Thesis reference number: XXXXX

University of Science and Technology of Hanoi

Doctoral School

Thesis title

Nearly zero Energy Buildings:

Energy modeling and Optimal control with Internet of Things application for Buildings

By: Vu Thi Tuyet Hong

Department of Technology and Engineering, USTH

The thesis has been successfully defended on Date 10/6/2024 in front of jury composed of:

Assoc. Prof. Tran Thanh Son, Electric Power University

Assoc. Prof. Nguyen Duc Tuyen, Hanoi University of Science and Technology

Prof. Chong Zhun Min, National University of Singapore

Dr. Hoang Trung Kien, University of Science and Technology of Hanoi

Dr. Pham Xuan Tung, University of Science and Technology of Hanoi

Prof. Nguyen Dinh Quang, Institution of Energy and Science, Supervisor

Prof. Benoit Delinchant, Université Grenoble Alpes, Co-Supervisor

Hanoi, 2024

Thesis title

Nearly zero Energy Buildings:

Energy modeling and Optimal control with Internet of Things application for Buildings

Table of contents

Acknowledgement

To complete this dissertation, I am particularly grateful to Prof. Benoit Delinchant and Dr Nguyen Dinh Quang for supervising me during my PhD research. They provided insightful suggestions on the structure and content of my thesis report, along with helpful feedback.

I want to say special thanks to Prof. Benoit Delinchant, who guided me in realizing my PhD plan and opened my mind to the ideas and the necessary groundwork for this dissertation, which I wish to develop further in my future research.

I would like to thank Dr Nguyen Dinh Quang, who gave me the necessary support, training, and shared his experiences in energy topics, for all the patience he had with me.

Many thanks to Dr Dang Hoang Anh for his advice, support, and cooperation in building energy projects in Vietnam.

Thanks to USTH, HAUI, and G2Elab for providing me with excellent research conditions to complete this dissertation.

Thanks to all my friends and colleagues in France and Vietnam.

Many thanks to Dr Nguyen Dieu Linh, Dr Nguyen Trung Kien, and Dr Phan Anh Tuan for their support, cooperation, and friendliness.

Many thanks to my friend Mr Nguyen Hoang Minh for his encouragement, helpful advice, and for sharing my study.

Thanks to my students for our meaningful time in the IoTeam lab. Many thanks to Tansi, Je Gome, Phuong Uyen, Quynh, Quang Anh, Quang Hung, and all my friends who supported my studies and lived for me in France.

My warm thanks to friends and colleagues at G2Elab in France, USTH, VHH, Vilandco, and IES in Vietnam for their support and collaboration.

I would like to thank my dissertation committee for accepting me to attend my dissertation defense and for their insightful comments on completing my research.

Special thanks to my wonderful parents, my husband (anh Cong), my daughter (Hoang Anh), my son (Hoang Bach), and other family members for their love, support, and motivation during my studies.

Hanoi, October 2024

Vu Thi Tuyet Hong

List of Figures

List of Tables

Acronyms

IEA	International Energy Agency
BMS	Building energy system
BEM	Building energy management
nZEB	Nearly Zero Energy Building
GDP	Gross domestic product
USAID	United States Agency for International Development
EUI	Energy use intensity
IoT	Internet of Things
RF	Radio Frequency
WSN	Wireless Sensor Network
LCS	Low cost sensor
SPI	Serial Peripheral Interface
I2C	Inter-Integrated Circuit
UART	Universal asynchronous receiver-transmitter
MQTT	Message Queuing Telemetry Transport
GPIO	General-purpose input/output
OpenHAB	Open Home Automation Bus
GPR	Gaussian process regression
MLR	Multi Linear Regression

Abstract

Nearly Zero-Energy Buildings (nZEBs) play an essential role in reducing energy consumption and CO_2 emissions. However, the nZEB application faces technical issues, investment costs, and the complexity of user behavior, building architecture, and grid constraints. Vietnam has a high potential for rooftop solar development. New energy policies encourage on-site self-consumption and expand the application of nZEBs. In the country, the construction sector has developed significantly over the years, but research on building energy has been limited due to insufficient building data.

Based on a practical approach, a roadmap for a low-cost IoT platform for Building Energy Management was developed. It uses IoT-based WSN architecture, open hardware and open-source solutions to reduce costs and technical barriers and explicitly considers user behavior, which strongly influences system adoption and performance.

This study proposed a two-level data quality framework for a low-cost IoT-BEMS to improve the reliability of analytics and control services. At the single-sensor level, it uses accuracy tests and virtual sensors, and at the network level, it applies AI to impute missing data in real time. Additionally, this work introduced lightweight energy models and nonlinear optimal control algorithms that can run on low-computing systems (such as Raspberry Pi). These solutions are tested in real buildings with technologies and grid constraints. The lessons from experimental case studies in France and Vietnam are summarized to guide the adaptation of IoT solutions to local contexts.

Overall, the study contributed to the development of low-cost IoT-BEMS platforms towards widespread nZEB applications in Vietnam.

Keywords: Buildings, nZEBs, WSN, IoT, energy modelling, optimal control.

Tóm tắt

Công trình tiêu thụ năng lượng tiệm cận không (nZEB) đóng vai trò quan trọng trong việc giảm mức tiêu thụ năng lượng và lượng khí thải CO_2. Tuy nhiên, việc ứng dụng nZEB đang gặp phải các vấn đề kỹ thuật, chi phí đầu tư, sự phức tạp trong hành vi người dùng, kiến trúc công trình và các hạn chế về lưới điện. Việt Nam có tiềm năng lớn để phát triển điện mặt trời áp mái. Các chính sách năng lượng mới khuyến khích tự tiêu thụ tại chỗ và mở rộng ứng dụng nZEB. Mặc dù, ngành xây dựng đã phát triển đáng kể trong những năm qua, nhưng nghiên cứu về năng lượng tòa nhà còn hạn chế do thiếu dữ liệu về tòa nhà.

Dựa vào tiếp cận thực tiễn, một lộ trình cho nền tảng IoT chi phí thấp dành cho Quản lý Năng lượng Tòa nhà đã được phát triển. Nó sử dụng kiến trúc mạng cảm biến không dây dựa trên công nghệ IoT, phần cứng mở và các giải pháp mã nguồn mở để giảm chi phí và rào cản kỹ thuật, đồng thời xem xét rõ ràng hành vi người dùng, yếu tố ảnh hưởng mạnh mẽ đến việc áp dụng và hiệu suất của hệ thống.

Nghiên cứu này đề xuất một khung chất lượng dữ liệu hai cấp cho hệ thống IoT-BEMS chi phí thấp nhằm cải thiện độ tin cậy của các dịch vụ phân tích và điều khiển. Ở cấp độ cảm biến đơn, nó sử dụng các bài kiểm tra độ chính xác và cảm biến ảo, và ở cấp độ mạng, nó áp dụng trí tuệ nhân tạo (AI) để điền dữ liệu bị thiếu trong thời gian thực. Ngoài ra, nghiên cứu này đã giới thiệu các mô hình năng lượng gọn nhẹ và các thuật toán điều khiển tối ưu phi tuyến có thể chạy trên các hệ thống tính toán thấp (như Raspberry Pi). Các giải pháp này được thử nghiệm trong các tòa nhà thực tế với các công nghệ và hạn chế về lưới điện. Các bài học từ các nghiên cứu trường hợp thực nghiệm ở Pháp và Việt Nam được tóm tắt để hướng dẫn việc áp dụng các giải pháp IoT vào bối cảnh địa phương.

Nhìn chung, nghiên cứu này đã đóng góp vào sự phát triển của các nền tảng IoT-BEMS chi phí thấp hướng tới ứng dụng rộng rãi trong các tòa nhà năng lượng gần bằng không (nZEB) tại Việt Nam.

Từ khóa: Tòa nhà, nZEB, WSN, IoT, mô hình hóa năng lượng, điều khiển tối ưu.

Part 1. Introduction

1.1. Energy context

1.1.1. Energy context in the world

An increase in global energy demand, resource depletion, energy supplies and environmental pressures are currently a concern for governments, organizations and individuals. In 2023, global energy demand is at a very high level, with total global primary energy demand around 640 EJ [1]. According to the APS scenario, fossil fuels (oil, coal and natural gas) still account for around 80% of total primary energy supply, while modern renewables provide only about 12% and nuclear 5% [1]. Although fossil fuels still account for a very high proportion, they are expected to be depleted in a short time, in detail, about 40–50 years for oil, 50–60 years for natural gas, and 100–120 years for coal [2]. The countries need to accelerate the energy transition and improve energy efficiency globally to achieve SDG7 targets and net-zero emissions [3].

