the other hand, the lowest load matching was observed in the district with an organic pattern. Façade and roof BIPV deployment together can cover only up to 30.57% of the electricity demands. It should be noted that electrification in the built environment can encourage more effective use of BIPV generations. Similarly, the energy generated from one district has the potential to support neighboring districts. In most cases, more than half of the energy generated is surplus and should be utilized more effectively considering grid operations.

What is the accuracy with which the solar energy potential is predicted using data-driven methods? One of the objectives of this thesis was the development and comparative analysis of ML models that can evaluate design decisions for BIPV energy generation capacity in urban contexts. Three ML methods, RF, XGB, and MLP, were utilized for training four prediction models (for facades and roofs with basic and complex predictors). Several metrics were calculated to evaluate the prediction performance of prediction models with the 10-fold CV method. For façade

predictions, the XGB model provided the most accurate predictions with respect to calculated performance metrics. On average, R^2 values were 0.97 and 0.91 with complex and basic predictors, respectively. For roof predictions, RF and XGB models outperform complex and basic predictors, respectively. While the RF model offered a mean R^2 value of 0.938 with complex predictors, the XGB model could only reach 0.735 on average. Overall, the XGB models were most suitable for utilizing the proposed method. The average nRMSE of models with the best prediction performance was less than or around 10% (except for facade prediction with basic predictors), which shows that the proposed method has the capacity to aid in early design phases.

6.2 Contributions and Final Words

Contributions of this thesis can be highlighted by the combination of the intervention and evaluation phases of the thesis methodology. Primary contributions include,

- Development of a method based on ML techniques that can run with two-level user-defined inputs, basic and complex, focusing on the architectural and urban design parameters. Developed models can be utilized in the prediction of BIPV generation from facades and rooftops in different cities all over the world.
- Prediction of façade BIPV generation in cities with a ML-based method.

In addition to primary contributions, minor contributions can be listed as follows,

- A classification method, inclusive of building and urban design parameters, in factors affecting BIPV performance: Climate and location, PV system, and building and context.
- A BIPV assessment method inclusive of topographical differences and vegetation.
- Prediction of BIPV performance with basic features that can be easily recorded by hand.

- Comparative analysis of four street network patterns with a controlled study design.
- Evaluation of four neighborhoods in Ankara in terms of total energy generation and load matching with simulations.

All in all, this research has shown that urban street network patterns impact solar generation through BIPV and data-driven methods have the potential to evaluate performance-driven decisions. The evaluation of BIPV potential can be used for the large-scale assessment of future energy systems with decentralized power grids. With possible extensions and improvements, the proposed tool has the potential to support regulations for the successful integration of photovoltaics into the built environment.

REFERENCES

Adres Sorgulama. (n.d.). Retrieved June 22, 2022, from https://adres.nvi.gov.tr/VatandasIslemleri/AdresSorgu

Aghamolaei, R., Ghaani, M. R., Tahsildoost, M., & Zandi, M. (2019). A comprehensive energy-oriented approach for optimization of solar potential in urban contexts: an application study for residential districts. *Advances in Building Energy Research*, *13*(2), 205–219. https://doi.org/10.1080/17512549.2018.1488613

Agrawal, T. (2021). Hyperparameter Optimization in Machine Learning. In *Hyperparameter Optimization in Machine Learning*. Apress. https://doi.org/10.1007/978-1-4842-6579-6

Ahmadian, E., Sodagar, B., Bingham, C., Elnokaly, A., & Mills, G. (2021). Effect of urban built form and density on building energy performance in temperate climates. *Energy and Buildings*, *236*. https://doi.org/10.1016/j.enbuild.2021.110762

Ahmed, R., Sreeram, V., Mishra, Y., & Arif, M. D. (2020). A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. In *Renewable and Sustainable Energy Reviews* (Vol. 124). Elsevier Ltd. https://doi.org/10.1016/j.rser.2020.109792

Akhter, M. N., Mekhilef, S., Mokhlis, H., & Shah, N. M. (2019). Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques. In *IET Renewable Power Generation* (Vol. 13, Issue 7, pp. 1009–1023). Institution of Engineering and Technology. https://doi.org/10.1049/iet-rpg.2018.5649

Akin, Ö. (2001). Variants in Design Cognition. In C. M. Eastman, W. M. McCracken, & W. C. Newstetter (Eds.), *Design Knowing and Learning: Cognition in Design Education*. Elsevier.

Anil, K. J., Mao, J., & Mohiuddin, K. M. (1996). Artificial Neural Networks: A Tutorial. *Computer*, *29*(3), 31–44.

Apt, J., & Jaramillo, P. (2014). *VARIABLE RENEWABLE ENERGY AND THE ELECTRICITY GRID*. Routledge.

