UNIVERSITÀ DEGLI STUDI DI GENOVA

SCUOLA POLITECNICA

DIME

Dipartimento di Ingegneria Meccanica, Energetica,

Gestionale e dei Trasporti

TESI DI LAUREA MAGISTRALE IN

INGEGNERIA MECCANICA - ENERGIA E AERONAUTICA

Assessing Household PV Potential in Ireland: A Data-Driven Techno-Economic Evaluation

Relatore: Chiar.mo Prof. Ing. Mattia De Rosa

Correlatore: Chiar.mo Prof. Ing. Fabiano Pallonetto

Allievi:

Davide Melone

Valentina Giussani

Luglio 2026

SOMMARIO

La crescente diffusione degli impianti fotovoltaici residenziali e la crescente complessità dei moderni mercati elettrici richiedono metodologie avanzate per l'analisi, la caratterizzazione e la gestione delle risorse energetiche distribuite. In questo contesto, la presente tesi sviluppa un framework integrato per la modellazione della produzione fotovoltaica, la calibrazione degli impianti e la valutazione economica basato su dati reali. In una prima fase è stato sviluppato in ambiente Python un modello di produzione fotovoltaica in grado di stimare l'energia elettrica generata a partire dai dati di radiazione solare. Il modello è stato inizialmente validato mediante dati di riferimento ottenuti da PVGIS. Successivamente, la metodologia è stata applicata a un dataset residenziale relativo alla località di Dingle, in Irlanda. Poiché non erano disponibili misure di radiazione con la stessa risoluzione temporale dei dati di produzione, è stata sviluppata una procedura inversa (backward) per la ricostruzione dei profili di radiazione a partire dai dati di produzione fotovoltaica misurati, consentendo la validazione del modello in condizioni operative reali. Partendo da tale framework, è stata quindi proposta una metodologia di calibrazione automatica degli impianti fotovoltaici basata esclusivamente sui flussi energetici misurati. Attraverso la ricostruzione dei profili di radiazione per ciascuna abitazione, è stato possibile stimare i principali parametri caratteristici degli impianti, tra cui la potenza nominale installata e i coefficienti di temperatura, senza disporre di informazioni tecniche dirette sugli stessi. L'ultima parte del lavoro è stata dedicata all'analisi economica delle utenze residenziali. Sono state valutate e confrontate diverse offerte realmente disponibili sul mercato elettrico irlandese, insieme a una tariffa dinamica denominata RENEW, sviluppata a partire dai prezzi del mercato elettrico Day-Ahead (DAM). L'analisi ha consentito di valutare l'influenza della generazione fotovoltaica e dei profili di consumo sulla convenienza economica delle diverse strutture tariffarie. Il framework proposto dimostra come i dati energetici possano essere utilizzati per stimare la produzione fotovoltaica, caratterizzare sistemi di generazione distribuita e supportare processi decisionali in ambito energetico, contribuendo alla digitalizzazione e all'ottimizzazione dei moderni sistemi elettrici.

ABSTRACT

The increasing diffusion of residential photovoltaic systems and the growing complexity of modern electricity markets require advanced methodologies for the analysis, characterization, and management of distributed energy resources. In this context, this thesis develops an integrated framework for photovoltaic production modeling, system calibration, and economic assessment based on real-world data. A photovoltaic production model was first developed in Python to estimate electrical generation from solar irradiance data. The model was initially validated using reference data obtained from PVGIS. Subsequently, the methodology was applied to residential datasets from Dingle, Ireland. Since high-resolution irradiance measurements were not directly available, a backward procedure was developed to reconstruct irradiance profiles from measured photovoltaic production data, enabling the validation of the model under real operating conditions. Building upon this framework, an automated calibration methodology was proposed to identify the main characteristics of residential photovoltaic systems using only measured energy flows. By reconstructing irradiance profiles for each dwelling, key system parameters such as installed peak power and temperature coefficients were estimated without requiring direct information about the photovoltaic installations. The final part of the work focused on the economic analysis of residential users. Real electricity tariffs available on the Irish market were evaluated and compared with a dynamic tariff, named RENEW, developed using Day-Ahead Market (DAM) electricity prices. This analysis enabled the assessment of the relationship between photovoltaic generation, consumption patterns, and tariff profitability. The proposed framework demonstrates how energy data can be used to estimate photovoltaic production, characterize distributed generation systems, and support economic decision-making, contributing to the digitalization and optimization of modern energy systems.

CONTEST

1. INTRODUCTION 9

1.1. Historical introduction 9

1.2. PV generations 10

1.2 Basic principles of photovoltaic cells 11

1.3 Structure of monocrystalline silicon cells 16

2. ENERGY CONTEXT 18

2.1 Irish energy context 18

2.2 Ireland 2024: energy flows analysis 19

2.3 Comparison with Italy 22

2.4 Households and energy communities 23

3. METHODOLOGY 27

3.1 General purpose 27

3.2 Forward PV performance model using PVGIS data 28

3.2.1 Model inputs and system configuration 28

3.2.3 Statistical validation framework and residual analysis 33

3.3 Inverse-forward model using StoreNet data 35

3.3.1 Script 1 36

3.3.1.1 Model inputs and system configuration 36

3.3.1.2 Core mathematical modeling of efficiency and production 38

3.3.2 Script 2 40

3.3.2.1 Model inputs and system configuration 40

3.3.2.2 Core mathematical modeling of efficiency and production 42

3.3.2.3 Statistical validation framework and residual analysis 44

3.4 Calibration and validation of PV system parameters 47

3.4.1 First Step: Backward estimation of the plane of array irradiance (POA) 48

3.4.1.1 Model inputs and system configuration 48

3.4.1.2 Core mathematical modeling 50

3.4.2 Second Step: Calibration of PV system parameters 54

3.4.2.1 Model inputs and system configuration 54

3.4.2.2 Core mathematical modeling 55

3.5 Comparative methodology for baseline tariffs and renew dynamic optimization 62
3.5.1 Model inputs and system configuration 62
3.5.1.1 Households load and generation profiles 62
3.5.1.2 Standard electricity tariffs (Baseline) 62
3.5.1.3 Dynamic tariff and day-ahead market prices 63
3.5.1.4 Fixed charges and network components 63
3.5.1.5 Battery storage system configuration 63
3.5.2 Core mathematical modeling 63
3.5.2.1 Baseline tariff cost calculation 63
3.5.2.2 Renew tariff optimization with battery storage 64
3.5.2.3 Forecast models 65
3.5.2.4 Temporal alignment and aggregation 65
3.5.3 Workflow and methodological comparison 66
3.5.3.1 Battery decision hierarchy 66
3.5.3.2 Intra-household comparison 66
3.5.3.3 Intra-category comparison 66
3.5.3.4 Inter-category comparison 66
4. RESULTS 68
4.1 PVGIS model results 68
4.1.1 Validation on monthly and annual scales 68
4.1.2 Error metrics on annual, monthly and hourly scales 70
4.1.3 Statistical tests for model vs. PVGIS comparison 73
4.1.4 Residual analysis 74
4.2 StoreNet results 78
4.2.1 Validation on annual and monthly scales 78
4.2.1.1 Annual validation 78
4.2.1.2 Monthly validation 79
4.2.2 Error metrics on annual, monthly and hourly scales 80
4.2.2.1 Monthly error metrics 82
4.2.2.2 Hourly error metrics 83
4.2.3 Residual analysis 85
4.3 Calibration results 87
4.3.1 Validation on annual scale 87

4.3.1.1 Annual energy comparison for H1 .... **87**
4.3.2 Errors metrics on annual, monthly and hourly scales **89**
4.3.3 Residual analysis **90**
4.4 Comparative methodology for baseline tariffs and renew dynamic optimization results 91
4.4.1 Baseline tariff analysis **91**
4.4.2 Renew tariff optimization results by household **92**
4.4.3 Multi-level comparative analysis **96**
4.4.4 Synthesis of key findings **103**
5. CONCLUSIONS 105
5.1 Future developments 106
6. BIBLIOGRAPHY 108
7. ACKNOWLEDGEMENTS 110

LIST OF TABLES

TABLE 3.1 PV SYSTEM PARAMETERS 30
TABLE 3.2 PRIMARY INPUT DATA 36
TABLE 3.3 FIXED SYSTEM PARAMETERS FOR INVERSE RECONSTRUCTION.. 37
TABLE 3.4 PRIMARY INPUT DATA 41
TABLE 3.5 FIXED SYSTEM PARAMETERS FOR FORWARD SIMULATION 41
TABLE 3.6 PRIMARY INPUT DATA FOR POA RECONSTRUCTION 48
TABLE 3.7 FIXED SYSTEM PARAMETERS FOR INVERSE RECONSTRUCTION.. 49
TABLE 3.8 PRIMARY INPUT DATA FOR CALIBRATION AND VALIDATION 54
TABLE 3.9 FIXED PARAMETERS FOR CALIBRATION 54
TABLE 4.1 MONTHLY PRODUCTION REPORT AND OPERATING INDICATORS 68
TABLE 4.2 MONTHLY ERROR METRICS 71
TABLE 4.3 ERROR METRICS ON HOURLY SCALE 72
TABLE 4.4 RESULTS OF STATISTICAL TESTS ON HOURLY SCALE 73
TABLE 4.5 PERCENTAGE ERROR BY HOUR OF THE DAY 74
TABLE 4.6 PERCENTAGE ERROR BY MONTH 75
TABLE 4.7 COMPARATIVE DAY/NIGHT METRICS 76
TABLE 4.8 HOURLY METRICS BY SEASON 76
TABLE 4.9 DAILY METRICS BY SEASON 76
TABLE 4.10 VERIFICATION OF HOURLY VALIDATION CRITERIA 77
TABLE 4.11 AGGREGATE COMPARISON OF ANNUAL PRODUCTION 78
TABLE 4.12 MONTHLY PRODUCTION REPORT 79
TABLE 4.13 VALIDATION METRICS FOR EACH HOUSEHOLD.. 82
TABLE 4.14 MONTHLY VALIDATION METRICS 83
TABLE 4.15 ANNUAL ENERGY PRODUCTION: REAL VS MODEL 88
TABLE 4.16 ESTIMATED PARAMETERS AND ACCURACY METRICS FOR THE 10 RESIDENTIAL SYSTEMS. 89
TABLE 4.17 BASELINE TARIFF RESULTS SUMMARY 91
TABLE 4.18 RENEW TARIFF RESULTS FOR H1 92
TABLE 4.19 RENEW TARIFF RESULTS FOR H3 93
TABLE 4.20 RENEW TARIFF RESULTS FOR H10 93
TABLE 4.21 RENEW TARIFF RESULTS FOR H8 94
TABLE 4.22 RENEW TARIFF RESULTS FOR H18 94
TABLE 4.23 RENEW TARIFF RESULTS FOR H19 95
TABLE 4.24 BASELINE VS RENEW COMPARISON SUMMARY 96
TABLE 4.25 HOUSEHOLD COMPARISON - BEST RENEW VS BASELINE 97
TABLE 4.26 NON-PV HOUSEHOLD COMPARISON - BEST RENEW VS BASELINE 99
TABLE 4.27 BATTERY IMPACT ON NON-PV HOUSEHOLDS (15 KWH CONFIGURATION).. 100
TABLE 4.28 INTER-CATEGORY COMPARISON SUMMARY 101

LIST OF FIGURES

Figure 1.1 PV generations [2] 11
Figure 1.2 P - N junction [3] 12
Figure 1.3 Solar spectral irradiance [3] 13
Figure 1.4 Illustration of a solar cell [4] 13
Figure 1.5 Cell behavior [3] 14
Figure 1.6 Characteristic curve of a PV cell [5] 14
Figure 1.7 I - V curve with incident irradiance [6] 15
Figure 1.8 Series connected PV cells [7] 15
Figure 1.9 Structure of Monocrystalline silicon cell [8] 16
Figure 2.2 Annual energy flows from raw materials to final consumption by sector [12] 19
Figure 2.3 Electricity generation broken-out by fuel type and energy source [13] 20
Figure 2.4 Renewable energy share in electricity [14] 21
Figure 2.5 Comparison of renewable electricity generation breakdown by technology between Italy and Ireland. 22
Figure 2.6 Renewable and citizen energy communities [18] 25
Figure 4.1 Daily production comparison by month 69
Figure 4.2 Hourly production parity plot 70
Figure 4.3 Monthly error metrics: MAPE and MBE 71
Figure 4.4 Distribution of hourly errors 72
Figure 4.5 Distribution of relative error by hour of day 75
Figure 4.6 Monthly production 80
Figure 4.7 Hourly comparison: model vs avarage of 6 houses 81
Figure 4.8 Hourly comparison: model vs individual houses 82
Figure 4.9 Hourly MBE: model vs reality 83
Figure 4.10 Hourly MAE: model vs reality 84
Figure 4.11 Average annual hourly profile 85
Figure 4.12 Average hourly profile: model vs individual houses 86
Figure 4.13 Comparison model vs reality H1 87
Figure 4.14 Residual analysis 90
Figure 4.15 Avarage day-ahead market price profile by hour 95
Figure 4.16 Distribution of day-ahead market prices 95
Figure 4.17 Annual electricity cost by household and tariff configuration 96
Figure 4.18 Diminishing returns of battery capacity-PV households 98
Figure 4.19 Feed-in credit evolution by configuration 98
Figure 4.20 Absolute vs percentage savings by battery configuration PV households 99
Figure 4.21 Cost structure comparison: baseline vs renew (non-PV households) 100
Figure 4.22 Avarage import price by household and configuration 101
Figure 4.23 Annual saving/loss vs baseline by household and battery capacity 102

1. INTRODUCTION

1.1. Historical introduction

The photovoltaic effect was first observed in 1839 by the French physicist Alexandre Edmond Becquerel while conducting electrochemical experiments. He noticed that silver and platinum electrodes exposed to sunlight generated an electric current. Later, in 1873, Willoughby Smith discovered that the electrical resistance of selenium varied significantly depending on the intensity of incident light, revealing its photoconductive properties. In 1877, William Grylls Adams and his student Richard Evans Day developed the first selenium-based solar cell, achieving an efficiency of about 0.5%. One year later, Charles Fritts improved this design by placing a selenium wafer between two thin metal layers, doubling the efficiency of the device. A major theoretical breakthrough occurred in 1905 when Albert Einstein proposed the theory of the photoelectric effect, building upon earlier concepts introduced by Max Planck. The theory explained that light consists of energy packets called photons, whose energy is proportional to their frequency. This discovery was fundamental for understanding the interaction between light and matter, and it later earned Einstein the Nobel Prize in Physics in 1921. The development of more efficient photovoltaic devices began with the introduction of silicon. In 1939, Russell Ohl discovered the existence of p-type and n-type regions in silicon and identified the photoelectric effect occurring in p–n junctions. Ohl's discovery laid the foundation for modern semiconductor technologies, including transistors and photovoltaic cells. Shortly after, Ohl developed one of the first silicon solar cells at Bell Laboratories. Early photovoltaic devices had very low efficiencies and were therefore mainly used as light sensors rather than power generators. For example, Werner von Siemens commercialized selenium-based light sensors used in photographic equipment. A major technological breakthrough occurred in 1954 when researchers at Bell Laboratories, including Calvin Fuller and Gerald Pearson, developed a silicon solar cell based on a p–n junction created through doping techniques. The resulting solar cell reached an efficiency of about 6%, representing a significant improvement compared to previous technologies. The first practical application of a solar cell occurred in 1955, when it was used as the power source for a telecommunications system in Americus, Georgia. Shortly afterward, photovoltaic technology found an important application in space exploration. In 1958, NASA launched the satellite Vanguard 1, which was equipped with silicon solar cells to power its transmitters. This successful mission demonstrated the reliability of solar power in space and encouraged further development of photovoltaic systems. Despite these advances, the high manufacturing costs initially limited the use of photovoltaic technology for large-scale public applications. Interest increased significantly during the 1970s following the global oil crisis. During this period, Joseph Lindmayer developed a process that improved silicon solar cell efficiency, contributing to the growth of the photovoltaic industry. In 1973, Lindmayer and Peter Varadi founded Solarex with the aim of promoting solar energy for broader commercial applications. From the 1970s onward, research focused on reducing production costs and improving efficiency. While early solar cells were mainly based on monocrystalline silicon, later developments introduced polycrystalline and amorphous silicon technologies, which allowed cheaper manufacturing processes. Over time, photovoltaic technologies have continued to evolve, leading to the classification of solar cells into different generations based on the materials and techniques used in their production. Advances in fields such as electronics, photonics, and quantum mechanics have also contributed to the development of new solutions, including flexible and innovative solar cells. Today, photovoltaic technology is not limited to solar cells alone but includes complete systems composed of inverters, batteries, and electrical connections. Improvements in these components have significantly increased the overall efficiency and reliability of photovoltaic energy systems [1].

1.2. PV generations

Photovoltaic technologies can be systematically classified into four generations according to their constituent materials, manufacturing processes, and overall cell performance. The first generation comprises crystalline silicon (c-Si) solar cells, including both monocrystalline silicon (sc-Si) and multicrystalline silicon (mc-Si) technologies. This generation currently dominates the global photovoltaic market, accounting for more than 90% of installed capacity, primarily due to the abundance of raw materials, mature manufacturing processes, and continuously decreasing production costs. Until 2018, aluminium back surface field (Al-BSF) technology represented the dominant c-Si architecture; however, it has progressively been replaced by passivated emitter and rear cell (PERC) technology, which provides efficiency gains of up to one percentage point through the introduction of an additional rear-side passivation layer that reduces recombination losses and improves light absorption. By 2023, PERC technology accounted for approximately 70% of the crystalline silicon market, whereas Al-BSF is expected to be almost completely phased out by 2026. Other advanced first-generation technologies include silicon heterojunction (SHJ) cells, which combine a crystalline silicon wafer with thin layers of amorphous silicon. SHJ devices have achieved laboratory efficiencies of approximately 26.7% and are characterized by low temperature coefficients and bifacial energy-harvesting capabilities. Furthermore, tunnel oxide passivated contact (TOPCon) and interdigitated back contact (IBC) technologies have attracted considerable industrial and academic interest, with reported efficiencies of 25.8% and 26.1%, respectively. Due to their improved performance and declining manufacturing costs, these technologies are expected to become increasingly widespread in the coming years. The second generation consists of thin-film photovoltaic technologies, which require absorber layers typically around 10 μm thick, significantly thinner than the 200–300 μm silicon wafers used in first-generation devices. Major thin-film technologies include amorphous silicon (a-Si), with efficiencies of approximately 14%, cadmium telluride (CdTe), reaching 22.1%, and copper indium gallium selenide (CIGS), with efficiencies up to 23.4%. III–V compound semiconductors, such as gallium arsenide (GaAs), indium phosphide (InP), aluminium gallium arsenide (AlGaAs), and indium gallium phosphide (InGaP), are also considered highly promising photovoltaic materials, with GaAs cells achieving efficiencies as high as 29.1%. Despite their advantages, second-generation technologies face several limitations, including relatively high manufacturing costs, challenges associated with bandgap engineering and multi-junction integration, long-term stability issues, and concerns related to material toxicity (e.g., cadmium and arsenic) or scarcity (e.g., indium and gallium). The third generation includes emerging photovoltaic technologies that are still largely in the research and development stage. These technologies comprise dye-sensitized solar cells (DSSCs), which have achieved efficiencies of approximately 13% but remain constrained by the instability of liquid electrolytes; organic solar cells (OSCs), which offer advantages such as low manufacturing costs, mechanical flexibility, and solution-processability but suffer from limited carrier mobility; and perovskite solar cells, which have experienced a remarkable increase in efficiency, rising from approximately 3% in 2009 to 25.5% in 2023. Despite their low production costs and rapid technological progress, third-generation solar cells have not yet achieved large-scale commercial deployment. Their main limitations include limited operational lifetimes, instability under real-world environmental conditions, and, in the case of perovskite devices, the presence of toxic lead compounds. Consequently, significant research efforts are currently focused on developing lead-free alternatives based on materials such as tin (Sn) and bismuth (Bi). The fourth generation refers to tandem or multijunction photovoltaic devices designed to overcome the fundamental efficiency limitations of conventional single-junction solar cells. Since crystalline silicon technology is approaching its theoretical Auger efficiency limit of approximately 29.43%, increasing attention has been devoted to tandem architectures. These devices integrate two or more photovoltaic absorbers with different bandgap energies, enabling more efficient utilization of the solar spectrum and reducing thermalization losses. Among the most promising configurations are III–V/Si and perovskite/silicon (PVK/Si) tandem cells. Laboratory-scale devices have demonstrated remarkable conversion efficiencies, reaching 32.5% for PVK/Si tandems in 2023 and up to 35.9% for III–V/Si structures fabricated through direct epitaxial growth. Furthermore, six-junction III–V tandem cells have achieved efficiencies of 39.2% under one-sun illumination and 47.1% under concentrated sunlight conditions equivalent to 143 suns. Despite their exceptional performance potential, tandem solar cells still face several technological and economic challenges. These include high fabrication costs, particularly for III–V devices, which were estimated at approximately 100 USD/W in 2020, as well as degradation mechanisms such as light- and elevated-temperature-induced degradation (LETID), the need for accurate current matching between subcells during operation, and, for perovskite-based tandem structures, concerns related to long-term stability and lead toxicity. Nevertheless, tandem photovoltaic technologies are widely regarded as the most promising candidates for next-generation solar energy conversion systems, owing to their exceptional efficiency potential and their capability to surpass the performance limits of conventional single-junction photovoltaic devices. [2]

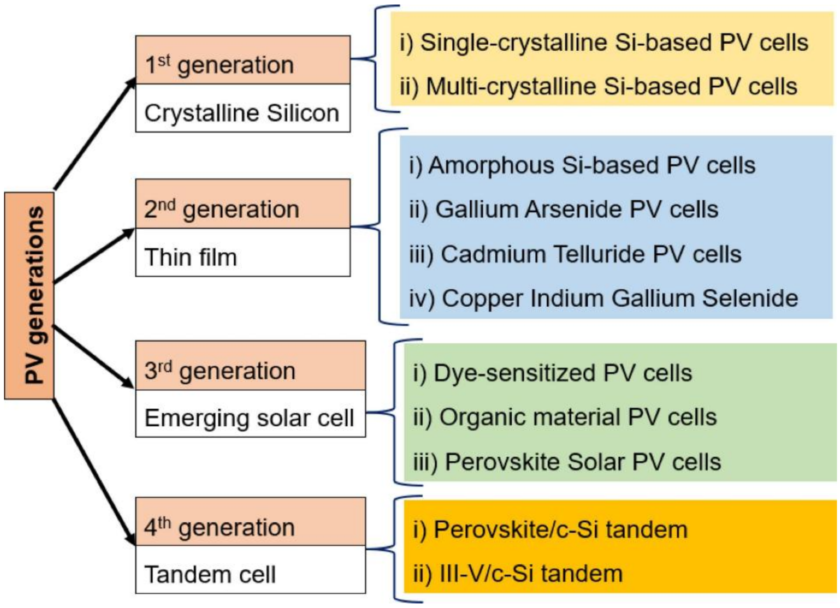


Figure 1.1 PV generations [2]

1.2 Basic principles of photovoltaic cells

Silicon is the material most commonly used in photovoltaic cells because of its semiconductor properties. Each silicon atom has four valence electrons that form bonds with neighboring atoms, creating a crystalline lattice. In this condition, the electrons are located in the so-called valence band and are not free to move. When a photon from solar radiation strikes the material and has sufficient energy, it can transfer this energy to an electron in the silicon. If the photon energy is at least equal to the band gap energy of silicon, the electron can move from the valence band to the conduction band.

During this process an electron–hole pair is generated: the electron moves to the conduction band, while a hole remains in the valence band, representing an empty state that can be occupied by another electron. In the absence of an internal electric field, the electron and the hole would quickly recombine, cancelling the effect. To prevent this recombination and exploit the generated charges, silicon is doped with small quantities of other elements. Typically, phosphorus is used to create n-type silicon, which contains a large number of free electrons, while boron is used to produce p-type silicon, characterized by a high concentration of holes. When a region of p-type silicon is brought into contact with a region of n-type silicon, a P-N junction is formed (Figure 1.2). Near this junction, an internal electric field is created spontaneously. This field separates the charges generated by light: electrons are driven toward the n-type region, while holes move toward the p-type region. The separation of charge carries prevents immediate recombination and allows the generation of an electric current when the circuit is closed. For this process to occur, photons must have energy at least equal to the band gap of silicon. Photons with lower energy cannot excite electrons and therefore do not contribute to current generation. Photons with higher energy transfer only the amount needed to promote the electron to the conduction band, while the excess energy is dissipated as heat. For silicon, the band gap corresponds to a wavelength of approximately 1100 nm (Figure 1.3). Finally, the amount of current generated by a photovoltaic cell also depends on the incident solar irradiance. As solar radiation increases, the number of photons striking the device increases, leading to the generation of more electron–hole pairs and therefore a higher current. This dependence can also be observed in the current–voltage characteristic of a photovoltaic cell, which changes depending on the level of irradiance [3].

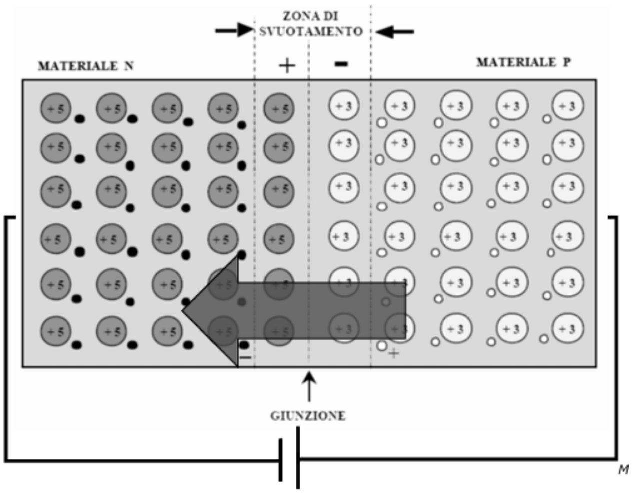


Figure 1.2 P - N junction [3]

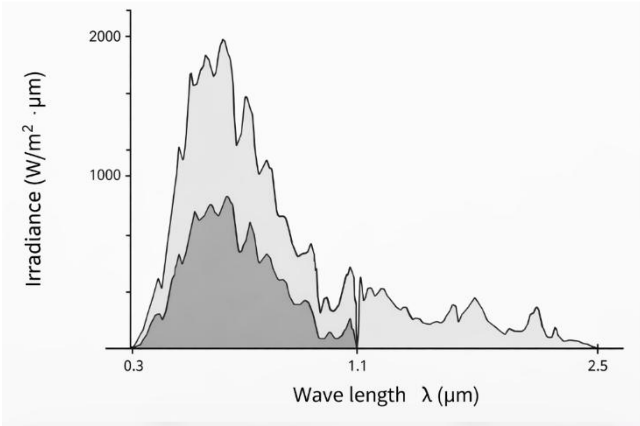


Figure 1.3 Solar spectral irradiance [3]

Figure 1.4 represents an illustration of a solar cell

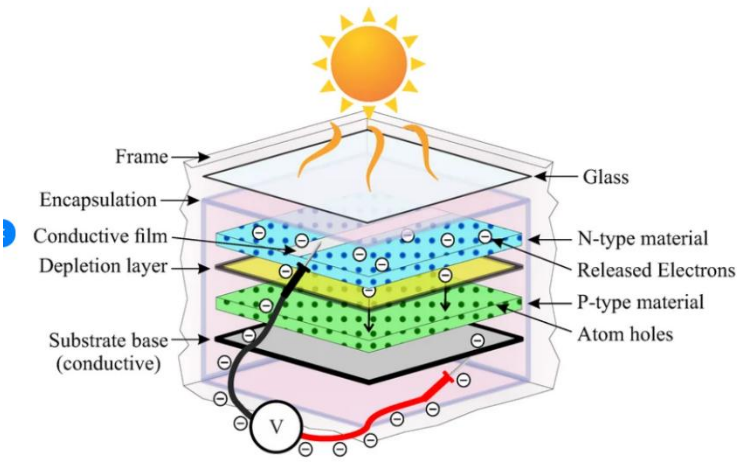


Figure 1.4 Illustration of a solar cell [4]

The photovoltaic cell can be electrically described as a semiconductor diode. Its behavior, represented in Figure 1.5, is therefore similar to that of a diode, whose current–voltage (I–V) characteristic curve has an exponential trend. When the cell is not illuminated (dark condition), its behavior coincides with that of a normal diode. In this case, the current–voltage characteristic is located in the first quadrant of the I–V plane. When the cell is illuminated, the absorption of solar radiation generates a photocurrent due to the creation of electron–hole pairs in the semiconductor material. This current modifies the diode characteristic, shifting the curve into the fourth quadrant of the current–voltage plane. [3]

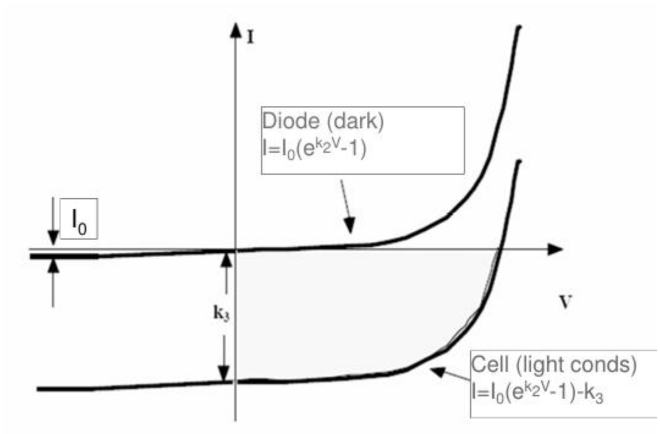


Figure 1.5 Cell behavior [3]

To analyze the operation of the photovoltaic cell as an energy generator, the curve in the fourth quadrant is often mirrored into the first quadrant. As a result, the electrical characteristic of the photovoltaic cell is obtained (Figure 1.6) and describes the relationship between the generated current and the voltage across the cell.

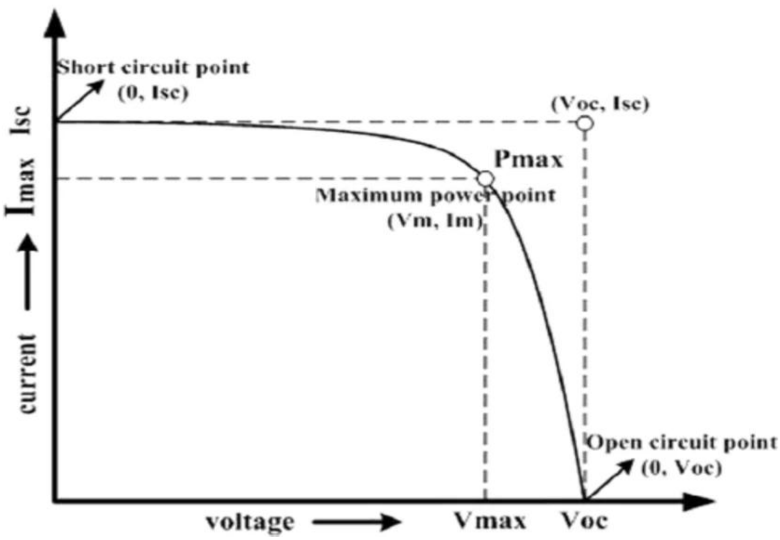


Figure 1.6 Characteristic curve of a PV cell [5]

The current produced by the cell mainly depends on solar irradiance as it’s represented in Figure 1.7. As the incident radiation increases, the generated photocurrent and therefore the available current also increase.