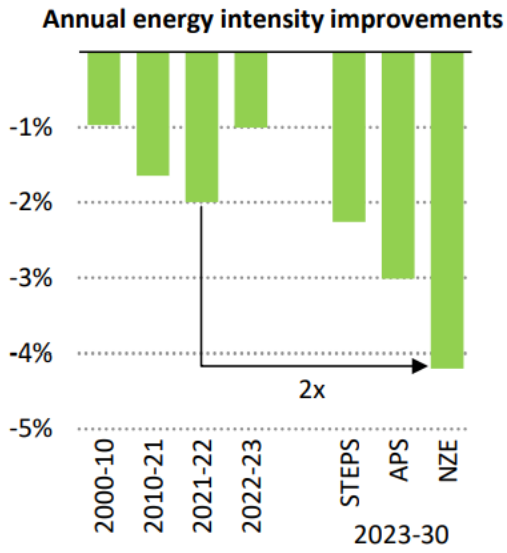


Figure 1. Global annual energy intensity improvements, 2000-2030 [1]

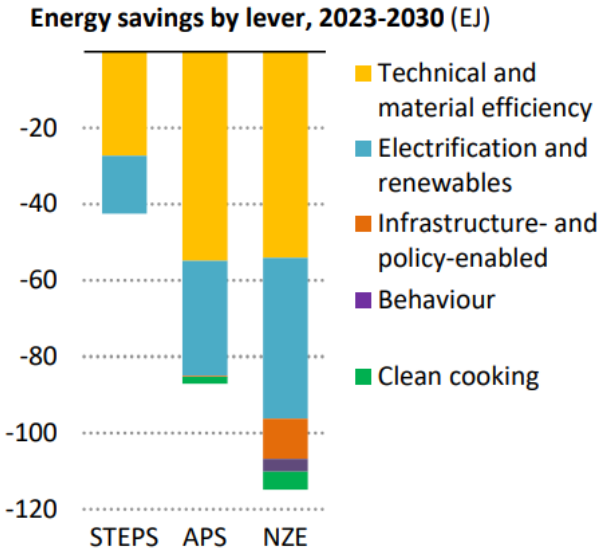


Figure 2. Cumulative energy savings by lever and scenario, 2023-2030 [1]

The data in the IEA report show global efforts in implementing energy efficiency solutions [1]. Figure 1 shows that the pace of energy intensity improvement has slowed to about 1% per year in 2022–2023 [1]. It is still far below the required scenarios (around -2% per year for STEPS, -3% per year for APS, and -4% per

year for NZE) [1]. These values reflect different levels of policy effort on energy savings. The scenarios also differ in their cumulative effect on energy demand. As shown in Figure 2, the total expected energy savings from 2023 to 2030 are about 40 EJ in STEPS, around 90 EJ in APS, and more than 100 EJ in the NZE scenario [1]. In all scenarios, "technical & material efficiency" and "electrification & renewables" make the largest contribution, while the other levers play a supporting role [1]. Together, they point to the key actions needed towards a net-zero energy pathway.

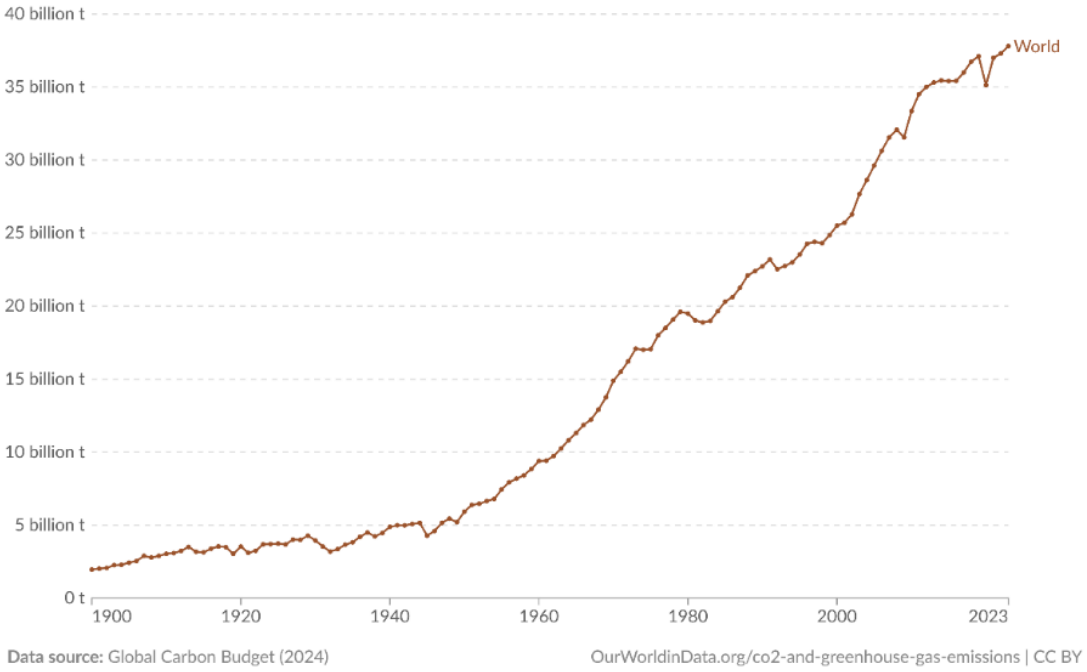


Figure 3. The annual change in CO_2 emissions from 1990-2023 [4]

Figure 3 shows global CO_2 emissions from 1990 to 2023 [4]. They rise almost continuously over this period, reaching about 40 billion tons per year in 2023 [4], with only a slight drop around 2020 due to COVID-19. Overall, emissions remain very high. According to the Energy Progress Report 2024 [5], the world remains far from achieving the 2030 sustainable energy targets. Furthermore, progress in renewable energy use and energy efficiency remains slow [5]. International finance for clean energy remains insufficient and unequally distributed [5]. Therefore, the countries need to deploy stronger solutions to achieve their net-zero energy targets.

1.1.2. Energy context in Vietnam

Vietnam's economy continues to consume energy at a relatively high level compared to the regional and global average [6]. In 2023, total final energy consumption (TFC) was around 2.9 EJ, with oil (~33%), electricity (~31%), and coal (~23%) accounting for the main shares [6]. Industrialization, urbanization, and the rapid expansion of buildings are driving up electricity demand. Between 2000 and 2023, per capita electricity consumption increased by 771% [7]. Electricity demand increased by more than 4% in 2023 and is projected to increase by approximately 7% annually from 2024 to 2026 [8]. Coal remains the primary fuel, accounting for about 45% of electricity production in 2023, leading to about 293 $MtCO_2$ in emissions (5.6 times higher than in 2000) [9], [10]. Renewable energy sources (excluding hydropower) accounted for only about 16% of electricity in 2023 and are expected to increase to 19% in 2026 [9]. However, Vietnam's energy intensity is still high, about 3.2 GJ/1 USD PPP (2015) in 2023, belonging to the high group in the Asia-Pacific region [6].

During 2019–2021, the FIT mechanism supported the development of large-scale solar projects, particularly in provinces with high resource potential, such as Ninh Thuan and Binh Thuan [9],[11],[12]. However, high-density renewable energy deployment has led to local grid overload and congestion, resulting in challenges in reduction and operation [12],[13]. These indicate that Vietnam needs to strengthen the grid promptly, operate the system more flexibly, and place greater emphasis on distributed renewable energy sources [11][12].

In 2024, the Government issued Resolution No. 135, which promotes rooftop solar for self-consumption [13] to reduce pressure on the power grid and emissions. Long-term strategies such as Power Development Plan VIII (PDP8) and national resolutions on climate and energy [14], [15], [16], [17], [18] have set a clear roadmap for Vietnam to build a safer, lower-carbon power system and move toward a net-zero emissions target by 2050.

1.2. Building Energy issues

1.2.1. Buildings energy consumption in the world

In 2022, energy use for operation in building sector accounted for approximately 30% of global final energy consumption (of which the residential sector accounts for 21%). Building floor area increased by approximating 25% between 2010 and 2022 and is expected to keep growing rapidly if no effective control measures [19].

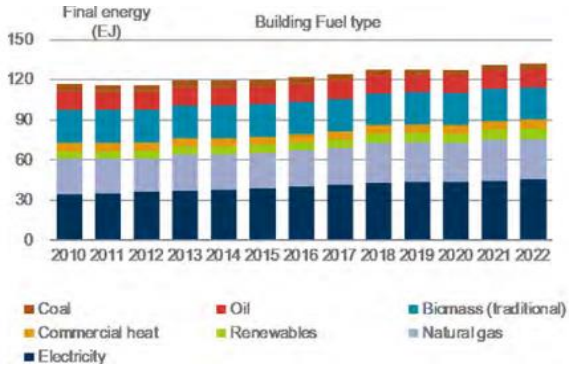


Figure 4. Energy consumption of buildings by fuel type (2010–2022) [20]

In Figure 4, during the period 2010-2022, total energy consumption in buildings increased from over 100 EJ to nearly 130 EJ (an increase of almost 30%) [20]. In this growth, the share of electricity has increased, but coal, oil, and traditional biomass still dominate, while modern renewables have only grown slowly.

Building and construction sector emissions (including "embodied" of new construction) accounted for about 37% of total global energy & industrial process emissions [20]. From 2010 to 2022, total CO_2 emissions from building operations fluctuated little, ranging from 9 to 10 $GtCO_2$ per year. This figure represents only a little more than half of the reduction needed by 2030 [19].

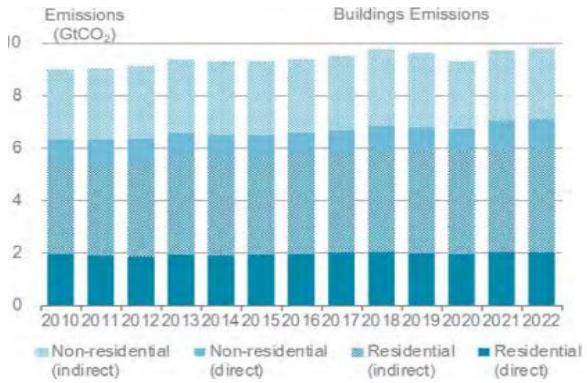


Figure 5. Demonstration of CO_2 emissions in buildings during 2010–2022 period [20].

Figure 5 shows that CO_2 emissions from the group using indirect energy (primarily electricity purchased from the grid) still account for a significantly higher proportion than the other group [20]. These observations indicate that reducing emissions in the building sector requires combining electricity savings with lower building energy intensity and a higher share of clean energy in the national power system.

However, energy intensity in the global building sector has improved very slowly. During 2015 – 2022, it decreased only slightly, from about 153 kWh/m² to ~145 kWh/m². The rate of energy efficiency improvement by 2022 was still less than 50% of the rate required for NZE in 2030 [20]. The contribution of renewables to building energy demand remains limited. In 2022, modern renewables supplied only about 6% of final energy use in buildings, which is roughly one third of the level required in the NZE scenario by 2030 [19].

Meanwhile, global floor area growth is expected to increase by approximately 15% by 2030, with more than half of that increase coming from emerging and developing economies (including Vietnam) [19]. This suggests that future energy demand in the building sector will rise primarily in these regions. Consequently, the policies and technological solutions adopted there will significantly influence global energy trends in the construction industry.