Assouline, D., Mohajeri, N., & Scartezzini, J. L. (2017). Quantifying rooftop photovoltaic solar energy potential: A machine learning approach. *Solar Energy*, *141*, 278–296. https://doi.org/10.1016/j.solener.2016.11.045

Bader, S., Ma, X., & Oelmann, B. (2019). One-diode photovoltaic model parameters at indoor illumination levels – A comparison. In *Solar Energy* (Vol. 180, pp. 707–716). Elsevier Ltd. https://doi.org/10.1016/j.solener.2019.01.048

BepTR. (n.d.). Retrieved June 22, 2022, from https://beptr.csb.gov.tr/bep-web/#/queryEkb

Berardi, U., & Graham, J. (2020). Investigation of the impacts of microclimate on PV energy efficiency and outdoor thermal comfort. *Sustainable Cities and Society*, *62*. https://doi.org/10.1016/j.scs.2020.102402

Boccalatte, A., Fossa, M., & Ménézo, C. (2020). Best arrangement of BIPV surfaces for future NZEB districts while considering urban heat island effects and the reduction of reflected radiation from solar façades. *Renewable Energy*, *160*, 686–697. https://doi.org/10.1016/j.renene.2020.07.057

Boccalatte, A., Thebault, M., Ménézo, C., Ramousse, J., & Fossa, M. (2022). Evaluating the impact of urban morphology on rooftop solar radiation: A new city-scale approach based on Geneva GIS data. *Energy and Buildings*, *260*. https://doi.org/10.1016/j.enbuild.2022.111919

Bredemeier, D., Schinke, C., Gewohn, T., Wagner-Mohnsen, H., Niepelt, R., & Brendel, R. (2021). Fast Evaluation of Rooftop and Façade PV Potentials Using Backward Ray Tracing and Machine Learning. *Conference Record of the IEEE Photovoltaic Specialists Conference*, 294–299. https://doi.org/10.1109/PVSC43889.2021.9518660

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5–32.

Brito, M. C., Freitas, S., Guimarães, S., Catita, C., & Redweik, P. (2017). The importance of facades for the solar PV potential of a Mediterranean city using LiDAR data. *Renewable Energy*, *111*, 85–94. https://doi.org/10.1016/j.renene.2017.03.085

Butti, K., & Perlin, J. (1980). *A Golden Thread: 2500 Years of Solar Architecture and Technology*. Cheshire Books.

CADMAPPER LLC. (2013). *CADMAPPER*. https://cadmapper.com/

Calcabrini, A., Ziar, H., Isabella, O., & Zeman, M. (2019). A simplified skyline-based method for estimating the annual solar energy potential in urban environments. *Nature Energy*, *4*(3), 206–215. https://doi.org/10.1038/s41560-018-0318-6

Cameron, C. P., Boyson, W. E., & Riley, D. M. (2008). Comparison of PV system performance-model predictions with measured PV system performance. *2008 33rd IEEE Photovoltaic Specialists Conference*.

Caruana, R., & Niculescu-Mizil, A. (2006). An Empirical Comparison of Supervised Learning Algorithms. *ICML '06: Proceedings of the 23rd International Conference on Machine Learning*. www.cs.cornell.edu

Chatzipoulka, C., Compagnon, R., & Nikolopoulou, M. (2016). Urban geometry and solar availability on façades and ground of real urban forms: using London as a case study. *Solar Energy*, *138*, 53–66. https://doi.org/10.1016/j.solener.2016.09.005

Chen, T., & Guestrin, C. (2016, March 8). XGBoost: A Scalable Tree Boosting System. *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. https://doi.org/10.1145/2939672.2939785

Chen, W., Wu, A. N., & Biljecki, F. (2021). *Classification of Urban Morphology with Deep Learning: Application on Urban Vitality*. https://doi.org/10.1016/j.compenvurbsys.2021.101706

Chokhachian, A., Perini, K., Giulini, S., & Auer, T. (2020). Urban performance and density: Generative study on interdependencies of urban form and environmental measures. *Sustainable Cities and Society*, *53*. https://doi.org/10.1016/j.scs.2019.101952

Cole, R., Purao, S., Rossi, M., & Sein, M. (2005). Being Proactive: Where Action Research Meets Design Research. *ICIS 2005 Proceedings*, 325–336. http://aisel.aisnet.org/icis2005/27

COST Action TU1205 (BISTS). (2015). *Overview of BISTS State of the Art, Models and Applications*. https://www.researchgate.net/publication/290429557

Cross, N. (2001). Design Cognition: Results From Protocol And Other Empirical Studies Of Design Activity. In C. M. Eastman, W. M. McCracken, & W. C. Newstetter (Eds.), *Design Knowing and Learning: Cognition in Design Education*. Elsevier.

Das, U. K., Tey, K. S., Seyedmahmoudian, M., Mekhilef, S., Idris, M. Y. I., van Deventer, W., Horan, B., & Stojcevski, A. (2018). Forecasting of photovoltaic power generation and model optimization: A review. In *Renewable and Sustainable Energy Reviews* (Vol. 81, pp. 912–928). Elsevier Ltd. https://doi.org/10.1016/j.rser.2017.08.017

Deason, J., Wei, M., Leventis, G., Smith, S., & Schwartz, L. (2018). *Electrification of buildings and industry in the United States Drivers, barriers, prospects, and policy approaches*.