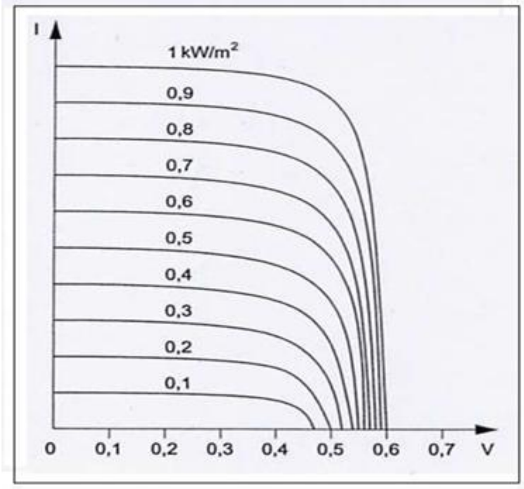


Figure 1.7 I - V curve with incident irradiance [6]

Since the voltage generated by a single cell is relatively low (about 0.5–0.6 V), photovoltaic cells are generally connected in series in practical applications (Figure 1.8). When multiple photovoltaic cells are connected in series, the positive terminal of one cell is connected to the negative terminal of the next, creating a single path for the flow of electrons. In this configuration, the current flowing through the circuit is the same in all the cells, since the electrons must pass sequentially through each of them. For this reason, the total current of the system remains equal to the current produced by a single cell. The voltage, on the other hand, increases because each cell introduces its own potential difference. When electrons pass through a cell, they experience an increase in energy associated with the voltage generated by that cell. By connecting several cells in series, these potential increases add together, resulting in a higher overall voltage across the system. Consequently, while the current remains the same throughout the circuit, the total potential difference is equal to the sum of the voltages generated by the individual cells. [3]

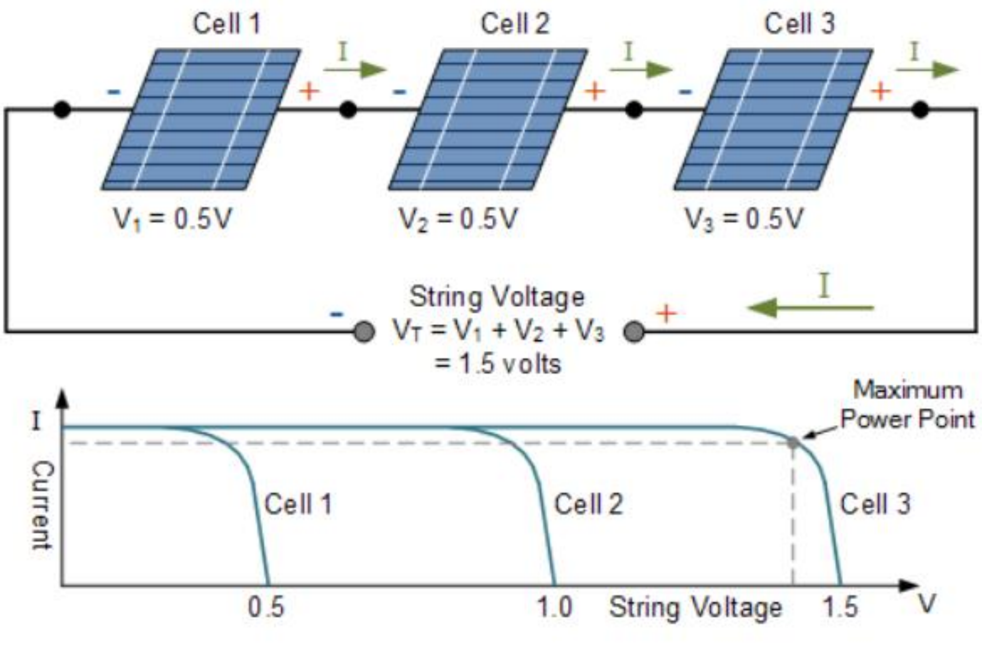


Figure 1.8 Series connected PV cells [7]

1.3 Structure of monocrystalline silicon cells

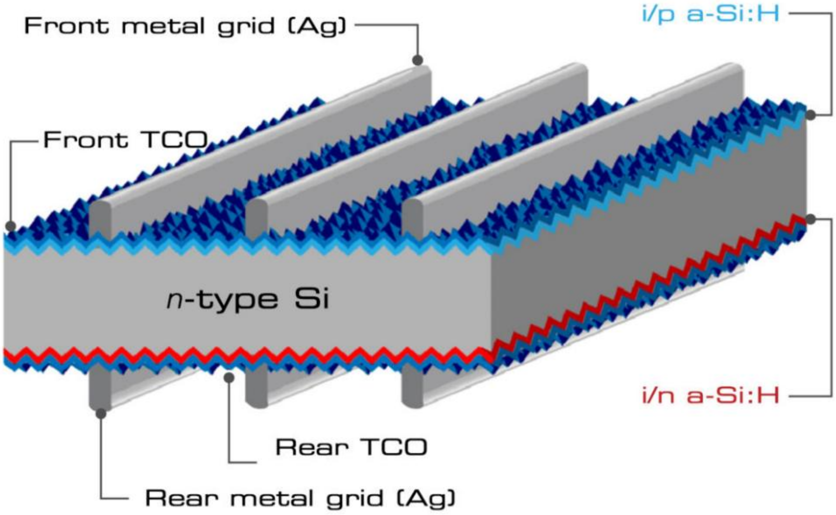


Figure 1.9 Structure of Monocrystalline silicon cell [8]

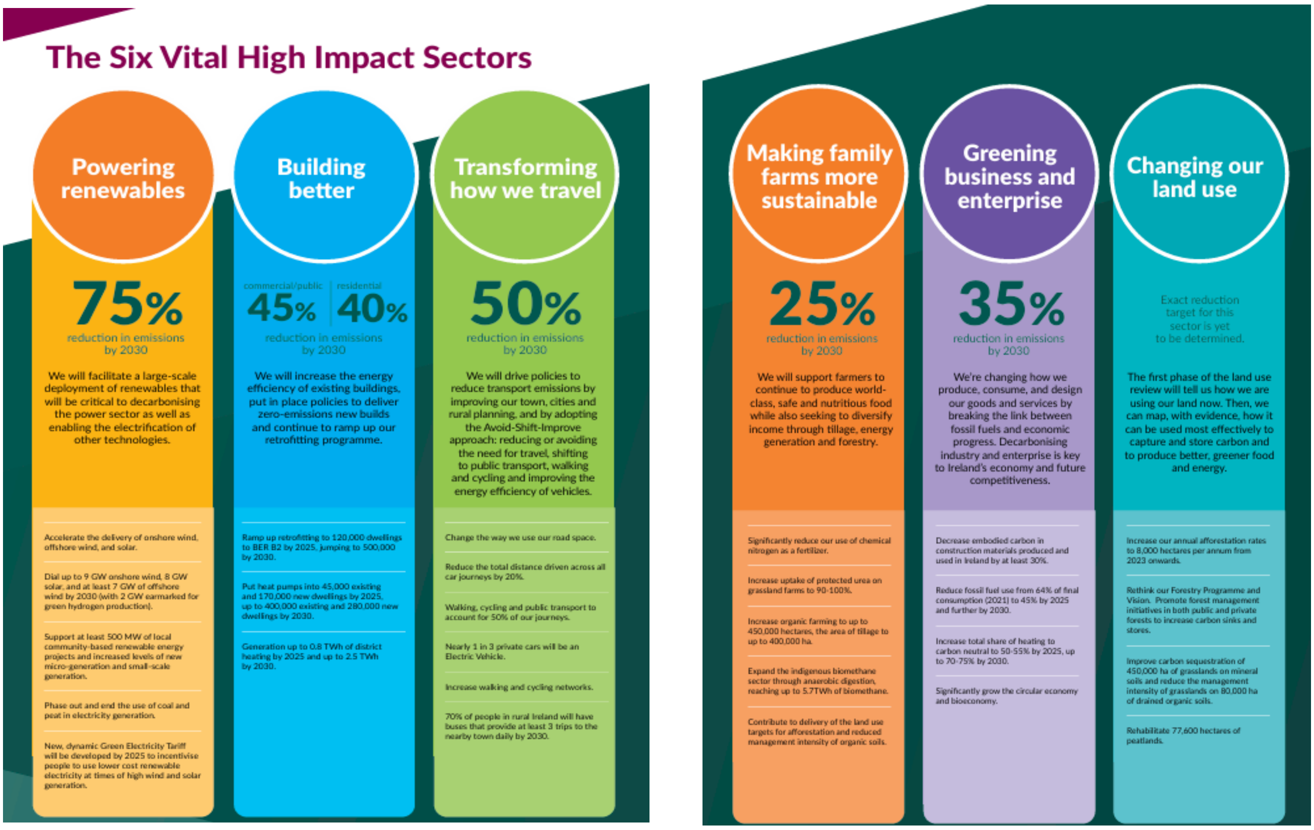
In Figure 1.9 monocrystalline silicon photovoltaic module is represented, which is one of the most mature and efficient technologies currently available in the photovoltaic sector. These modules are manufactured from single-crystal silicon wafers, which provide a highly ordered crystal structure and contribute to superior electrical performance compared with other silicon-based technologies. The structure of a monocrystalline silicon solar cell is based on a P-N junction formed within a silicon wafer. The front surface of the cell typically includes an anti-reflective coating and a metallic grid that collects the generated current while minimizing optical losses. The rear surface is equipped with a metallic contact that completes the electrical circuit. When solar radiation reaches the cell, photons with sufficient energy generate electron–hole pairs within the semiconductor material. The internal electric field established at the P-N junction separates these charge carriers, enabling the generation of electrical current. Monocrystalline photovoltaic modules generally exhibit conversion efficiencies ranging from 20% to 22%, making them among the highest-performing commercially available photovoltaic technologies. Their high efficiency is primarily attributed to the purity of the silicon and the uniformity of the crystal lattice, which reduce charge-carrier recombination losses and enhance electrical conductivity. As a result, a greater amount of electrical energy can be generated from a given installation area compared with polycrystalline or thin-film technologies. Because of their high energy density, monocrystalline modules are particularly suitable for applications where the available installation area is limited, such as residential rooftops and commercial buildings. Typical module power ratings range from approximately 300 to 500 Wp, depending on the module size, cell configuration and manufacturing technology. Another important characteristic of monocrystalline photovoltaic modules is their long operational lifetime. Under normal operating conditions, these systems can maintain satisfactory performance for 25–30 years, with relatively low annual degradation rates. Furthermore, continuous improvements in manufacturing processes and cell design have contributed to enhanced reliability and durability, making monocrystalline silicon one of the most widely adopted photovoltaic technologies worldwide. Despite their technical advantages, monocrystalline modules generally involve higher manufacturing costs than some alternative photovoltaic technologies due to the complexity of the crystal growth process and the high purity requirements of the silicon material. Nevertheless, their combination of high efficiency, long-term reliability and mature manufacturing processes has contributed to their widespread deployment in both residential and utility-scale photovoltaic installations.

2. ENERGY CONTEXT

2.1 Irish energy context

Over the past two decades, Ireland has made significant investments in the development of wind energy, taking advantage of the particularly favorable geographical conditions of the country. The island's position, exposed to strong and consistent winds coming from the Atlantic Ocean, makes wind power the main renewable energy source in the country. According to recent data, this technology covers more than 35% of the national energy demand, and the Irish government has set the goal of reaching around 70% by 2030. Alongside the development of wind energy, solar power has also been gradually increasing, although its contribution is still lower. In recent years, the government has introduced several incentives to encourage the installation of photovoltaic systems both in the industrial and residential sectors, with the aim of promoting energy self-consumption and reducing dependence on fossil fuels. One of the key factors that has supported the development of renewable energy in Ireland is the strong political and social support [9].

Figure 2.1 The six high impact sectors [11]



The Climate Action Plan 2023 includes significant investments in energy infrastructure, such as the strengthening of the electrical grid to facilitate the integration of renewable energy sources and the development of energy storage systems aimed at managing the variability of energy production. At the same time, Ireland has defined an ambitious strategy to significantly increase the production of energy from renewable sources, with particular attention to the development of wind energy, both onshore and offshore. Among the main objectives is the goal of reaching 80% of electricity generated from renewable sources by 2030 (5 GW of offshore wind power and 2.5 GW of photovoltaic capacity) [10]. In the Figure 2.1 the six vital high impact sectors are reported.

2.2 Ireland 2024: energy flows analysis

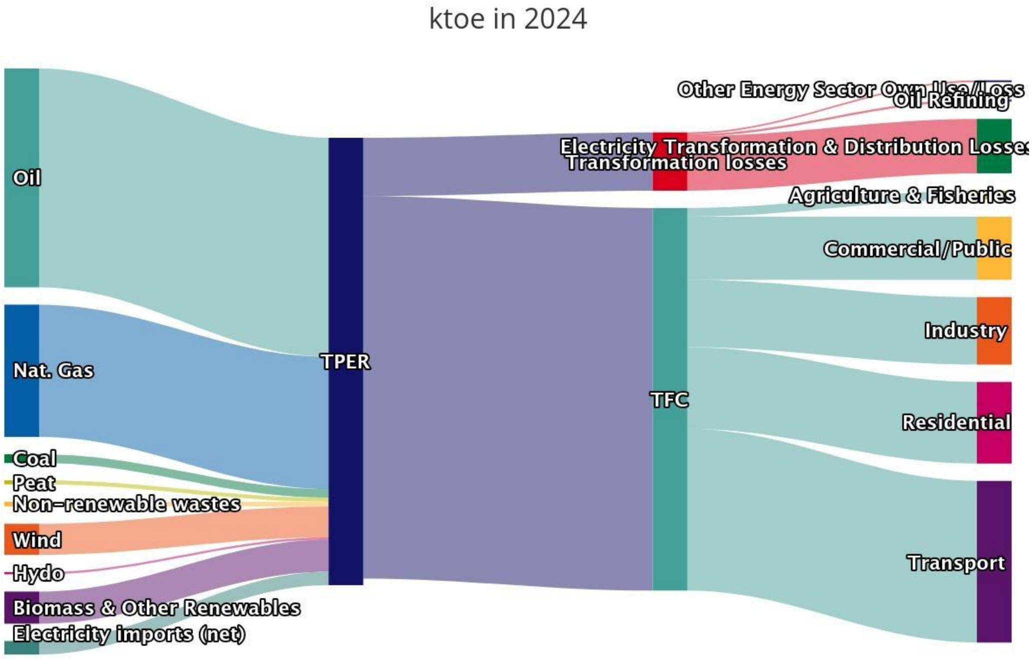


Figure 2.1 Annual energy flows from raw materials to final consumption by sector [12]

The total energy flow in Ireland in 2024 illustrates how primary energy resources are supplied, transformed, and finally consumed across the main sectors of the economy (Figure 2.1). The Total Primary Energy Requirement (TPER) amounts to 14,450 ktoe, which represents the total quantity of energy entering the Irish energy system from domestic production and imports. The primary energy mix is largely dominated by fossil fuels, particularly oil and natural gas, while smaller contributions come from coal, petroleum products, non-renewable wastes, wind, hydro, biomass, other renewable sources, and energy imports. Despite the increasing deployment of renewable energy technologies in recent years, fossil fuels still represent a significant share of the primary energy supply, reflecting the structural characteristics of Ireland’s energy system and its dependence on imported fuels. Before reaching the final consumption sectors, part of this primary energy is lost during the energy transformation processes, which include electricity generation, fuel refining, and other conversion activities. In 2024, these processes account for approximately 1,883 ktoe of energy losses, which correspond to inefficiencies inherent in thermal conversion technologies and other transformation stages. As a consequence, the energy available for final consumption across all sectors is reduced to approximately 12,360 ktoe. Among the final demand sectors, the residential sector represents a significant share of energy consumption. In 2024, households account for approximately 2,636 ktoe of final energy use. Such consumption levels highlights the importance of residential energy demand within the national energy balance. A large portion of this consumption is typically associated with space heating, water heating, cooking, and household electricity use, with heating needs playing a particularly dominant role due to Ireland's climatic conditions and building stock characteristics. From a systemic perspective, the residential sector therefore represents a key area for energy efficiency improvements and decarbonisation strategies. Measures such as building insulation upgrades, electrification of heating through heat pumps, increased adoption of renewable electricity such as rooftop photovoltaic systems, and improvements in energy efficiency standards could significantly reduce the energy demand of households while also lowering greenhouse gas emissions. Therefor, analysing the residential energy demand within the national energy flow provides valuable insights into where policy interventions and technological innovations can most effectively contribute to Ireland's broader energy transition objectives. The analysis of final energy consumption provides an overview of the total energy demand across the different sectors of the Irish energy system. However, in order to better understand the role of specific energy carriers and the ongoing transition of the energy sector, it is also important to examine electricity consumption separately. Electricity represents a key component of final energy use and plays a central role in the decarbonisation process, particularly through the integration of renewable energy sources in power generation.

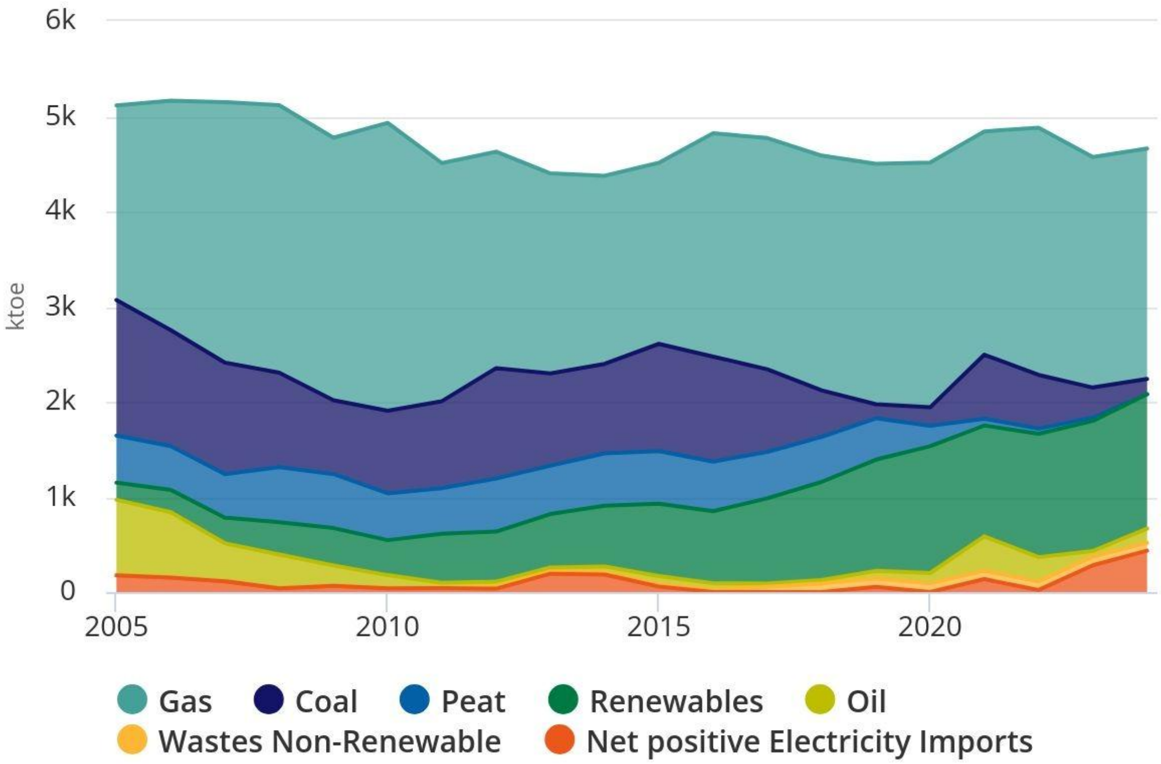


Figure 2.2 Electricity generation broken-out by fuel type and energy source [13]

The Figure 2.2 illustrates the evolution of the primary energy inputs used for electricity generation in Ireland over the period 2006–2024. The data highlight a clear predominance of fossil fuels in the electricity generation mix, particularly natural gas and coal, which represent the largest shares of primary inputs throughout the considered time span. Natural gas, in particular, plays a central role in the Irish power system due to the extensive use of gas-fired power plants, which provide a flexible and reliable source of electricity generation. One of the most notable trends observable in the figure is the progressive phase-out of peat as a fuel for electricity generation. Historically, peat represented a relevant domestic energy resource in Ireland; however, its contribution has steadily declined over the years as a consequence of environmental policies and the gradual closure of peat-fired power plants. The reduction in peat-fired generation has been partially offset by a steady and gradual increase in renewable energy sources, reflecting the growing deployment of renewable generation technologies within the Irish electricity sector. In contrast, other energy sources such as oil, non-renewable wastes, and electricity imports maintain relatively small and almost constant shares over the analysed period. Their contribution to the total primary inputs remains limited and does not exhibit significant variations compared with the major energy sources. Overall, the figure highlights a transitional phase in the Irish electricity generation system. While fossil fuels, particularly natural gas, still dominate the primary energy inputs, the gradual reduction of peat and the increasing penetration of renewable energy sources indicate a structural shift towards a more sustainable electricity generation mix over the long term. Building on the observation of evolving fuel shares in electricity generation, it is important to emphasise the role of renewable energy sources. Despite the gradual increase in renewables over the period 2006–2024, their overall contribution to primary inputs for electricity generation remains modest relative to dominant fossil fuels such as gas and coal. The continued dominance of fossil fuels underscores the need for a deeper investigation into the factors that have influenced the pace of renewable deployment in Ireland, and the policies and market conditions that could accelerate the energy transition towards a cleaner generation mix. [13]

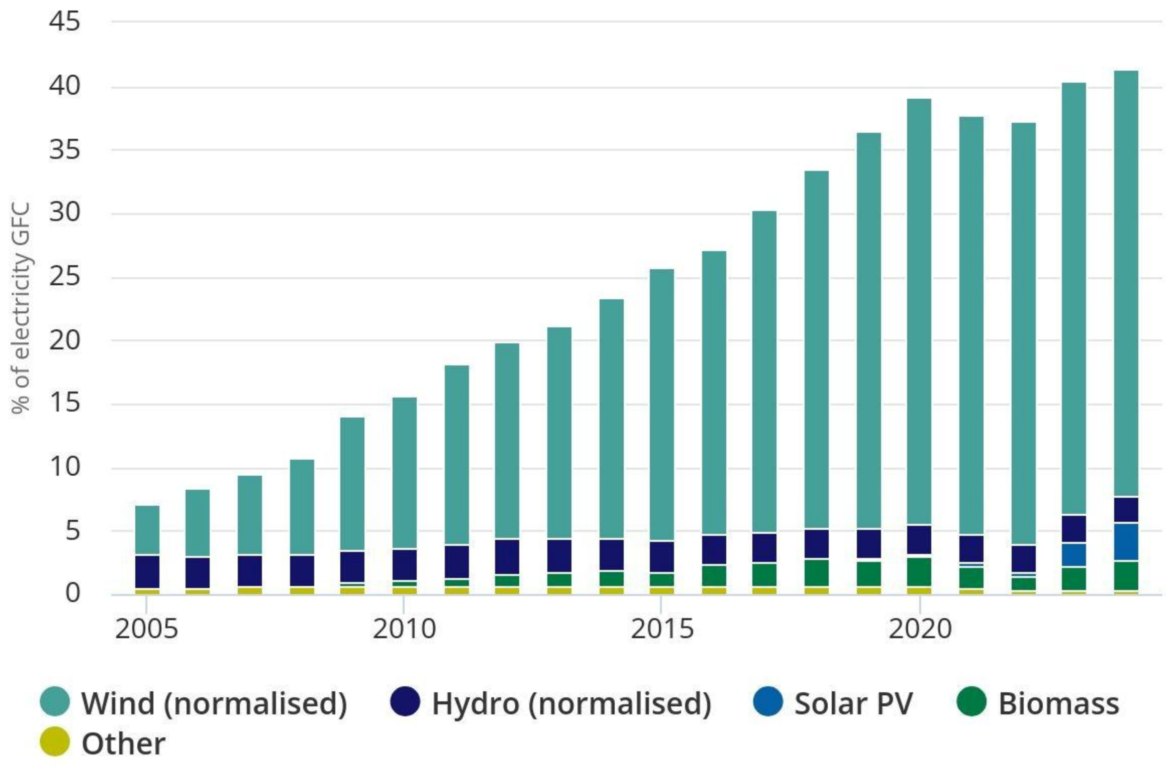


Figure 2.3 Renewable energy share in electricity [14]

The Figure 2.3 shows the evolution of the share of electricity generation from renewable sources in Ireland over the period 2005–2024. The data are expressed as normalized shares, indicating each technology's contribution relative to the total renewable electricity generation in percentage terms. Normalization allows for a clear comparison of the relative importance of different renewable technologies over time, independently of the absolute growth of total renewable generation. Overall, there is a clear upward trend in renewable electricity generation, reflecting the gradual integration of these technologies into Ireland's power system. Wind energy consistently dominates throughout the entire period, with normalized values remaining well above 30 in recent years, underscoring its central role in the renewable electricity mix. Hydropower maintains a relatively constant share, showing little change over time due to geographic and resource limitations. Solar photovoltaic generation appears around 2018 and gradually increases its share from approximately 0.1 to 3 by 2024. Despite this growth, solar remains a small fraction of the total renewable generation. Biomass-based electricity generation has been introduced from around 2010 and stabilizes at a share of approximately 2.3, reflecting its role in diversifying Ireland's renewable energy portfolio. In summary, while the overall share of renewable electricity has steadily increased over the past two decades, the growth is largely driven by wind energy. Other technologies, including hydro, solar, and biomass, play complementary but smaller roles, highlighting wind as the cornerstone of Ireland's renewable electricity strategy, with emerging technologies gradually expanding their presence [15].

2.3 Comparison with Italy

To provide a clearer perspective on the role of renewable energy in electricity generation, a comparison has been made between Ireland and Italy using consistent datasets from TERNA and ISPRA. The corresponding histogram (Figure 2.4) reveal that in Italy, renewable sources contribute a substantially larger share to the total electricity production compared to Ireland. A similar trend can be observed across all major renewable technologies, including wind, hydro, biomass, and notably solar photovoltaic generation. In particular, solar energy in Italy exhibits a significantly higher utilization relative to Ireland, reflecting the more favorable climatic conditions and the extensive deployment of solar installations, which together result in a more prominent role of renewables within the Italian electricity mix. [16]

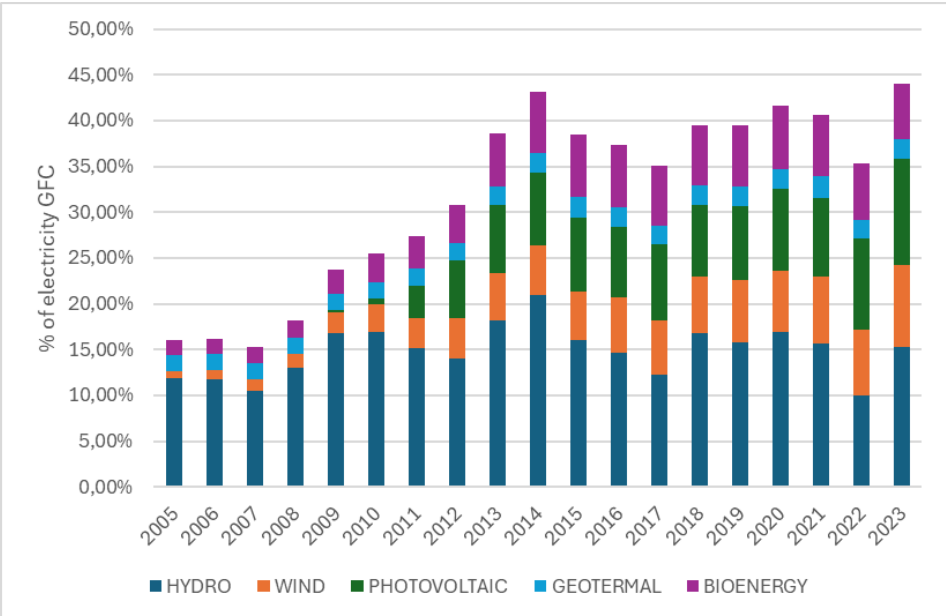


Figure 2.4 Comparison of renewable electricity generation breakdown by technology between Italy and Ireland.

For instance, in 2023 solar photovoltaic generation accounted for approximately 11.6 % of total electricity production in Italy, whereas in Ireland it represented only 1.9 % of the corresponding national generation. The observed difference highlights the much more significant role that solar energy plays in the Italian electricity mix, reflecting both the country's more favorable solar resources and the greater deployment of photovoltaic capacity compared to Ireland.