1.2.2. Buildings energy consumption in Vietnam

According to the energy efficiency diagnostic report, buildings in Vietnam are among the most significant sources of electricity consumption. Buildings accounted for about 39% of national electricity consumption, with most coming from residential and administrative buildings (approximately 34%) and commercial buildings (5%) [21].

The IEA's PDP8 scenarios forecast Vietnam's electricity demand to continue to increase sharply, from about 250 TWh in 2022 to 415 TWh in 2030 and 860 TWh in 2050 [12], with the building sector accounting for more than 40% [12]. This shows that buildings are the main component of Vietnam's total energy demand

in the future.

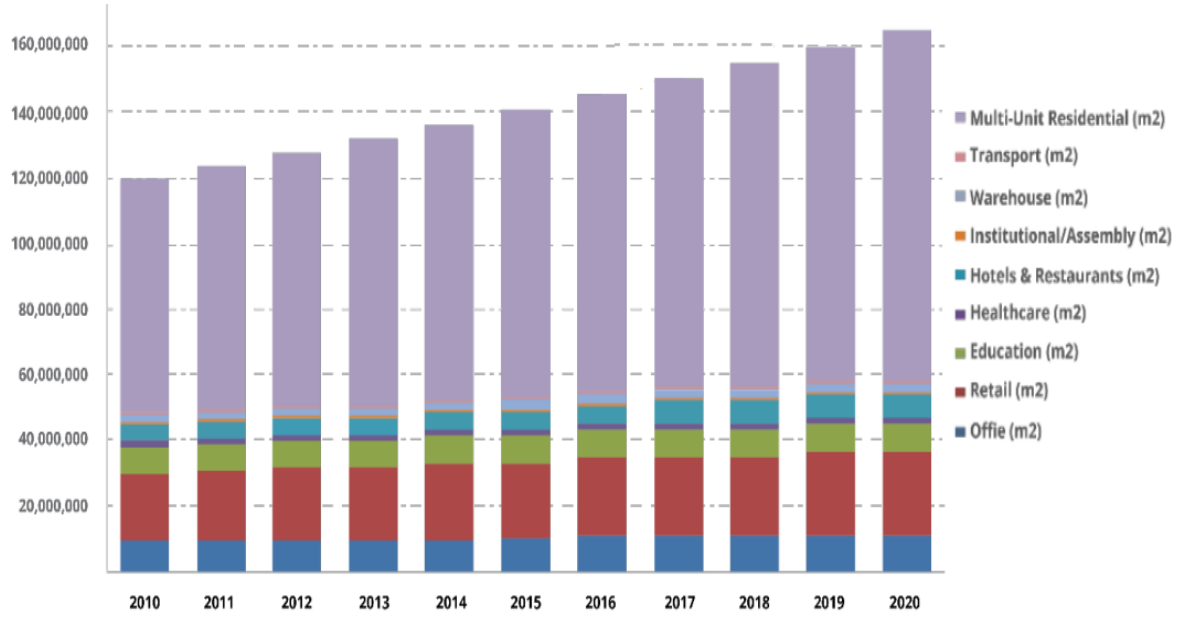


Figure 6. Buildings stock projections growth rate in Vietnam [21].

During 2010-2020, the construction industry grew at nearly 6%/year, higher than GDP growth (~5.5%) [21]. The fact that the construction growth rate has been increasing faster than GDP indicates that the building stock is expanding more rapidly than the overall economy. In particular, multi-unit residential buildings account for the largest share of the floor area (see Figure 6).

Vietnam has issued the National Standard on Energy Efficiency in Construction QCVN 09:2017/BXD and other policies to promote energy-saving and low-carbon construction projects. However, implementation and enforcement still face many challenges [22]. In recent project reports, audit results indicate that many office and commercial buildings operate above recommended standards, indicating significant energy-saving potential [21],[23]. However, data on building energy intensity (EUI) are sparse and incomplete. They mostly come from a few case studies and outdated datasets. This indicates the urgent need to improve measurement, monitoring, and data collection to quantify savings better and design future NZE solutions.

1.2.3. nZEBs – For sustainable energy development in Vietnam

Globally, nZEBs are understood as buildings with very high energy efficiency,

where the remaining energy demand is supplied mainly by on-site or nearby renewable sources [24], [25]. In Vietnam's building energy sector, analyses indicate that building energy consumption is currently high but can be reduced through practical solutions, such as nZEBs. Furthermore, Vietnam has the potential to develop nZEB thanks to its rich solar resources. The average solar radiation is up to 5 kWh/m²/day in many regions, and the technical potential of solar power is estimated at hundreds of GW [26].

In the ASEAN region, Vietnam is among the leading countries in deploying rooftop PV in the building sectors. Vietnam has issued numerous policies, including the QCVN 09:2017/BXD standard for energy-efficient buildings, as well as national strategies and programs prioritizing energy saving, renewable energy, and net-zero emissions targets [14], [15], [16], [17], [18]. In particular, Decree No. 135/2024/NĐ-CP, issued in 2024, creates favorable conditions for Vietnamese buildings to self-supply part of their energy demand through on-site consumption [8]. These provide an essential foundation for the future deployment of nZEBs in Vietnam. However, QCVN 09:2017/BXD sets only minimum requirements for energy efficiency and has not yet established a dedicated nZEB standard with target EUI values and mandatory shares of renewable energy. The reason is that the data are incomplete, scattered, and very outdated, making it challenging to design and evaluate nZEBs. Meanwhile, Europe has a relatively clear nZEB definition framework, specified by type and climate zone [26]. Many studies on nZEBs have been conducted successfully in Europe, where buildings are subject to strict energy-efficiency regulations [25], [27]. These studies highlight that the nZEB concept must be adapted to local climatic and economic conditions in each country.

IEA studies [28] indicate that the efficiency of Net ZEB also strongly depends on the level of load and source matching and how the building interacts with the grid. In the studies, calculating metrics such as the Load Match Index and Load/Supply Cover Factor requires detailed measurement data at hourly or even minute intervals.

Overall, to achieve the goal of (nZEB) in Vietnam, building a reliable building monitoring and data collection system is essential. This system will serve as a foundation for assessing the current operational status and proposing effective building energy management solutions.

1.2.4. Building energy management

1.2.4.1. Modeling and optimal control – Need monitoring parts in Buildings

In building energy management, monitored data can be used to more efficiently coordinate energy use from the grid, on-site batteries, and other distributed renewable sources.

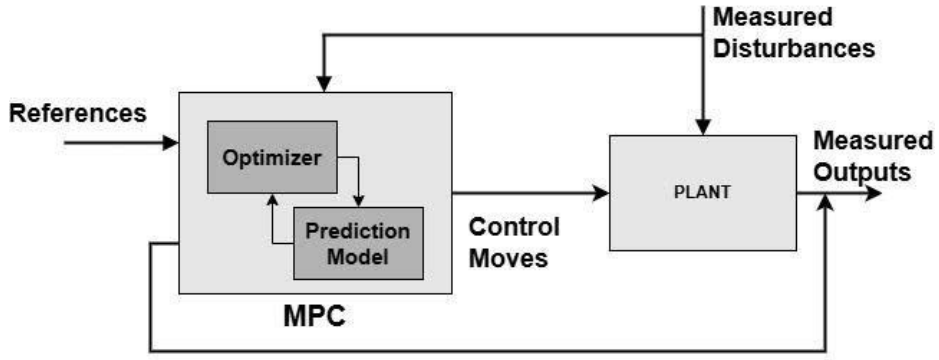


Figure 7. MPC Basic Control Loop (*Source: https://www.mathworks.com*)

In Figure 7, the Model Predictive Control (MPC) principle indicated the need for measured data. Modelling methods and optimization techniques all require sensor data in buildings [29], [30]. The study in Artiges (2016) shows that applying Model Predictive Control (MPC) to improve building energy performance requires high-quality monitoring and measurement devices in the building [30].

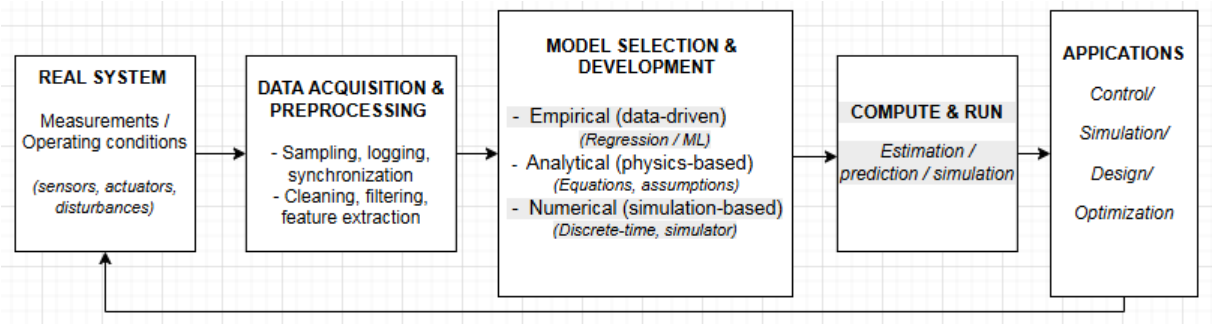


Figure 8. Modeling Topology for Building Design and Operation [2] [29], [30].

Figure 8 illustrates the process of building a model from a real system to application objectives. Data is collected, preprocessed, then the model is selected and developed in three directions (experimental, analytical, and numerical model) [2]. The numerical model poses practical challenges because it is complex and requires difficult to obtain physical parameters. The data-driven models, including both experimental and analytical models, are preferred for their simplicity. The statistical data in [31] indicated that most of the energy models studies use a real monitoring system.

Then, the results are used for computing and applying control, simulation, design, and optimization in practical. There is an iterative process to improve the accuracy and reliability of the real system. In some studies, the monitoring components could enable a specific diagnosis to optimize the building's renovation [32]. Furthermore, it also allows the realization of occupants' behavior models [33]. Monitoring indoor conditions (T, RH, CO_2, energy sensors) and weather conditions should help better understand behavior regarding consumption, consequently, the level of use of PV generation and battery energy storage [34].