Dobos, A. P. (2014). *PVWatts Version 5 Manual*. www.nrel.gov/publications.

Dresch, A., & Valle Jr., D. L. P. A. J. A. (2015). *Design Science Research A Method for Science and Technology Advancement*. Springer International.

Dubey, S., Sarvaiya, J. N., & Seshadri, B. (2013). Temperature dependent photovoltaic (PV) efficiency and its effect on PV production in the world - A review. *Energy Procedia*, *33*, 311–321. https://doi.org/10.1016/j.egypro.2013.05.072

Duffie, J. A., & Beckman, W. A. (1980). *Solar Engineering of Thermal Processes* (3rd Edition). Wiley and Sons.

EU. (2010). The Directive 2010/31 of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings (recast). *Official Journal of the European Union*.

Ghosh, A. (2020). Potential of building integrated and attached/applied photovoltaic (BIPV/BAPV) for adaptive less energy-hungry building's skin: A comprehensive review. In *Journal of Cleaner Production* (Vol. 276). Elsevier Ltd. https://doi.org/10.1016/j.jclepro.2020.123343

Google. (n.d.). *Google Earth*. Retrieved May 11, 2022, from https://earth.google.com/web/

Gutierrez, L., & Durrant, P. (2018). *Global Energy Transformation: A Roadmap to 2050*. https://doi.org/10.13140/RG.2.2.21732.60807

Hachem-Vermette, C. (2020). *Solar Buildings and Neighborhoods Design Considerations for High Energy Performance*. Springer. http://www.springer.com/series/8059

Haklay, M., & Weber, P. (2008). *openstreetMap: User-Generated street Maps*. www.openstreetmap.

Hegger, M. (2003). From Passive Utilization to Smart Solar Architecture. In C. Schittich (Ed.), *In Detail: Solar Architecture*. Birkhäuser Architecture.

Hestnes, A. G. (1999). *BUILDING INTEGRATION OF SOLAR ENERGY SYSTEMS*. *67*, 181–187. www.elsevier.com/locate/solener

IEA. (2012). *IEA SHC Task 41 Solar energy and Architecture - SOLAR ENERGY SYSTEMS IN ARCHITECTURE integration criteria and guidelines*.

IEA. (2015). *Towards net zero energy solar buildings. Task 40 (EBC Annex 52) – Net Zero Energy Solar Buildings SHC Position Paper*. OECD/IEA and AFD.

IEA SHC. (1996). *Photovoltaics in buildings : a design handbook for architects and engineers* (F. Sick & T. Erge, Eds.). James & James (Science Publishers) Ltd.

IEC. (2016). *IEC 61215-2:2016: Terrestrial photovoltaic (PV) modules–Design qualification and type approval–Part 2: Test procedures*. IEC-IECEE.

Irena. (2021). *Renewable Power Generation Costs 2020*. www.irena.org

Jakica, N., Yang, R. J., & Eisenlohr, J. (Eds.). (2019). *BIPV Design and Performance Modelling: Tools and Methods BIPV Design and Performance Modelling, Tools and Methods, IEA PVPS Task 15, Subtask E, Report IEA-PVPS T15-09: 2019*.

Jakubiec, J. A., & Reinhart, C. F. (2013). A method for predicting city-wide electricity gains from photovoltaic panels based on LiDAR and GIS data combined with hourly Daysim simulations. *Solar Energy*, *93*, 127–143. https://doi.org/10.1016/j.solener.2013.03.022

Jordan, D. C., & Kurtz, S. R. (2012). *Photovoltaic Degradation Rates -- An Analytical Review: Preprint*. http://www.osti.gov/bridge

Kalogirou, S. A. (2014). Photovoltaic systems. In *Solar Energy Engineering: Processes and Systems* (2nd Edition). Academic Press. https://doi.org/10.1016/b978-0-12-397270-5.00009-1

King, D. L., Boyson, W. E., & Kratochvil, J. A. (2003). *PHOTOVOLTAIC ARRAY PERFORMANCE MODEL*.

Kohavi, R. (1995). *A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection*. http://robotics.stanford.edu/~ronnyk

Lau, K. K. L., Lindberg, F., Johansson, E., Rasmussen, M. I., & Thorsson, S. (2017). Investigating solar energy potential in tropical urban environment: A case study of Dar es Salaam, Tanzania. *Sustainable Cities and Society*, *30*, 118–127. https://doi.org/10.1016/j.scs.2017.01.010

Lobaccaro, G., Lisowska, M. M., Saretta, E., Bonomo, P., & Frontini, F. (2019). A methodological analysis approach to assess solar energy potential at the neighborhood scale. *Energies*, *12*(18). https://doi.org/10.3390/en12183554