2.4 Households and energy communities

The traditional centralized electricity system has progressively evolved towards a more decentralized model in which households are no longer only energy consumers but can also act as producers. This phenomenon often referred to as "prosumerism", is driven by the rapid diffusion of rooftop photovoltaic systems, residential battery storage, and electrification technologies such as heat pumps. As a result, an increasing number of households are able to partially satisfy their own electricity demand through self-generation, reducing their dependence on the electricity grid while contributing to a more distributed and flexible energy system. The European Union's long-term climate neutrality targets require a profound transformation of the energy system. By 2050, at least 75% of Europe's energy demand should come from renewable sources, with a growing share of electricity generated through collective initiatives. Achieving this transition requires a shift from centralized systems to decentralized renewable-based structures, where citizens actively participate in energy generation, consumption, and management. In the context of the energy transition, the role of households (residential units or families) is becoming increasingly important. Thanks to the diffusion of small-scale renewable technologies, such as rooftop photovoltaic systems, households can produce part of the energy needed to meet their own energy demand. As a result, households can reduce their dependence on the traditional electricity grid and to increase the self-consumption of locally generated energy. When energy production exceeds domestic consumption, the surplus electricity can be fed into the grid or shared with other users, for example within an energy community. Consequently, households take on a more active role in the energy system and become prosumers, meaning actors who not only consume energy but also contribute to its production and potentially to its sale. In the literature, four main types of households can be distinguished according to their role in the energy system. The first category is consumer households, which only consume electricity supplied by the grid. A second category is prosumer households, which both produce and consume energy, for example through residential photovoltaic systems. A third type is represented by self-sufficient households, which, thanks to local generation and storage systems, are able to cover most of their energy demand and therefore become almost independent from the grid. Finally, there are community households, which participate in energy communities and collaborate with other users to produce, share, and manage renewable energy at the local level. Households are increasingly becoming prosumers, producing and consuming their own electricity through small-scale technologies like rooftop photovoltaics. Prosumers reduce grid withdrawals, help flatten peak demand, and can provide flexibility services. Falling costs and grid parity have further encouraged self-consumption of locally generated energy. Although households represent the basic unit of residential energy consumption, increasing attention is now being given to energy communities, which enable collective participation in decentralized energy systems. Energy communities are not a completely new concept. Although they have become more popular in recent years, their origins date back to the early 20th century. At that time many infrastructures, such as the electricity grid, were not yet developed. For this reason, groups of citizens organized themselves to obtain essential services such as electricity, water, or banking services. In the energy sector in particular, people joined together in cooperatives or communities to build and manage their own electricity networks. During the 20th century these initiatives mainly contributed to the electrification of rural areas, improving energy security and fighting energy poverty. In recent decades, however, the focus of energy communities has increasingly shifted toward environmental issues and the fight against climate change. An important moment for the development of energy communities occurred in 2018, when the European Union officially recognized them as a key tool for the energy transition. In that year two main types were defined: renewable energy communities and citizen energy communities. Thanks to this legal definition, energy communities gained new rights and a more active role in the energy system. For example, they can produce and self-consume energy, trade it among members, participate in energy markets, and offer flexibility services to the grid. Energy communities can mainly be divided into two types recognized by European legislation: renewable energy communities and citizen energy communities. Both allow citizens to actively participate in the energy system, but they differ in terms of structure, objectives, and the activities they can perform. Renewable energy communities are defined by European legislation on renewable energy and are mainly based on the production and sharing of energy generated from renewable sources, such as photovoltaic, wind, or biomass. An important characteristic of these communities is their strong connection with the local territory: members must be located close to the energy generation facilities in order to promote the local production and consumption of energy. Participants can include citizens, small and medium-sized enterprises, or local authorities, and the main objective is not economic profit but the creation of environmental, economic, and social benefits for the local community. Citizen energy communities, on the other hand, are defined by European legislation on the internal electricity market and adopt a broader approach. Unlike renewable energy communities, they are not limited exclusively to renewable energy sources and do not necessarily require geographical proximity between members and generation facilities. In addition, they can carry out a wider range of activities within the electricity market, such as energy generation, distribution, supply, or participation in flexibility services. Participation must also be voluntary and based on democratic governance, with the aim of providing benefits for members and the community rather than maximizing profit. In summary, renewable energy communities are generally local initiatives focused on the production and sharing of renewable energy, whereas citizen energy communities represent more flexible structures with a broader scope of action within the energy market (Figure 2.5) [17]

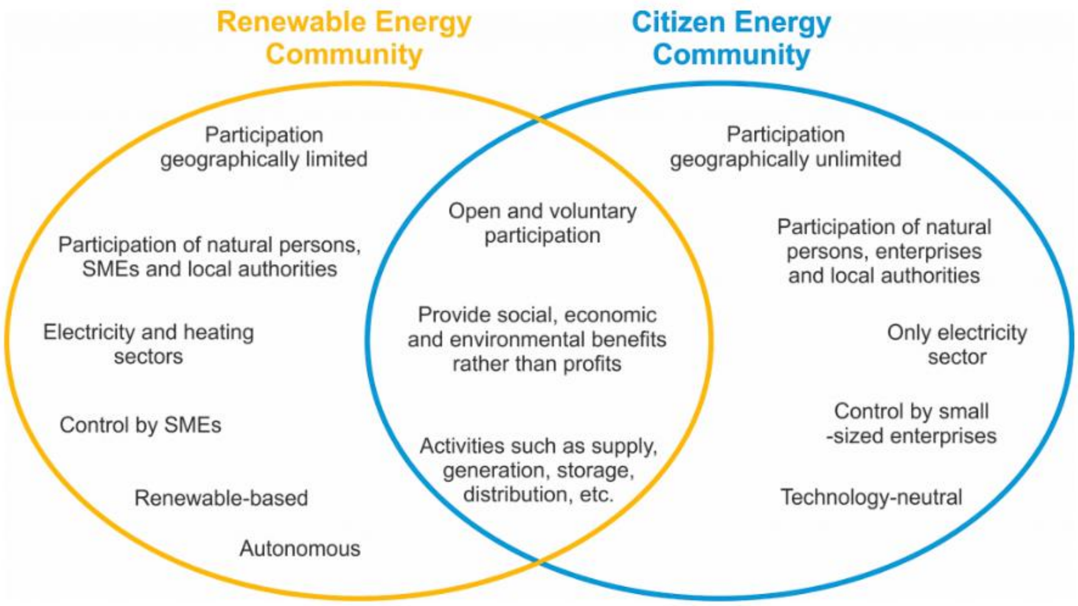


Figure 2.5 Renewable and citizen energy communities [18]

In general, an energy community is a legal entity formed by citizens, small and medium-sized enterprises, or local authorities. Participation must be open and voluntary, and the main objective is not financial profit but the creation of social, environmental, and economic benefits for the community. Energy communities may operate in different parts of the energy system, such as energy production, distribution, supply, and consumption. They also represent a fair and inclusive way to achieve the energy transition because they allow citizens to actively participate in the energy system [19]. The main benefits include inproved energy efficiency, reduced energy bills, lower greenhouse gas emissions, and support for the local economy through the creation of new job opportunities. Finally, energy communities can also attract private investment in the clean energy transition, for example by using citizens' savings to finance renewable energy projects. Energy communities represent an important instrument for the transformation of the energy system, as they enable citizens to actively participate in energy production and management. They can provide economic benefits by reducing energy costs for members, encouraging local investments, and supporting job creation. In addition, they contribute to increasing citizens' awareness and understanding of renewable energy and environmental issues. Energy communities also encourage greater citizen participation. Members tend to be more engaged with the energy transition, climate change, and energy consumption behaviors, which can lead to more sustainable choices and environmentally conscious lifestyles. Energy communities may also generate important social benefits. Energy communities strengthen ties among members, improve the sense of belonging, and can contribute to social and psychological well-being. They also help achieve renewable energy production targets by involving citizens directly in investments and energy projects. Finally, they foster innovation, both technological and social, changing the way energy is produced, shared, and used [20]. However, energy communities also face several barriers. These include organizational, legal, and bureaucratic challenges, which can make project management difficult, especially when relying on volunteer work. The structure of the energy market can also be an obstacle, as it is primarily designed for large energy companies, making it harder for smaller communities to participate or access funding. Another challenge is political and institutional support, which is not always sufficient. In some cases, there is also social opposition to renewable energy projects, for example when people do not want wind or solar installations near their homes. Energy communities may also struggle to secure and maintain long-term financial and technical resources, as consulting and professional costs can be high. Finally, saturation barriers may occur, for example when there are no more spaces available for new installations or when interested participants have already joined other energy communities. In conclusion, energy communities are an important element of the energy transition due to the many benefits they provide, but their development requires addressing and overcoming various economic, social, political, and organizational challenges [21]. The evolution of the Irish electricity sector, together with the increasing adoption of residential photovoltaic systems and the emergence of prosumer-based energy models, forms the basis for the present study. In this framework, accurate photovoltaic production modelling, system calibration, and economic assessment play a crucial role for evaluating the performance and profitability of distributed energy resources. The following chapters therefore focus on the development and validation of methodologies for analysing residential photovoltaic systems using real-world Irish datasets.

3. METHODOLOGY

3.1 General purpose

This chapter, presents the methodology adopted to develop, validate, and calibrate the mathematical–computational model. The objective is to describe the logical and analytical framework followed to transform raw data into an accurate predictive tool, capable of evaluating the performance of a residential photovoltaic system under real operating conditions.

The methodological approach is structured around four main phases:

- The Forward Model (Predictive): The first part of the methodology relies on a forward approach. The model takes as inputs the geographical coordinates, theoretical or satellite-derived solar radiation data (e.g., PVGIS), and the technical characteristics of the components (panel efficiency, tilt, orientation, and inverter losses). Using these physical and environmental inputs, the model calculates the theoretical hourly and daily energy production, subsequently compared with PVGIS reference estimates.

- The Inverse Model (POA Estimation): The second phase arises from the need to address real-world data from the StoreNet pilot project. As pyranometers were not available to measure the actual irradiance striking the panels (Plane of Array – POA), an inverse procedure was applied. In this case, the analysis is performed using a high-resolution, minute-by-minute dataset. Once this high-resolution irradiance profile is reconstructed, it is fed back into the forward model to estimate the expected electricity generation. This reconstructed production profile is then directly compared with the actual, measured real-world data to verify the model's accuracy and complete the cross-validation loop.

- Calibration and Validation: This methodological phase involves the calibration of the system. To feed this model, we use both environmental variables, such as ambient temperature from PVGIS, and the plane-of-array (POA) irradiance estimated indirectly, alongside all the energy flows recorded by the smart meters, including household consumption, grid exchanges, and storage dynamics. The ultimate objective is to calibrate three key parameters: peak power, Performance Ratio, and the temperature coefficient.

- Economic Analysis, Optimization, and Energy Storage Study: The final part of the methodology extends the technical analysis to an economic and management feasibility assessment tailored to the Irish energy market. Starting from the collection and classification of a comprehensive dataset containing all fixed electricity tariffs currently available on the market in Ireland, a ranking was established to identify the most advantageous contractual solutions for the residential user. This scenario was subsequently compared with the new Renew tariff (a dynamic tariff based on half-hourly wholesale market variations) which is expected to be introduced at national level. Based on this combined tariff structure, a consumption optimization algorithm was implemented to evaluate the user's economic flexibility (load shifting and smart storage management). Finally, to quantify the impact of hardware investment, the methodology integrates a dimensional sensitivity analysis aimed at verifying how economic flows, payback periods, and grid independence levels change when varying the capacity of the residential battery, specifically comparing three storage sizes: 5 kW, 10 kW, and 15 kW.

3.2 Forward PV performance model using PVGIS data

The model was developed through a sequence of progressively refined stages aimed at continuously improving its accuracy and its consistency with real operating conditions. In its initial version, the model was able to simulate photovoltaic energy production with a relatively coarse temporal resolution, relying on solar radiation data obtained from PVGIS (Photovoltaic Geographical Information System). PVGIS is an online tool developed by the European Commission that provides free access to solar radiation and photovoltaic performance data based on geographical location. It combines satellite observations, meteorological databases, and physical models to estimate solar irradiance and potential energy production for photovoltaic systems across different regions. While highly reliable and widely used for preliminary assessments, PVGIS data are typically provided at daily or monthly resolutions, which can limit the level of detail in time-dependent analyses. To overcome this limitation, the model was progressively refined by increasing its temporal resolution, moving from a daily-based approach to an hourly one. This improvement allowed for a more accurate representation of the variability in solar radiation and system performance throughout the day, leading to more precise estimates of energy production and a better alignment with real consumption patterns. Once the intermediate version of the model had been validated against PVGIS estimates, a further and more rigorous validation phase was carried out using real measured data. The key electrical parameters of the PV panel (such as rated power, voltage, and current) were taken directly from Vertex datasheet [22] to ensure an accurate representation of the specific system under study**.** The strong agreement between simulated and PVGIS data confirms the reliability of the model and its suitability as a tool for analyzing and designing residential photovoltaic systems, supporting its applicability to residential PV performance analysis.

3.2.1 Model inputs and system configuration

The model requires three categories of inputs: meteorological data, geometric and electrical characteristics of the PV system, and empirical/physical coefficients.

- **Meteorological inputs:**

For the location of interest, the model queries the PVGIS (Photovoltaic Geographical Information System) database of the European Union, downloading hourly data over a period of sixteen years (2005–2020). The following variables are acquired for each hour:

- Global Horizontal Irradiance (GHI) [W/m²]
- Direct Normal Irradiance (DNI) [W/m²]
- Diffuse Horizontal Irradiance (DHI) [W/m²]
- Ambient air temperature T_{amb} [°C]

From these, the model constructs a Typical Meteorological Year (TMY) by averaging, for each hour of the day and each day of the year, the corresponding values across the 16-year period:

$$TMY_{h,d} = \frac{1}{N} \sum_{y=2005}^{2020} Data_{y,h,d} \quad (3.1)$$

where N = 16 represents the number of years, h the hour of the day (0-23), and d the day of the year (1-365). February 29th is also removed to avoid distortions due to leap years.

- **Wind speed estimation**

Since PVGIS does not provide wind speed data in its standard TMY export, the model estimates wind speed v_w [m/s] seasonally:

- Summer months (May–August): $v_w = 0.8$ m/s
- Winter months (November–February): $v_w = 1.5$ m/s
- Intermediate months (March, April, September, October): $v_w = 1.1$ m/s.

- **PV system configuration inputs**

The model simulates a fixed-tilt PV system with the following parameters reported in Table 3.1

Table 3.1 PV system parameters

PARAMETER	SYMBOL	VALUE
Number of panels	N_{panels}	5
Panel peak power (STC)	$P_{peak,panel}$	435 Wp
Total installed power	P_{tot}	2,175 kWp
Panel area	A_{module}	Derived from efficiency (see below)
Tilt angle	β_{tilt}	30° from horizontal
Azimuth angle	γ	180° (south-facing)
Nominal efficiency at STC	η_{STC}	21,8%
Temperature coefficient of P_{max}	β	-0,0034 1/°C
Reference performance ratio	PR_{ref}	0,94
Inverter efficiency	η_inv	0,96

From the nominal efficiency and peak power, the total collecting surface area is:

$$A_{tot} = N_{panel} * A_{module} \tag{3.2}$$

- **Empirical coefficients for thermal model**

The Sandia cell temperature model uses two empirical coefficients specific to the module technology:

$$a = -3.48, b = -0.080$$

- **Latitude-dependent coefficients for dynamic Performance Ratio**

The coefficients a and b in the dynamic PR formulation depend on the site latitude ϕ:

- If $\phi < 40°$: $a = 0.92,\ b = 0.08$
- If $40° \leq \phi \leq 50°$: $a = 0.90,\ b = 0.10$
- If $\phi > 50°$: $a = 0.88,\ b = 0.12$

All these inputs are defined before any production calculation begins.

3.2.2 Core mathematical modeling of efficiency and production

- **Step 1- Radiation on the tilted plane (POA)**

Once the meteorological data is acquired, the model calculates the solar radiation incident on the plane of the panels, known as POA (Plane of Array). This quantity is given by the sum of three contributions: the direct component, the diffuse sky component, and the ground-reflected component.

$$G_{POA} = G_{beam} + G_{sky\ diff} + G_{ground\ diff} \tag{3.3}$$

where:

- G_{beam} is the direct radiation, calculated from Direct Normal Irradiance (DNI) and the angle of incidence;
- $G_{sky\ diff}$ is the diffuse sky radiation, modeled according to isotropic or anisotropic models depending on PVGIS settings;
- $G_{ground\ diff}$ is the ground-reflected radiation, a function of the local albedo.

This value represents the fundamental energy input for calculating electrical production.

- **Step 2 – Cell temperature modeling (Sandia model)**

A crucial aspect for an accurate production estimate is the correct modeling of the operating temperature of the photovoltaic cells, as panel efficiency decreases with increasing temperature. The model adopts the Sandia model, which expresses cell temperature as:

$$T_{cell} = T_{amb} + G_{POA} * exp(a + b * v_w) \tag{3.4}$$

where:

- T_{amb} is the ambient air temperature [°C];
- G_{POA} is the incident radiation on the panel plane [W/m²];
- v_w is the wind speed [m/s];
- a = -3.48 and b = -0.080 are empirical parameters characteristic of the photovoltaic module.

Wind speed, not being directly available from PVGIS, is estimated seasonally: 0.8 m/s in summer months (May-August), 1.5 m/s in winter months (November-February), and 1.1 m/s in the intermediate months.

- **Step 3 – Panel efficiency correction with temperature**

Once the incident radiation and the operating temperature of the cells are known, the model proceeds to calculate the hourly electricity production according to a chain of physical relationships that account for the efficiencies of the various components. The efficiency of the photovoltaic panel is corrected based on the actual cell temperature relative to standard reference conditions (25°C). The relationship used is as follows:

$$\eta_{panel}(T_{cell}) = \eta_{STC} * [1 + \beta * (T_{cell} - 25)] \tag{3.5}$$

where:

- η_{STC} = 0.218 is the nominal efficiency under standard conditions (21.8%);
- β = -0.0034 1/°C is the temperature coefficient of maximum power;
- T_{cell} is the cell temperature calculated in the previous step.

As can be observed, the term beta is negative, therefore as temperature increases, efficiency decreases linearly.

- **Step 4 – Clearness index and dynamic performance ratio**

The Performance Ratio (PR) represents the set of system losses not directly related to panel efficiency, such as soiling losses, shading, mismatch, wiring, and inverter losses. In the developed model, the PR is not considered constant but is made dynamic as a function of sky clearness through the clearness index K_t, defined as:

$$K_t = \frac{G_{STC}}{G_{POA}} \tag{3.6}$$

with G_{POA} expressed in W/m².

This formula is acceptable because Kt is used exclusively as a linear scaling factor to modulate the Performance Ratio as a function of irradiance, and not for radiometric purposes or comparison with theoretical physical models, making normalization to the STC reference value of 1000 W/m² sufficient, which represents the maximum irradiance typically observable in real operating conditions for the considered system.

K_t = 1 sunny

K_t = 0 cloudy

The dynamic PR is then calculated as:

$$PR(K_t) = PR_{ref} * (a + b * K_t) \tag{3.7}$$

This formulation accounts for the fact that on very clear days high K_t, relative losses tend to be slightly higher than on cloudy days.

- **Step 5 – Overall system efficiency**

The overall system efficiency is given by the product of the panel efficiency (temperature-corrected), the dynamic Performance Ratio, and the inverter efficiency:

$$\eta_{system} = \eta_{panel}(T_{cell}) * PR(K_t) * \eta_{inv} \tag{3.8}$$

with $\eta_{inv} = 0.96$ (constant, assumed inverter efficiency at partial load is representative).

- **Step 6 – Hourly energy production**

Finally, the hourly electricity production expressed in kWh is calculated as:

$$E_{hourly} = \frac{G_{POA}}{1000} * A_{tot} * \eta_{system} \tag{3.9}$$

where:

- The factor $\frac{G_{POA}}{1000}$ converts radiation from W/m² to kW/m²
- A_{TOT} is the total panel area in m²
- η_{system} is the dimensionless overall efficiency. Daily production is then obtained by summing the contributions of the 24 hours:

$$E_{daily} = \sum_{h=0}^{23} E_{hourly}(h) \tag{3.10}$$

Similarly, monthly and annual production are obtained by aggregating daily values.

3.2.3 Statistical validation framework and residual analysis

A fundamental aspect of this work is the rigorous validation of the developed model. To this end, the estimated production is systematically compared with the reference values provided by the same PVGIS database, which constitutes the validation benchmark. The validation is conducted on multiple temporal scales (hourly, daily, monthly, and seasonal) using a comprehensive set of statistical metrics. The main metrics adopted to quantify the model's accuracy are:

- **Mean Absolute Percentage Error (MAPE):**

$$MAPE = \frac{100\%}{n} * \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{3.11}$$

where y_i is the reference production (PVGIS) and $\hat{y}\_i$ is the production estimated by the model. This metric expresses the magnitude of the error in percentage terms, providing an intuitive measure of prediction error.

- **Mean Bias Error (MBE):**

$$MBE = \frac{1}{n} * \sum_{i=1}^{n} \hat{y}_i - y_i \tag{3.12}$$

MBE indicates whether the model tends to overestimate (positive MBE) or underestimate (negative MBE) production systematically. Values close to zero indicate no significant bias.

- **Root Mean Square Error (RMSE):**

$$RMSE = \sqrt{\frac{1}{n} * \sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \tag{3.13}$$

RMSE penalizes large errors more heavily than MAE, being particularly useful for identifying the presence of outliers.

- **Mean Absolute Error (MAE):**

$$MAE = \frac{1}{n} * \sum_{i=1}^{n} |\hat{y}_i - y_i| \tag{3.14}$$

- **R^2 (Coefficient of Determination)**

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{3.15}$$

where $\hat{y}\_i$ is the value estimated by the model and $\bar{y}$ is mean of observed values calculated as $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y\_i$

R^2 expresses the proportion of the variance in the reference data explained by the model. A value close to 1 indicates very strong predictive agreement.

- **Statistical distribution tests**

In addition to descriptive metrics, rigorous statistical tests were conducted to verify the similarity between the distributions of estimated and reference production:

- Wilcoxon signed-rank test for paired samples, to test the null hypothesis that the two distributions are equal;
- Spearman's correlation, to assess the strength of the monotonic relationship between model and reference;
- Student's t-test on the null hypothesis that the mean difference between model and reference is zero;
- Kolmogorov-Smirnov test, to compare the cumulative distribution functions of the two samples.

- **Residual analysis**

A detailed residual analysis was also conducted, defined as the difference between estimated and reference production:

$$\varepsilon_i = \hat{y}_i - y_i \tag{3.16}$$

Residuals were analyzed as a function of:

- the hour of the day, to identify any systematic biases in specific time slots;
- the month of the year, to verify the presence of seasonal errors;
- the magnitude of production, to ensure error homogeneity.

The results confirmed the absence of significant bias and an error distribution centered around zero, with a contained standard deviation.

3.3 Inverse-forward model using StoreNet data

To validate the proposed model under real-world operating conditions, this study integrates high-resolution experimental data from the StoreNet project, conducted on the Dingle Peninsula, Ireland. These data were provided by Maynooth University and include detailed information on solar irradiance, photovoltaic energy production, and electricity consumption for a sample of twenty residential buildings. The availability of high-resolution, real-world measurements made it possible to directly compare the model outputs with actual system performance under realistic conditions. This final validation step represents a crucial aspect of the work, as it demonstrates the model's ability to accurately reproduce real energy production and consumption dynamics. The dataset [23] includes measurements from 20 residential households, 10 of which are equipped with rooftop photovoltaic systems (2.0–2.2 kWp, south-facing, ~35° tilt) and Sonnen batteries. The temporal resolution is one minute for the entire year 2020, recording variables such as PV generation, consumption, and battery state of charge. A critical limitation of the dataset is that solar irradiance the primary input for conventional PV modeling was not measured at each rooftop. Instead, irradiance data were collected at a single, centralized weather station. This spatial mismatch prevents traditional, direct validation where on-site pyranometer data drive the model. Consequently, a direct comparison between simulated and measured production based solely on external irradiance data would yield a relatively low correlation, not due to model inadequacy, but due to unrepresentative input data. To overcome this limitation, this study develops an inverse–forward modeling framework. The framework assumes that all analyzed households are exposed to similar irradiance conditions, being located in the same geographical area. Six houses are selected for the analysis (H1, H2, H3, H4, H5, H7). Instead of using external irradiance measurements, the Plane-of-Array (POA) irradiance is reconstructed directly from the measured electrical production of a reference house (H2). This reconstructed irradiance then becomes the common input for simulating all other systems, while external data (temperature, wind speed) are still used for thermal modeling.

3.3.1 Script 1

The validation procedure is implemented through two sequential computational scripts, where the output of the first script constitutes the fundamental input for the second.

3.3.1.1 Model inputs and system configuration

The objective is to reconstruct a site-specific, high-resolution (minute-by-minute) POA irradiance time series from the measured production of a reference house, by inverting the photovoltaic conversion process. The first script requires two categories of input data, summarized in Table 3.2 and in Table 3.3.

Table 3.2 Primary input data

PARAMETER	SYMBOL	UNIT	SOURCE	DESCRIPTION
Measured energy production	$E_{measured}$	Wh/min	Reference house smart meter	Raw PV generation at 1-minute resolution
Ambient temperature	T_{air}	°C	PVGIS (satellite-derived)	Hourly air temperature, expanded to minutes
Wind speed	v_{wind}	m/s	PVGIS (satellite-derived)	Hourly wind speed, expanded to minutes
PVGIS POA irradiance	POA_{PVGIS}	W/m^2	PVGIS	Auxiliary irradiance (used only for thermal model, not as primary input)

Table 3.3 Fixed system parameters for inverse reconstruction

PARAMETER	SYMBOL	VALUE	UNIT	DESCRIPTION
Module area	A_{m2}	10	m^2	Total active area of the reference PV system
STC efficiency	η_{STC}	0,21	-	Module efficiency at Standard Test Conditions (25°C, 1000 W/m²)
Temperature coefficient	γ	-0,0045	1/°C	Power temperature coefficient for crystalline silicon
Sandia coefficient (intercept)	a	-3,56	-	Empirical coefficient for wind cooling effect
Sandia coefficient (slope)	b	-0,075	-	Empirical coefficient for wind cooling effect
Maximum Wh per minute	$MAX_{Wh-per-min}$	100	Wh/min	Physical upper bound for production (equivalent to 6000 W)
Noise threshold	$RUMOR_{THRESHOLD}$	0,00167	Wh/min	Minimum detectable production (equivalent to 0.1 W)
PR range	PR_{min}, PR_{max}	0,3;1,0	-	Physically plausible Performance Ratio bounds

POA range	POA_{min}, POA_{max}	0;1200	W/m^2	Physically plausible irradiance bounds
Smoothing window	W_{PR}	180	Min	Centered moving average window for PR stabilization

3.3.1.2 Core mathematical modeling of efficiency and production

The inverse reconstruction proceeds through a sequence of well-defined mathematical steps.

- **Step 1: Unit conversion from energy to power**

The raw measured data are provided as energy produced per minute ($E_{measured}$ in Wh/min). To obtain average power over that minute, the following conversion is applied:

$$P_{measured} = E_{measured} * 60 \tag{3.17}$$

The factor 60 derives from the conversion between minutes and hours: $P = \frac{E}{\Delta t}$ with $\Delta t = \frac{1}{60}$ hour, thus $P = E * 60$.

- **Step 2: Data cleaning and preprocessing**

The measured power time series is filtered to remove non-physical or unreliable values:

- Nighttime forcing: For hours between 23:00 and 04:00, production is set to zero, as no solar generation is physically possible.
- Outlier removal: Any value exceeding $MAX\_Wh\_per\_min \times 60 = 6000$ W is flagged as an outlier and set to NaN.
- Spike detection: Isolated non-zero observations surrounded by zero values (both preceding and following values equal to zero) are identified as measurement spikes and set to zero.
- Interpolation: Missing values (NaNs) resulting from outlier removal are filled using linear interpolation, limited to a maximum of 5 consecutive minutes to avoid introducing artificial trends.
- Noise thresholding: Values below $RUMOR_{THRESHOLD} \times 60 = 0.1$ W are considered instrumental noise and set to zero.
- Negativity correction: Any negative values resulting from previous operations are set to zero.

- **Step 3: Cell temperature calculation**

The operating temperature of the PV cells is estimated using the Sandia empirical model, which accounts for the cooling effect of wind:

$$T_{cell} = T_{air} + POA_{PVGIS} \cdot e^{(a+b \cdot v_{wind})} \tag{3.18}$$

where T_{air} is the ambient temperature (°C), POA_{PVGIS} is the auxiliary irradiance from PVGIS (W/m²), v_{wind} is the wind speed (m/s), and $a = -3.56$, $b = -0.075$ are empirically derived coefficients. The resulting temperature is constrained to the physically plausible range $[-20,85]°C$ to prevent numerical extremes.

- **Step 4: Module efficiency calculation**

The efficiency of the PV module at operating conditions deviates from its STC value due to temperature effects. The linear temperature correction model is applied:

$$\eta = \eta_{STC} \cdot [1 + \gamma \cdot (T_{cell} - 25)] \tag{3.19}$$

where $\eta_{STC} = 0.21$ (21%), $\boldsymbol{\gamma}$ = -0.0045 1/°C, and T_{cell} is the cell temperature from Step 3. The resulting efficiency is further constrained to the range [0,05;0,25] to avoid physically implausible values arising from extreme temperature inputs.

- **Step 5: Theoretical power calculation**

The theoretical DC power that would be produced under the PVGIS irradiance, given the actual cell temperature and module efficiency, is computed as:

$$P_{theoretical} = POA_{PVGIS} \cdot A_{m2} \cdot \eta \tag{3.20}$$

where $A_{m2} = 10$ m² is the total module area. This represents the theoretical DC power output before accounting for system-level losses (other than those implicitly captured by the efficiency model).

- **Step 6: Dynamic Performance Ratio calculation**

The raw Performance Ratio (PR_{raw}) is defined as the ratio of measured power to theoretical power:

$$PR_{raw} = \frac{P_{measured}}{P_{theoretical}} \tag{3.21}$$

This calculation is performed only for time steps where $P_{theoretical} > 0$ to avoid division by zero. The PR_{raw} values are then clipped to the physically meaningful interval $[PR_{min}, PR_{max}] = [0.3; 1.0]$, as values outside this range would imply unrealistic loss or gain mechanisms.

To mitigate short-term variability and measurement noise, a smoothed Performance Ratio (PR_{din}) is obtained by applying a centered moving average:

$$PR_{din}(t) = \frac{1}{W_{PR}} \sum_{k=-W_{PR}/2}^{W_{PR}/2} P\,R_{raw}(t + k \cdot \Delta t) \tag{3.22}$$

where $W_{PR} = 180$ minutes (3 hours) is the smoothing window and $\Delta t = 1$ minute is the temporal resolution. This smoothing window is chosen to preserve the diurnal trend while eliminating high-frequency fluctuations.

- **Step 7: POA irradiance reconstruction (inversion)**

The core step of the inverse procedure consists of rearranging the photovoltaic power equation to estimate the actual POA irradiance that would have produced the measured power output, given the estimated efficiency and dynamic PR:

$$POA_{estimated} = \frac{P_{measured}}{A_{m2} \cdot \eta \cdot PR_{din}} \tag{3.23}$$

This equation is derived from $P_{measured} = POA \cdot A_{m2} \cdot \eta \cdot PR$, solved for POA. The calculation is performed for all time steps; for steps where $P_{measured} = 0$, the resulting $POA_{estimated}$ is also set to zero.

The reconstructed irradiance is then subjected to physical constraints:

- Clipping to the range $[POA_{min}, POA_{max}] = [0,1200]$ W/m², as values above 1200 W/m² are considered physically unrealistic under normal operating conditions.
- Nighttime forcing: for hours between 23:00 and 04:00, the irradiance is explicitly set to zero, regardless of the calculated value.

  The output of Script 1 is a complete time series of reconstructed POA irradiance, aggregated to hourly resolution (by averaging the minute-by-minute values) and exported for use as the primary input to Script 2.

3.3.2 Script 2

The objective of the second script is to use the reconstructed POA irradiance as input to simulate AC power production for all six residential PV systems, and to validate the simulation against measured real production data using a comprehensive statistical framework.

3.3.2.1 Model inputs and system configuration

The second script loads two primary data sources and defines the parameters of the forward PV simulation model, summarized in Table 3.4 and in

Table *3.5*.

Table 3.4 Primary input data

PARAMETER	SYMBOL	UNIT	SOURCE	DESCRIPTION
Reconstructed POA irradiance	$POA_{estimated}$	W/m^2	Output of Script 1	Site-specific irradiance reconstructed from reference house production
Measured real production (6 houses)	P_{real}	Wh	Raw household data (cleaned)	Hourly measured PV generation for houses H1–H7
Ambient temperature	T_{air}	°C	PVGIS	Hourly air temperature for thermal modeling
Wind speed	v_{wind}	m/s	PVGIS	Hourly wind speed for thermal modeling

Table 3.5 Fixed system parameters for forward simulation

CATEGORY	PARAMETER	SYMBOL	VALUE	UNIT
Geometric	Latitude	$\boldsymbol{\phi}$	52.141	°
	Longitude	$\boldsymbol{\lambda}$	-10,271	°
	Tilt angle	β	35	°
	Azimuth angle	α	195	°
Electrical	Nominal power	P_{kWp}	2,0	kW

	Inverter max AC power	$P_{inv,max}$	2,0	kW
	Number of panels	N_{panels}	4	-
	Panel dimension (L x W)	-	1,762 x 1,134	m
	Total active area	A_{TOT}	4 x 1,762 x 1,134	m^2
	STC efficiency	η_{STC}	0,218	-
Temperature	Temperature coefficient	$\boldsymbol{\gamma}$	-0,0045	1/°C
	Sandia coefficient (intercept)	a	-3,56	-
	Sandia coefficient (slope)	b	-0,075	-
Losses	Aging loss	L_{aging}	0,02	-
	Wiring loss	L_{wiring}	0,02	-
	Mismatch loss	$L_{mismatch}$	0,01	-
Operational	POA threshold	POA_{th}	1,0	W/m^2

3.3.2.2 Core mathematical modeling of efficiency and production

The forward simulation follows a sequential physical model to convert the reconstructed POA irradiance into AC power output.

- **Step 1: Cell temperature model.**

The cell operating temperature is computed using the same Sandia model as in Script 1, but now employing the reconstructed POA as the irradiance input:

$$T_{cell} = T_{air} + POA_{estimated} \cdot e^{(a+b \cdot v_{wind})} \tag{3.24}$$

The result is clipped to the range $[-10{,}85]°C$ for physical consistency.

- **Step 2: Module efficiency model.** The temperature-corrected module efficiency is calculated as:

$$\eta = \eta_{STC} \cdot [1 + \gamma \cdot (T_{cell} - 25)] \tag{3.25}$$

where $\eta_{STC} = 0.218$ and $\boldsymbol{\gamma}$ = -0.0045 1/°C.

- **Step 3: DC power calculation.**

The DC power output at the module terminals is computed as the product of incident irradiance, total active area, temperature-corrected efficiency, and the comulative effect of fixed system losses.

$$P_{dc} = POA_{estimated} \cdot A_{tot} \cdot \eta \cdot (1 - L_{aging}) \cdot (1 - L_{wiring}) \cdot (1 - L_{mismatch}) \tag{3.26}$$

with $L_{aging} = 0.02$, $L_{wiring} = 0.02$, and $L_{mismatch} = 0.01$. The product of the three loss factors yields a total derating factor of approximately 0.94.

- **Step 4: Inverter efficiency model.**

The inverter efficiency is modeled as a piecewise function of the load ratio $P_{ratio} = \frac{P_{dc}}{P_{kwp}}$

$$\eta_{inv} = \begin{cases} 0.85, & P_{ratio} \leq 0 \\ 0.85 + 0.9 \cdot P_{ratio}, & 0 < P_{ratio} < 0.1 \\ 0.94 + 0.025 \cdot \left[1 - \left(\frac{P_{ratio} - 0.35}{0.5}\right)^2\right], & 0.1 \leq P_{ratio} < 0.6 \\ 0.965 - 0.0125 \cdot (P_{ratio} - 0.6), & P_{ratio} \geq 0.6 \end{cases} \tag{3.27}$$

The first regime (idle) represents standby consumption. The second regime (start-up) is a linear interpolation reflecting the rapid efficiency increase as power becomes detectable. The third regime (mid-range) uses a quadratic function that peaks at approximately 96.5% around $P_{ratio} = 0.35$. The fourth regime (high power) models a slight efficiency decay due to increased thermal losses near nominal capacity.