1.2.4.2. Monitoring platforms- The state of the art

Monitoring platforms are increasingly fundamental components of building energy management. Recent studies show that energy and environmental measurement data contribute to performance assessment, near-zero energy building (nZEB) design, and advanced control [35], [36].

The current trend is shifting from single-point measurement to multi-point, multi-sensor, high-resolution monitoring methods control [35]. Many buildings now utilize Wireless Sensor Networks (WSNs) and smart electricity meters to monitor energy and indoor conditions. Such systems include sensor nodes, gateways, a central data system, and a user interface [35], [36]. For buildings to interact effectively with the grid, the monitoring systems require an additional bidirectional communication infrastructure with the grid [37].

Numerous studies have proposed “low-cost” monitoring platforms using standard microcontrollers (Arduino, ESP32 [38], Raspberry Pi) and inexpensive sensors [35], [36]. This approach demonstrates the benefits of extensive monitoring across multiple areas within a building at an acceptable budget. Typically, a low-cost system has a total deployment cost under a few thousand US [36], suitable for small to medium sized building applications.

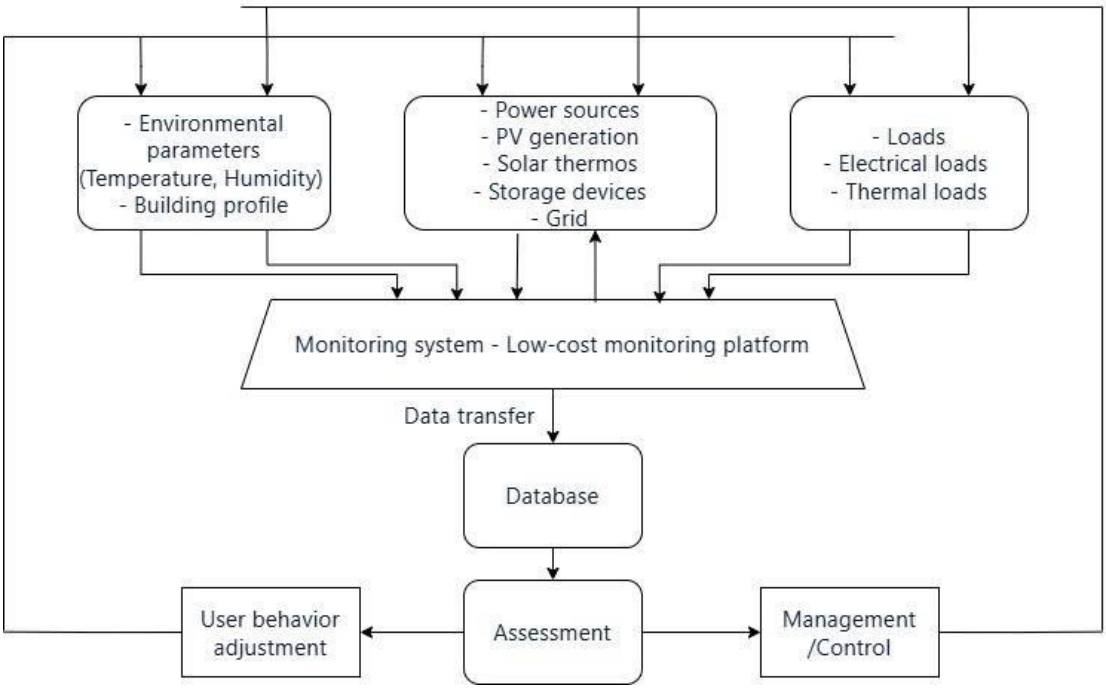


Figure 9. Schematic Diagram of Energy-Efficiency Management [36]

Figure 9 [36] shows that effective energy management requires a “Low-cost monitoring platform” to collect and store data on building operations. Continuous monitoring of energy and environment creates a database that allows the construction of a characteristic consumption profile over time, weather and operating schedule. From there, managers can plan more rational equipment operation and maintenance, improve efficiency, extend equipment lifespan and reduce costs.

However, the use of low-cost monitoring technologies poses challenges for data quality, requiring additional machine learning techniques [39], [40]. Another challenge is that building performance is shaped by occupant behavior. As monitoring systems combine sensor data with user actions, user-centered strategies that engage occupants become essential for practical energy

management [35], [36], [41].

Overall, monitoring systems are providing the primary data source for analyzing, forecasting, and optimizing building energy consumption.

1.2.4.3. Building energy management barriers

Despite recent advanced in monitoring platform, building energy management are still limited adoption in Vietnam. Multiple barriers to energy efficiency technologies have been identified in previous studies [42], and broader challenges have been highlighted in the National Energy Efficiency Program 2019–2030 [43].

The lack of building-level energy management tools is one reason the energy-saving potential has not been fully realized [43]. A significant barrier is the missing and scattered building measurement data. It is available only for a few large buildings. In most small to medium sized buildings, users can access only total meter data and bills through an EVN application. However, they cannot query their energy data in real time. The critical data, such as indoor conditions, equipment status, and user behavior, have not been systematically collected, thereby limiting the ability to evaluate and optimize building operations [36].

The installation of monitoring systems in existing buildings remains constrained by investment costs and user participation [36]. The IEA report [37] shows that in ASEAN (including Vietnam), smart sensors, automation systems, and energy management are now implemented only in pilot projects. The lack of data and user feedback mechanisms remains a significant barrier to transferring information into practical actions (such as adjusting operations or changing energy use habits) [35], [36].

The advancement of IoT technologies provides a practical way to overcome several of these barriers by supporting open ecosystems, low-cost sensing, and improved connectivity in buildings [44], [45]. However, it is still difficult to deploy in practice because the available technology, user skills, and budgets differ cross country.

Therefore, the following sections examine how to deploy a building energy monitoring system towards nZEB in Vietnam.

1.3. Objectives of the thesis

The aim of thesis will focus on Nearly Zero Energy Buildings applications in Vietnam. There are two main objectives of the thesis as below:

Objective 1: Proposing solutions to launch an IoT platform in Buildings at low-cost

- Proposing the low-cost monitoring platform bases on IoT technologies. Allowing low-skill users could easily develop and operate platform through open-hardware and open-sources.
- Proposing methodologies for ensuring measured data quality.
- Public energy efficiency awareness in community by cooperation projects.

Objective 2: Practical energy management for nZEBs:

- The IoT platform implementation.
- Proposing models and optimization technique adapted to the low-cost platform.
- Proposing optimal energy management strategies.

1.4. Challenges to achieve the goals

This work must deal with some challenges such as:

- Low-cost IoT platform and system quality.
- Complicated energy modeling and optimized control methods.
- Energy management strategies adapt to technical constraints, users and energy conditions in local regions.

1.5. Structure of the thesis report

Part 1 – Introduction

- Presents the energy context, with a focus on the building sector, which contributes to energy consumption and CO_2 emissions.
- Review on Building Energy Management Systems to find gaps and challenges.
- States the objectives of the thesis.
- Summarizes the structure of the manuscript.

Part 2 – A practical route for launching an IoT platform for building energy management

- Analyses IoT and wireless sensor network (WSN) technologies for building applications and proposes a low-cost IoT-BEMS architecture.
- Proposes a data quality assurance framework (including single sensor level and sensor network level), including accuracy testing, correlation analysis, fault detection and data compensation using machine-learning models.
- Describes the design, deployment of the IoT-BEMS platform on real testbeds, and provides practical guidelines for non-expert users in Vietnam to implement energy monitoring and basic analytics.

Part 3 – Implementation of the IoT-BEMS platform - A case study

- Implements the proposed IoT platform for BEMS on a real office building (VHH) in Vietnam, using low-cost hardware and time-series data.
- Develops data-driven analyses for environment, user behavior, and Energy.
- Proposes energy strategies relating behavior and analysis.
- Estimating energy efficiency of project

Part 4 – Optimal energy management strategies toward nZEBs

- Introduces simplified energy models for PV production, battery storage and loads.
- Presents optimization problems and algorithms adapted to low-cost computation (e.g. SQP-based methods and the NoLoad tool).
- Applies the modelling and optimization framework to two case studies: an aquaponics greenhouse testbed in France and the VHH office building in Vietnam.
- Investigates energy balance, PV–battery system sizing and operation strategies with TOU, and quantifies self-consumption, electricity bills and CO_2 emissions.

Final part – Conclusions of Thesis

- Summarizes the main scientific and practical contributions of the thesis in IoT-BEMS design, data quality assurance and energy management.
- Discusses the limitations of the current work and future deployment.
- Highlights the collaborations with Vietnamese and French universities and research projects that supported to the thesis.

Part 2. A practical route for launch an IoT platform for building energy management

2.1. Overview

2.1.1. IoT and IoT-BEMS: definition and scope

According to the report of Asian development bank [44], the definition of Internet of Things (IoT) identified as: "*An IoT is a network that connects uniquely identifiable 'Things' to the Internet. The 'Things' have sensing/actuation and potential programmability capabilities. Through the exploitation of unique identification and sensing, information about the 'Thing' can be collected and the state of the 'Thing' can be changed from anywhere, anytime, by anything."*

The term IoT-enabled BEMS (IoT-BEMS) refers to Building Energy Management Systems using connected sensors/actuators and gateways to monitor and control energy related subsystems (including Loads, on-site PV and storage) [45], [46].

The present work focuses on small and medium-sized buildings, optionally equipped with PV and/or battery storage, and typically operated by low-skilled users. In this scope, IoT-BEMS deployments must remain cost-effective and simple enough to be installed and maintained by non-experts, while providing sufficient data for Building energy services. To ground the analysis, several IoT-BEMS case studies in France and Vietnam are considered. These deployments provide the empirical basis for defining and evaluating the platform architecture (Section 2.2) and data quality (Section 2.3).

2.1.2. IoT-BEMS platform architecture: opportunities and challenges

In the BEMS, IoT platforms are organized into multiple layers that connect smart sensors, meters, and controllers to back-end services. Reference IoT architectures for the building sector, such as those discussed in [45], [47], [48] illustrate a familiar pattern: sensors convert analog signals into digital signals via ADC/DAC; gateways collect and pre-process data and communicate with data centers or cloud services, and send back commands to actuators using MQTT; block data center is

configured to manage, analyze, and security in real-time.