Mahaya, C., Zemmouri, N., Benharra, H., & Elnokaly, A. (2022). Solar Access Assessment in Semi-Arid Urban Context: An Application Study for Ten Urban Forms of Existing Apartment Buildings Districts in Batna City, Algeria. *Sustainable Cities and Society*, *83*, 103909. https://doi.org/10.1016/j.scs.2022.103909

Markovics, D., & Mayer, M. J. (2022). Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction. *Renewable and Sustainable Energy Reviews*, *161*, 112364. https://doi.org/10.1016/j.rser.2022.112364

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. https://doi.org/10.1016/j.enbuild.2010.12.022

Martín-Chivelet, N., Kapsis, K., Wilson, H. R., Delisle, V., Yang, R., Olivieri, L., Polo, J., Eisenlohr, J., Roy, B., Maturi, L., Otnes, G., Dallapiccola, M., & Upalakshi Wijeratne, W. M. P. (2022). Building-Integrated Photovoltaic (BIPV) products and systems: A review of energy-related behavior. *Energy and Buildings*, *262*. https://doi.org/10.1016/j.enbuild.2022.111998

Martins, T. A. de L., Adolphe, L., Bastos, L. E. G., & Martins, M. A. de L. (2016a). Sensitivity analysis of urban morphology factors regarding solar energy

potential of buildings in a Brazilian tropical context. *Solar Energy*, *137*, 11–24. https://doi.org/10.1016/j.solener.2016.07.053

Martins, T. A. de L., Adolphe, L., Bastos, L. E. G., & Martins, M. A. de L. (2016b). Sensitivity analysis of urban morphology factors regarding solar energy potential of buildings in a Brazilian tropical context. *Solar Energy*, *137*, 11–24. https://doi.org/10.1016/j.solener.2016.07.053

Mohajeri, N., Assouline, D., Guiboud, B., Bill, A., Gudmundsson, A., & Scartezzini, J. L. (2018). A city-scale roof shape classification using machine learning for solar energy applications. *Renewable Energy*, *121*, 81–93. https://doi.org/10.1016/j.renene.2017.12.096

Mohajeri, N., Gudmundsson, A., Kunckler, T., Upadhyay, G., Assouline, D., Kämpf, J. H., & Scartezzini, J. L. (2019). A solar-based sustainable urban design: The effects of city-scale street-canyon geometry on solar access in Geneva, Switzerland. *Applied Energy*, *240*, 173–190. https://doi.org/10.1016/j.apenergy.2019.02.014

Mohajeri, N., Upadhyay, G., Gudmundsson, A., Assouline, D., Kämpf, J., & Scartezzini, J. L. (2016). Effects of urban compactness on solar energy potential. *Renewable Energy*, *93*, 469–482. https://doi.org/10.1016/j.renene.2016.02.053

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2018). *Foundations of Machine Learning* (F. Bach, Ed.; 2nd ed.). The MIT Press.

Morganti, M., Salvati, A., Coch, H., & Cecere, C. (2017). Urban morphology indicators for solar energy analysis. *Energy Procedia*, *134*, 807–814. https://doi.org/10.1016/j.egypro.2017.09.533

Musall, E., Weiss, T., Voss, K., Lenoir, A., Donn, M., Cory, S., & Garde[3], F. (2010). *Net Zero Energy Solar Buildings: An Overview and Analysis on Worldwide Building Projects*.

Nakazato, R., Yokogawa, S., Ichikawa, H., Ushirokawa, T., & Takeda, T. (2021). Compact model for estimating area-level photovoltaic power generation on facade surface using 3D city model and solar radiation simulation. *2021 IEEE PES Innovative Smart Grid Technologies - Asia (ISGT Asia)*, 1–5. https://doi.org/10.1109/ISGTAsia49270.2021.9715273

Natanian, J., Aleksandrowicz, O., & Auer, T. (2019). A parametric approach to optimizing urban form, energy balance and environmental quality: The case of Mediterranean districts. *Applied Energy*, *254*. https://doi.org/10.1016/j.apenergy.2019.113637

Natanian, J., & Wortmann, T. (2021). Simplified evaluation metrics for generative energy-driven urban design: A morphological study of residential blocks in Tel Aviv. *Energy and Buildings*, *240*. https://doi.org/10.1016/j.enbuild.2021.110916

NREL. (2014). *PVWatts Calculator*. https://pvwatts.nrel.gov/

Oke, T. R. (1981). Canyon Geometry and The Nocturnal Urban Heat Island: Comparison of Scale Model and Field Observations. In *JOURNAL OF CLIMATOLOGY* (Vol. 1).

Peng, C., Huang, Y., & Wu, Z. (2011). Building-integrated photovoltaics (BIPV) in architectural design in China. *Energy and Buildings*, *43*(12), 3592–3598. https://doi.org/10.1016/j.enbuild.2011.09.032

Peronato, G. (2019). *Urban planning support based on the photovoltaic potential of buildings: a multi-scenario ranking system*.