- **Step 5: AC power calculation.**

The final AC power output is obtained by multiplying the DC power by the inverter efficiency and applying the inverter's maximum power constraint:

$$P_{ac} = min(P_{dc} \cdot \eta_{inv}, P_{inv,max}) \qquad (3.28)$$

The result is expressed in kW per hour; for comparison with measured data (provided in Wh), the simulation output is converted to Wh by multiplying by 1000.

- **Step 6: Data alignment and preprocessing.**

For each of the six houses, the measured real production data are independently cleaned using the identical procedure described in Script 1 (nighttime forcing, outlier removal, spike detection, interpolation, noise thresholding). The simulated hourly production time series is then temporally aligned with each cleaned real production series. A unified dataset is constructed containing:

- Simulated production (P_{sim}),
- Measured production for each individual house ($P_{real}^{(i)}$),
- The average measured production across all houses ($\bar{P}_{real}^{all}$).

This structure enables both per-system and aggregate validation.

3.3.2.3 Statistical validation framework and residual analysis

The performance of the complete inverse-forward modeling framework is evaluated by comparing the simulated production (P_{sim}) against the measured production (P_{real}) for each house and for the average of all houses. A comprehensive set of statistical indicators is calculated over the temporally aligned hourly time series where both values are valid (n observations).

Core error metrics:

- **Mean Absolute Error (MAE):** Measures the average magnitude of the prediction errors, without considering their direction.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | P_{sim,i} - P_{real,i} | \qquad (3.29)$$

- **Root Mean Square Error (RMSE):** Measures the standard deviation of the residuals, assigning greater weight to large deviations, thus penalizing significant deviations more heavily.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_{sim,i} - P_{real,i})^2} \quad (3.30)$$

- **Mean Bias Error (MBE):** Measures the average tendency of the model to overestimate (positive) or underestimate (negative) the observed values.

$$MBE = \frac{1}{n} \sum_{i=1}^{n} (P_{sim,i} - P_{real,i}) \quad (3.31)$$

The relative MBE is also computed as:

$$MBE\% = \frac{MBE}{\bar{P}_{real}} \times 100 \quad (3.32)$$

where $\bar{P}_{real}$ is the mean measured production over the validation period.

- **Mean Absolute Percentage Error (MAPE):** Measures the average magnitude of the errors as a percentage of the measured values, providing a scale-independent metric.

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{P_{real,i} - P_{sim,i}}{P_{real,i}} \right| \times 100\% \quad (3.33)$$

This metric is computed only for time steps where $P_{real,i} > 0$ to avoid division by zero.

- **Correlation Coefficient (r):** Measures the strength and direction of the linear relationship between simulated and measured values. Values range from -1 (perfect negative correlation) to +1 (perfect positive correlation).

$$r = \frac{\sum_{i=1}^{n} (P_{real,i} - \bar{P}_{real})(P_{sim,i} - \bar{P}_{sim})}{\sqrt{\sum_{i=1}^{n} (P_{real,i} - \bar{P}_{real})^2} \sqrt{\sum_{i=1}^{n} (P_{sim,i} - \bar{P}_{sim})^2}} \quad (3.34)$$

- **Coefficient of Determination (R^2):** Indicates the proportion of the variance in the measured data that is predictable from the simulated data. It is interpreted as the fraction of variability explained by the model.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (P_{real,i} - P_{sim,i})^2}{\sum_{i=1}^{n} (P_{real,i} - \bar{P}_{real})^2} \quad (3.35)$$

- **Aggregate performance indicators:**

Beyond the time-series metrics, additional system-level indicators are calculated to assess long-term consistency:

- Total annual energy: $E_{tot,sim} = \sum P_{sim}$ and $E_{tot,real} = \sum P_{real}$ (kWh)
- Specific yield: $Y_{sim} = E_{tot,sim}/P_{kwp}$ and $Y_{real} = E_{tot,real}/P_{kwp}$ (kWh/kWp)
- Total annual irradiance: $H_{tot} = \sum POA_{estimated}/1000$ (kWh/m²)
- Operating hours: Number of hours with $P > 0$ for both simulation and real data
- Daily and monthly profiles: Average production per hour-of-day and per month-of-year to assess diurnal and seasonal patterns.

- **Residual analysis – time-of-day specific diagnostics:**

To diagnose systematic model behavior at different times of the day and identify potential physical processes not fully captured by the model, the error metrics are disaggregated by hour. The hourly MBE and hourly MAE are computed by grouping all residuals and absolute residuals according to the hour of the day $h \in \{0,1,\dots,23\}$:

$$MBE(h) = \frac{1}{n_h} \sum_{i \in \Omega_h} (P_{sim,i} - P_{real,i}) \quad (3.36)$$

$$MAE(h) = \frac{1}{n_h} \sum_{i \in \Omega_h} | P_{sim,i} - P_{real,i} | \quad (3.37)$$

where Ω_h is the set of time indices with hour-of-day equal to h, and $n_h = | \Omega_h |$ is the number of observations for that hour. This analysis is typically restricted to daytime hours (e.g., 06:00 to 18:00) and to periods with production above a minimal threshold (e.g., >20 Wh) to avoid distortion from nighttime zeros.

The hourly bias profile reveals systematic overestimation or underestimation during specific phases of the daily cycle:

- Morning hours (06:00–09:00): May indicate model accuracy during ramp-up and low-angle conditions
- Midday hours (10:00–14:00): Reflect performance at peak irradiance and maximum temperatures
- Afternoon hours (15:00–18:00): Capture model behavior during declining irradiance and cooling

This diagnostic approach heps identify whether discrepancies arise from thermal dynamics (temperature coefficient accuracy), solar geometry (angle-of-incidence effects), or inverter behavior (low-power efficiency).

3.4 Calibration and validation of PV system parameters

In the specific context of this thesis, real production data are available at minute-by-minute [23] resolution for several residential units (H1, H2, H3, H4, H5, H7, H10, H11, H13, H17), each equipped with a photovoltaic system and battery storage. The data include not only PV production but also battery charge/discharge flows, household consumption, and grid exchange. However, these data present two main critical issues.

- The first issue relates to the presence of the battery. The production measured directly by sensors does not represent the instantaneous output of the PV panels, but rather the net production after some energy has been sent to or drawn from the battery. To obtain the true PV generation, it is necessary to "decouple" the battery contribution through an energy balance of the entire household system.
- The second issue concerns measurement quality. Current and voltage sensors used to monitor the system can suffer from drifts, systematic offsets, or synchronisation errors between different measured quantities. In particular, during nighttime hours, when PV production should be zero, small positive values are often observed due precisely to these instrumental offsets. If not corrected, such offsets introduce systematic errors in the estimation of real production and, consequently, in model calibration.

A further fundamental challenge is the absence of direct irradiance measurements at the installation sites. The physical model that relates PV production to incident radiation requires, as input, the Plane of Array irradiance (POA) — that is, the solar power (in W/m²) actually striking the tilted surface of the panels. Since no on-site pyranometers or reference cells are available, the POA must be estimated indirectly from the available data.

In this work, a two-step sequential procedure was therefore adopted:

- **First step – Backward estimation of the POA:** Starting from the cleaned real production data and using a simplified forward model with assumed reference parameters (area, base efficiency, temperature coefficient, and a dynamic Performance Ratio), the POA is reconstructed minute by minute by inverting the physical relationship between production and irradiance.
- **Second step – Calibration of the PV system parameters:** Using the estimated POA as input to the full physical model, a numerical optimisation algorithm is applied to estimate the unknown parameters of each system: peak power (P_{peak}), Performance Ratio (PR), and temperature coefficient (γ). The optimisation minimises the Root Mean Square Error (RMSE) between the model output and the reconstructed real production obtained from the meter data via energy balance.

The two steps are conceptually and operationally distinct: the first step produces an irradiance time series that serves as the essential input for the second step. The same core physical model (cell temperature calculation, temperature correction factor, and power equation) is used in both directions, ensuring consistency between the estimated POA and the subsequently calibrated parameters.

3.4.1 First Step: Backward estimation of the plane of array irradiance (POA)

The first step aims to reconstruct the minute-by-minute POA (in W/m²) starting from the cleaned production data, in the complete absence of direct irradiance measurements.

3.4.1.1 Model inputs and system configuration

The first script requires two distinct categories of input data, summarised in Table 3.6 and in

Table 3.7.

Table 3.6 Primary input data for POA reconstruction

PARAMETER	SYMBOL	UNIT	SOURCE	DESCRIPTION
Measured energy production	$E_{measured}$	Wh/min	Reference house smart meter	Raw PV generation at 1-minute resolution
Ambient temperature	T_{air}	°C	PVGIS (satellite-derived)	Hourly air temperature, expanded to minutes
Wind speed	v_{wind}	m/s	PVGIS (satellite-derived)	Hourly wind speed, expanded to minutes
PVGIS POA irradiance	POA_{PVGIS}	W/m^2	PVGIS	Auxiliary irradiance (used only for thermal model, not as primary input)

Table 3.7 Fixed system parameters for inverse reconstruction

PARAMETER	**SYMBOL**	**VALUE**	**UNIT**	**DESCRIPTION**
Module area	A_{m2}	10	m^2	Total active area of the reference PV system
STC efficiency	η_{STC}	0,21	-	Module efficiency at Standard Test Conditions (25°C, 1000 W/m^2)
Temperature coefficient	γ	-0,0045	1/°C	Power temperature coefficient for crystalline silicon
Sandia coefficient (intercept)	a	-3,56	-	Empirical coefficient for wind cooling effect
Sandia coefficient (slope)	b	-0,075	-	Empirical coefficient for wind cooling effect
Maximum Wh per minute	$MAX_{Wh-per-min}$	100	Wh/min	Physical upper bound for production (equivalent to 6000 W)

Noise threshold	$RUMOR_{THRESHOLD}$	0,00167	Wh/min	Minimum detectable production (equivalent to 0.1 W)
Nighttime interval	$NIGHT_{START}, NIGHT_{END}$	23;4	Hour	Hours when production is forced to zero
PR range	PR_{min}, PR_{max}	0,3;1,0	-	Physically plausible Performance Ratio bounds
POA range	POA_{min}, POA_{max}	0;1200	W/m^2	Physically plausible irradiance bounds
Cell temperature bounds	$T_{cell,min}, T_{cell,max}$	-20;85	°C	Physically plausible cell temperature range
Smoothing window	W_{PR}	180	Min	Centered moving average window for PR stabilization

3.4.1.2 Core mathematical modeling

- **Step 1: Unit conversion from energy to power**

The raw measured data are provided as energy produced per minute ($E_{measured}$ in Wh/min). To obtain the corresponding average power in Watts (W), the following conversion is applied:

$$P_{measured} = E_{measured} * 60 \quad (3.38)$$

The factor 60 derives from the number of minutes in one hour: $P = \frac{E}{\Delta t}$ with $\Delta t = \frac{1}{60}$ hour, thus $P = E \times 60$. This conversion is essential because the physical model works with instantaneous power, not with accumulated energy.

- **Step 2: Data cleaning and preprocessing**

The measured power time series is filtered to remove non-physical or unreliable values:

- Nighttime forcing: For hours between 23:00 and 04:00, production is set to zero, as no solar generation is physically possible.
- Outlier removal: Any value exceeding $MAX\_Wh\_per\_min \times 60 = 6000$ W is flagged as an outlier and set to NaN.
- Spike detection: Isolated non-zero values surrounded by zeros (both preceding and following values equal to zero) are identified as measurement spikes and set to zero.
- Interpolation: Missing values (NaNs) resulting from outlier removal are filled using linear interpolation, limited to a maximum of 5 consecutive minutes to avoid introducing artificial trends.
- Noise thresholding: Values below $RUMORE_{THRESHOLD} \times 60 = 0.1$ W are considered instrumental noise and set to zero.
- Negativity correction: Any negative values resulting from previous operations are set to zero.

- **Step 3: Loading and expansion of PVGIS data**

Hourly meteorological data (POA irradiance, air temperature, wind speed) are downloaded from the PVGIS website for the location of interest (latitude 52.139°N, longitude 10.278°W). Since the PV production data are available at minute resolution, the hourly PVGIS data are expanded to one-minute resolution by assuming that all minutes within the same hour share the same constant value:

$$X\_min(t) = X_{hourly}(t_{ora}) \; \forall t \; [h, h+1) \quad (3.39)$$

where X represents any of the meteorological variables (POA_{PVGIS}, T_{amb}, v_{wind}).

- **Step 4: Cell temperature calculation**

The operating temperature of the PV cells is estimated using the Sandia empirical model, which accounts for the cooling effect of wind:

$$T_{cell} = T_{amb} + POA_{PVGIS} * e^{(A_{sandia} + B_{sandia} * v_{wind})} \quad (3.40)$$

where $a = -3.56$ and $b = -0.075$ are empirically derived coefficients. The result is clipped to the physically plausible range $[-20{,}85]°C$ to prevent numerical extremes:

$$T_{cell} = max\ (min(T_{cell}, 85)\ ; -20) \quad (3.41)$$

- **Step 5: Module efficiency calculation**

The efficiency of the PV module at operating conditions deviates from its STC value due to temperature effects. The linear temperature correction model is applied:

$$\eta = \eta_{STC} * [1 - \ \gamma * (T_{cell} - 25)] \quad (3.42)$$

where:

- $\eta_{\mathrm{STC}} = 0.21$ (21%)
- $\boldsymbol{\gamma}$ = -0.0045 1/°C
- T_{cell} is the cell temperature from Step 4

The resulting efficiency is further constrained to the range $[0.05, 0.25]$ to avoid physically implausible values arising from extreme temperature inputs.

- **Step 6: Theoretical power calculation**

The theoretical DC power that would be produced under the PVGIS irradiance, given the actual cell temperature and module efficiency, is computed as:

$$P_{theoretical} = POA_{PVGIS} \cdot A_{m2} \cdot \eta \quad (3.43)$$

where:

- $A = 10\ \mathrm{m}^2$ is the total module area

This represents the expected power output in the absence of any system losses (other than those implicitly captured by the efficiency model).

- **Step 7: Dynamic Performance Ratio calculation**

The raw Performance Ratio (PR_{raw}) is defined as the ratio of measured power to theoretical power:

$$PR_{raw} = \frac{P_{measured}}{P_{theoretical}} \tag{3.44}$$

This calculation is performed only for time steps where $P_{\text{theoretical}} > 0$ to avoid division by zero. The PR_{raw} values are then clipped to the physically meaningful interval $[0.3, 1.0]$, as values outside this range would imply unrealistic loss or gain mechanisms.

To reduce short-term variability and measurement noise, a smoothed Performance Ratio (PR_{din}) is obtained by applying a centered moving average:

$$PR_{din}(t) = \frac{1}{W_{PR}} \sum_{k=-W_{PR}/2}^{W_{PR}/2} P\,R_{raw}(t + k \cdot \Delta t) \tag{3.45}$$

where:

- $W_{PR} = 180$ minutes (3 hours) is the smoothing window
- $\Delta t = 1$ minute is the temporal resolution

This smoothing window is chosen to preserve the diurnal trend while eliminating high-frequency fluctuations.

- **Step 8: POA irradiance reconstruction (inversion)**

The fundamental step of the inverse procedure consists of inverting the photovoltaic power equation to estimate the actual POA irradiance that would have produced the measured power output, given the estimated efficiency and dynamic PR:

$$POA_{estimated} = \frac{P_{measured}}{A_{m2} \cdot \eta \cdot PR_{din}} \tag{3.46}$$

This equation is derived from:

$$P_{measured} = POA \cdot A \cdot \eta \cdot PR \tag{3.47}$$

solved for POA. The calculation is performed for all time steps; for steps where $P_{\text{measured}} = 0$, the resulting $POA_{\text{estimated}}$ is also set to zero.

The reconstructed irradiance is then subjected to physical constraints:

- Clipping to the range $[0,1200]$ W/m^2, as values above 1200 W/m² exceed the solar constant corrected for atmospheric effects.
- Nighttime forcing: for hours between 23:00 and 04:00, the irradiance is explicitly set to zero, regardless of the calculated value.

The output of the first step is a complete time series of reconstructed POA irradiance at minute resolution, which serves as the essential input for the second step.

3.4.2 Second Step: Calibration of PV system parameters

The second step uses the estimated POA (obtained from the first step) to calibrate the unknown parameters of each PV system and validate the model against real production data.

3.4.2.1 Model inputs and system configuration

The second script loads three primary data sources and defines the parameters for optimisation, summarised in Table 3.8 and in Table 3.9.

Table 3.8 Primary input data for calibration and validation

PARAMETER	SYMBOL	UNIT	SOURCE	DESCRIPTION
Reconstructed POA irradiance	$POA_{estimated}$	W/m^2	Output of Step 1	Site-specific irradiance from inverse reconstruction
Raw household meter data	*Various*	Wh	Excel file	Battery, grid, consumption, and production flows
Ambient temperature	T_{amb}	°C	PVGIS	Hourly air temperature for thermal modeling

Table 3.9 Fixed parameters for calibration

CATEGORY	PARAMETER	SYMBOL	VALUE	UNIT
Thermal model	NOCT	$NOCT$	45	°C
	Reference temperature	T_{NOCT}	20	°C
	Reference irradiance	G_{NOCT}	800	W/m^2
Operational	POA threshold	$POA_{threshold}$	10	W/m^2

	Assumed panel efficiency	η_panel	0,20	-
Optimisation bounds	Peak power	P_{peak}	[0,5;10]	kWp
	Performance Ratio	PR	[0,6;0,95]	-
	Temperature coefficient	$\boldsymbol{\gamma}$	[0,002;0,006]	1/°C
Initial guesses	Peak power	$P_{peak,0}$	2,0	kWp
	Performance Ratio	PR_0	0,80	-
	Temperature coefficient	γ_0	0,004	1/°C

Critical methodological note on production data: The raw Excel file contains a column labelled Production_Wh, which represents the net production measured by the smart meter. However, for the purpose of this analysis, this value is deliberately not assumed as known a priori. Instead, the analysis adopts a different approach: starting from the raw meter readings (discharge, charge, consumption, feed-in, grid import), the true PV generation is reconstructed via an energy balance that decouples the battery contribution. This choice is motivated by the desire to develop a methodology that is independent of the direct PV production reading, allowing the model to be applied even in contexts where such a direct measurement is unavailable or unreliable. The reconstructed production is then used as the target for the optimisation of the PV system parameters. Subsequently, the calibrated model is validated against the original Production_Wh column, providing an independent check of the methodology's accuracy.

3.4.2.2 Core mathematical modeling

The second script implements a sequence of operations divided into three distinct phases: (A) reconstruction of true PV production from meter data, (B) optimisation of system parameters using the reconstructed production as target, and (C) forward simulation and validation against the original production data.

- **Phase A: Reconstruction of true PV production from meter data**
  - **Step A.1: Meter data loading and preliminary corrections**
    The raw Excel file is loaded for the specific household. The dataset contains eight columns: timestamp, discharge energy (Wh), charge energy (Wh), PV production

(Wh), consumption (Wh), feed-in energy (Wh), energy from grid (Wh), and state of charge (%). A critical energy balance correction is applied to ensure consistency among the measured flows. For each minute, the net energy balance is computed as:

$$Balance = P_{prod} + P_{from\ grid} + P_{discahrge} - P_{consumption} - P_{feed\ in} - P_{charge} = 0 \quad (3.48)$$

if the absolute balance exceeds a tolerance of 0.01 Wh, the discrepancy is corrected by adjusting the feed-in and grid import values:

$$P_{feed\ in\ corr} = P_{feed\ in} - \frac{Balance}{2}\ ;\ P_{from\ grid\ corr} = P_{from\ grid} + \frac{Balance}{2} \quad (3.49)$$

The corrected values are then clipped to non-negative values.

- **Step A.2: Daily battery offset correction**
  A daily correction is applied to ensure that the net energy variation of the battery over each day is consistent with the charge and discharge flows. For each day, the net battery energy is:

$$E_{batt,net} = \sum(P_{charge} - P_{discharge}) \quad (3.50)$$

This net energy is distributed uniformly across all minutes of that day as a correction offset:

$$offset = \frac{E_{batt,net}}{N_{day}} \quad (3.51)$$

The feed-in and grid import are then adjusted accordingly:

$$P_{feed\_in,corr} = P_{feed\_in} - \frac{offset}{2}\ ;\ P_{from\_grid,corr} = P_{from\_grid} + \frac{offset}{2} \quad (3.52)$$

- **Step A.3: Disentangling true PV production from battery flows**
  The true PV generation is reconstructed by performing an energy balance around the entire household system. The fundamental principle is that the power generated by the PV panels must equal the sum of all power flows leaving the generation point:

$$P_{gen} = P_{feed\_in} + P_{charge} + P_{consumption} - P_{from\_grid} - P_{discharge} \quad (3.53)$$

Negative values resulting from this calculation are clipped to zero:

$$P_{gen} = max\ (P_{gen}, 0) \tag{3.54}$$

- **Step A.4: Standby offset removal**
  A standby offset is computed as the median of nighttime production values, where the POA is below 10 W/m²:

$$offset_{standby} = median(P_{gen} \mid POA < 10) \tag{3.55}$$

  This offset represents systematic measurement errors (sensor drifts, calibration issues) and is subtracted from the entire production series:

$$P_{gen,clean} = max\ (P_{gen} - offset_{standby}, 0) \tag{3.56}$$

  The cleaned production is then converted from Wh/min to Watts:

$$P_{reconstructed}(W) = P_{gen,clean}(Wh/min) \times 60 \tag{3.57}$$

  This reconstructed production time series ($P_{\text{ricostruita}}$) serves as the ground truth target for the optimisation phase.

- **Step A.5: Merging with POA and temperature data**
  The estimated POA from the first step is loaded and merged with the household data on the timestamp. Ambient temperature data from PVGIS are also loaded, expanded to minute resolution, and merged. Missing temperature values are interpolated linearly and then forward- and backward-filled.

- **Phase B: Optimisation of PV system parameters**

  - **Step B.1: Physical model definition**

    The physical model used in the second step is similar to that used in the first step, with the crucial difference that now the POA is known (estimated from Step 1) and the parameters P_{peak}, PR, and γ are unknown and will be optimised.

    - **Cell temperature model (NOCT Model):** The cell temperature is calculated using the simplified NOCT (Nominal Operating Cell Temperature) model, which is widely used for standard crystalline silicon modules:

$$T_{cell} = T_{amb} + \frac{NOCT - T_{NOCT}}{G_{NOCT}} \tag{3.58}$$

      where:

      NOCT = 45 °C

$T_{NOCT} = 20$ °C

$G_{NOCT} = 800$ W/m²

- **Temperature correction factor:** The temperature correction factor accounts for the reduction in efficiency as cell temperature increases above STC (25°C):

$$f_{temp} = 1 - \gamma \cdot (T_{cell} - 25) \tag{3.59}$$

where γ is the temperature coefficient to be optimised. The correction factor is clipped to the range $[0.70, 1.05]$ to avoid physically implausible values.

- **Power production model**: The AC power produced by the PV system is calculated as:

$$P_{modello}(W) = P_{peak} \cdot \frac{POA}{1000} \cdot f_{temp} \cdot PR \cdot 1000 \tag{3.60}$$

which simplifies to:

$$P_{modello}(W) = P_{peak} \cdot POA \cdot f_{temp} \cdot PR \tag{3.61}$$

where:

- Ppeak is the nominal peak power in kWp
- POA is the plane-of-array irradiance in W/m²
- f_{temp} is the dimensionless temperature correction factor
- PR is the dimensionless Performance Ratio

The resulting power is constrained to non-negative values:

$$P_{modello} = max\,(P_{modello}, 0) \tag{3.62}$$

- **Step B.2: Objective function definition**
The three unknown parameters (P_{peak}, PR, γ) are estimated by minimising the Root Mean Square Error (RMSE) between the model output and the reconstructed real production ($P_{\text{ricostruita}}$ from Phase A). The objective function is:

$$RMSE(\theta) = \sqrt{\frac{1}{N_{valid}} \sum_{i \in \Omega_{valid}} \left(P_{model,i}(\theta) - P_{reconstructed,i}\right)^2} \tag{3.63}$$

where:

- $\theta = (P_{\text{peak}}, PR, \gamma)$ is the parameter vector
- Ω_{valid} is the set of time steps satisfying:

- $POA > POA_{\text{soglia}}$ (where $POA_{\text{soglia}} = 10$ W/m^2)
- $P_{\text{ricostruita}} > 0$

These restrictions ensure that the optimisation focuses on physically meaningful conditions where production is actually occurring.

- **Step B.3: Optimisation algorithm**

  The minimisation is performed using the L-BFGS-B algorithm (Limited-memory Broyden-Fletcher-Goldfarb-Shanno with box constraints), which is well suited for bound-constrained optimisation problems. The algorithm iteratively updates the parameter estimates to minimise the RMSE, using gradient information approximated from past iterations.

  The bounds are set as follows:

  $P_{\text{peak}} \in [0.5, 10.0]$ kWp

  $PR \in [0.6, 0.95]$

  $\gamma \in [0.002, 0.006]$ (equivalent to 0.20–0.60%/°C)

  The initial guess is:

$$P_{\text{peak}} = 2.0\ \text{kWp}, PR = 0.80, \gamma = 0.004$$

- **Step B.4: Convergence check**

If the optimisation converges successfully (result.success = True), the optimal parameters are stored as:

$$P_{\text{peak,opt}}, PR_{\text{opt}}, \gamma_{\text{opt}}$$

If convergence is not achieved, the initial guesses are retained as fallback values.

- **Step B.5: Derived parameters**

  Once the optimal parameters are found, the following derived quantities are computed:

  - **Estimated module area:**

$$A_{estimated} = \frac{P_{peak,opt}}{\eta_{pannel}} \tag{3.64}$$

  where $\eta_{\text{pannello}} = 0.20$ is the assumed panel efficiency.

  - **System efficiency:**

$$\eta_{system} = PR_{opt} \cdot \eta_{pannel} \tag{3.65}$$

- **Phase C: Forward simulation and validation against original production data**

  - **Step C1: Full forward simulation**

Using the optimal parameters obtained from Phase B, the model computes the simulated production for the entire time series. The same physical equations are applied:

$$T_{cell}(t) = T_{amb}(t) + \frac{NOCT - T_{NOCT}}{G_{NOCT}} \cdot POA(t) \quad (3.66)$$

$$f_{temp}(t) = 1 - \gamma_{opt} \cdot (T_{cell}(t) - 25) \quad (3.67)$$

$$f_{temp}(t) = max\left(min\left(f_{temp}(t), 1.05\right), 0.70\right) \quad (3.68)$$

$$P_{model}(t) = P_{peak,\ opt} \cdot POA(t) \cdot f_{temp}(t) \cdot PR_{opt} \quad (3.69)$$

$$P_{model}(t) = max\ (P_{model}(t), 0) \quad (3.70)$$

The simulated power is then converted to Wh/min for comparison with the original data format:

$$P_{model, Wh/min} = \frac{P_{model}}{60} \quad (3.71)$$

- **Step C.2: Validation against original production data**

The model output is validated against the original production data from the dataset ($P_{\text{originale}}$), i.e., the Production_Wh column from the Excel file (colonna D), converted to Watts:

$$P_{original}(W) = P_{original} \times 60 \quad (3.72)$$

This is a critical methodological choice: the optimisation uses the reconstructed production ($P_{\text{ricostruita}}$) as the target to find the optimal parameters, but the final validation is performed against the original measured production ($P_{\text{originale}}$). This two-step approach allows an independent assessment of the model's ability to reproduce the actual meter readings.
The comparison is performed over the valid time steps where POA > 10W/m^2 and $P_{\text{originale}} > 0$.

The following error metrics are computed:

- **Root Mean Square Error (RMSE):**

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(P_{model,i} - P_{original,i}\right)^2} \tag{3.73}$$

- **Mean Absolute Error (MAE):**

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | P_{model,i} - P_{original,i} | \tag{3.74}$$

- **Mean Bias Error (Bias):**

$$Bias = \frac{1}{n}\sum_{i=1}^{n}\left(P_{model,i} - P_{original,i}\right) \tag{3.75}$$

- **Correlation Coefficient (r):**

$$r = \frac{\sum_{i=1}^{n}\left(P_{original,i} - \bar{P}_{original}\right)\left(P_{model,i} - \bar{P}_{model}\right)}{\sqrt{\sum_{i=1}^{n}\left(P_{original,i} - \bar{P}_{original}\right)^2} \cdot \sqrt{\sum_{i=1}^{n}\left(P_{model,i} - \bar{P}_{model}\right)^2}} \tag{3.76}$$

- **Coefficient of Determination (R^2):**

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(P_{original,i} - P_{model,i}\right)^2}{\sum_{i=1}^{n}\left(P_{original,i} - \bar{P}_{original}\right)^2} \tag{3.77}$$

○ **Step C.3: Total annual energy comparison**

The total annual energy produced according to the original data and by the calibrated model are compared:

$$E_{original} = \frac{\sum P_{original}}{1000} \tag{3.78}$$

$$E_{model} = \frac{\sum P_{model}}{1000} \tag{3.79}$$

$$\Delta E = E_{modello} - E_{originale} \tag{3.80}$$

3.5 Comparative methodology for baseline tariffs and renew dynamic optimization

This study consists of two distinct yet complementary phases. The first phase involves an ex-post evaluation of standard electricity tariffs available on the Irish market, establishing a baseline for comparison. The second phase assesses the impact of a dynamic tariff (Renew) combined with battery storage systems employing advanced predictive control logics. The entire analysis is replicated across twenty household case studies, six of which are selected for in-depth investigation: three equipped with photovoltaic systems (H1, H3, H10) and three without distributed generation (H8, H18, H19). For each case, a high-resolution historical dataset is available, recording minute-by-minute energy flows over a full annual cycle.

3.5.1 Model inputs and system configuration

The modelling architecture integrates four principal data sources, temporally aligned to ensure consistency and replicability of results.

3.5.1.1 Households load and generation profiles

For each selected case study, the dataset provides a complete annual record of electricity flows at one-minute resolution. The recorded variables include photovoltaic production, household consumption, energy drawn from the grid, energy injected into the grid, battery charging and discharging power, and the battery's state of charge. For households without photovoltaics, the production and feed-in values are naturally zero throughout the entire year.

3.5.1.2 Standard electricity tariffs (Baseline)

The baseline analysis employs a comprehensive database provided by Maynooth University [24] of all electricity tariffs currently offered in Ireland. For each tariff, the database specifies the provider, the tariff name, the applicable meter type, the annual standing charge differentiated by urban or rural location, any welcome bonus, and the presence or absence of a feed-in credit mechanism. Crucially, each tariff is accompanied by a detailed schedule of pricing intervals. These intervals define the purchase price of electricity as a function of the time of day and the type of day (weekday, weekend, or all days). Some tariffs also specify a feed-in price for surplus energy injected into the grid. The dataset includes both flat-rate tariffs (single price for all hours) and time-of-use tariffs with multiple daily bands, including night storage rates and weekend special offers. All tariffs are validated prior to inclusion: those with overlapping pricing intervals or with zero or null standing charges are automatically excluded from the analysis. This filtering ensures that only commercially valid and practically applicable tariffs are considered.

3.5.1.3 Dynamic tariff and day-ahead market prices

The Renew tariff, by contrast, is not based on fixed pricing intervals but on the hourly prices of the Irish Day-Ahead Market (DAM). The DAM dataset [25], provided by Maynooth University too, covers a full year at half-hourly resolution and includes the market clearing price for each interval, expressed in euros per megawatt-hour. Additional metadata indicates whether each day is a weekday, weekend, or public holiday. For the Renew tariff, the purchase price paid by the household at any given time is the DAM price plus applicable network charges, while the feed-in credit for surplus energy injected into the grid is calculated at the pure DAM price without additional levies. The methodology to follow for the Renew tariff calculation starting from DAM data was supplied by the University of Maynooth. This asymmetry incentivizes self-consumption and strategic battery operation.