From building owners and operators' point of view, the primary motivation is to reduce energy costs while maintaining continuous monitoring and control of their facilities. IoT technologies enable this by connecting physical components through sensors, actuators, and software to exchange data with other systems and end users on a platform. At a larger scale, such platforms also support smart-city integration and emerging energy services [45]. Wireless sensor networks (WSNs) are a core enabler in many IoT architectures, with documented benefits in recent studies [49], [50]: Lower deployment effort and cost; installation flexibility to existing infrastructure; reduced node power consumption thanks to advances in low-power electronics and communication protocols [30], [49].

Future buildings are expected to increase device connectivity to IoT, widely deploy WSN, functionalize components, and adopt advanced control to improve overall performance [44], [45]. However, three practical challenges in the adoption of WSN-based IoT-BEMS in real buildings, especially in small and medium-sized premises operated by low-skilled users:

- Cost: hardware, installation, operation and maintenance must remain acceptable for small and medium-sized buildings.
- Reliability: measurement accuracy, wireless coverage, interference, fault tolerance, battery lifetime and interoperability all affect the stability of data flows and control actions [45], [49]. Poor data quality directly degrades the performance of forecasting, control and energy services.
- Low-skilled users: limited expertise in IoT, technologies and data analytics, which constrains how complex the platform, configuration procedures and diagnostic tools can be in practice [46].

These considerations motivate the simple, interoperable IoT-BEMS architecture adopted in this thesis, based on low-cost WSNs. The following section (Section 2.2) details the system architecture and (Section 2.3) introduces the data quality framework.

2.2. Development of IoT-BEMS platform

This section presents the overall workflow for implementing the proposed low-cost IoT-BEMS, and shows how it is applied in practice in the case studies (see Figure 10).

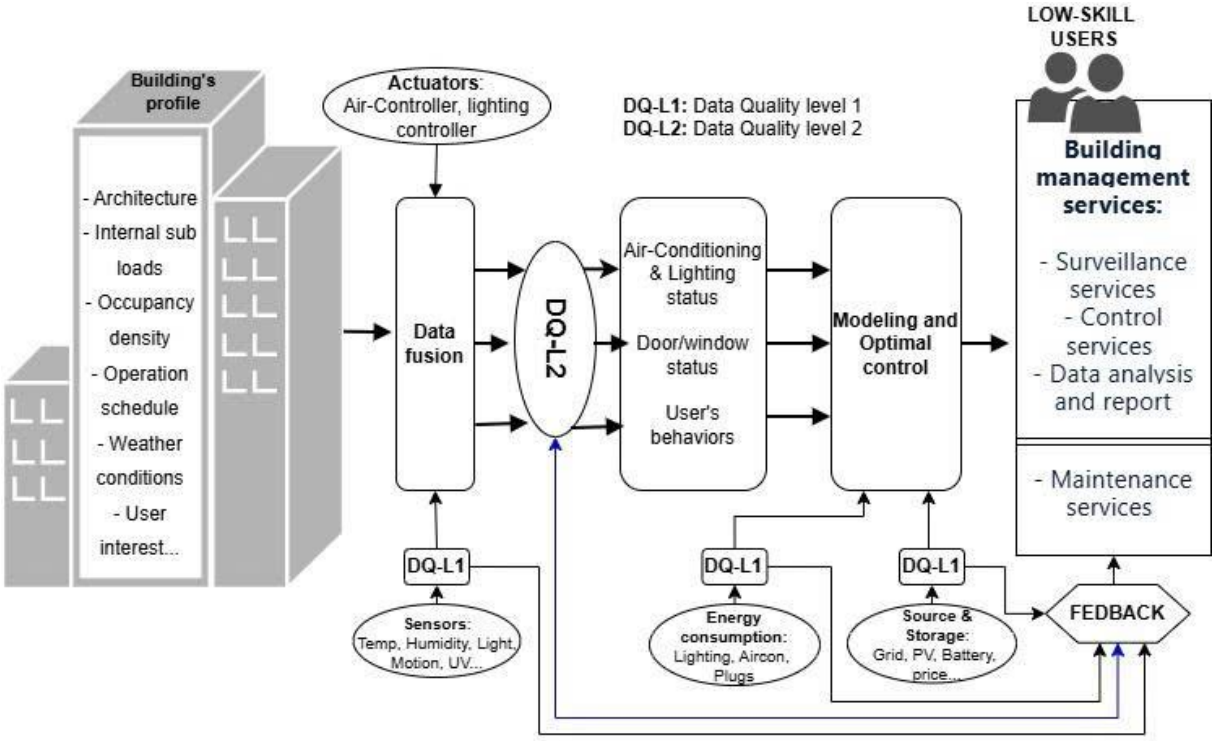


Figure 10. Overall workflow of the proposed low-cost IoT-BEMS implement [36]

In Figure 10, first, the building profile is collected (location, architecture, internal loads, occupancy schedule, weather conditions and user preferences). Low-cost sensors then acquire environmental and energy data such as temperature, humidity, light, motion, PV, battery and grid power. A data-fusion layer aggregates these raw measurements into higher-level states (user behavior, HVAC and lighting status, door/window status). These variables are sent to the modelling and optimal-control layer, which computes commands for actuators (air-conditioner and lighting controllers). In the proposed workflow, the data quality (DQ) framework is embedded in the feedback loop to support maintenance. DQ-L1 outputs (per-sensor fault labels and health indicators) are aggregated in the feedback block and forwarded to DQ-L2 for cross-sensor consistency checks and fault diagnosis. DQ-L2 results then drive maintenance actions

(test/calibrate/replace sensors). After DQ-L2, the fused data become a cleaner, validated stream and are sent to the next block for more reliable processing. The platform finally provides building energy services: monitoring dashboards, maintenance alerts, basic control suggestions and energy reports.

The workflow shows that most of the complexity is handled automatically by the platform. Low-skill users without specialized technologies or energy management interact with the right-hand “services” block, while others run in the background.

2.2.1. System Design

In most projects, the building’s profile is required as the first step in defining the platform's boundaries and conditions. It includes the following information fields: project information includes targets, timeline and budget; Building data includes occupancy, operation schedules, energy consumption/production, and weather conditions; Building architecture includes location, floor area, type, age and function; User priorities includes cost, life-cycle, comfort and environmental concerns; Standard references: international standards and national building energy-efficiency codes.

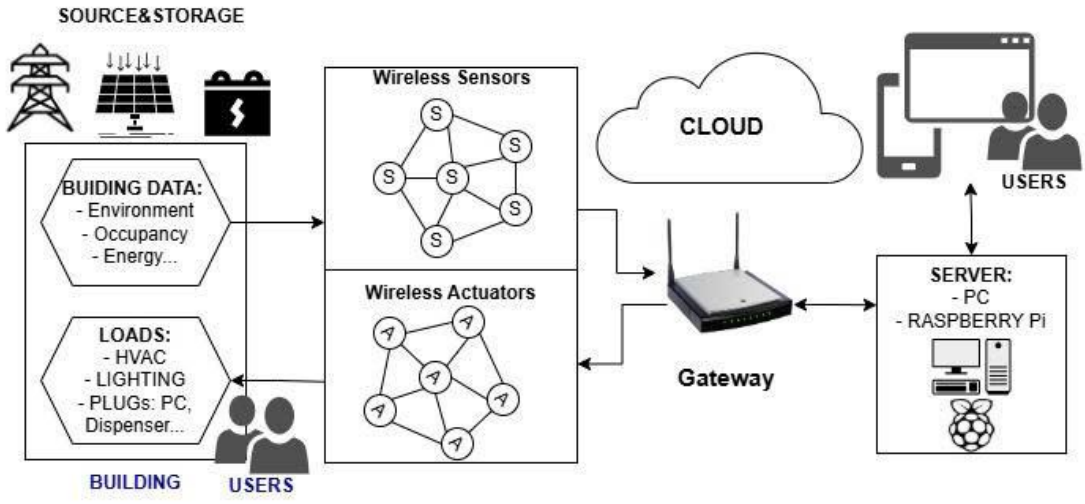


Figure 11. System Architecture for Building Monitoring and Control [43].

Figure 11 shows the IoT platform architecture tailored to these dynamic energy and environmental management requirements. Based on the building profile, the proposed architecture comprises [43]:

- Smart sensors for monitoring energy and environmental conditions (energy

meters, temperature and light sensors, motion sensors); Smart actuators for controlling electrical devices (lighting, air-conditioning and plug loads);

- A gateway that communicates over multiple interfaces (Wi-Fi, RF24, ZigBee, Z-Wave, Bluetooth, etc.). In our implementation, the MQTT protocol is used to minimize network bandwidth;
- A computer (Raspberry Pi) acting as a data center and hosting embedded algorithms, complemented by a cloud component for database management;
- Open-source tools for time-series storage and visualization (InfluxDB, Grafana) and user interfaces (OpenHAB);
- Constraint inputs related to electricity price, sources, and storage and user preferences.

The next subsection presents the IoT-based wireless sensor network, from sensor choice to network design and communication protocols.

2.2.2. IoT-based wireless sensors network

2.2.2.1.Sensor selection

Wireless sensor nodes available on the market provide a wide range of functions and communication technologies. Sensor selection should consider the physical quantity to be measured, required accuracy, building type and installation location. In the IoT-BEMS platform, sensors are grouped into three main categories:

(1) Environmental sensors (Temperature, humidity, and luminosity sensors) to adjust cooling, heating and lighting according to ambient conditions;

(2) Occupancy sensors (motion detectors, door contact sensors) to infer building occupancy and usage patterns;

(3) Energy sensors (Current, voltage, power, and energy) to monitor electrical consumption and power sources, including the grid and photovoltaic (PV) systems.