Poon, K. H., Kämpf, J. H., Tay, S. E. R., Wong, N. H., & Reindl, T. G. (2020). Parametric study of URBAN morphology on building solar energy potential in Singapore context. *Urban Climate*, *33*. https://doi.org/10.1016/j.uclim.2020.100624

Ramkumar, T., Gamage, H., Xiao, E. W., & Cassat, M. (2019). Analysis of solar PV potential for roofs and faades in high dense residential urban scenario of

Singapore. *Journal of Physics: Conference Series*, *1343*(1). https://doi.org/10.1088/1742-6596/1343/1/012050

Raza, M. Q., Nadarajah, M., & Ekanayake, C. (2016). On recent advances in PV output power forecast. In *Solar Energy* (Vol. 136, pp. 125–144). Elsevier Ltd. https://doi.org/10.1016/j.solener.2016.06.073

REN21. (2020). *RENEWABLES 2020 GLOBAL STATUS REPORT*.

Romero Rodríguez, L., Duminil, E., Sánchez Ramos, J., & Eicker, U. (2017). Assessment of the photovoltaic potential at urban level based on 3D city models: A case study and new methodological approach. *Solar Energy*, *146*, 264–275. https://doi.org/10.1016/j.solener.2017.02.043

Roudsari, M. S., & Pak, M. (2013). *Ladybug: A Parametric Environmental Plugin for Grasshopper to Help Designers Create an Environmentally-Conscious Design*.

Rutten, D. (n.d.). *Grasshopper 3D* (1.0). Robert McNeel and associates.

Salom, J., Widén, J., Candanedo, J., Sartori, I., Voss, K., & Marszal, A. (2011). UNDERSTANDING NET ZERO ENERGY BUILDINGS: EVALUATION OF LOAD MATCHING AND GRID INTERACTION INDICATORS. *Proceedings of Building Simulation 2011*.

Saretta, E., Bonomo, P., & Frontini, F. (2020). A calculation method for the BIPV potential of Swiss façades at LOD2.5 in urban areas: A case from Ticino region. *Solar Energy*, *195*, 150–165. https://doi.org/10.1016/j.solener.2019.11.062

Sarralde, J. J., Quinn, D. J., Wiesmann, D., & Steemers, K. (2015). Solar energy and urban morphology: Scenarios for increasing the renewable energy potential of neighbourhoods in London. *Renewable Energy*, *73*, 10–17. https://doi.org/10.1016/j.renene.2014.06.028

Sartori, I., Napolitano, A., & Voss, K. (2012). Net zero energy buildings: A consistent definition framework. *Energy and Buildings*, *48*, 220–232. https://doi.org/10.1016/j.enbuild.2012.01.032

Schittich, C. (2012). Toward Solar Architecture. In *Solar Architecture : Strategies, Visions, Concepts*. Birkhäuser.

Schön, D. A. (1983). *The Reflective Practitioner*. Basic Books, Inc.

Shafique, M., Luo, X., & Zuo, J. (2020). Photovoltaic-green roofs: A review of benefits, limitations, and trends. *Solar Energy*, *202*, 485–497. https://doi.org/10.1016/j.solener.2020.02.101

Shi, Z., Fonseca, J. A., & Schlueter, A. (2017). A review of simulation-based urban form generation and optimization for energy-driven urban design. In *Building and Environment* (Vol. 121, pp. 119–129). Elsevier Ltd. https://doi.org/10.1016/j.buildenv.2017.05.006

Shi, Z., Fonseca, J. A., & Schlueter, A. (2021). A parametric method using vernacular urban block typologies for investigating interactions between solar energy use and urban design. *Renewable Energy*, *165*, 823–841. https://doi.org/10.1016/j.renene.2020.10.067

Shubbak, M. H. (2019). Advances in solar photovoltaics: Technology review and patent trends. *Renewable and Sustainable Energy Reviews*, *115*. https://doi.org/10.1016/j.rser.2019.109383

Silverman, B. W. (1961). Density Estimations for Statistics and Data Analysis. In *J.Sl Mariti*.

Snellen, D., Borgers, A., & Timmermans, H. (2002). Urban form, road network type, and mode choice for frequently conducted activities: A multilevel analysis using quasi-experimental design data. *Environment and Planning A*, *34*(7), 1207–1220. https://doi.org/10.1068/a349

Swierc, A., Gruenhagen, P., Kubiniec, A., Dise, J., Keelin, P., & Atkins, A. (2020). Capital cost and demand charge reduction through improved battery sizing and scheduling with high resolution irradiance data: Case study of a commercial building. *Proceedings of the ISES Solar World Congress 2019 and IEA SHC International Conference on Solar Heating and Cooling for Buildings and Industry 2019*, 2283–2289. https://doi.org/10.18086/swc.2019.45.12

T.C. Tarım ve Orman Bakanlığı. (2020). *2020 Türkiye Orman Varlığı*.

Tian, J., & Xu, S. (2021). A morphology-based evaluation on block-scale solar potential for residential area in central China. *Solar Energy*, *221*, 332–347. https://doi.org/10.1016/j.solener.2021.02.049

TW SOLAR. (2022). *Specifications of TW325MWP-60-H Monocrystalline solar module*.