3.5.1.4 Fixed charges and network components

For both baseline tariffs and the Renew tariff, the total annual cost includes fixed components that are independent of consumption. These comprise the standing charge (specific to each tariff and user location), the Public Service Obligation levy (a universal charge applied to all Irish electricity customers), and the ESB Networks distribution tariff, which differs between urban and rural connections. Value Added Tax is applied at the standard Irish rate of 9 percent on the subtotal of all energy costs and fixed charges after deducting any feed-in credits.

3.5.1.5 Battery storage system configuration

For the optimization phase, a grid-connected lithium-ion battery system is modelled with realistic technical parameters. Three nominal capacities are tested: 5 kWh, 10 kWh, and 15 kWh. The battery has a maximum charge and discharge power of 3.3 kW, which limits the energy that can be transferred in any half-hour interval. Round-trip efficiency is set at 90 percent. The initial state of charge is configured differently depending on the presence of photovoltaics: for solar-equipped households, the battery begins at 50 percent capacity; for non-PV households engaged in pure price arbitrage, the battery starts empty. Two price thresholds govern the optimization logic: the battery charges when the forecasted price falls below 0.08 €/kWh and discharges when the forecasted price rises above 0.14 €/kWh.

3.5.2 Core mathematical modeling

This section formalises the mathematical framework used to compute total annual costs and to optimise battery scheduling under the dynamic Renew tariff.

3.5.2.1 Baseline tariff cost calculation

For each standard tariff, the total annual cost is computed as the sum of four components:

$$C_{tot}^{std} = C_{fix} + C_{var} - C_{feedin} - Bonus \quad (3.81)$$

where:

- C_{fix} is the annual standing charge (€/year), which depends solely on the user's urban or rural location. For the Renew tariff, this term also includes the fixed Renew fee, the PSO Levy, and the ESB distribution tariff.

- C_{var} is the variable cost for energy drawn from the grid, calculated minute by minute as:

$$C_{var} = \sum_{t=1}^{T} E_{grid}(t) * P_{tariff}(t) \quad (3.82)$$

Where P_{feedin} is the specific feed-in rate of the tariff (if such a mechanism exists).

- $Bonus$ is any welocme bonus offered by the tariff provider

The determination of $P_{tariff}(t)$ at each minute is handled by a temporal membership algorithm that identifies which pricing interval contains the given time. Intervals that cross midnight are correctly managed by normalising time on a 24-hour basis.

3.5.2.2 Renew tariff optimization with battery storage

The Renew tariff optimization employs a predictive control strategy that simulates a real battery system without any look-ahead bias. At each decision step, the algorithm accesses only information realistically available: historical data and forecasts for the next 4 hours.

For households equipped with photovoltaics, the net surplus at each half-hour interval is defined as:

$$S(t) = P_{PV}(t) - C_{load}(t) \quad (3.83)$$

Where $P_{PV}(t)$ is the photovoltaic production and $C_{load}(t)$ is the household consuption. The battery control logic follows these rules:

- If $S(t) > 0$ (surplus): the battery charges up to its remaining capacity, limited by the maximum power constraint. Any excess surplus beyond the battery's charging capability is exported to the grid, generating a feed-in credit at the DAM price.

- If $S(t) < 0$ (deficit) and the forecasted purchase price exceeds the discharge threshold (0.14 €/kWh), the battery discharges to cover as much of the deficit as possible, limited by its available energy and maximum power. Any remaining deficit is met by grid imports at the DAM price plus network charges.

For households without photovoltaics, the logic simplifies to pure price arbitrage. The battery charges when the forecasted price falls below 0.08 €/kWh and discharges when the price exceeds 0.14 €/kWh, regardless of the household's concurrent consumption. The discharged energy first serves the household's load, reducing grid imports; any surplus beyond immediate needs is exported to the grid, earning revenue at the high DAM price. This strategy effectively uses the battery as a financial arbitrage device.

3.5.2.3 Forecast models

To enable realistic decision-making without look-ahead bias, two forecasting models are implemented using the XGBoost machine learning algorithm. Both models operate on a rolling training window of the previous 30 days and produce predictions for the next 4 hours at half-hourly steps.

- Solar radiation forecast: The model predicts solar radiation $R(t+1)$ [W/m²] using temporal features (hour of the day, day of the year, transformed via sine and cosine to capture cyclical patterns) and lagged radiation values. From the predicted radiation, photovoltaic production is derived as:

$$P_{PV}^{pred}(t+1) = \eta * A * R_{pred}(t+1) * 0{,}5 * 10^{-3} \tag{3.84}$$

where η =0.18 is the panel efficiency, A=16.7 m^2 is the estimated panel area for a typical 3 kWp system, and the factor 0.5 converts from power to energy over a half-hour interval.

- Consumption forecast: The model predicts household consumption $C_{load}^{pred}(t+1)$ using temporal features (hour, day of week, month, weekend indicators) and lagged consumption values extending back up to 24 hours. This allows the model to capture daily and weekly patterns as well as autocorrelation structures typical of residential load profiles. To ensure statistical reliability, both forecasting models undergo a warm-up period of seven days. During this initial phase, when insufficient training data is available, forecasts are approximated using the naive persistence method, which assumes that the next value will equal the current value:

$$\hat{X}(t+1) = X(t) \tag{3.85}$$

After the warm-up period, the XGBoost models are retrained at each time step, continuously adapting to recent patterns.

3.5.2.4 Temporal alignment and aggregation

While the original load data is recorded at one-minute resolution, the DAM prices and most tariff intervals are defined at half-hourly or hourly resolution. To align these different temporal scales, the minute-by-minute data is aggregated into half-hourly intervals. For each half-hour, the total energy flows are summed and converted from watt-hours to kilowatt-hours. The timestamp is then rounded down to the nearest half-hour boundary and localised to the Dublin time zone. This aggregation preserves the essential dynamics of household energy use while maintaining compatibility with market price signals.

3.5.3 Workflow and methodological comparison

This section describes the battery decision-making hierarchy and the multi-level comparative structure designed to ensure consistent and interpretable results across diverse case studies.

3.5.3.1 Battery decision hierarchy

At each half-hour step, the battery control algorithm evaluates a sequence of conditions. First, it examines the current and forecasted DAM prices, comparing them against the charge and discharge thresholds. Second, for photovoltaic households, it computes the expected net surplus based on the forecasted production and consumption. Third, it checks the current state of charge and the remaining energy capacity. Fourth, for pure arbitrage households, it estimates the potential future value of discharging during upcoming price peaks, using this estimate to avoid charging when there is insufficient profitable discharge opportunity ahead. This hierarchical logic ensures that the battery operates efficiently: it does not charge when prices are moderately low if no significant price peaks are expected in the following hours, and it does not discharge when prices are moderately high if the state of charge is low and future peaks may be even higher. The analysis is organised into three distinct comparative levels, each designed to isolate different effects and answer specific research questions.

3.5.3.2 Intra-household comparison

At the most granular level, each individual household is analysed across multiple scenarios. Specifically, for every household in both categories, the total annual cost is computed under four distinct configurations: the baseline scenario without battery storage, and three optimisation scenarios with battery capacities of 5 kWh, 10 kWh, and 15 kWh. This intra-household comparison allows quantification of the marginal benefit of adding storage capacity to a specific household, holding constant all other variables such as load profile, production pattern (if any), and location. For example, comparing the 5 kWh scenario against the baseline reveals the benefit of installing a small battery, while comparing the 10 kWh against the 5 kWh reveals the incremental value of doubling the capacity for the same household.

3.5.3.3 Intra-category comparison

At the intermediate level, households belonging to the same technological category are compared against one another. The three photovoltaic-equipped households (H1, H3, H10) are analysed together using identical model configurations, allowing the identification of how differences in load profiles and production patterns affect the performance of the Renew tariff and the battery optimisation. For instance, a household with a favourable alignment between solar production and consumption may achieve higher self-consumption rates and greater savings compared to another with less favourable timing. Similarly, the three non-photovoltaic households (H8, H18, H19) are compared among themselves to assess how pure consumption patterns influence the profitability of price arbitrage. All models are applied with identical parameters across households within each category, ensuring that any observed differences are attributable solely to the intrinsic characteristics of each household rather than to variations in model configuration.

3.5.3.4 Inter-category comparison

At the highest level, the analysis compares the aggregated results of the photovoltaic category against those of the non-photovoltaic category. This comparison serves to evaluate the synergistic role of distributed generation when combined with battery storage and dynamic pricing. For each category, summary statistics such as average total annual cost, average savings relative to baseline, and average battery utilisation metrics are computed across the three representative households. The inter-category comparison reveals whether households with photovoltaics benefit more or less from storage than those without, and whether the optimal battery capacity differs between the two categories. This level of analysis addresses the broader policy question of whether storage should be incentivised primarily for solar-equipped households or for all consumers regardless of generation status. All three comparative levels are implemented using consistent metrics and visualisation techniques, including cost tables, percentage savings, battery cycle counts, and state-of-charge profiles, ensuring that the results are directly comparable across households, categories, and scenarios.

4. RESULTS

4.1 PVGIS model results

This chapter presents the results obtained from the application of the model described in Chapter 3. The validation was conducted by comparing the hourly, daily, and monthly production estimates of the model with the reference data provided by PVGIS, based on the Typical Meteorological Year (TMY) constructed over the period 2005-2020. The chapter structure follows the model architecture: first, the configuration parameters of the simulated system are reported, then the energy production is analyzed on monthly, hourly, and seasonal scales, and finally the statistical tests and residual analysis are presented according to the validation protocol defined in the methodology. The model was configured to simulate a 2.17 kWp photovoltaic system, composed of 5 panels with individual peak power of 435 Wp and nominal efficiency of 21.8%. The panels, have a total active area of 9.99 m², are south-facing (azimuth 180°) with a tilt angle of 30° relative to the horizontal, an optimal value for the installation site latitude (Maynooth, Ireland, 53.383° N). The reference Performance Ratio was set to 0.94, while the inverter efficiency is 0.96. These parameters, consistent with the selected panel datasheet, form the basis for all subsequent simulations.

1. The analysis is structured into three main phases:
2. Validation on monthly and annual scales
3. Error metrics on annual, monthly and hourly scales
4. Residual analysis

4.1.1 Validation on monthly and annual scales

The processing of meteorological data downloaded from PVGIS for the period 2005-2020 yielded a Typical Meteorological Year characterized by an annual average plane-of-array (POA) irradiance of 132.4 W/m² and an average ambient temperature of 9.7°C. These values constitute the climatic context in which the model operated. The application of the Sandia thermal model produced an average cell temperature of 13.5°C, 3.8°C higher than the ambient temperature, an effect mainly due to the incident irradiance. This temperature corresponds to an average panel efficiency of 22.66%, slightly higher than the nominal STC value (21.8%) due to the relatively low average temperatures of the Irish site. The overall system efficiency, which also incorporates the dynamic Performance Ratio and inverter efficiency, stands at 18.31%. In the following Table 4.1 the monthly production and operating indicators are illustrated.

Table 4.1 Monthly production report and operating indicators

MONTH	MODEL (kWh)	PVGIS (kWh)	DIFF (kWh)	DIFF %	T_AMB (°C)	T_CELL (°C)	EFF_PAN (%)	EFF_SYS (%)
January	70,10	68,90	1,21	1,74	5,0	6,3	23,18	18,54
February	111,7	109,5	2,14	1,94	4,9	7,2	23,12	18,57
March	174,9	172,1	2,70	1,58	6,1	9,7	22,94	18,52
April	242,6	240,1	2,52	1,05	8,5	13,6	22,64	18,42
May	281,7	281,4	0,29	0,11	11,3	17,3	22,37	18,25
June	266,4	268,7	-2,31	-0,87	14,0	19,9	22,17	18,09

July	260,8	264,5	-3,65	-1,38	15,5	21,2	22,08	17,99
August	235,7	238,5	-2,82	-1,19	14,9	19,9	22,18	18.02
September	190,7	191,9	-1,20	-0,62	13,1	17,3	22,37	18,11
October	136,3	136,2	0,12	0,09	10,5	13,4	22,66	18,24
November	87,00	86,00	1,04	1,22	6,90	8,70	23,01	18,43
December	61,60	60,60	1,00	1,63	5,30	6,50	23,17	18,51
TOTAL	2119,4	2118,3	1,05	0,05	-	-	-	-

The total annual production estimated by the model is 2119.4 kWh, compared to 2118.3 kWh from PVGIS, with a percentage difference of 0.05%, demonstrating excellent accuracy on the annual scale. Monthly errors are contained within ±2%, with positive values (overestimation) in winter months and negative values (underestimation) in summer months. This trend is consistent with the cell temperature modeling: in cold months, panel efficiency is higher (23.18% in January), while in warm months it decreases (22.08% in July) due to the negative temperature coefficient β = -0.0034 1/°C. In the following Figure 4.1 the daily production comparison by month is shown

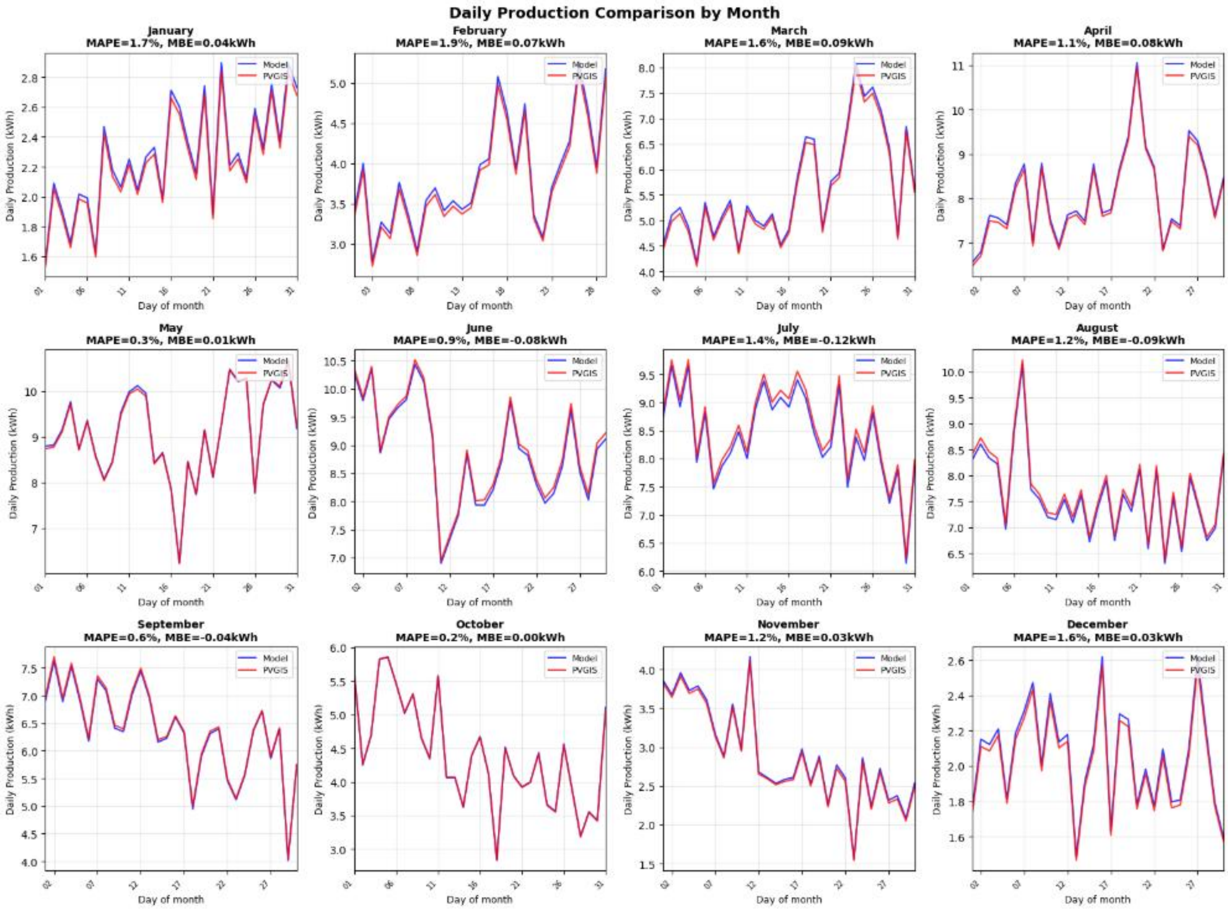


Figure 4.1 Daily production comparison by month

4.1.2 Error metrics on annual, monthly and hourly scales

The accuracy metrics calculated on monthly totals and aggregated hourly values confirm the high predictive quality of the model (Figure 4.1).

On the annual scale, the following values are obtained:

- MAPE (Mean Absolute Percentage Error): 1.146%
- MBE (Mean Bias Error): 0.0029 kWh (practically zero)
- RMSE (Root Mean Square Error): 0.0697 kWh
- MAE (Mean Absolute Error): 0.0600 kWh
- NRMSE (Normalized RMSE): 1.202%
- R^2 (Coefficient of determination): 0.9993

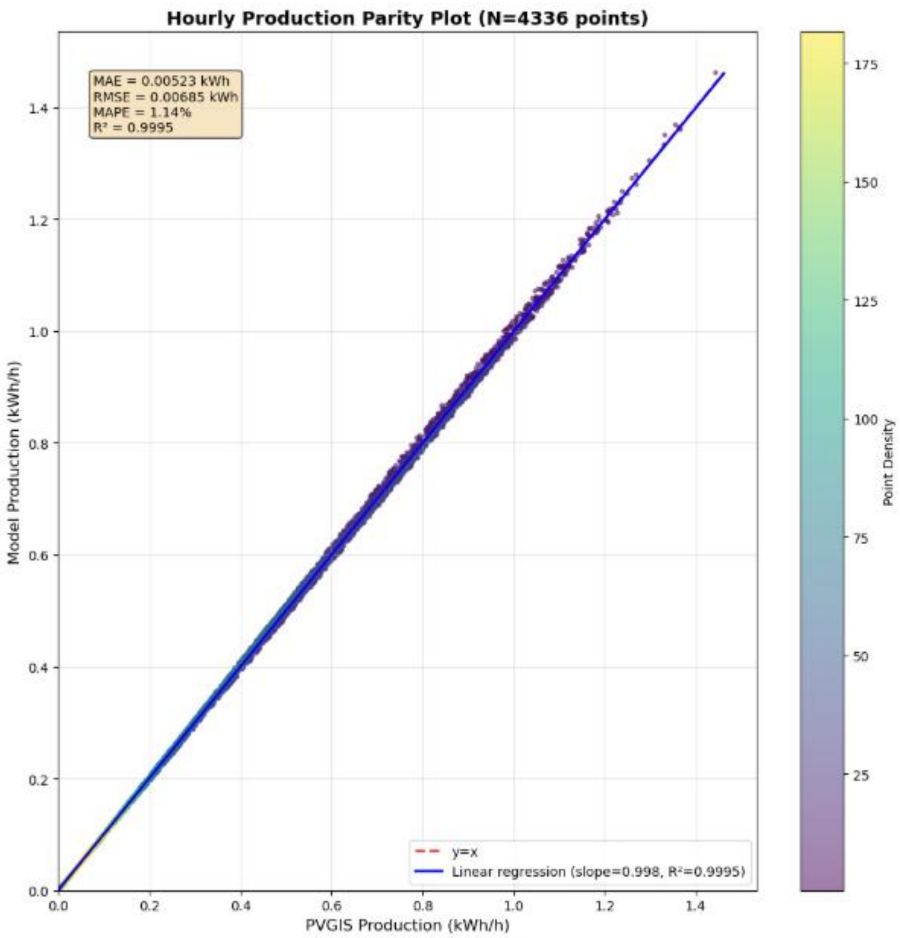


Figure 4.2 Hourly production parity plot

On the monthly scale (Figure 4.2), MAPE ranges from a minimum of 0.25% (October) to a maximum of 1.94% (February). The monthly MBE is always less than 0.12 kWh in absolute value, and RMSE never exceeds 0.12 kWh. The following Table 4.2 shows the monthly error metric.

Table 4.2 Monthly error metrics

MONTH	MAPE %	MBE [kWh]	RMSE [kWh]	MAE [kWh]
January	1.73	0.039	0.040	0.039
February	1.94	0.073	0.075	0.073
March	1.57	0.087	0.090	0.087
April	1.05	0.083	0.086	0.083
May	0.33	0.009	0.037	0.030
June	0.87	-0.076	0.079	0.076
July	1.38	-0.117	0.119	0.117
August	1.18	-0.091	0.092	0.091
September	0.61	-0.039	0.043	0.040
October	0.24	0.003	0.014	0.010
November	1.22	0.034	0.036	0.034
December	1.62	0.033	0.034	0.033

The following Figure 4.3 shows the trend of MAPE and MBE.

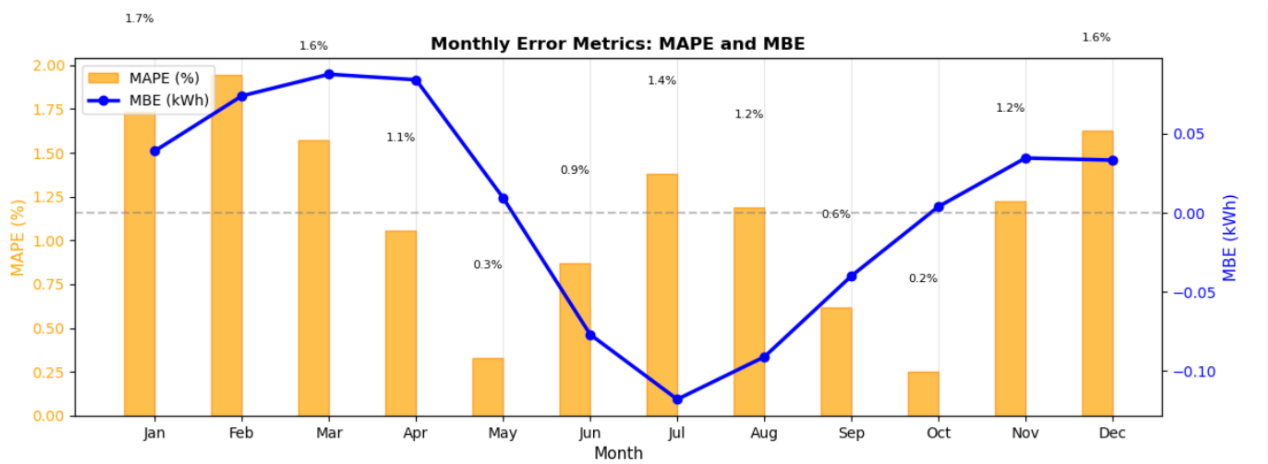


Figure 4.3 Monthly error metrics: MAPE and MBE

Hourly analysis represents the most stringent validation level, as it evaluates the model's ability to reproduce the intra-day dynamics of photovoltaic production. A total of 4305 hours were analyzed (excluding nighttime hours with zero production) and in the following Table 4.3 the error metrics on hourly scale are reported

Table 4.3 Error metrics on hourly scale

METRIC	VALUE
MBE (Mean Bias Error)	0.00024 kWh
MAE (Mean Absolute Error)	0.00523 kWh
RMSE (Root Mean Square Error)	0.00685 kWh
MAPE (Mean Absolute Percentage Error)	1.14%
MdAPE (Median Absolute Percentage Error)	1.10%
NRMSE (Normalized RMSE)	1.40%
R^2	0.9995

The results show a negligible mean hourly bias (0.00024 kWh) and a mean percentage error of 1.14%. The error distribution is extremely concentrated: 44.5% of the analyzed hours exhibit an error below 1%, 87.9% less than 2%, and 100% of hours have an error less than 3%, far exceeding the predefined validation thresholds. The Figure 4.4 Distribution of hourly errors illustrates the distribution of hourly errors.

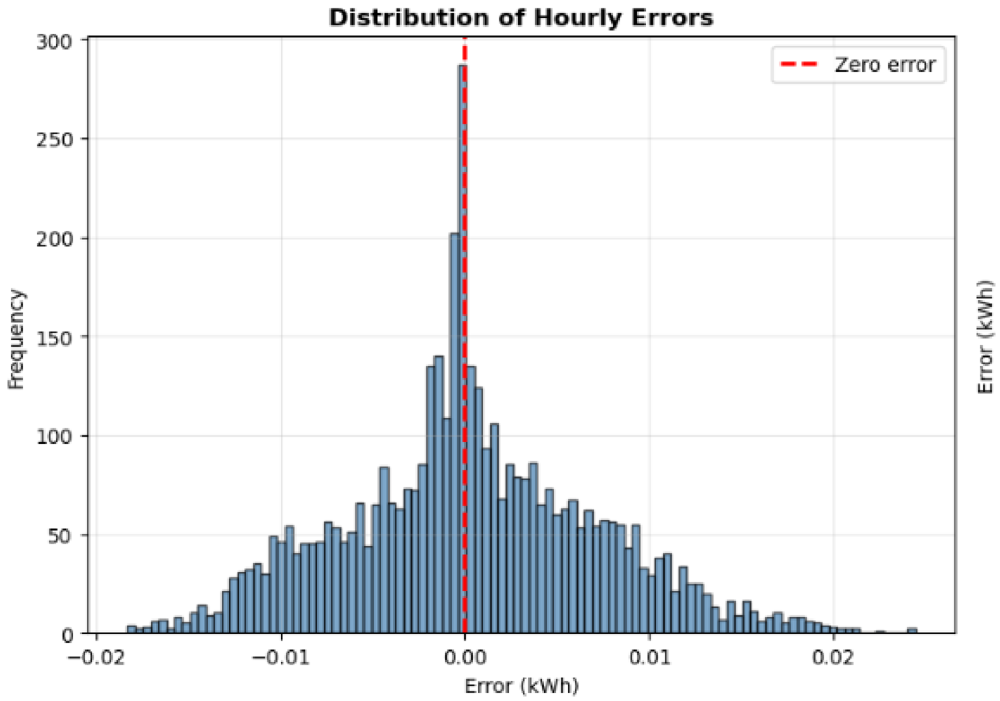


Figure 4.4 Distribution of hourly errors

4.1.3 Statistical tests for model vs. PVGIS comparison

To rigorously verify the similarity between the production distributions estimated by the model and the reference, the statistical tests described in the methodology were applied. The Table 4.4 represents the results of statistical tests on hourly scale.

Table 4.4 Results of statistical tests on hourly scale

TEST	STATISTIC	P-VALUE	INTERPRETATION
Wilcoxon signed-rank	4,546,958.0	0.284	H0 not rejected: no statistically significant difference between model and PVGIS
Spearman correlation	0.999807	< 0.001	Near perfect correlation
Student t-test (diff=0)	2.33	0.0197	Reject H0 (mean difference non-zero but negligible)
Kolmogorov-Smirnov	0.0056	1.000	Same distribution
Shapiro-Wilk (normality of errors)	0.9931	< 0.001	Non-normal errors (acceptable for large n)

The extremely high Spearman correlation coefficient ($\rho = 0.9998$) indicates an almost perfect relationship between model and reference. The Wilcoxon test and the Kolmogorov-Smirnov test confirm the null hypothesis of equal distributions ($p > 0.05$). The Student's t-test indicates a statistically significant difference between the means, but the corresponding bias is only 0.00024 kWh and is therefore negligible for an engineering perspective. Regarding the Shapiro-Wilk test, the non-normality of residuals is expected given the large sample size (n=4305) and does not invalidate the model.

4.1.4 Residual analysis

In-depth residual analysis, defined as the difference between estimated and reference production, was conducted according to the methodology, broken down into three dimensions: hour of the day, month of the year, and production magnitude. In the following Table 4.5 the percentage error by hour of the day is reported.

Table 4.5 Percentage error by hour of the day

HOUR	MEAN (%)	MEDIAN (%)	STD DEV (%)	Q1 (%)	Q3 (%)
5	-0,64	-0,79	0,68	-1,22	-0,09
6	-0,40	0,76	0,87	-1,12	0,39
7	0,04	-0,32	0,95	-0,83	0,82
8	0,34	0,19	1,01	-0,57	1,29
9	0,72	0,91	1,07	-0,29	1,61
10	0,76	0,94	1,18	-0,36	1,78
11	0,73	0,94	1,23	-0,45	1,84
12	0,58	0,73	1,22	-0,57	1,68
13	0,41	0,57	1,23	-0,79	1,53
14	0,18	0,33	1,19	-0,98	1,30
15	0,00	0,19	1,15	-1,12	1,07
16	-0,44	-0,53	1,13	-1,51	0,67
17	-0,97	-1,26	1,07	-1,91	0,11
18	-1,46	-1,82	0,95	-2,23	-0,68
19	-1,80	-2,14	0,76	-2,39	-1,11
20	-2,06	-2,15	0,45	-2,37	-1,75

A systematic pattern is observed: the model tends to slightly overestimate production during the central hours of the day (from 8 to 14, positive error up to +0.76%) and to underestimate it in the early morning hours (5-7) and late afternoon (16-20), with a maximum negative deviation of -2.06% at 20:00 (Figure 4.5). This pattern is attributable to the simplified wind speed modeling, assumed constant by seasonal band, while in reality wind can vary significantly during the day, affecting panel cooling. However, the maximum hourly bias magnitude (2.06%) is perfectly acceptable for engineering applications.

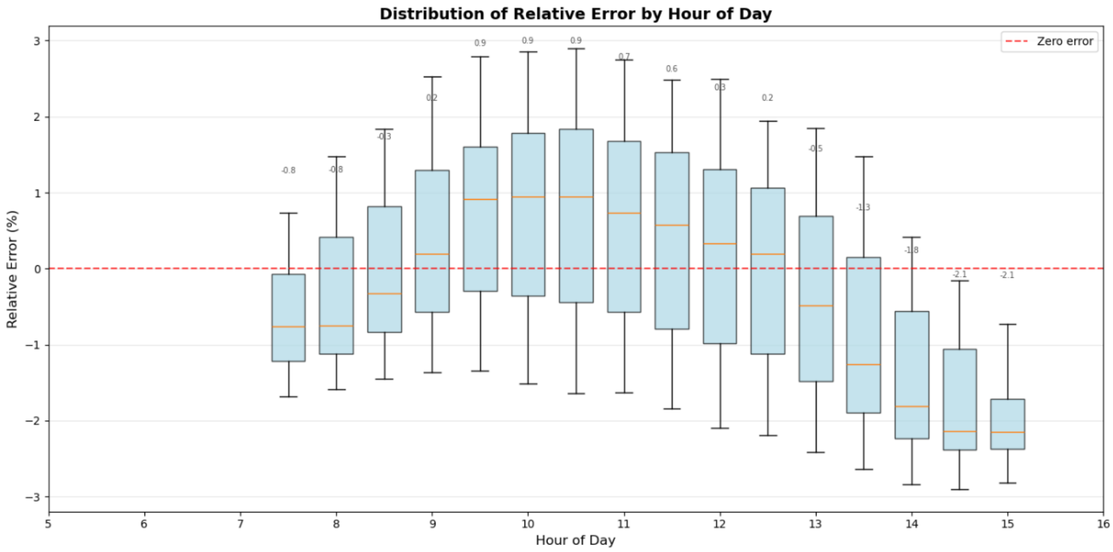


Figure 4.5 Distribution of relative error by hour of day

The monthly error shows a clear seasonal dependence: overestimation in winter months (up to +1.80% in February) and underestimation in summer months (up to -1.55% in July). This trend reflects the model's sensitivity to cell temperature and seasonal wind modeling. In summer months, actual wind may be stronger than the assumed value (0.8 m/s), resulting in greater cooling and therefore higher efficiencies than simulated; the opposite occurs in winter. In the following Table 4.6 the percentage error by hour month is reported.