Table 1. Main characteristics of sensors deployed in the IoT-BEMS platform

Location	Variables	Sensor type	Range	Accuracy	Resolution	Interface	Category
Indoor room (BME280)	Temperature, relative humidity	Environmental sensor	−40÷80 °C; 0÷100 %RH	±0.5 °C; ±2 %RH	0.1 °C; 0.1 %RH	I²C/ RF24	Low-cost, self-developed
Indoor wall (M1, M2)	Temperature, relative humidity	Environmental sensor	-20÷50°C; 0÷100 %RH	±0.3°C ±3% RH	0.1 °C; 0.1 %RH	Zigbee	Low-cost commercial
Indoor wall (Z1, Z2, Z3, Z4)	Temperature, relative humidity	Environmental sensor	- 10÷ 50°C 0÷100 %RH	±1°C ± 3% RH	0.1 °C; 0.1 %RH	Zwave	Low-cost commercial
Door/ window	Occupancy / motion	Occupancy sensor	0/1	–	–	Zwave / Zigbee	Low-cost commercial
Main AC panel	Active power/ active energy	AC energy sensor	0÷23 kW 0 ÷ 9999 kWh	±0.5% ±0.5%	0.1W 1Wh	RF24	Low-cost, self-developed
PV-DC board	DC voltage, DC current	DC energy sensor	0.05÷300V, 0.02÷300A	±1.0% ±1.0%	0.01V 0.01A	RS485/ Wifi	Low-cost, self-developed

The main sensors and meters used in the IoT-BEMS deployment are summarized in Table 1. The table reports the measured variables, nominal range and accuracy, sampling period, communication interface, and low-cost category (self-develop or commercial sensor).

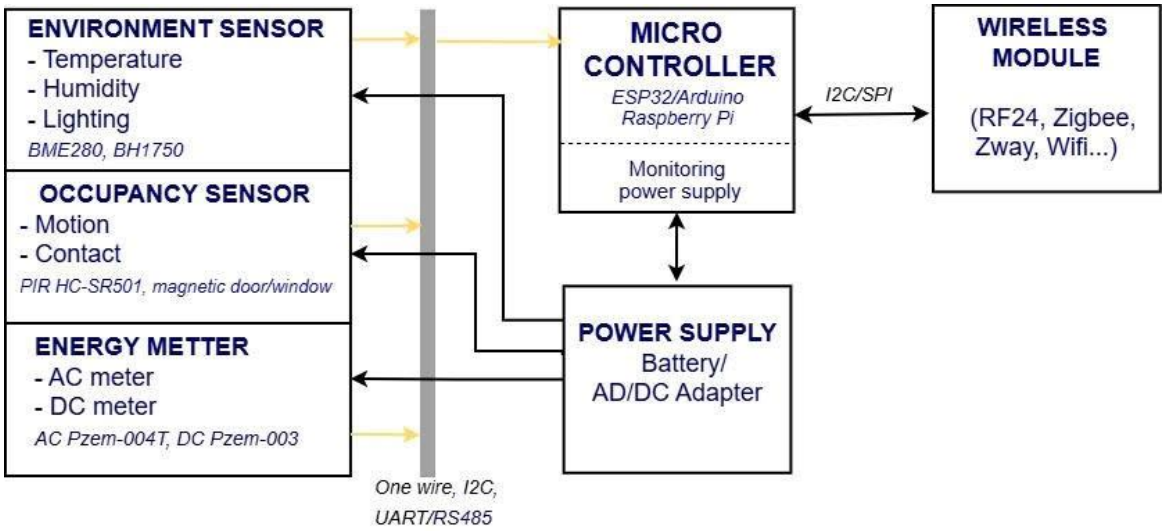


Figure 12. Architecture of the Environmental – Occupancy – Energy Sensors

Figure 12 illustrates the sensor's architecture within the IoT-BEMS platform, comprising a power supply unit, a sensor module, a microcontroller that integrates algorithms for optimal sensor operation, and a wireless transceiver for data communication.

To ensure operational reliability, sensor development follows four steps: (1)

hardware design, (2) algorithm development and firmware programming, (3) final assembly, and (4) packaging and functional testing.

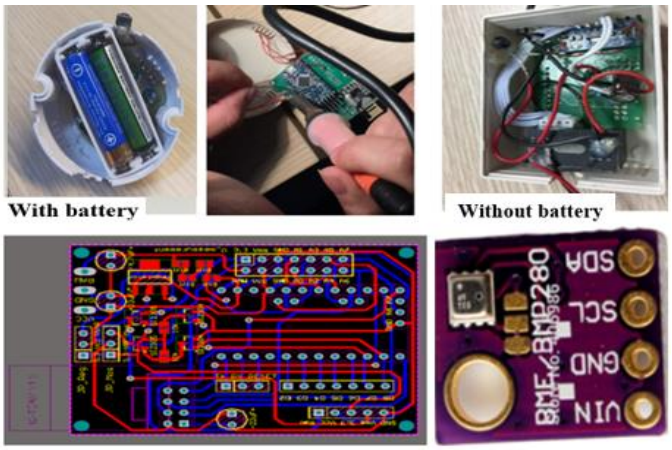


Figure 13. Development of Multi-sensors in our projects

There are two types of multi-sensor, depending on the power supply (with/without a battery). Figure 13 illustrates the physical implementation (design, assembly, and package) of the multi-sensor node (measuring relative humidity, barometric pressure, and ambient temperature) used in our projects.

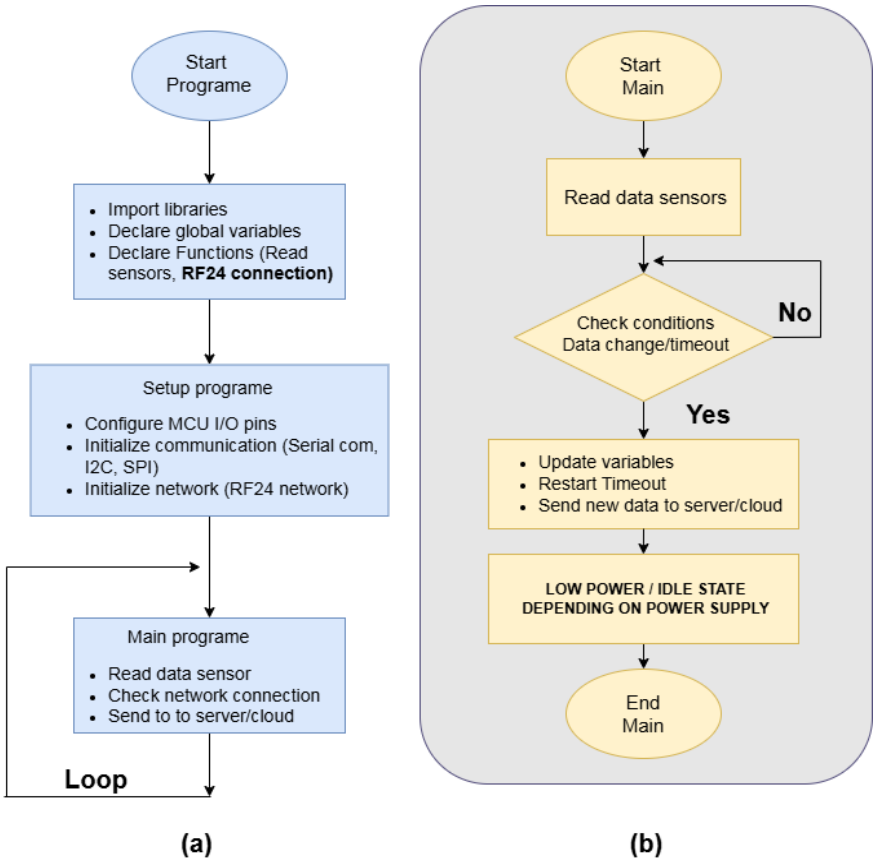


Figure 14. Programing flowchart of RF24-based sensors

In Figure 14, the firmware architecture is standardized for all RF24-based sensor nodes. Figure 14.a describes the generic setup phase: importing libraries, declaring global variables and functions, configuring MCU I/O pins, initializing communications (Serial, I²C, and SPI) and starting the RF24 network. Figure 14.b details the main loop: reads measure data, checks whether a value change or timeout occurs, updates variables, resets the timer and sends a new data packet to the server/cloud. At the end of each cycle, the node enters a low-power or idle state (depending on the power supply), which limits energy consumption while keeping the network connection active. The flowchart simplifies firmware development and maintenance. It therefore fits the requirements of the proposed low-cost IoT-BEMS, allowing pre-configured sensor nodes to operate reliably.

Experiment 1: Developing wireless sensors with a battery power supply

The experiment details how a battery-powered RF24 wireless sensor node was developed for building environmental monitoring. To achieve reliable data acquisition, sensors are deployed at designed locations to collect critical environmental parameters. The technical design needs to use low-power hardware (BME280 low-power sensor) and an optimal program (deep sleep, control sampling rate and sampling frequency). For instance, low-power mode is to put the sensor in sleep with a time sleep relating to a minimal sampling frequency.

Table 2: Main technical requirements for Battery-Powered Sensor

No	Requirement	Description
1	Location	Installed in critical areas such as around wall, doors and windows.
2	Power Supply	Minimal reliance on external electricity.
3	Energy Efficiency	Minimize working frequency (nodes only activate for schedule or data transmission).
4	Battery Lifetime	A battery can supply power for the sensor's long-time operation.
5	Communication	Support low-power wireless protocols (e.g., RF24, ZigBee …) for efficient data transfer and energy consumption.
6	Data quality	Maintain sensor precision and stability.
7	Reliability and Maintenance	Reliable operation on diverse environmental conditions; easy to replace or maintain.

In Table 2, basing on practical work, seven technical requirements for Battery-Powered Sensors should be considered.

Experiment 2: Developing a RF24 wireless sensor with the Grid power supply

This experiment presents the development of an RF24-based wireless sensor node powered by an AC power supply. From a design perspective, it is necessary to ensure a power supply for the controllers and the critical sensors that run continuously in the network. Therefore, these nodes must be located near the building's electrical grid. Enhancing the data sampling rate and implementing dynamic routing updates are key to improving network reliability. The sensor's hardware includes a compact 5VDC power module to minimize design complexity; an Arduino Pro Mini 5V module serves as the central processing unit; and an RF24 transceiver module enables wireless communication (see Figure 15).