Twidell, J., & Weir, T. (2006). *Renewable Energy Resources* (2nd Edition). Taylor & Francis.

United Nations. (2019). *World Population Prospects 2019*.

United Nations. (2020). *THE STRATEGIC PLAN 2020-2023*. www.unhabitat.org

Vermeulen, T., Merino, L., Knopf-Lenoir, C., Villon, P., & Beckers, B. (2018). Periodic urban models for optimization of passive solar irradiation. *Solar Energy*, *162*, 67–77. https://doi.org/10.1016/j.solener.2018.01.014

Voyant, C., Notton, G., Kalogirou, S., Nivet, M. L., Paoli, C., Motte, F., & Fouilloy, A. (2017). Machine learning methods for solar radiation forecasting: A review. In *Renewable Energy* (Vol. 105, pp. 569–582). Elsevier Ltd. https://doi.org/10.1016/j.renene.2016.12.095

Vulkan, A., Kloog, I., Dorman, M., & Erell, E. (2018). Modeling the potential for PV installation in residential buildings in dense urban areas. *Energy and Buildings*, *169*, 97–109. https://doi.org/10.1016/j.enbuild.2018.03.052

Wall, M., Munari Probst, M. C., Roecker, C., Dubois, M. C., Horvat, M., Jørgensen, O. B., & Kappel, K. (2012). Achieving solar energy in architecture - IEA SHC Task 41. *Energy Procedia*, *30*, 1250–1260. https://doi.org/10.1016/j.egypro.2012.11.138

Wang, M., Peng, J., Luo, Y., Shen, Z., & Yang, H. (2021). Comparison of different simplistic prediction models for forecasting PV power output: Assessment with experimental measurements. *Energy*, *224*. https://doi.org/10.1016/j.energy.2021.120162

Wang, W., Liu, K., Zhang, M., Shen, Y., Jing, R., & Xu, X. (2021). From simulation to data-driven approach: A framework of integrating urban morphology to low-energy urban design. *Renewable Energy*, *179*, 2016–2035. https://doi.org/10.1016/j.renene.2021.08.024

Wang, Y., Tian, W., Zhu, L., Ren, J., Liu, Y., Zhang, J., & Yuan, B. (2006). Interactions between building integrated photovoltaics and microclimate in urban environments. *Journal of Solar Energy Engineering, Transactions of the ASME*, *128*(2), 168–172. https://doi.org/10.1115/1.2188533

Wengenmayr, R., & Bührke, T. (Eds.). (2011). *Renewable Energy: Sustainable Energy Concepts for the Future*. Wiley-VCH.

Yang, D., Alessandrini, S., Antonanzas, J., Antonanzas-Torres, F., Badescu, V., Beyer, H. G., Blaga, R., Boland, J., Bright, J. M., Coimbra, C. F. M., David, M., Frimane, Â., Gueymard, C. A., Hong, T., Kay, M. J., Killinger, S., Kleissl, J., Lauret, P., Lorenz, E., … Zhang, J. (2020). Verification of deterministic solar forecasts. *Solar Energy*, *210*, 20–37. https://doi.org/10.1016/j.solener.2020.04.019

Zhang, J., Heng, C. K., Malone-Lee, L. C., Hii, D. J. C., Janssen, P., Leung, K. S., & Tan, B. K. (2012). Evaluating environmental implications of density: A comparative case study on the relationship between density, urban block

typology and sky exposure. *Automation in Construction*, *22*, 90–101. https://doi.org/10.1016/j.autcon.2011.06.011

Zhang, J., Xu, L., Shabunko, V., Tay, S. E. R., Sun, H., Lau, S. S. Y., & Reindl, T. (2019). Impact of urban block typology on building solar potential and energy use efficiency in tropical high-density city. *Applied Energy*, *240*, 513–533. https://doi.org/10.1016/j.apenergy.2019.02.033

Zhu, D., Song, D., Shi, J., Fang, J., & Zhou, Y. (2020). The effect of morphology on solar potential of high-density residential area: A case study of Shanghai. *Energies*, *13*(9). https://doi.org/10.3390/en13092215

Zhu, R., Wong, M. S., You, L., Santi, P., Nichol, J., Ho, H. C., Lu, L., & Ratti, C. (2020). The effect of urban morphology on the solar capacity of three-dimensional cities. *Renewable Energy*, *153*, 1111–1126. https://doi.org/10.1016/j.renene.2020.02.050