Table 4.6 Percentage error by month

MONTH	MEAN (%)	MEDIAN (%)	STD DEV (%)
January	1,61	1,65	0,42
February	1,80	1,86	0,53
March	1,43	1,50	0,61
April	0,85	0,98	0,70
May	-0,09	0,04	0,73
June	-1,04	-0,96	0,64
July	-1,55	-1,40	0,60
August	-1,35	-1,20	0,58
September	-0,77	-0,68	0,58
October	-0,05	-0,04	0,53
November	1,04	1,12	0,56
December	1,54	1,54	0,36

Nighttime hours have an extremely low mean absolute error (0.00006 kWh) and a slightly higher MAPE (1.54%) due to zero or near-zero production, which amplifies percentage errors even for negligible absolute deviations. The following Table 4.7 shows a comparison between day and night metrics.

Table 4.7 Comparative day/night metrics

METRIC	ALL HOURS	DAYTIME (6-18)	NIGHTTIME (19-5)
MAE (kWh)	0,00523	0,00473	0,00006
MAPE (%)	1,14	1,11	1,54
MBE (kWh)	0,00024	0,00027	-0,00006
Number of hours	4336	4745	4015

To evaluate the model's behavior under different climatic conditions, a seasonal analysis was performed on both hourly and daily scales and the metrics are respectively reported in Table 4.8and in Table 4.9

Table 4.8 Hourly metrics by season

SEASON	HOURS	MBE (kWh)	MAE (kWh)	RMSE (kWh)	MAPE (%)	Err < 3%
Winter (Dec-Feb)	731	0,00595	0,00595	0,00734	1,648	100,0
Spring (Mar-May)	1255	0,00440	0,00536	0,00731	0,951	100,0
Summer (Jun-Aug)	1420	-0,00618	0,00619	0,00760	1,318	100,0
Autumn (Sep-Nov)	930	0,00004	0,00303	0,00405	0,739	100,0

Table 4.9 Daily metrics by season

SEASON	DAYS	MODEL MEAN PROD (kWh)	PVGIS MEAN PROD (kWh)	MAPE (%)	T_AMB MEAN (°C)	T_CELL MEAN (°C)	η_{PANEL} (%)
Winter	90	2,70	2,66	1,77	5,04	6,70	23,2
Spring	92	7,59	7,54	0,99	8,62	13,5	22,7
Summer	92	8,29	8,38	1,15	14,8	20,3	22,1
Autumn	91	4,55	4,55	0,69	10,2	13,1	22,7

The highest mean daily production occurs in summer (8.29 kWh), although panel efficiency is at its minimum (22.14%) due to high temperatures. Winter has the lowest production (2.70 kWh/day) but the highest efficiency (23.16%) due to module cooling. Autumn shows the best accuracy (MAPE 0.69%), while winter shows the worst (1.77%). The parameters for verification of hourly validation are illustrated in Table 4.10.

Table 4.10 Verification of hourly validation criteria

VALIDATION CRITERION	THRESHOLD	OBTAINED VALUE	OUTCOME
HOURLY MAPE	< 2%-3%	< 1,14%	PASSED
SYSTEMATIC BIAS (MBE)	< 0,01 kWh	< 0,00024 kWh	PASSED
R^2 COEFFICIENT	> 0,99	> 0,9995	PASSED
SPEARMAN CORRELATION	> 0,999	> 0,999807	PASSED
WILCOXON TEST	P > 0,05	0,284	PASSED
KOLMOGROV-SMIRNOV TEST	P > 0,05	1,00	PASSED
HOURS WITH ERROR < 1%	> 40%	> 44,5%	PASSED
HOURS WITH ERROR < 2%	> 60%	> 87,9%	PASSED
HOURS WITH ERROR < 3%	> 80 %	> 100%	PASSED

4.2 StoreNet results

Based on the inverse-forward methodology described in the previous chapter, this section presents the validation results for the residential photovoltaic system simulation model. The objectives are twofold: (1) to verify the model's accuracy in reconstructing Plane-of-Array (POA) irradiance from the measured production of a reference house (H2), and (2) to evaluate the model's capability to simulate hourly electricity production for six different households (H1, H2, H3, H4, H5, H7), comparing the results with real measured data from the StoreNet project on the Dingle Peninsula, Ireland.

The analysis is structured into three main phases:

1. Validation on monthly and annual scales
2. Error metrics on annual, monthly and hourly scales
3. Residual analysis

All results refer to hourly aggregated data for the entire year 2020, after appropriate cleaning procedures (nighttime forcing to zero, spike removal, noise threshold of 0.1 Wh).

4.2.1 Validation on annual and monthly scales

4.2.1.1 Annual validation

To validate the model's performance on an annual scale, a comparison was conducted between the total simulated production and the average real production of the six households over the entire year 2020. The average production of the six houses was adopted as a representative indicator of the photovoltaic cluster, reducing the influence of household-specific anomalies, local shading effects, and measurement noise. This approach allows the validation to focus on the model's ability to reproduce the typical behaviour of the study area rather than the characteristics of a single installation. The following Table 4.11 reports the aggregate annual performance metrics. The model simulates a total production of 1863.7 kWh, compared to an average measured production of 1842.9 kWh, corresponding to a percentage difference of only +1.1%. The average Mean Bias Error (MBE) across the six households is +1.3%, indicating a slight tendency towards overestimation, while the average Mean Absolute Error (MAE) is 85.4 Wh. Overall, these results demonstrate a very good agreement between simulated and measured annual production.

Table 4.11 Aggregate comparison of annual production

METRIC	VALUE
Total Simulated Production (kWh)	1863,7
Average Real Production (kWh)	1842,9
Absolute Difference (kWh)	+20,8
Percentage Difference (%)	+1,10%
Average MAE (Wh)	85,4
Average MBE (%)	+1,30%

4.2.1.2 Monthly validation

To verify the long-term robustness of the model, a comparison was conducted between the monthly energy production predicted by the simulation (Simulated Model, in pink) and the average real production recorded from the dataset of six households (Average Real Houses, in orange). The most prominent feature of the graph is the almost perfect overlap with which the model replicates the seasonal profile. Both data series follow a bell-shaped seasonal profile, showing a steady increase in production from the winter months toward the spring and summer peaks. The model's ability to faithfully follow the transition months (such as March and October) confirms that the solar angle modeling and the management of the panels' thermal efficiency have been correctly implemented. The discrepancies between simulation and reality are minimal, with differences rarely exceeding 10%. The disaggregated monthly analysis, reported in the following table, highlights that the largest discrepancies are concentrated in the months of April (+11.6%), August (+7.9%), and May (+7.0%), where the model systematically overestimates production. Conversely, the winter months (January: -9.1%; November: -6.4%; December: -5.0%) show an underestimation, likely due to a greater component of diffuse radiation not fully captured by the model. The Table 4.12 shows the monthly production report.

Table 4.12 Monthly production report

MONTH	MODEL (kWh)	AVG_REAL (kWh)	DIFF (kWh)	DIFF %
January	92,60	101,9	-9,22	-9,05
February	108,4	110,1	-1,70	-1,55
March	169,0	171,9	-2,84	-1,65
April	212,9	190,8	22,1	11,6
May	245,8	229,7	16,1	7,02
June	189,4	182,7	6,66	3,64
July	182,8	190,5	-7,69	-4,04
August	192,0	178,0	14,1	7,90
September	172,2	179,0	-6,77	-3,78
October	127,0	126,5	0,45	0,35
November	88,00	94,00	-6,00	-6,37
December	83,50	87,90	-4,39	-5,00

The following Figure 4.6 shows the comparison of monthly production between the simulated model and the average real houses.

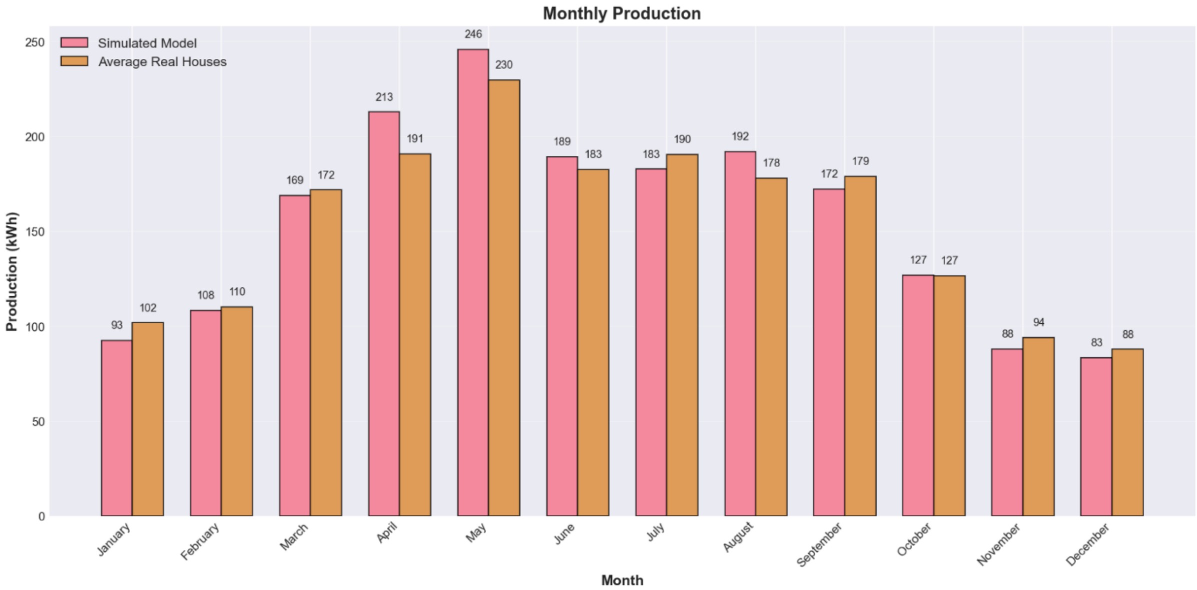


Figure 4.6 Monthly production

4.2.2 Error metrics on annual, monthly and hourly scales

The following Figure 4.7 compares the average measured production of the six selected houses (x-axis) with the output simulated by the model (y-axis).

- X-axis (Average Real Production): The measured hourly average (Wh).
- Y-axis (Simulated Production): The value predicted by the model (Wh).

Darker colors (red/maroon) indicate where the highest concentration of annual data points occurs. An R^2 value exceeding 0.71 is a very positive result for a validation based on real-world residential monitoring data. It indicates that the model is capable of explaining over 71% of the variability observed in actual production. In contexts where variables such as module soiling, local shading, or inverter tolerances are unknown, this value confirms that the inverse-forward methodology is effective.

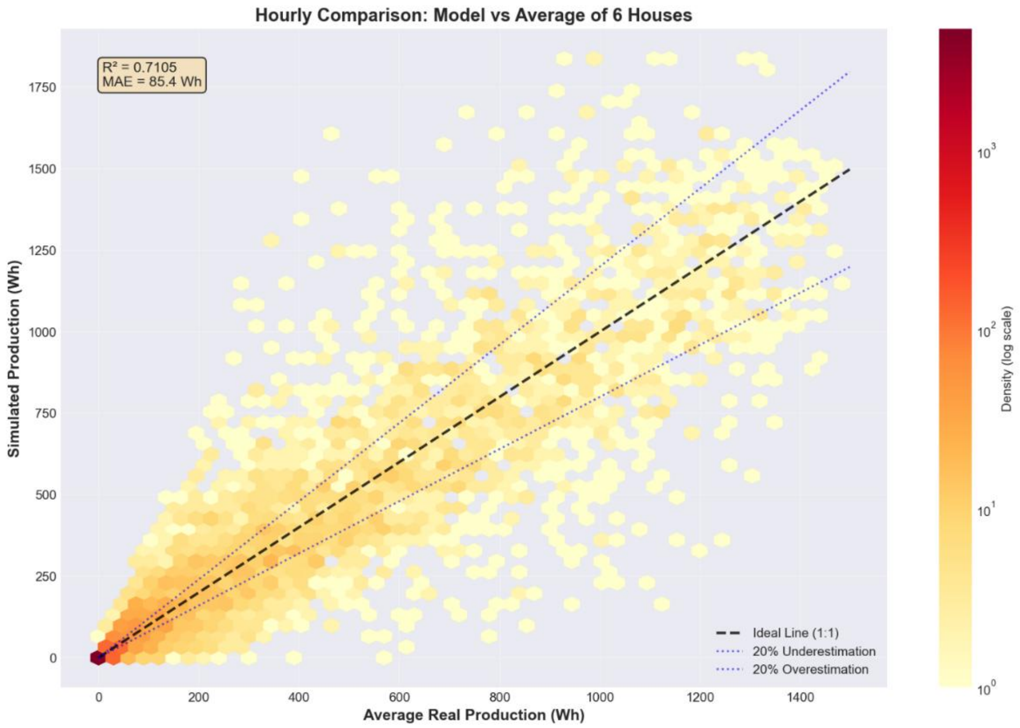


Figure 4.7 Hourly comparison: model vs avarage of 6 houses

The simulation was tested individually for each household in the dataset (H1, H2, H3, H4, H5, and H7). This approach allows for a direct verification of whether the irradiance reconstructed from the reference house (H2) is truly representative of the solar conditions across the entire study area. As expected, the comparison for H2 yields the best results, achieving an R^2 of 0.8876 and the lowest MAE of the group (54.8 Wh). This result confirms the successful calibration of the inverse modeling procedure. The households H1, H3, H5, and H7 exhibit remarkable statistical consistency, with R^2 values ranging between 0.70 and 0.81. This alignment provides further support for the validity of the methodology: despite the physical distance between the houses, the reconstructed irradiance remains a reliable input. Specifically, H1 and H3 show a highly compact distribution of data points along the ideal diagonal. This suggests that these properties share not only the same microclimate as H2 but also very similar exposure conditions (lack of shading) and system orientations. In the cases of H5 and H7, while a strong correlation is maintained, a slightly higher dispersion is observed. This is likely due to small local variations in module operating temperatures or slight differences between the efficiency of the real-world inverters and the simulated model. The case of house H4 represents the most interesting aspect from a diagnostic perspective. Here, the model shows a clear decline in performance, with the R^2 dropping to 0.33 and the average error doubling (MAE = 148.3 Wh). A visual inspection of the corresponding scatter plot reveals a much "wider" cloud of points that is less aligned with the diagonal. This clearly indicates that H4 is subject to phenomena not experienced by the reference house H2. These may include systematic partial shading (caused by trees or chimneys), a panel orientation significantly different from the one reported, or an anomalous degradation of the PV strings. All the trend are reported in Figure 4.8. These findings suggest that the reduced performance observed for H4 is more likely associated with local system-specific conditions than with a limitation of the proposed methodology. The satisfactory agreement obtained for the remaining households supports the assumption that the irradiance reconstructed from H2 is representative of the wider study area.

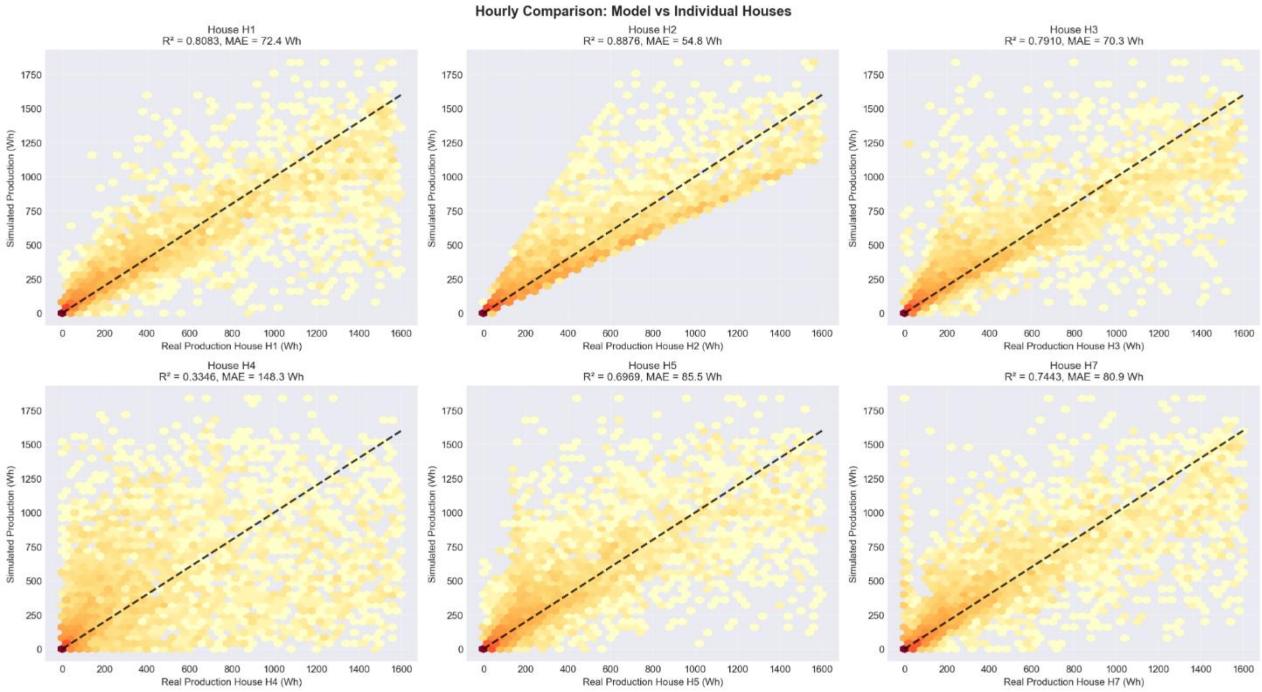


Figure 4.8 Hourly comparison: model vs individual houses

The following Table 4.13 summarizes the validation metrics for each household.

Table 4.13 validation metrics for each household

HOUSE	R^2	MAE (Wh)	MBE (%)	MAPE (%)	CORRELATION
H1	0,80	72,37	-6,34	18,4	0,89
H2	0,88	54,78	1,44	12,7	0,94
H3	0,79	70,31	1,54	21,3	0,89
H4	0,33	148,3	2,57	42,1	0,65
H5	0,69	85,51	6,11	24,6	0,85
H7	0,74	80,85	2,29	17,9	0,86
TOTAL AVERAGE	0,71	85,35	1,27	22,8	0,85

4.2.2.1 Monthly error metrics

The monthly coefficient of determination R^2, reported in the following Table 4.14, ranges from a minimum of 0.73 (April) to a maximum of 0.92 (March and November), confirming the model's good predictive capability on a monthly scale. The monthly MAE values show that the spring and summer months (April–August) present the highest absolute errors (74.4–98.7 Wh), while the winter months (November–January) show the lowest (30.2–35.5 Wh).

Table 4.14 Monthly validation metrics

MONTH	R^2	MAE (Wh)	MBE (Wh)	MBE (%)	CORRELATION
January	0,88	35,5	-12,4	-9,11	0,94
February	0,85	44,1	-2,40	-1,50	0,92
March	0,91	51,8	-3,80	-1,70	0,95
April	0,73	98,7	30,7	11,6	0.90
May	0,79	97,3	21,7	7,00	0,91
June	0,82	74,4	9,20	3,60	0,92
July	0,85	77,2	-10,3	-4,00	0,92
August	0,79	74,9	18,9	7,90	0,91
September	0,90	53,9	-9,40	-3,80	0,95
October	0,87	44,1	0,60	0,40	0,93
November	0,91	30,2	-8,30	-6,40	0,95
December	0,85	34,8	-5,90	-5,00	0,92

4.2.2.2 Hourly error metrics

The Hourly Bias (MBE) plot represented in Figure 4.9 provides a detailed view of the model's accuracy throughout the operational day (from 6:00 AM to 6:00 PM). The hourly bias identifies whether the model tends to overestimate or underestimate production at specific times, revealing the system's response to varying solar incidence angles and temperature conditions.

- Underestimation during transition phases (Sunrise and Sunset) A systematic underestimation is observed (red bars) in the early morning hours (6:00–7:00 AM) and late afternoon (4:00–6:00 PM). During these intervals, the sun is low on the horizon, and direct radiation is minimal compared to the diffuse and reflected radiation.
- Overestimation during peak hours (9:00 AM–1:00 PM) During the central hours of the day, the bias sign reverses (green bars), indicating a tendency of the model to overestimate real production. The peak positive bias is recorded at 11:00 AM. Interestingly, at 3:00 PM, the bias is nearly zero (+1.5 Wh). In this time slot, the underestimations due to the solar angle and the overestimations due to temperature appear to perfectly compensate for each other, making the model's prediction extremely adherent to the real-world data.

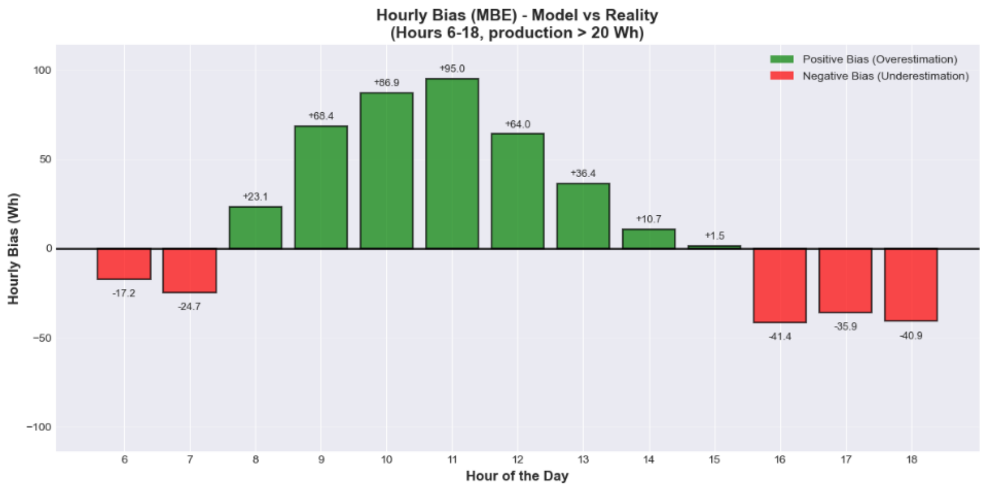


Figure 4.9 Hourly MBE: model vs reality

This following graph (Figure 4.10) shows the hourly Mean Absolute Error (MAE) between the model's predictions and actual observed values, limited to hours 6–18 and only for periods with production > 20 Wh.

The X-axis represents the hour of the day, and the Y-axis represents the MAE in Wh/h.

The graph shows a progressive increase from early morning until midday:

- At 6:00, the error is at its minimum (19 Wh/h)
- The error rises steadily, reaching a peak between 12:00 and 14:00, with maximum values of 173–175 Wh/h
- It then gradually decreases to 97 Wh/h at 18:00.

A red dashed line indicates the overall average MAE over the considered period, which is 121.7 Wh/h. Most of the central hours of the day (10–16) are above this average, while the early morning and late afternoon hours (6–9 and 17–18) are below it. This suggests that the model is more accurate in the early morning and late afternoon, but less accurate during the central hours of the day (likely when production is more variable or higher).

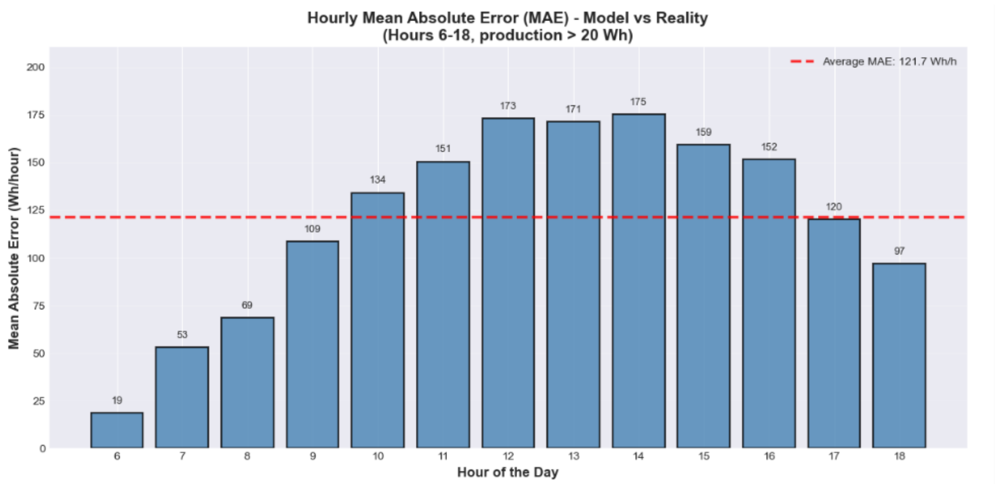


Figure 4.10 Hourly MAE: model vs reality

4.2.3 Residual analysis

To complete the model validation, an analysis of the average annual hourly profile was conducted to verify the simulation's ability to replicate a "typical day" of production. This graph illustrated in Figure 4.11 relates energy production (Wh) to the hours of the day. The most prominent feature is the model's extreme precision in following the temporal dynamics of production. The simulated model curve and the actual average curve are nearly perfectly aligned throughout the 24-hour cycle. The model accurately captures the onset of production at dawn, the morning ramp-up, and the peak generation occurring, as expected, around 1:00 PM. A highly significant technical detail is the shaded area, which represents the standard deviation of the actual production among the different houses in the cluster. The fact that the simulated curve consistently remains at the center of this area confirms that the model is not only accurate but also statistically representative of the expected value for the entire geographical area considered. The slight discrepancy visible during the late morning hours (between 10:00 AM and 12:00 PM), where the model tends to slightly overestimate real production, can be attributed to localized morning shading or intermittent cloud cover affecting the cluster but not the reference house. It may also reflect a real-world module thermal response that differs slightly from the theoretical model during the heating phase. During the afternoon decline, the correspondence returns to near-total agreement, demonstrating that the model correctly handles the solar angle of incidence and system efficiency variations even under decreasing irradiance. In conclusion, this graph provides strong evidence that the adopted methodology is capable of reconstructing not only the total volume of energy produced but also its hourly distribution a crucial parameter for any analysis aimed at optimizing self-consumption or integrating photovoltaic systems into the local power grid.

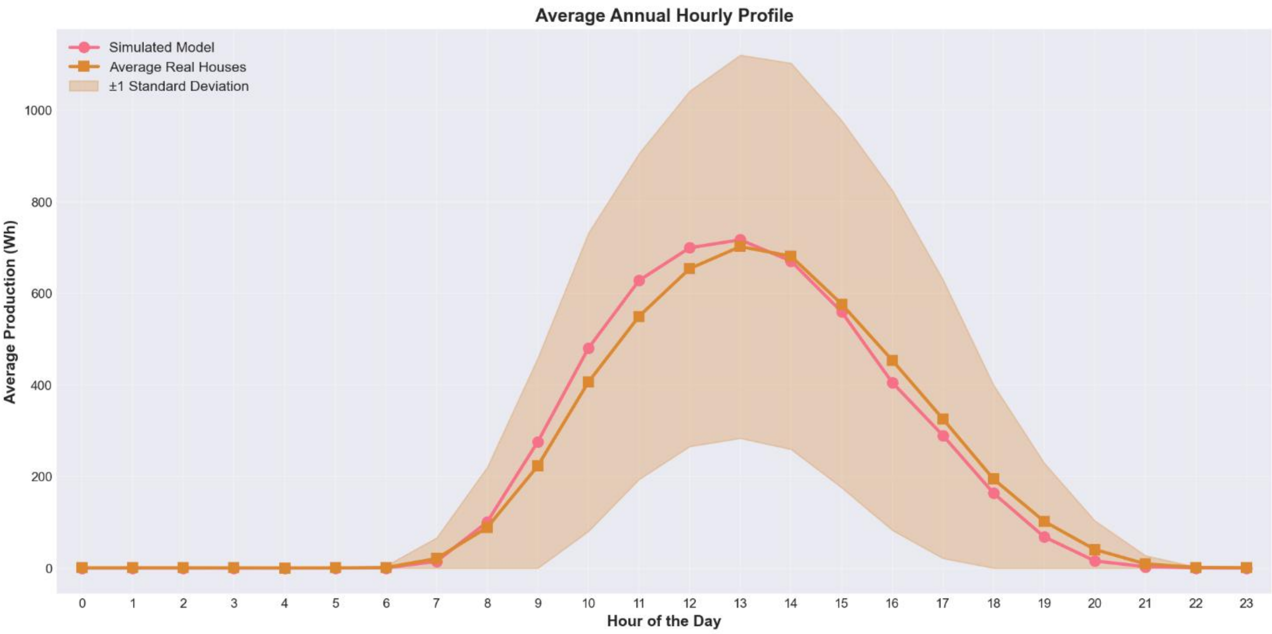


Figure 4.11 Average annual hourly profile

The trend graphs were also obtained for individual houses as it’s shown in Figure 4.12

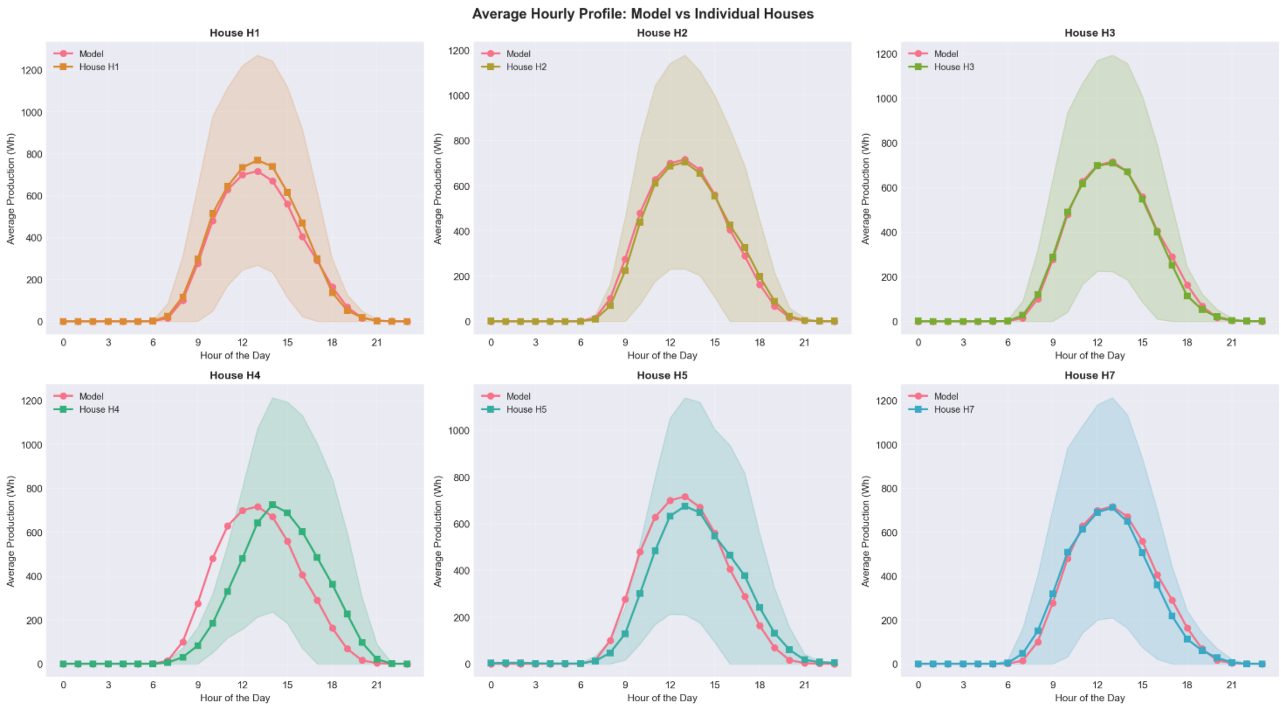


Figure 4.12 Average hourly profile: model vs individual houses

These profiles evaluate the model's ability to replicate the production "shape" during a typical day for each specific system. For the majority of the households (H1, H2, H3, and H7), an almost perfect overlap is observed between the model curve (pink line) and the actual production (orange line). The production peak is correctly centered between 12:00 PM and 1:00 PM, confirming that the orientation (azimuth) used in the model faithfully reflects reality. Furthermore, the model's ability to follow the morning ramp-up and afternoon ramp-down indicates that solar angles of incidence and the system's response to low irradiance levels have been modeled with high precision. As expected, the profile for H2 shows the lowest residual error, as this house was used as the reference to reconstruct the irradiance data.