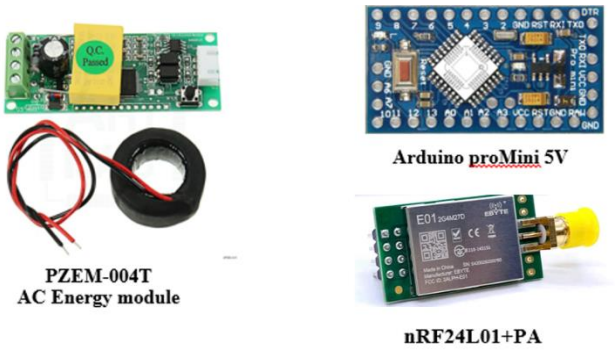


Figure 15. Main components of Energy Metter

Figure 16. Energy meter in VHH (RF24 Mesh-Network).

In Figure 16, a wireless energy meter was installed in the VHH platform. The experiment design maintains devices over the long term by using modular hardware and an open library [51].

2.2.2.2.Developing a wireless sensor network

a. A wireless sensor network (WSN) architecture

A wireless sensor network typically comprises sensor nodes, actuators, gateways, and data processing and storage centers. Sensor nodes collect data and transmit it through coordinator nodes to a gateway. From the gateway, data is forwarded via the Internet connections to the data processing and storage center. Users can manage and monitor the network in real time [49].

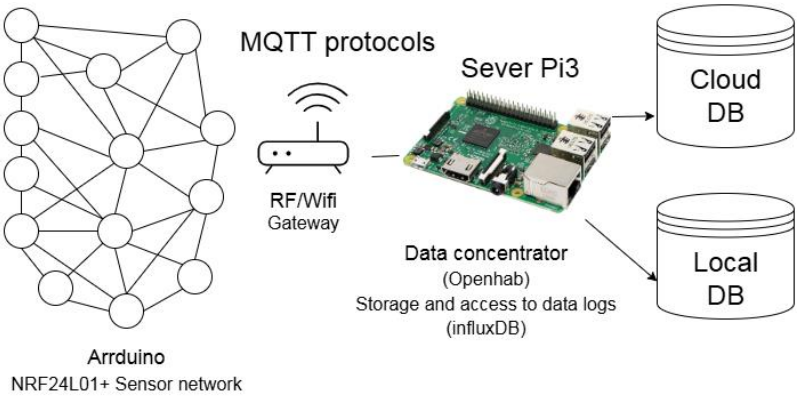


Figure 17. Description of a RF24-Based Wireless sensor network

Figure 17 illustrates an example of an RF24-Based WSN on real platforms, which integrates open-hardware (Arduino and RF24 modules), a Nano-Computer (Raspberry Pi), and open-source tools (OpenHAB and InfluxDB) for data collection, processing, and visualization.

b. Network Topology

Selecting a suitable network topology is essential for improving network performance regarding energy efficiency, frequency band, or deployment cost. Practically, there are typical topology types such as (Flat topology, Cluster topology).

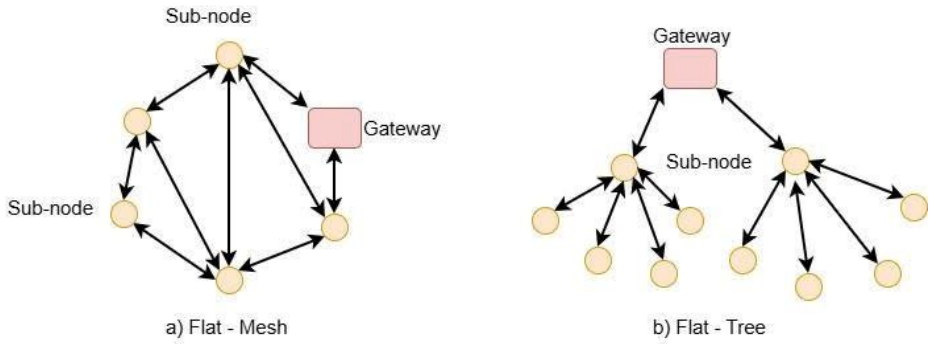


Figure 18. Flat topology (Mesh vs Tree) on HaUI's platform [37]

Figure 18 shows the flat structures used to develop the self-developed RF24 network on HaUI's platform. This topology enabled simple, low-cost deployment with direct communication paths and minimal management overhead, suitable for small to medium size building sections.

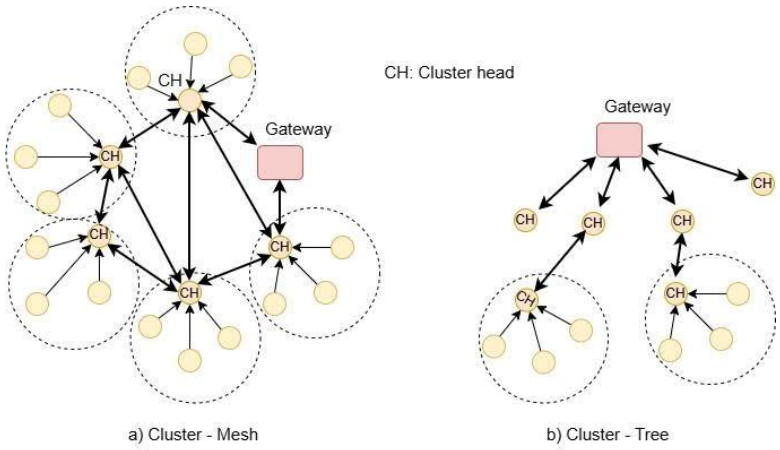


Figure 19. Cluster topology (Mesh vs Tree) on VHH's platform [37]

Figure 19 shows Cluster topologies implemented at VHH's platform. An integrated system based on a hybrid WSN network, such as ZigBee (Xiaomi), Z-Wave AEONTEC, Wi-Fi (ESP32/8266), and nRF24L01+ energy meters. Clustered organization improved bandwidth utilization and reduced energy consumption. The combining protocols allowed functional optimization across different building zones. From practical deployment, key factors requires to consider when selecting network topologies as follows:

- Using multi-topology integration to meet various monitoring needs for the system's flexibility.
- Low-power networks (ZigBee, Z Wave, and nFR24L01+) to reduce maintenance.
- Multi-protocol integration requires robust gateways and data center management.
- Using hybrid networks increases reliability and potential system expansion.

Experiment 3: Gateway Development

In the RF24 network experiment, the gateway (node 00) collects sensor data and effectively sends control commands to each actuator node. Working as a bridge

between the RF24 and Wi-Fi communication networks, the gateway is necessary for maintaining data flow and system coordination.

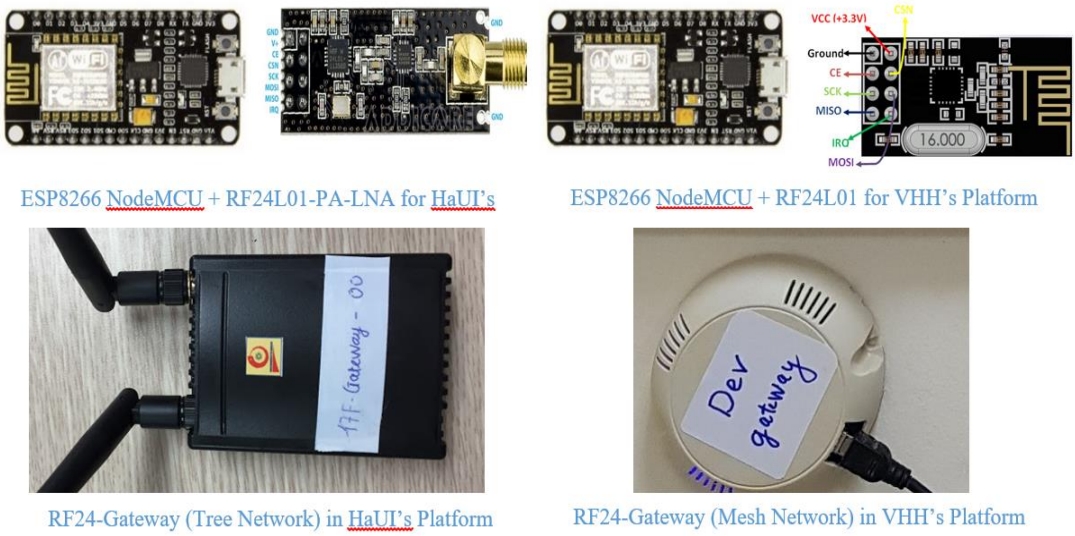


Figure 20. Development of gateway devices in real projects

To ensure consistent network routing, the gateway must be supplied by an uninterrupted power source. Gateway devices should be sited in areas with reliable access to the electrical grid and robust communication links to network nodes. Figure 20, the gateway is based on a Wi-Fi module (ESP8266) and an RF24 module, which operates in the 2.4-2.5 GHz frequency band and supports 125 channels, so providing flexibility and scalability for the network.