APPENDICES

A. Utilized ML Models and Hyperparameter Grids

Model	Info	Description
RF	Family	Ensemble Models
	Name (scikit-learn)	Random Forest Regressor
	Tuned Hyper-parameters	n_estimators: [200, 288, 377, 466, 555, 644, 733, 822, 911, 1000], max_depth: [10, 21, 32, 43, 54, 65, 76, 87, 98, 110], min_samples_leaf: [1, 2, 4, 5, 6, 7], min_samples_split: [2, 5, 10, 20],
XGB	Family	Ensemble Models
	Name (scikit-learn)	XGB Regressor
	Tuned Hyper-parameters	learning_rate: np.arange(0.0005,0.2,0.0005), max_depth: [3, 5, 7, 9, 11], min_child_weight: [1, 3, 5], n_estimators: [25, 75, 125, 175, 225, 275, 325, 375, 425, 475], subsample: [0.6, 0.7, 0.8, 0.9], colsample_bytree: [0.6, 0.7, 0.8, 0.9] gamma: [0, 0.1, 0.2, 0.3, 0.4]
MLP	Family	Neural Networks
	Name (scikit-learn)	MLP Regressor
	Tuned Hyper-parameters	hidden_layer_sizes: [(5, 10), (5, 50), (10, 5), (10, 50), (50, 5), (50, 10), 5, 10, 50], learning_rate: ['constant', 'adaptive', 'invscaling'], solver: ['adam', 'sgd'], batch_size: [16, 32, 64], alpha: [0, 0.0001, 0.001, 0.01, 0.1]

Number of neurons vary according to predictor sets: Complex-Façade: (4,8,32), Basic-Façade: (22, 44, 88), Complex-Roof: (6,18,36), Basic-Roof: (11,21,84).

B. Statistical Description Building PV Generation Data

İ. Total PV Generation (kWh)

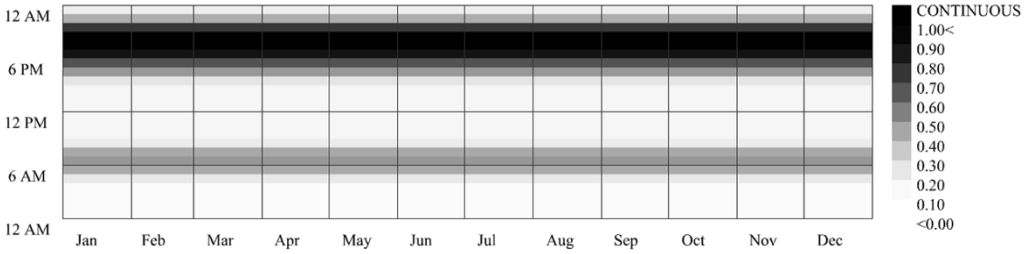
	Gridiron		Organic		Radial		No Pattern	
	Roof	Facade	Roof	Facade	Roof	Façade	Roof	Facade
Count	640	640	525	525	475	297	297	297
Mean	45816.43	28702.44	44276.94	31375.29	67437.01	90639.18	68676.37	90639.18
Std	22626.86	12909.32	18070.37	8459.25	70646.22	61910.42	68302.89	61910.42
Min	10036.42	4615.18	16600.36	10154.94	6351.74	12919.39	21936.63	12919.39
25%	32940.31	20271.46	35785.46	25624.49	35523.02	45640.74	40135.28	45640.74
50%	43334.27	28049.75	41387.78	30684.30	45374.06	73111.43	52826.47	73111.43
75%	53461.43	36109.74	47808.69	35832.58	63323.18	117014.67	76883.83	117014.67
Max	246076.76	92348.92	319270.14	64819.77	619531.95	324612.22	899899.79	324612.22

II. Unit Area PV Generation (kWh/m^2)

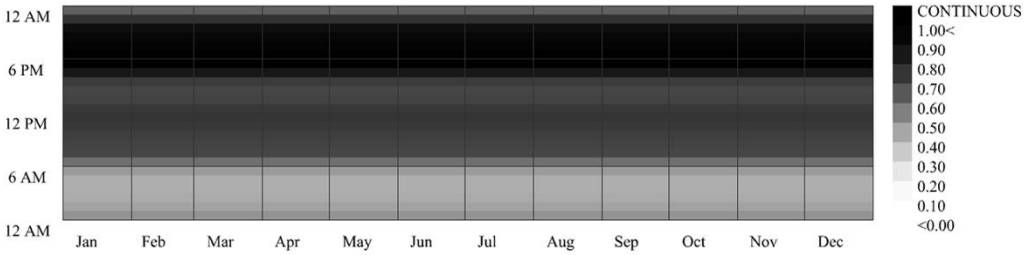
	Gridiron		Organic		Radial		No Pattern	
	Roof	Facade	Roof	Facade	Roof	Façade	Roof	Facade
Count	640	640	525	525	475	475	297	297
Mean	128.46	33.63	133.91	37.69	127.44	29.91	133.20	46.52
Std	11.22	9.03	3.76	6.64	15.35	10.22	7.41	9.39
Min	74.87	8.74	97.47	20.17	42.38	4.11	99.03	21.15
25%	125.78	28.19	132.65	32.82	123.45	22.12	132.05	39.98
50%	132.90	34.50	134.30	37.40	133.93	30.10	135.68	47.79
75%	135.78	39.37	135.92	41.70	136.75	36.85	137.51	53.20
Max	141.84	59.85	140.00	65.15	141.84	62.56	141.84	68.44