The profiles for H4 and H5, however, reveal systematic discrepancies that offer important diagnostic insights:

- Case H4: This profile exhibits the most significant deviation. Real production (green line) is consistently lower than the simulated output between 10:00 AM and 2:00 PM, though it appears to "recover" or shift slightly in the late afternoon. This suggests a potential systematic error in the declared orientation of the panels or, more likely, the presence of significant shading (trees or buildings) affecting the system during peak hours.
- Case H5: A slight but systematic underestimation by the model is noted during peak hours. This could indicate that the actual system at H5 has an overall efficiency (or module ventilation) slightly higher than the standard parameters set in the model, or that it benefits from a cooler local microclimate that enhances thermal performance.

The shaded areas surrounding the curves represent the standard deviation of actual production. The fact that the model curve remains almost entirely within this band demonstrates that the simulation produces physically plausible and "normal" values within the given climatic context.

4.3 Calibration results

Based on the calibration procedure, this section presents the results of the parameter optimisation for the residential photovoltaic systems. The objectives are twofold: (1) to estimate the unknown parameters (peak power P_peak, Performance Ratio PR, and temperature coefficient γ) for each dwelling using the reconstructed POA irradiance, and (2) to validate the calibrated model against the original measured production data from the smart meters.

The analysis is structured into three main phases:

1. Validation on monthly and annual scales
2. Error metrics on annual, monthly and hourly scales
3. Residual analysis

The calibration was performed on ten residential units (H1, H2, H3, H4, H5, H7, H10, H11, H13, H17) using minute-by-minute data for the entire year 2020. The optimisation algorithm minimised the Root Mean Square Error (RMSE) between the model output and the reconstructed real production obtained via energy balance, with bounds: P_peak $\in$ [0.5, 10.0] kWp, PR $\in$ [0.6, 0.95], γ $\in$ [0.002, 0.006] (0.20–0.60%/°C).

By way of example, the case of dwelling H1, which represents the only significant outlier in the sample, is analysed in detail first, followed by the aggregate results for all ten systems.

4.3.1 Validation on annual scale

4.3.1.1 Annual energy comparison for H1

The following Figure 4.13 presents a direct comparison between the temporal trend of the actual production of the photovoltaic plant (reported in column D of the dataset) and that obtained through the calibrated mathematical model, characterised by a peak power of 2.21 kWp.

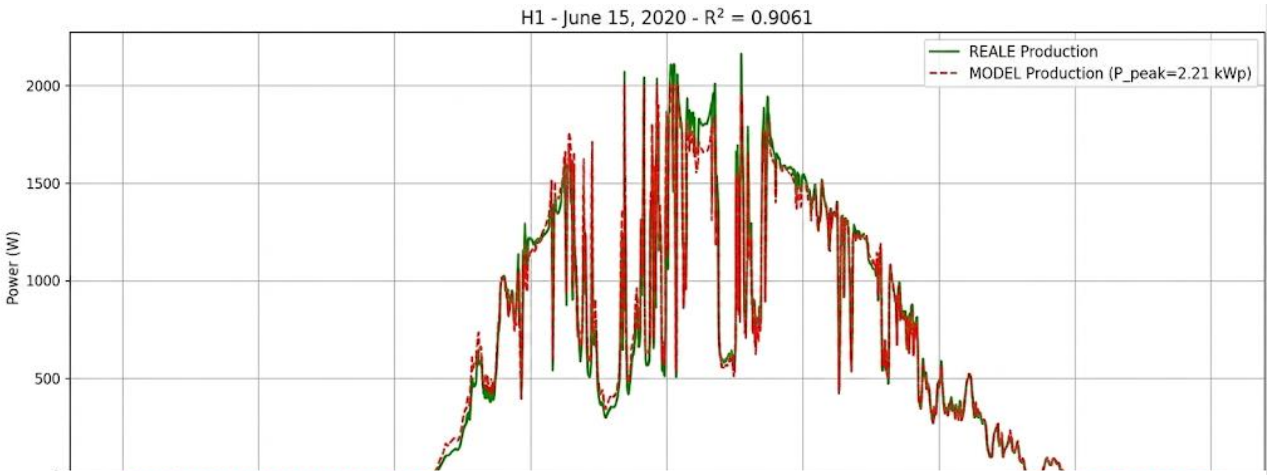


Figure 4.13 Comparison model vs reality H1

The time interval considered covers approximately 24 hours, from midnight on June 15 to midnight the following day, with a temporal resolution of three hours. Both curves exhibit the typical bell-shaped trend of a day characterized by stable irradiance conditions, likely clear or with little cloud cover. Production is zero during nighttime hours, increases rapidly from early morning, reaches a maximum around noon, and gradually decreases in the afternoon until dropping to zero after sunset. A relevant aspect concerns the peak power value declared for the model, equal to 2.21 kWp. In the graph, however, the peak actually reached by the model is around 1.7–1.8 kW. This indicates that the model does not correspond to a simple multiplication of nominal power by irradiance, but already includes a series of system losses, such as inverter efficiency, the effect of cell temperature, possible soiling, or non-optimal module orientation. In other words, the 2.21 kWp represents a reference value for system sizing, while the simulated curve returns the expected actual AC production under real conditions. The actual production, on the other hand, shows a slightly lower peak, approximately 1.5–1.6 kW. From a quantitative point of view, the goodness of fit between model and measurements is expressed by a coefficient of determination $R^2 = 0.9061$, a rather high value indicating that over 91% of the variance observed in actual production is explained by the model. This is an excellent result for a simplified physical model, often based exclusively on irradiance data and panel tilt angle. However, visual analysis of the graph reveals the presence of a non-negligible systematic deviation. In the morning hours, between 6 and 9, the model almost perfectly overlaps the actual data, suggesting good accuracy during periods of low solar elevation. Around noon, on the other hand, the model tends to clearly overestimate production, with a difference exceeding 0.2 kW. This overestimation gradually attenuates during the afternoon, until disappearing entirely in the evening hours, where the two curves once again coincide. In the following Table 4.15 is reported the comparison between real data and the model about the annual energy production. H1 represents the only dwelling for which the calibrated model does not achieve near-perfect agreement, suggesting the presence of system-specific factors not captured by the optimization process.

Table 4.15 Annual energy production: real vs model

HOUSE	REAL ENERGY (MWh)	MODEL ENERGY (MWh)	ENERGY ERROR (%)
H1	2,08	2,15	+2,8
H2	1,92	1,95	+1,6
H3	1,92	1,95	+1,4
H4	1,87	1,89	+1,4
H5	1,81	1,82	+0,7
H7	1,84	1,86	+1,1
H10	1,93	1,95	+1,1
H11	1,75	1,77	+1,1
H13	1,85	1,880	+1,6
H17	2,08	2,11	+1,5
MEAN	1,91	1,93	+1,5
STD DEV	113,0	118,0	0,60

Annual real energy production ranges from 1,748 MWh (H11) to 2,084 MWh (H1), with a mean of 1,905 MWh and a standard deviation of 113 kWh (5.9% of the mean). The equivalent full-load hours average 918 hours (range: 853–1,005 hours), values consistent with the typical annual solar irradiation at Irish latitude (approximately 900–1,000 kWh/m²). The energy error shows a systematic positive bias for all dwellings, averaging +1.5% (range: +0.7% for H5 to +2.8% for H1).

4.3.2 Errors metrics on annual, monthly and hourly scales

The Table 4.16 reports the optimized parameters for each dwelling, together with the accuracy indicators.

Table 4.16 Estimated parameters and accuracy metrics for the 10 residential systems.

HOUSE	P_peak [kWp]	AREA [m2]	PR	Gamma γ [%/°C]	R^2	RMSE [W]	MAE [W]	BIAS [W]
H1	2,21	11	0,804	0,0020	0,9061	173,0	113,2	+16,8
H2	2,07	10,4	0,81	0,0020	0,9960	33,4	16,6	+8,6
H3	2,07	10,3	0,81	0,0020	0,9926	46,6	15,8	+7,5
H4	2,06	10,3	0,81	0,0020	0,9966	30,5	14,2	+7,6
H5	2,05	10,3	0,81	0,0020	0,9921	44,2	13,0	+4,3
H7	2,05	10,2	0,81	0,0020	0,9971	26,3	10,8	+5,8
H10	2,05	10,3	0,81	0,0020	0,9985	20,1	10,8	+5,7
H11	2,05	10,2	0,81	0,0020	0,9979	21,9	10,1	+5,5
H13	2,07	10,3	0,81	0,0020	0,9976	25,3	13,7	+7,5
H17	2,07	10,3	0,81	0,0020	0,9964	32,9	15,2	+8,0
MEAN	2,07	10,4	0,801	0,0020	0,9921	45,4	23,3	+7,8
STD DEV	0,04	0,2	0,001	0,000	0,0270	46,0	31,0	3,5

The calibration results reveal a remarkable consistency across the ten residential systems. The average peak power is 2.07 ± 0.04 kWp, ranging from 2.05 to 2.21 kWp, with nine out of ten systems showing a very homogeneous estimated peak power between 2.05 and 2.07 kWp (standard deviation of only 0.04 kWp). The only exception is dwelling H1, with a value of 2.21 kWp (6.8% above the sample mean), which is consistent with its slightly larger module area (11.0 m² vs. 10.3 m² mean) and is accompanied by a higher energy error (+2.8%). This uniformity suggests that the systems may have been installed as part of a coordinated intervention, implying a standard power configuration for the analysed residential units. The average Performance Ratio is 0.801 ± 0.001 (0.804 for H1), a value that indicates total system losses of 20% and is fully compatible with literature expectations for small residential systems without significant shading (typical range 0.75–0.85). The temperature coefficient assumed the lower bound of the optimisation interval (γ = 0.20 %/°C) for all dwellings. It is likely that the optimisation compensated for an overestimation of cell temperature by the NOCT model, or that the ventilation of the systems (typically installed on well-ventilated roofs) keeps the cells cooler than the model assumes; an alternative explanation is that the correlation between production and temperature in the available data is weak, and the optimisation algorithm therefore tends to "offload" the explanation of losses onto the PR, leaving γ at its minimum. The model explains on average 99.2% of the variance in real production (mean R^2 = 0.9921), with excellent quality of fit for almost all systems: nine out of ten houses have an R^2 above 0.99, the best performances being H10 (R^2 = 0.9985, RMSE = 20.1 W), H11 (R^2 = 0.9979, RMSE = 21.9 W) and H7 (R^2 = 0.9971, RMSE = 26.3 W). The only significant outlier is H1, with $R^2 = 0.9061$ and RMSE = 173 W. The mean RMSE of 45.4 W corresponds to approximately 2.2% of the mean peak power (2,070 W); excluding the anomalous H1 case, the model error falls below 2% of installed capacity – a remarkable level of precision for a simplified physical model fed with minute-by-minute data. Among special cases, dwelling H5 is the only one for which the optimisation detected a non-zero standby offset (7.8 W), plausible for the self-consumption of a small inverter in the absence of production, while the houses with the largest balance corrections are H11 (0.87% of corrected data), H13 (0.85%) and H17 (0.62%), probably due to unsynchronised sampling or quantisation errors. The high accuracy achieved by the model calibrated ($R^2 > 0.99$ for nine out of ten systems) provides additional confidence in both the reconstructed POA irradiance dataset and the adopted data cleaning procedure. The close agreement between simulated and measured production indicates that the irradiance reconstruction based on the assumed PR value of 0.80 yelds physically consistent and reliable results for the analyzed residential systems.

4.3.3 Residual analysis

The Figure 4.14 presents a residuals analysis. The differences between the measured values (actual production) and those estimated by the model, plotted as a function of the measured power. This type of representation is particularly useful for evaluating the presence of systematic errors or patterns in the error distribution, which the coefficient of determination R^2 alone cannot highlight.

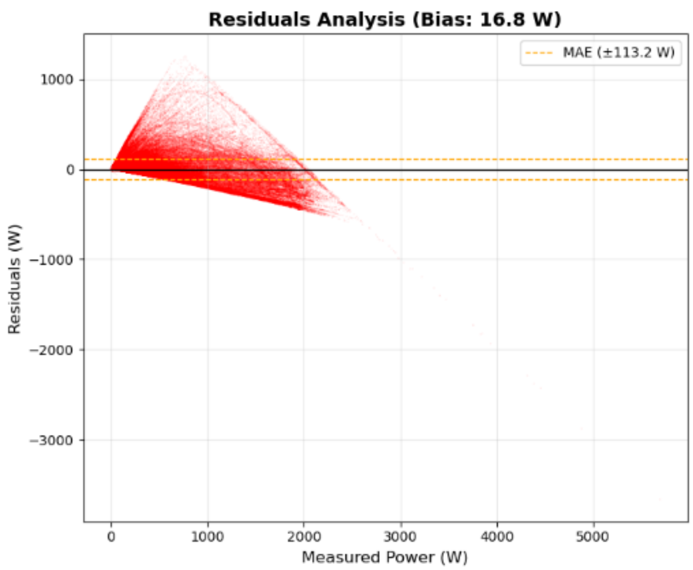


Figure 4.14 Residual analysis

The most immediate element is the bias value, equal to 16.8 W, which represents the average systematic error of the model. This is a positive value, meaning that, on average, the model tends to overestimate the actual production by about 17 watts. Alongside the bias, the mean absolute error (MAE) is reported, equal to ±113.2 W. This value indicates that, on average, the absolute deviation between model and measurement is about 113 watts, regardless of the sign of the error. This represents a relatively small percentage error, approximately 7–8% of the peak production, which is still a satisfactory level of accuracy for many engineering applications. Observing the distribution of residuals along the measured power axis, it can be noted that the points are mainly concentrated in the range between 0 and about 1500 W, which corresponds to the actual power output range of the system. The cloud of points appears relatively compact and symmetric around zero for most of the range, with a slight prevalence of positive residuals (overestimation) at intermediate and high power levels

4.4 Comparative methodology for baseline tariffs and renew dynamic optimization results

This chapter presents the results of the two-phase analysis described in the methodology. The first section reports the baseline evaluation of standard tariffs across all six case studies. The second section presents the Renew tariff optimization results for each household. The third section develops the multi-level comparative analysis following the framework established in the methodology.

4.4.1 Baseline tariff analysis

The baseline analysis evaluated 52 standard tariffs for each household after excluding tariffs with overlapping intervals or zero standing charges. The following Table 4.17 summarizes the key results. All values reflect the real battery's operation where present.

Table 4.17 Baseline tariff results summary

HOUSE HOLD	PV	REAL BATTERY	GRID IMPORT (kWh)	GRID EXPORT (kWh)	BEST TARIFF COST (€/yr)	PROVIDER	TARIFF TYPE
H1	Yes	Yes	2384	875	738.26	Flogas Ireland	Flat with feed-in
H3	Yes	Yes	4104	131	1,202.1	Electric Ireland	Night Boost (no feed-in)
H10	Yes	Yes	6986	323	1,846.6	Flogas Ireland	EV Night with feed-in
H8	No	Yes	7582	0	1,822.9	Electric Ireland	Night Boost
H18	No	Yes	6526	0	1,839.5	Electric Ireland	Night Boost
H19	No	Yes	5712	0	1,663.8	Electric Ireland	Night Boost

- For H1, the optimal tariff is Flogas Ireland's Smart 24Hr Electricity 29% Loyalty Discount, a flat-rate tariff with a feed-in credit of €0.185/kWh. The standing charge is €270/year, variable costs are €630/year, and the feed-in credit on 875 kWh of exports amounts to approximately €162, yielding a net cost of €738. The effective average price of imported energy before the feed-in credit is €0.31/kWh, or €0.26/kWh after accounting for the credit across net consumption.

- For H3, the optimal tariff is Electric Ireland's Home Electric+ Night Boost, which offers reduced night-time rates (23:00-08:00) and standard day rates with no feed-in mechanism. The standing charge is €329/year, variable costs are €874/year, and exports of 131 kWh receive no credit, yielding a net cost of €1,202 and an effective average price of €0.293/kWh.
- For H10, the optimal tariff is Flogas Ireland's Smart EV Night Charge 29% Electricity Loyalty Discount, which combines a favourable night-time rate with a feed-in credit of €0.185/kWh. The standing charge is €349/year, and the feed-in credit on 323 kWh of exports amounts to approximately €60, yielding a net cost of €1,847 and an effective average price of €0.264/kWh.
- For the three non-photovoltaic households, the optimal tariff is uniformly Electric Ireland's Home Electric+ Night Boost, with a standing charge of €329/year, a night rate of approximately €0.09/kWh, and a day rate of approximately €0.18/kWh. The effective average price varies inversely with consumption: €0.240/kWh for H8 (7,582 kWh), €0.282/kWh for H18 (6,527 kWh), and €0.291/kWh for H19 (5,713 kWh).

4.4.2 Renew tariff optimization results by household

The Renew tariff optimization replaced the real battery with a virtual battery controlled by predictive XGBoost logic (4-hour forecasts, charge threshold €0.08/kWh, discharge threshold €0.14/kWh, round-trip efficiency 90 per cent). Three virtual battery capacities were tested: 5 kWh, 10 kWh, and 15 kWh. The "Renew no battery" scenario represents the same household with the battery completely removed, isolating the battery's value contribution. Results are presented for urban users as the reference case and are presented in the following Tables from Table 4.18 to Table 4.23 for the three houses with the photovoltaic system (H1, H3, H10) and for the three houses without a PV system (H8, H18, H19).

Table 4.18 Renew tariff results for H1

CONFIGURATION	**ANNUAL COST (€)**	**Δ FROM BASELINE (€)**	**GRID IMPORT (kWh)**	**FEED-IN CREDIT (€)**	**AVG. IMPORT PRICE (€/kWh)**
Baseline (best standard)	738,2	-	2385	161,95	0,264
Renew no battery	901,8	+163,5	2377	92,3	0,31
Renew +5 kWh	831,7	+93,4	2194	130,1	0,30
Renew +10 kWh	823,6	+85,3	2150	122,5	0,30
Renew +15 kWh	820,2	+81,9	2132	120,2	0,31

The baseline outperforms all Renew configurations primarily because of the feed-in credit differential. The baseline offers €0.185/kWh for exports, while the Renew tariff offers feed-in at the DAM price, which averaged €0.109/kWh over the study period. For H1's export volume of 875 kWh, this differential alone accounts for €66.50 of the €163.54 gap between baseline and Renew no battery.

The remaining gap is explained by higher average import prices under Renew (€0.310/kWh versus €0.264/kWh), which on 2,377 kWh of imports adds approximately €109. The virtual battery reduces but does not eliminate the gap: the 5 kWh configuration reduces the deficit to €93, and further capacity increases yield only marginal improvements, with the 15 kWh configuration still €82 above the baseline.

Table 4.19 Renew tariff results for H3

CONFIGURATION	ANNUAL COST (€)	Δ FROM BASELINE (€)	GRID IMPORT (kWh)	FEED-IN CREDIT (€)	AVG. IMPORT PRICE (€/kWh)
Baseline (best standard)	1202,1	-	4104	0,000	0,29
Renew no battery	1226,1	+23,9	4091	14,8	0,29
Renew +5 kWh	1123,2	-78,7	3669	72,4	0,28
Renew +10 kWh	1115,9	-86,2	3638	69,7	0,29
Renew +15 kWh	1109,9	-92,2	3619	68,6	0,28

For H3, the baseline tariff has no feed-in credit, and the Renew no battery case adds a small feed-in credit (€15) at the cost of slightly higher import prices, resulting in a modest disadvantage of €24. The addition of a virtual battery transforms the comparison. The 5 kWh configuration reduces grid imports by 422 kWh (10.3 per cent) and increases feed-in credits to €72, yielding a net saving of €79 relative to the baseline. The 5 kWh battery captures 85 per cent of the total saving achievable with the 15 kWh configuration, indicating diminishing returns beyond the smallest capacity tested.

Table 4.20 Renew tariff results for H10

CONFIGURATION	ANNUAL COST (€)	Δ FROM BASELINE (€)	GRID IMPORT (kWh)	FEED-IN CREDIT (€)	AVG. IMPORT PRICE (€/kWh)
Baseline (best standard)	1846,6	-	6986	59,8	0,26
Renew no battery	1913,8	+67,2	6959	34,6	0,27
Renew +5 kWh	1772,7	-73,9	6198	57,6	0,26
Renew +10 kWh	1754,1	-92,6	6124	51,9	0,26
Renew +15 kWh	1744,2	-102,4	6087	49,5	0,26

H10 exhibits the largest absolute saving of any household. The 5 kWh virtual battery reduces grid imports by 761 kWh (10.9 per cent) and increases feed-in credits by €23 relative to the Renew no battery case, yielding a net saving of €74 relative to the baseline. The 15 kWh configuration saves €102 annually, with marginal benefits larger than those observed for H1 or H3, reflecting the higher consumption volume that amplifies arbitrage opportunities. The battery cycles 101 times annually for the 5 kWh configuration, the highest among all PV households, indicating frequent surplus-deficit alternations.

Table 4.21 Renew tariff results for H8

CONFIGURATION	ANNUAL COST (€)	Δ FROM BASELINE (€)	GRID IMPORT (kWh)	FEED-IN CREDIT (€)	AVG. IMPORT PRICE (€/kWh)
Baseline (best standard)	1822,9	-	7582	0,00	0,24
Renew no battery	1929,6	+106,7	7548	0,00	0,26
Renew +5 kWh	1974,3	+151,5	7121	45,1	0,27
Renew +10 kWh	1964,9	+142,0	7373	82,0	0,26
Renew +15 kWh	1953,5	+130,6	7537	108,3	0,25

For H8, the baseline Night Boost tariff offers an average price of €0.240/kWh, which is lower than the Renew no battery average of €0.256/kWh. The differential of 1.6 cents per kWh, multiplied by the annual import volume of approximately 7,500 kWh, accounts for most of the €107 disadvantage. Adding a battery increases costs further: even the 15 kWh configuration, which generates €108 in feed-in credits, results in a net cost €131 above the baseline. The battery cycles 61 times annually for the 15 kWh configuration, with an average state of charge of only 2.8 kWh, indicating that the arbitrage opportunities captured are insufficient to overcome the 10 per cent round-trip efficiency penalty.

Table 4.22 Renew tariff results for H18

CONFIGURATION	ANNUAL COST (€)	Δ FROM BASELINE (€)	GRID IMPORT (kWh)	FEED-IN CREDIT (€)	AVG. IMPORT PRICE (€/kWh)
Baseline (best standard)	1839,5	—	6526	0,00	0,28
Renew no battery	1858,4	+18,9	6500	0,00	0,29
Renew +5 kWh	1926,2	+86,7	6060	43,5	0,31
Renew +10 kWh	1915,4	+75,8	6307	79,8	0,29
Renew +15 kWh	1902,3	+62,8	6464	104,45	0,28

H18 follows the same pattern as H8 but with smaller absolute gaps. The Renew no battery disadvantage is only €19, compared to €107 for H8, because H18's higher baseline average price (€0.282/kWh versus €0.240/kWh) makes the DAM-based Renew tariff relatively more competitive. However, adding a battery still increases costs, with the 5 kWh configuration losing €87 relative to the baseline and even the 15 kWh configuration losing €63.

Table 4.23 Renew tariff results for H19

CONFIGURATION	ANNUAL COST (€)	Δ FROM BASELINE (€)	GRID IMPORT (kWh)	FEED-IN CREDIT (€)	AVG. IMPORT PRICE (€/kWh)
Baseline (best standard)	1663,9	—	5712	0,00	0,29
Renew no battery	1688,4	+24,5	5700	0,00	0,29
Renew +5 kWh	1742,0	+78,1	5402	44,4	0,31
Renew +10 kWh	1731,3	+67,4	5651	80,9	0,29
Renew +15 kWh	1719,5	+55,6	5815	107,0	0,28

H19 exhibits the same pattern as H8 and H18. The Renew no battery disadvantage is €24, and adding a battery increases costs further, with the 5 kWh configuration losing €78 relative to the baseline and the 15 kWh configuration losing €56.

In the following charts in Figure 4.15 and Figure 4.16 the average day-ahead market price profile by hour and the distribution of day-ahead market prices are reported respectively

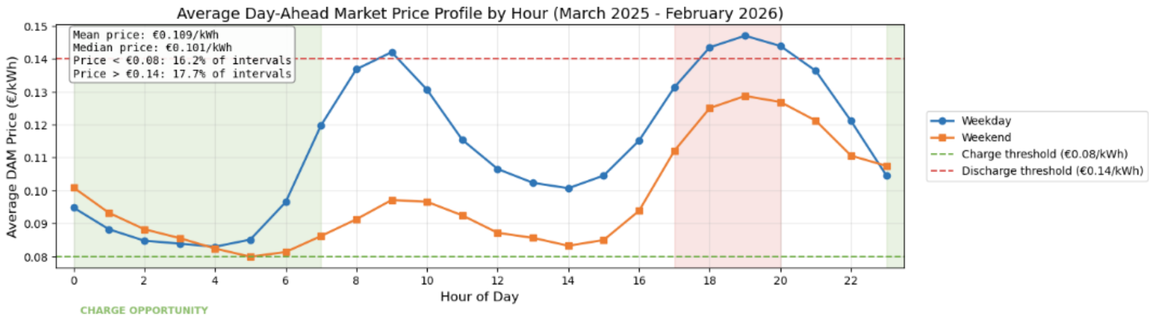


Figure 4.15 Avarage day-ahead market price profile by hour

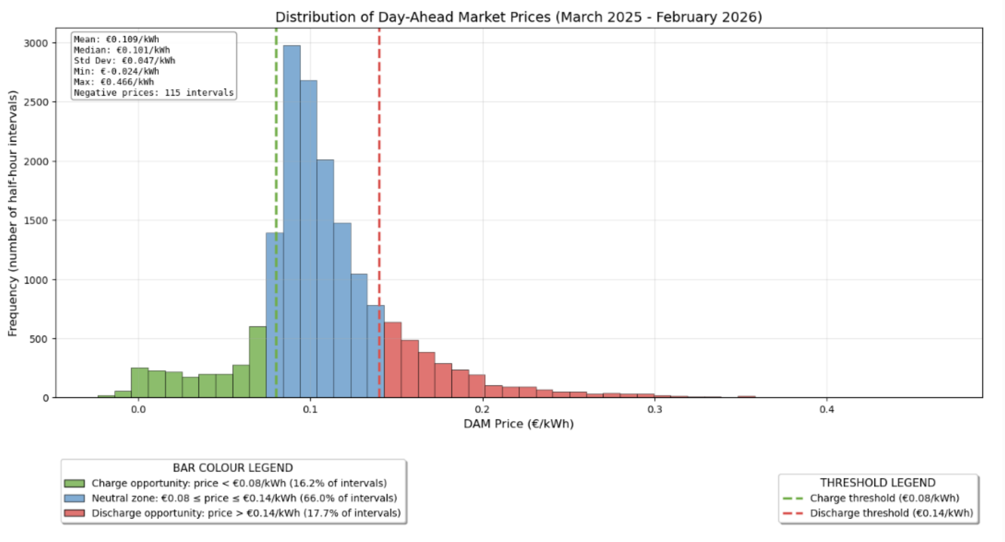


Figure 4.16 Distribution of day-ahead market prices

4.4.3 Multi-level comparative analysis

This section develops the three-level comparison framework established in the methodology: intra-household, intra-category, and inter-category comparisons. Each level examines the economic mechanisms driving the observed outcomes. The following Table 4.24 summarizes the comparison between the best baseline tariff and the best Renew configuration for each household and in Figure 4.17 the annual electricity cost by household and tariff configuration is shown.

Table 4.24 Baseline vs Renew comparison summary

HOUSEHOLD	BASELINE COST (€)	BEST RENEW COST (€)	BEST RENEW CONFIGURATION	DIFFERENCE (€)	RENEW WINS
H1	738,3	820,2	Renew + 15 kWh	+81,9	No
H3	1202,1	1109,9	Renew + 15 kWh	-92,2	Yes
H10	1846,6	1744,2	Renew + 15 kWh	-102,4	Yes
H8	1822,9	1929,6	Renew no battery	+106,7	No
H18	1839,5	1858,4	Renew no battery	+18,9	No
H19	1663,9	1688,4	Renew no battery	+24,5	No

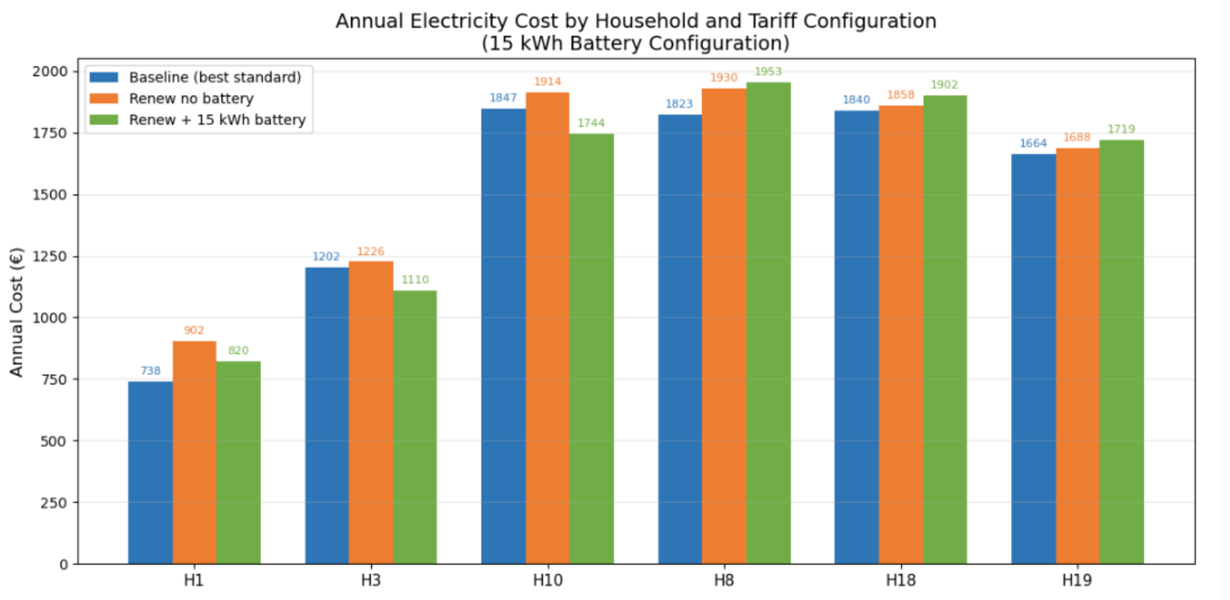


Figure 4.17 Annual electricity cost by household and tariff configuration

- For H1, the Renew tariff with battery is €82 more expensive than the baseline. The main economic driver is the feed-in credit differential: the baseline offers €0.185/kWh for exports, while Renew offers feed-in at the DAM price (€0.109/kWh average). On H1's export volume of 875 kWh, this sacrifice amounts to €67 annually. Additionally, the baseline standing charge (€270) is €85 lower than Renew's fixed charges (€355). The combined disadvantage of €152 exceeds the battery's ability to generate value through self-consumption enhancement and arbitrage, which is limited by the battery's cycling frequency of 86 times per year for the 5 kWh configuration.

- For H3, the Renew tariff with battery is €92 less expensive than the baseline. The baseline has no feed-in credit, so switching to Renew adds a new revenue stream without sacrificing any existing benefit. The 5 kWh virtual battery reduces grid imports by 422 kWh (10.3 per cent) and increases feed-in credits by €57, yielding a net saving of €79. The 5 kWh battery captures 85 per cent of the total saving achievable with the 15 kWh configuration, indicating diminishing returns beyond the smallest capacity.
- For H10, the Renew tariff with battery is €102 less expensive than the baseline. Although the baseline has a feed-in credit (€0.185/kWh), the export volume is only 323 kWh, so the forgone feed-in value (€60) is modest relative to H10's high consumption (6,986 kWh imported). The high consumption amplifies the benefits of battery arbitrage, with the 5 kWh battery cycling 101 times annually — the highest among all PV households. The net saving of €102 is the largest absolute saving observed.
- For the three non-photovoltaic households, the Renew tariff without battery is the best Renew configuration, but it is still more expensive than the baseline in all cases. The disadvantage ranges from €19 for H18 to €107 for H8. Adding a battery increases costs further for all three households. The fundamental obstacle is the absence of solar generation, which eliminates the self-consumption enhancement channel and leaves only pure arbitrage, which is unprofitable under 2025-2026 DAM conditions due to the 10 per cent round-trip efficiency penalty and insufficient price differentials.

The following Table 4.25 presents the comparison of the three photovoltaic households, highlighting the variables that determine whether Renew with battery outperforms the baseline.

Table 4.25 household comparison - best renew vs baseline

HOUSEHOLD	BASELINE FEED-IN RATE	EXPORT VOLUME (kWh)	BASELINE FEED-IN VALUE	RENEW BEST *Δ*	RENEW WINS
H1	0,185	875	162	+82	No
H3	0,00	131	0	-92	Yes
H10	0,185	323	60	-102	Yes

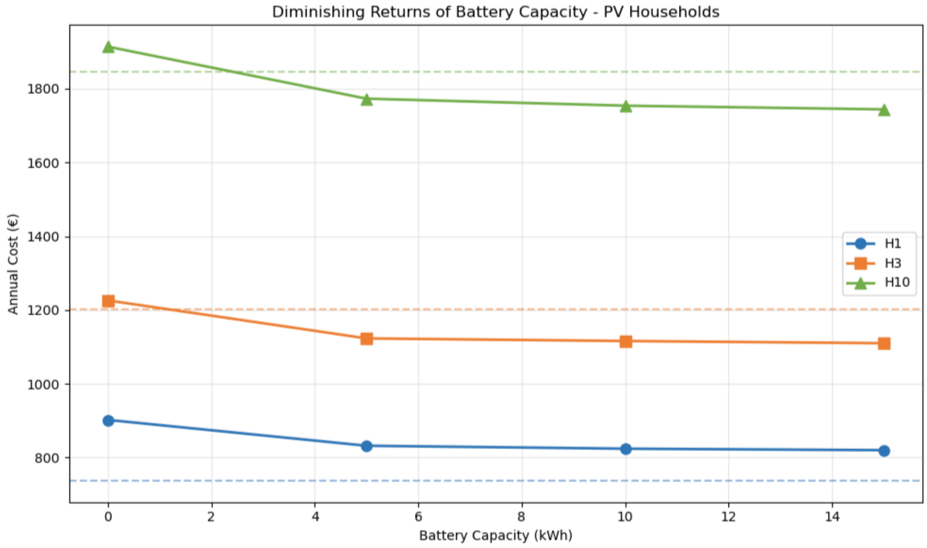


Figure 4.18 Diminishing returns of battery capacity-PV households

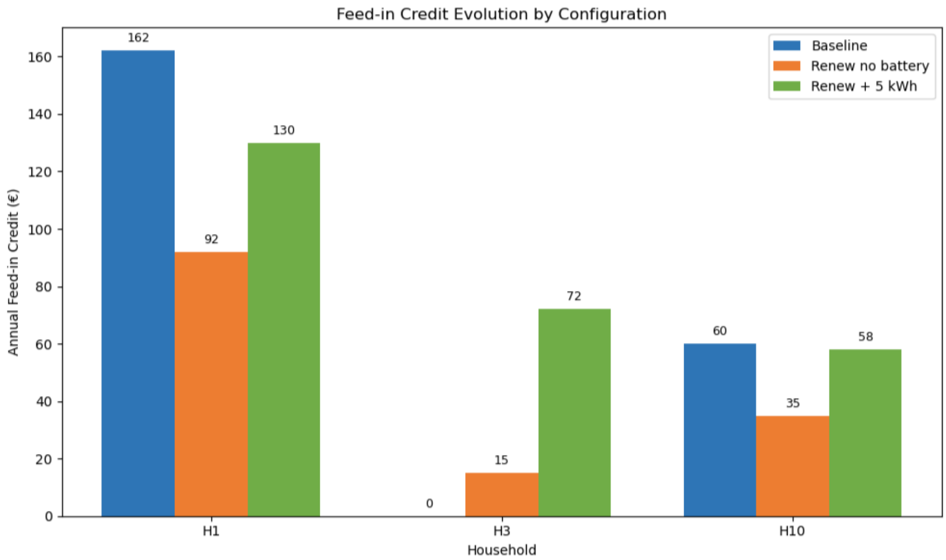


Figure 4.19 Feed-in credit evolution by configuration

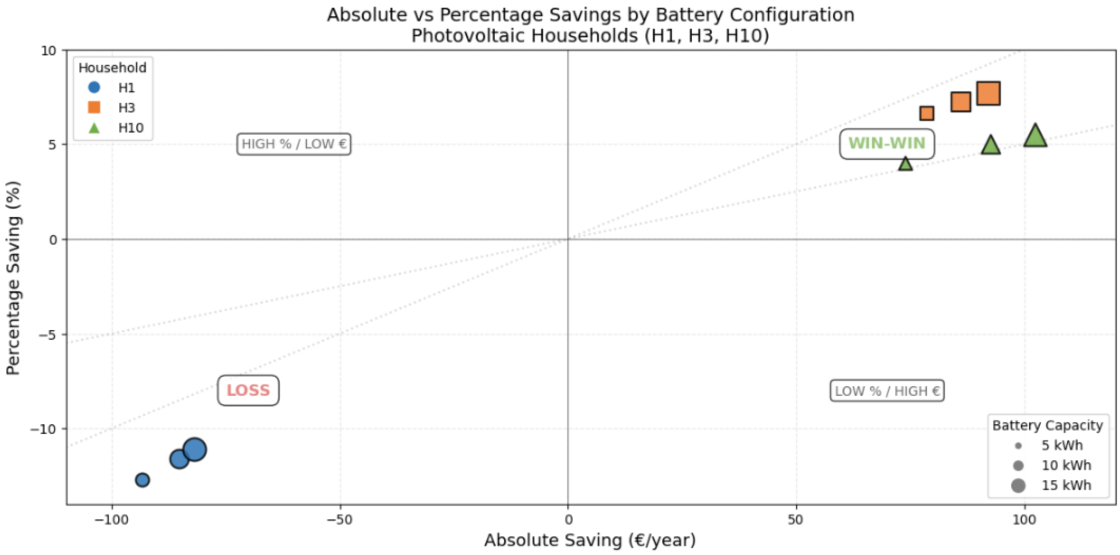
The comparison that comes from Figure 4.18 and Figure 4.19 reveals that the presence of photovoltaics alone is insufficient to guarantee the economic superiority of the Renew tariff with battery. The critical discriminating factor is the export volume and, consequently, the value of feed-in credits forgone when switching from the optimal static tariff to the dynamic Renew tariff.

- For H1, the feed-in value forgone (€162) represents 22 per cent of the baseline cost. The sacrifice of this value, combined with a standing charge that is €85 lower than Renew's fixed charges, creates a disadvantage that the battery cannot overcome. The battery's cycling frequency of 86 times per year indicates limited surplus available for storage, which further constrains its ability to recover the deficit.
- For H3, the baseline has no feed-in credit, so no value is forgone. The switch to Renew adds a new revenue stream (€72 in feed-in credits from the 5 kWh battery) while reducing grid imports by 422 kWh. The battery cycles 87 times per year, indicating effective utilisation of

available surplus. The net saving of €79 demonstrates that households with low export volumes benefit most from switching to dynamic pricing with storage.

- For H10, although the baseline has a feed-in credit, the export volume is modest (323 kWh) and the forgone value (€60) represents only 3 per cent of the baseline cost. The high consumption volume (6,986 kWh) amplifies the arbitrage benefits, with the 5 kWh battery cycling 101 times annually — the highest among all PV households. The net saving of €102 demonstrates that households with high consumption benefit from dynamic pricing even when some feed-in value is sacrificed.

The following Figure 4.20 illustrates absolute vs percentage savings by battery configuration.



Household	Capacity	Abs Saving (€)	Pct Saving (%)
H1	5 kWh	-93.4	-12.7%
H1	10 kWh	-85.3	-11.6%
H1	15 kWh	-81.9	-11.1%
H3	5 kWh	+78.7	+6.6%
H3	10 kWh	+86.2	+7.2%
H3	15 kWh	+92.2	+7.7%
H10	5 kWh	+73.9	+4.0%
H10	10 kWh	+92.6	+5.0%
H10	15 kWh	+102.4	+5.5%

Figure 4.20 Absolute vs percentage savings by battery configuration PV households

The following Table 4.26 presents the comparison between Renew tariff and the baseline for the three houses without photovoltaic system.

Table 4.26 Non-PV household comparison - best renew vs baseline

HOUSEHOLD	BASELINE COST (€)	BASELINE AVG. PRICE (€/kWh)	RENEW NO BATTERY COST (€)	RENEW NO BATTERY AVG. PRICE (€/kWh)	DIFFERENCE (€)	RENEW WINS
H8	1822,9	0,24	1929,6	0,26	+106,7	No
H18	1839,5	0,28	1858,4	0,28	+18,8	No
H19	1663,9	0,29	1688,4	0,29	+24,5	No

All three non-photovoltaic households exhibit a consistent pattern: the Renew tariff, whether with or without battery storage, fails to outperform the optimal static tariff. The consistency is striking given the variation in consumption levels and baseline average prices. The loss from switching to Renew no battery is approximately the product of the average price differential and the consumption volume. For H8: (0.256 - 0.240) × 7,548 ≈ €121 (close to the observed €107). For H18: (0.286 - 0.282) × 6,500 ≈ €26 (close to €19). For H19: (0.296 - 0.291) × 5,700 ≈ €29 (close to €24). The high-consumption household (H8) is penalised more heavily because even a small price differential multiplies across a large energy base. The battery impact on non PV households in shown in Table 4.27 while the cost structure comparison is presented in Figure 4.21.

Table 4.27 Battery impact on non-PV households (15 kWh configuration)

HOUSEHOLD	RENEW NO BATTERY COST (€)	RENEW +15 kWh COST (€)	CHARGE (€)	BATTERY CYCLES	FEED-IN CREDIT (€)
H8	1929,6	1953,5	+23,9	61	108
H18	1858,4	1902,3	+43,9	61	104
H19	1688,4	1719,5	+31,0	61	107

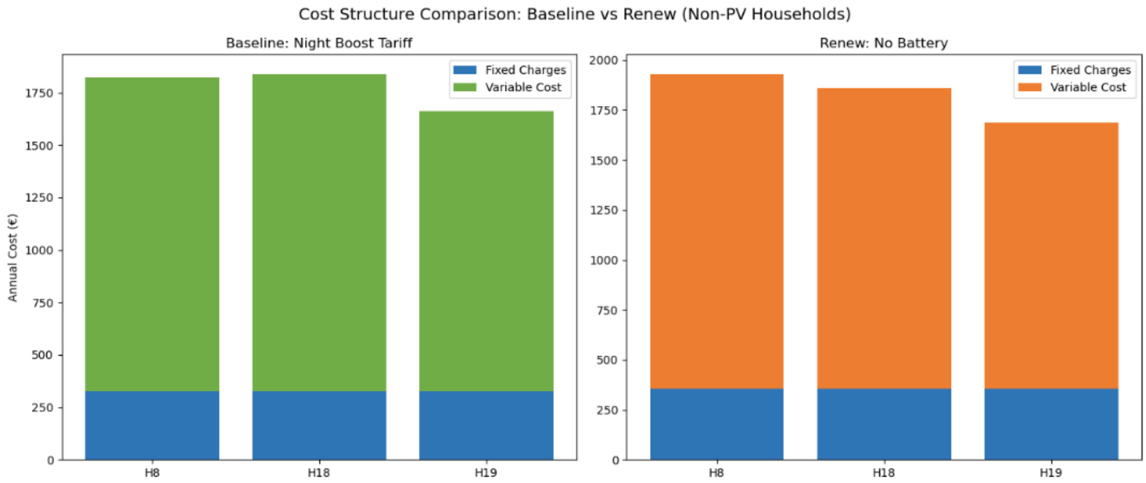


Figure 4.21 Cost structure comparison: baseline vs renew (non-PV households)

In all three cases, adding a battery increases annual costs relative to the Renew no-battery scenario. The battery generates feed-in credits of €104-108 through discharging during high-price periods, but these credits are insufficient to offset the additional costs incurred during charging. The fundamental obstacle is the round-trip efficiency of 90 per cent, which imposes an 11.1 per cent energy penalty on every cycle (since 1/0.9 ≈ 1.111). For arbitrage to be profitable, the average discharge price must exceed the average charge price by more than 11.1 per cent. Analysis of the 2025-2026 DAM price series reveals that while such differentials occur on specific days, the weighted average over the full year falls below this threshold. In the following Table 4.28 the inter-category comparison is reported.

Table 4.28 Inter-category comparison summary

CATEGORY	HOUSEHOLD	BASELINE FEED-IN VALUE	RENEW BEST Δ	RENEW WINS	MAIN ECONOMIC DRIVER
PV	H1	162	+82	No	High feed-in value sacrificed
PV	H3	0	-92	Yes	No feed-in to sacrifice; battery adds value
PV	H10	60	-102	Yes	Modest feed-in; high consumption amplifies arbitrage
Non-PV	H8	0	+107	No	No solar; pure arbitrage unprofitable
Non-PV	H18	0	+19	No	No solar; pure arbitrage unprofitable
Non-PV	H19	0	+24	No	No solar; pure arbitrage unprofitable

The following chart (Figure 4.22) presents the average import price by household and configuration.

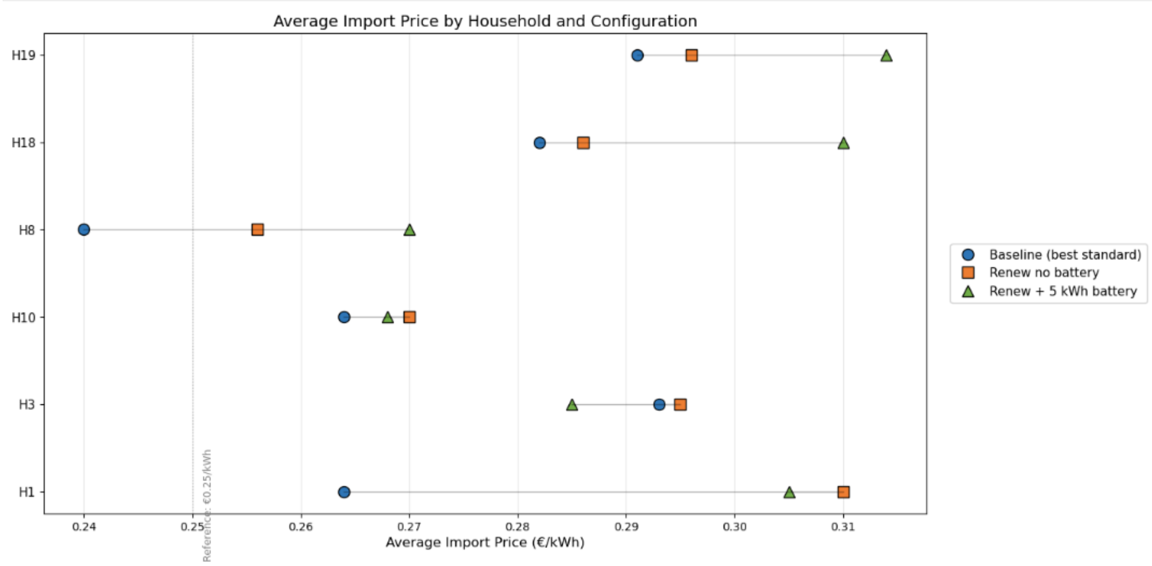


Figure 4.22 Avarage import price by household and configuration

The inter-category comparison reveals that the presence of photovoltaics is a necessary but not sufficient condition for the Renew tariff with battery to outperform the best static tariff. Two of the three PV households benefit from switching, while the third does not. All three non-PV households are worse off under Renew regardless of battery configuration. The fundamental difference lies in the number of value streams available to the battery. For photovoltaic households, battery storage serves two economic functions. The first is self-consumption enhancement: surplus solar production that would otherwise be exported at the DAM price (approximately €0.109/kWh) is stored and subsequently used to avoid imports that would otherwise be purchased at the retail rate (DAM price plus network charges, approximately €0.155-0.180/kWh). The net benefit per kWh stored through this channel is the difference between the avoided import cost and the foregone export revenue, approximately €0.046-0.071/kWh, even after accounting for the 10 per cent round-trip efficiency loss. This channel is reliably profitable because the network charges provide a buffer that exceeds the efficiency penalty. The second function is price arbitrage: the battery charges from the grid during periods of low DAM prices and discharges during periods of high DAM prices, capturing the price differential. This channel is profitable only when the discharge price exceeds the charge price by more than 11.1 per cent to overcome the efficiency loss. For non-photovoltaic households, only the second channel (price arbitrage) is available, as the first channel requires surplus solar production. The analysis of 2025-2026 DAM price data demonstrates that the second channel alone is not sufficiently profitable under current Irish market conditions. The price differentials necessary for profitable arbitrage occur too infrequently or are too small to overcome the cumulative efficiency loss over a full year. The exceptional case of H1 demonstrates that even within the PV category, the self-consumption enhancement channel can be insufficient to overcome large sacrifices in feed-in value. H1's baseline feed-in credit of €0.185/kWh is substantially higher than the DAM price of €0.109/kWh, creating a forgone value of €0.076 per exported kWh. On an export volume of 875 kWh, this sacrifice amounts to €67 annually. When combined with a standing charge that is €85 lower than Renew's fixed charges, the total disadvantage of €152 exceeds the combined benefit that the battery can provide through both channels, given its limited cycling frequency of 86 times per year. The annual saving/loss vs baseline household and battery has been reported in Figure 4.23.

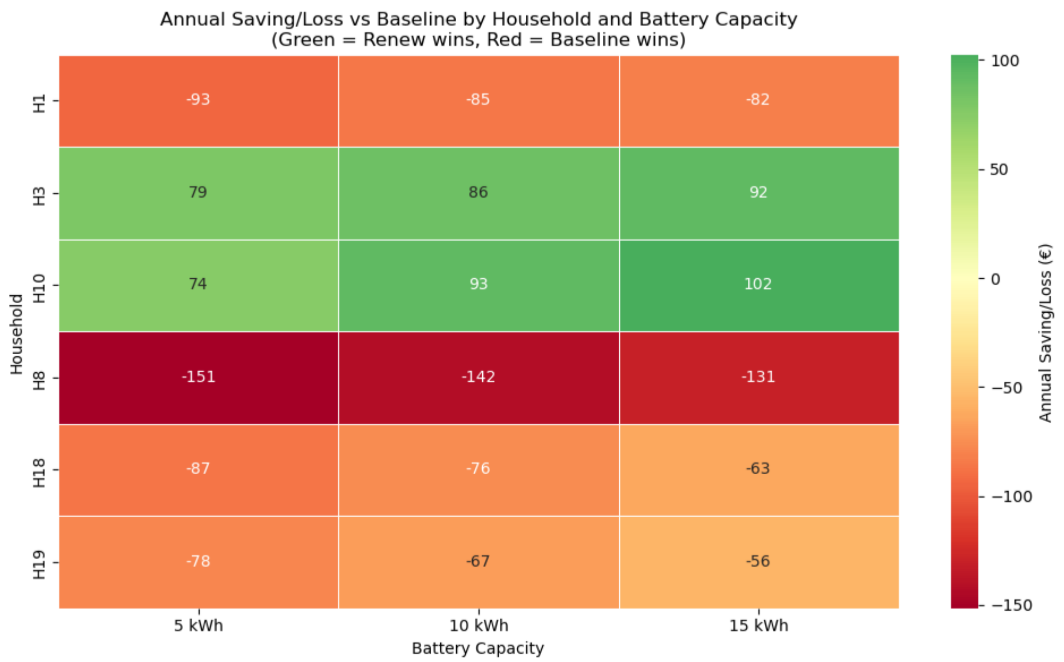


Figure 4.23 Annual saving/loss vs baseline by household and battery capacity

4.4.4 Synthesis of key findings

The analysis yields five principal findings:

- First, standard tariff optimisation alone generates substantial savings. Across all six households, selecting the optimal static tariff reduces annual costs by €512 to €1,575 compared to the most expensive tariff available. For non-photovoltaic households, this is the single most effective economic intervention, requiring no technology investment.
- Second, the Renew tariff without battery is never optimal for any household in the study. The volatility of Day-Ahead Market prices, combined with network charges and fixed fees, makes the dynamic tariff less competitive than well-designed static time-of-use tariffs for households without storage. The average import price under Renew no battery exceeds the baseline average price for all six households, with the disadvantage ranging from €18 to €164 annually.
- Third, battery storage provides positive economic returns for photovoltaic households with low export volumes (H3 and H10) but not for those with high export volumes and generous feed-in rates (H1). The critical discriminating variable is the feed-in value that would be sacrificed by switching from a static feed-in tariff to the DAM-based feed-in of Renew. For households like H3 that have no feed-in credit in their optimal static tariff, the switch to Renew adds a new revenue stream without sacrificing any existing benefit. For households like H1 that have generous feed-in credits, the sacrifice can exceed the combined benefit of self-consumption enhancement and arbitrage.
- Fourth, for non-photovoltaic households, battery storage for pure arbitrage is economically unprofitable under 2025-2026 Irish Day-Ahead Market conditions. The 90 per cent round-trip efficiency penalty requires price differentials exceeding 11.1 per cent to break even, and such differentials do not occur sufficiently frequently or with sufficient magnitude to overcome the cumulative efficiency loss over a full year. Adding a battery to the Renew tariff increases annual costs by €24-44 relative to the no-battery Renew case, and all Renew configurations lose relative to the optimal static tariff.
- Fifth, the presence of photovoltaics is a necessary but not sufficient condition for the economic viability of battery storage under dynamic pricing. Two of the three PV households in the study benefit from Renew with battery, while the third does not. The successful cases are characterised by low export volumes (and therefore low feed-in value to sacrifice) or high consumption volumes that amplify arbitrage benefits. The unsuccessful PV case is characterised by high export volumes and an exceptionally generous feed-in rate under the static tariff, which creates a large disadvantage that the battery cannot overcome.

Overall, the results indicate that the economic value of residential battery storage depends more strongly on the interaction between tariff structure and export behavior than on battery capacity itself. Across all analyzed households, increasing battery capacity beyond 5kWh produced progressively smaller benefits, highlighting that tariff design and household consumption patterns are primary determinants of economic performance. These findings have clear implication for policy and future research. Policy incentives for residential battery storage should be targeted at households with existing photovoltaic systems and low export volumes, as these are the contexts in which storage provides positive economic returns under current market conditions. For non-photovoltaic households, the optimal intervention is the promotion of time-of-use tariffs with low night rates, which provide load-shifting benefits without the capital cost or efficiency losses associated with storage. Future research should examine whether increasing renewable penetration and the resulting growth in price volatility could improve the profitability of pure arbitrage, and whether alternative battery control strategies could capture additional value.

5. CONCLUSIONS

The increasing penetration of distributed generation and the ongoing evolution of electricity markets toward more dynamic structures are creating the need for tools capable not only of estimating photovoltaic energy production, but also of characterizing system performance, identifying key technical parameters, and assessing their economic impact within modern energy systems. Within this context, the present thesis has developed an integrated methodology aimed at the modeling, calibration, and economic assessment of residential photovoltaic systems using real-world data. The first significant contribution of this work concerns the development of a physics-based model capable of estimating the electrical output of a photovoltaic system starting from solar irradiance data. Validation against the results provided by PVGIS demonstrated a high degree of accuracy, as confirmed by the obtained correlation coefficients and error metrics. These results indicate that the proposed model is able to reliably reproduce the energy behavior of photovoltaic systems, thus providing a robust foundation for more advanced applications. A second contribution concerns the validation of the model under real operating conditions. Unlike simulation environments or synthetic datasets, the case study based on data collected in Dingle, Ireland, introduced challenges that are typical of real-world applications, including the incomplete or non-uniform availability of the variables required for the analysis. In particular, while high-resolution measurements of photovoltaic generation were available, irradiance data with the same temporal granularity were not directly accessible. This limitation was addressed through a backward reconstruction procedure, which enabled the derivation of an irradiance profile from the measured production of a reference photovoltaic system. This reconstructed irradiance profile was subsequently used as input to the model, allowing the estimation of photovoltaic production for the other dwellings included in the analysis. The value of this result lies not only in the accuracy achieved, but also in demonstrating that the proposed model maintains its predictive capability even when applied to real datasets affected by the imperfections and uncertainties that inevitably characterize field measurements. The robustness observed during this validation phase represents one of the key elements supporting the practical applicability of the methodology. Perhaps the most significant contribution of the present work is the development of an automated calibration procedure for photovoltaic systems. In practice, information regarding installed photovoltaic systems is often incomplete, outdated, or entirely unavailable. The possibility of identifying the characteristics of a photovoltaic installation using only measured energy data is therefore of considerable interest for both network operators and energy management stakeholders. The proposed methodology demonstrated that it is possible to reconstruct, for each individual dwelling, the incident irradiance profile through a backward approach, without relying on the simplifying assumption that all users experience exactly the same irradiance conditions. Starting from these reconstructed irradiance profiles, the main characteristic parameters of each photovoltaic system were calibrated, including nominal peak power and temperature-related coefficients, yielding physically meaningful and mutually consistent results. The most important aspect of this contribution is not merely the estimation of system parameters, but rather the ability of the procedure to automatically identify the presence and characteristics of real photovoltaic installations using only limited information. From a broader perspective, such an approach could be applied on a large scale to characterize distributed photovoltaic assets, helping to address one of the major information gaps that still affect modern electricity networks. The final stage of the research extended the analysis from the technical domain to the economic one, with the objective of investigating how the energy characteristics of residential users influence the profitability of different electricity procurement strategies. The evaluation of real electricity tariffs available on the Irish retail market made it possible to identify the most economically advantageous options for each dwelling, highlighting how tariff convenience strongly depends on the specific characteristics of the underlying energy profile. Particularly noteworthy was the comparison between conventional fixed-price tariffs and the dynamic tariff RENEW, developed using hourly prices from the Irish Day-Ahead Market (DAM). The results revealed that the attractiveness of a dynamic tariff is not determined solely by market price levels, but rather by the user's ability to interact with those price fluctuations through local energy production. The households equipped with photovoltaic systems showed a greater capacity to benefit from the opportunities offered by dynamic pricing mechanisms, making RENEW competitive for several photovoltaic households, particularly those characterized by low export volumes or high electricity consumption. Conversely, for households without distributed generation, this advantage was largely absent. This finding highlights a particularly relevant aspect: the economic value of energy flexibility increases with the user's ability to generate, self-consume, and actively manage energy resources. Taken together, the results obtained throughout this thesis reveal a common underlying theme. The main novelty of the proposed framework lies in the integration of physical modeling, inverse parameter identification and economic assessment within a single workflow applied to real residential data. Photovoltaic generation should not be viewed solely as a renewable source of electricity, but also as a valuable source of information from which knowledge can be extracted regarding system behavior, user characteristics, and opportunities within modern energy markets. In this sense, the contribution of the present work extends beyond the mere prediction of photovoltaic production and becomes part of the broader process of energy system digitalization. The proposed methodology demonstrates how properly processed energy data can transform seemingly disconnected measurements into a powerful tool for analysis, system identification, and decision support. The ability to estimate photovoltaic production, reconstruct system characteristics, and assess the economic performance of different tariff structures provides a concrete example of how information itself can become a key resource in the management of future electricity systems. In a context characterized by the growing penetration of renewable energy sources, increasing electrification of energy demand, and the emergence of increasingly dynamic market structures, methodologies of this kind may play a strategic role in enabling a more efficient, sustainable, and informed management of energy resources.

5.1 Future developments

Several research directions emerge naturally from the work presented in this thesis and align with the ongoing transformation of the energy sector. A first development concerns the extension of the proposed methodology to larger datasets encompassing different user categories, photovoltaic technologies, and climatic conditions. Such an extension would allow for a more comprehensive assessment of the generalizability and scalability of the proposed approach. Another promising area of research involves the integration of battery energy storage systems. The growing adoption of residential batteries introduces an additional degree of flexibility in energy management, enabling users to better adapt their consumption patterns to market conditions while maximizing the self-consumption of renewable energy production. Further developments may involve the application of advanced artificial intelligence and machine learning techniques to the joint forecasting of photovoltaic generation, electricity demand, and market prices. The combination of physics-based models, such as those developed in this work, with data-driven approaches could lead to even higher levels of predictive accuracy and support the implementation of fully automated energy management systems. Finally, the transition toward renewable energy communities, local energy markets, and smart grids will increasingly require tools capable of automatically identifying and characterizing distributed energy resources. In this scenario, the methodologies developed in this thesis could serve as a foundation for advanced monitoring and optimization platforms, contributing to the transformation of the growing availability of energy data into tangible value for the transition toward a more sustainable, resilient, and participatory energy system. More broadly, this work demonstrates how the combination of physical modeling, data analytics and economic evaluation can support the transition from passive electricity consumers to active participants in increasingly decentralized and data-driven energy systems.

6. BIBLIOGRAPHY

[1] Photovoltaic Technology Review: History, Fundamentals and Applications. (https://www.mdpi.com/1996-1073/15/5/1823).

[2] Progress of PV cell technology: Feasibility of building materials, cost, performance, and stability, M.A. Fazal and Saeed Rubaiee.

[3] Marco Fossa. Conversione fotovoltaica, appunti del corso, Genova.

[4] Illustration of a solar cell. A depletion layer separates two layers of... | Download Scientific Diagram

[5] https://www.researchgate.net/figure/characteristics-curve-of-a-PV-cell_fig2_313450163

[6] Monitoraggio di impianti fotovoltaici innovativi e tradizionali installati sul territorio italiano, tesi_SusannaGhelfi.pdf

[7] Bypass Diode for Solar Panel Protection (https://www.alternative-energy-tutorials.com/photovoltaics/bypass-diode.html).

[8] Solar Cells: Size, Process and Technology Explained (https://www.maysunsolar.com/solar-cells-size-process-and-technology-explained/)

[9] Energia Sostenibile in Irlanda: un Modello da Seguire di Beatrice Maggi (Energia Sostenibile in Irlanda: un Modello da Seguire | Radio Dublino)

[10] In Irlanda 80% di elettricità rinnovabile e 600.000 pompe di calore entro il 2030 | QualEnergia.it

[11] Climate Action Plan 2023 - Summary Document from Department of Climate, Energy and the Environment (Climate Action Plan 2023, climate-action-plan-2023-summary-document-a52e0ba9-c8b7-4766-8ad4-785c25eda404.pdf)

[12] Annual energy flow (https://www.seai.ie/data-and-insights/seai-statistics/energy-use-overview)

[13] Primary inputs to electricity generation (https://www.seai.ie/data-and-insights/seai-statistics/electricity)

[14] Renewable energy share in electricity https://www.seai.ie/data-and-insights/seai-statistics/renewables

[15] https://www.seai.ie/data-and-insights/seai-statistics

[16] Statistical Data of Energy Generation| Terna Driving Energy data

[17] International Journal of Energy Economics and Policy, Markus Flaute, Anett Großmann, Christian Lutz, Anne Nieters

[18] Renewable and citizen energy communities (https://www.ped-interact.eu/about/)

[19] Energy community concept RES4CITY Project (MC19.1.3.1 - Energy community concept)

[20] Energy community concept RES4CITY Project (MC19.1.3.2 - Main benefits of energy communities)

[21] Energy community concept RES4CITY Project (MC19.1.3.3 - Barriers for their formation and resilience)

[22] Vertex S backsheet monocrystalline module

[23] Rohit trivedi , Mohamed Bahloul, aziz Saif, Sandipan Patra & Shafi Khadem 'Comprehensive Dataset on Electrical Load Profiles for Energy Community in Ireland' and file excel Dati case minuto per minuto completo 1.xlsx, Maynooth University

[24] File excel renew-current-live-tariffs-2026-05-15-05-02-37-manual-edits.xlsx, Maynooth University

[25] File excel ie_dam_30min_consolidated_2025-03-01_to_2026-02-28_fully_filled_enriched_holidays_via_cli 2.xlsx, Maynooth University

7. ACKNOWLEDGEMENTS

Ringrazio i relatori Prof. Mattia De Rosa e Prof. Fabiano Pallonetto per il supporto fornitomi durante la stesura della tesi e per avermi dato l’opportunità di poter scrivere la tesi in Irlanda, un’esperienza che porterò con me per sempre.