C. Residential and Non-residential Equipment and Lighting Schedules

Residential

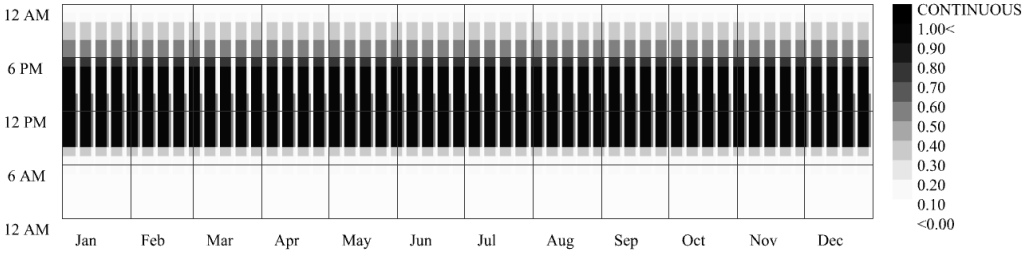


MidriseApartment Apartment Light (CONTINUOUS) - Hourly
schedule:year
1 JAN 1:00 - 31 DEC 24:00

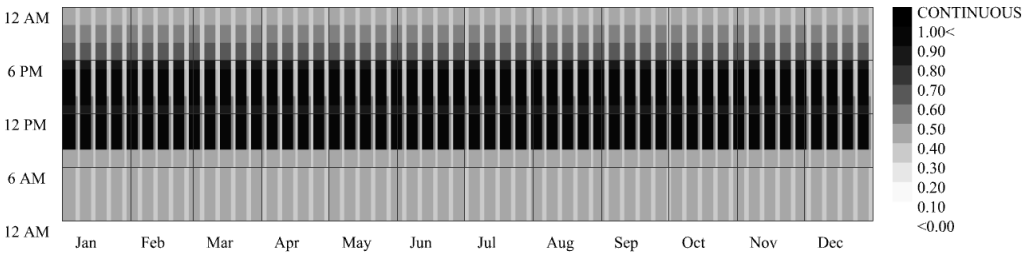


MidriseApartment Apartment Equip (CONTINUOUS) - Hourly
schedule:year
1 JAN 1:00 - 31 DEC 24:00

Non-residential



Office Bldg Light (CONTINUOUS) - Hourly
schedule:year
1 JAN 1:00 - 31 DEC 24:00



Medium Office Bldg Equip (CONTINUOUS) - Hourly
schedule:year
1 JAN 1:00 - 31 DEC 24:00

D. Fine-tuned ML model hyperparameters

		RF	XGB	MLP
Facade	Complex	n_estimators: 644, min_samples_split: 5, min_samples_leaf: 1, max_features: 'auto', max_depth: 54, bootstrap: True	subsample: 0.6, n_estimators: 325, min_child_weight: 3, max_depth: 7, learning_rate: 0.0595, gamma: 0.0, colsample_bytree: 0.9	solver: 'sgd', shuffle: True, max_iter: 400, activation: 'relu' learning_rate: 'adaptive', hidden_layer_sizes: (8, 32), early_stopping: True, batch_size: 16, alpha: 0, momentum: 0.910
	Basic	n_estimators: 644, min_samples_split: 2, min_samples_leaf: 1, max_features: 'auto', max_depth: 87, bootstrap: True	subsample: 0.7, n_estimators: 375, min_child_weight: 5, max_depth: 7, learning_rate: 0.038, gamma: 0.0, colsample_bytree: 0.8	solver: 'sgd', shuffle: True, max_iter: 400, activation: 'relu' learning_rate: 'adaptive', hidden_layer_sizes: (44, 88), early_stopping: True, batch_size: 16, 'alpha': 0.001, momentum: 0.905
Roof	Complex	n_estimators: 644, min_samples_split: 5, min_samples_leaf: 1, max_features: 'auto', max_depth: 54, bootstrap: True	subsample: 0.6, n_estimators: 375, min_child_weight: 1, max_depth: 9, learning_rate: 0.037, gamma: 0.0, colsample_bytree: 0.7	solver: 'adam', shuffle: True, max_iter: 400, activation: 'relu' learning_rate: 'adaptive', hidden_layer_sizes: 18, 6, early_stopping: True, batch_size: 4, alpha: 0.001,
	Basic	n_estimators: 1000, min_samples_split: 2, min_samples_leaf: 3, max_features: 'auto', max_depth: 43, bootstrap: True	subsample: 0.9, n_estimators: 475, min_child_weight: 3, max_depth: 5, learning_rate: 0.095, gamma: 0.2, colsample_bytree: 0.8	solver: 'adam', shuffle: True, max_iter: 400, activation: 'relu' learning_rate: 'adaptive', hidden_layer_sizes: (84, 11, 21), early_stopping: True, batch_size: 4, alpha: 0.0001,