Experiment 4: RF24-Flat topology development

The topology networks enable real-time data acquisition and direct control of lighting and air-conditioning systems. This contributes to operational efficiency and enhanced user comfort. In this work, the topology is deployed following three steps [52], including (1) Address format in network; (2) Routing Mechanism; (3) Installation

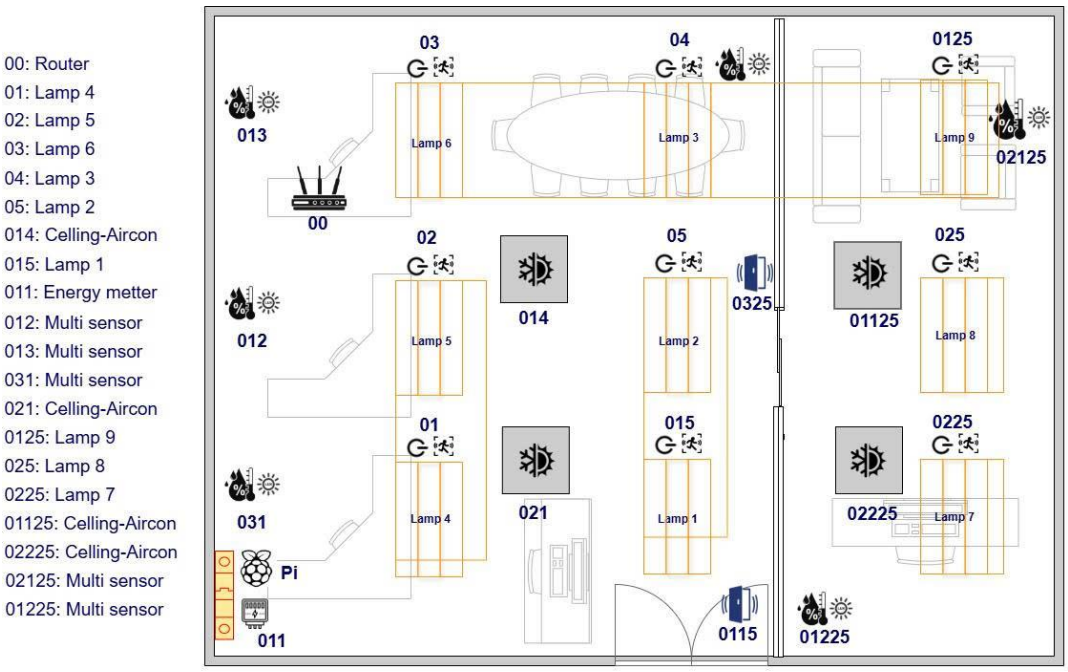


Figure 21. Description of RF24-Tree network configuration in a real project

Figure 21 illustrates the RF24-Tree topology configuration implemented for the office. In the RF24 network, the address format uses a 15-bit octal (base-8) scheme as follows:

- Router: address 00 is the Base node.
- Lamps in the office: addressed 01-05 are first-level children
- Other devices: addressed 011, 012, 013, 014, 015, 021, 025, and 031 are second-level child nodes. Each additional octal digit represents a deeper level in the network hierarchy. For example: Multi sensor nodes (01225, 0225); AC controller nodes (Celling-Aircon: 01125, 0225)

The network architecture allows for reducing routing complexity. However, in the topology, if a parent node fails, all its descendant nodes will lose connectivity.

2.2.2.3.Communication technologies

a. Wireless communication protocols

Wireless communication technologies are essential enablers of efficient monitoring and control in modern buildings. They simplify the installation of sensors, actuators and controllers. In smart buildings, devices and end users interact remotely via wireless systems to receive and transmit control commands

in real time. When selecting a communication technology, several key criteria should be considered: low cost, low power consumption, ease of use, security, low interference, interoperability, scalability and accessibility to local markets [49].

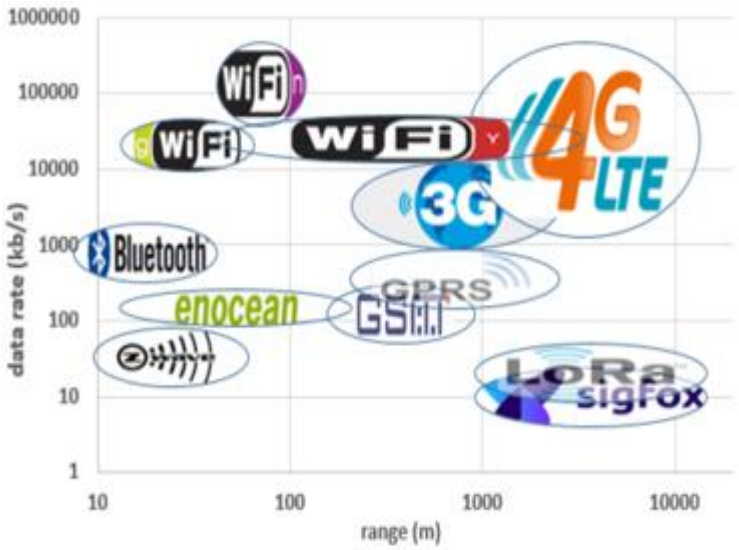


Figure 22. Flow of the main radio frequency protocols [53]

Figure 22 shows the standard RF protocols in IoT along with key characteristics such as frequency, range, data rate, power consumption and network structure [53]. In practice, buildings often use many different communication standards as users choose devices based on brand and application needs. Each protocol, such as ZigBee, Z-Wave, RF24, Bluetooth and Wi-Fi, has its own advantages.

On the VHH platform, the system combines ZigBee, Z-Wave, Wi-Fi (ESP32), and RF24 to increase flexibility and support energy management. Open standards such as Wi-Fi, Bluetooth, and ZigBee have great potential for building management, but selecting the proper protocol remains difficult due to frequency conflicts, node count, cost, flexibility, and power consumption.

b. MQTT protocol

MQTT is an open connectivity protocol for the Internet of Things (IoT) messaging submitted to OASIS in 2013. It is a lightweight message transport for connecting IoT devices with limited network bandwidth.

Experiment 5: MQTT application in WSN's communication

MQTT is used for communication between gateways and sensor and actuator nodes. In the experiment, MQTT was applied to facilitate communication between gateway devices and sensor/actuator nodes.

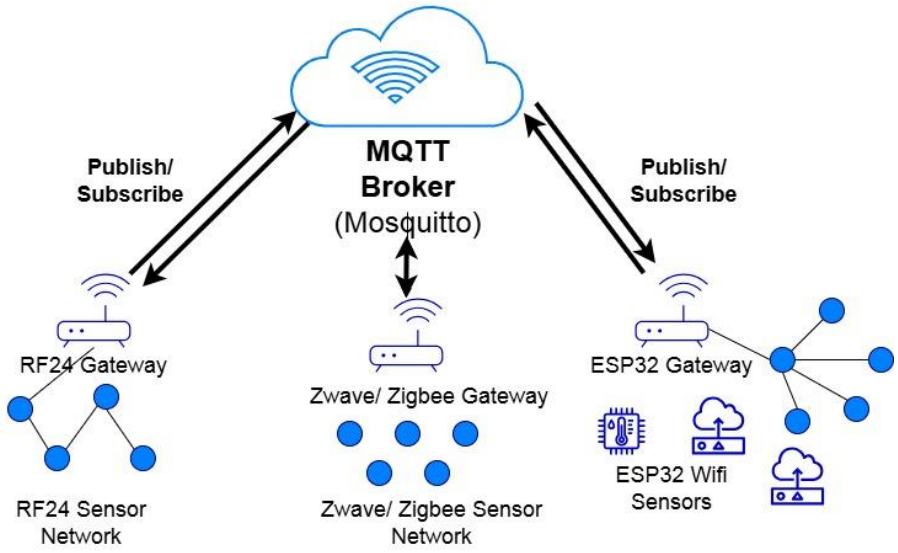


Figure 23. MQTT application communication in VHH's project

Practical implementation demonstrates that the MQTT publish/subscribe architecture provides efficient, flexible, and scalable real-time data management. In the system, an MQTT broker is installed on the Nano computer to handle all communication. Each MQTT message consists of four main parts: information about the system layer, a node that sends data, node that receives data and the type of exchanging message.

Layer network/sending node/receiving node/message type

Practical observations highlighted that careful hierarchy design is essential for maintaining system clarity, scalability, and ease of maintenance as the network expands.

Figure 23 illustrates the structure of MQTT application within the deployed VHH's platform. Sensor nodes act as publishers, sending environmental data (such as temperature, humidity, and light intensity) to subscribed topics on the MQTT broker. Simultaneously, gateways subscribe to command topics then send to actuator nodes.

2.2.3. Data Visualization and Users interaction

This work used InfluxDB, a time series database suitable for data management, Grafana facilitates data querying, visualization, and alerting. Multi-sensor data is combined for data analyzing reliability [53]. The IoT-BEMS platform provides data to operate Buildings through a user interface or via a Web service.

Experiment 6: Energy Data analysis in VHH's platform

The energy consumption are often linked to the user's behaviors.

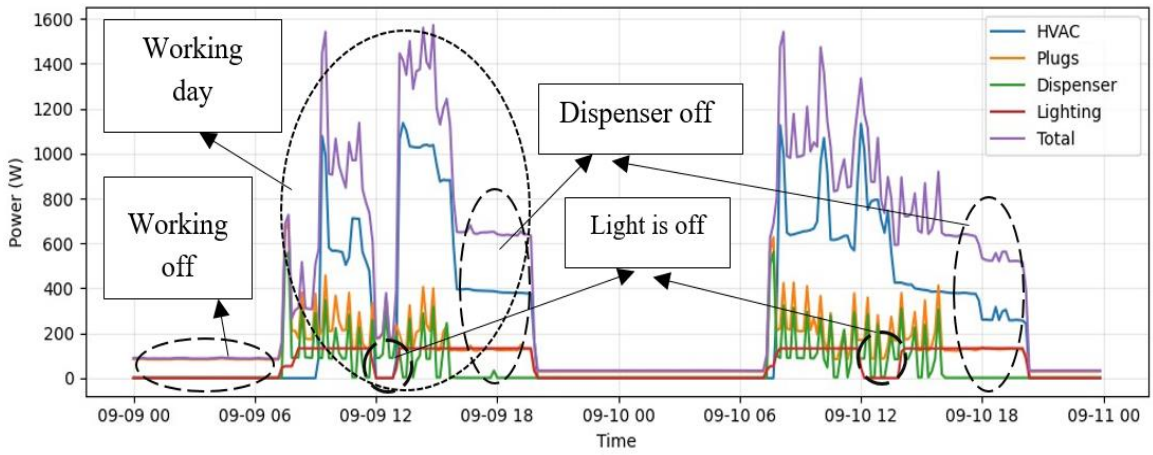


Figure 24. Correlation between consumption and occupancy in VHH′s office

Figure 24 shows that the Dispenser is set to operate according to the working schedule of the day. The lighting system will turn off at the end of working hours, during lunch break. When analyzing carefully based on other additional data from the system, this action can be identified in more detail, specifically:

- On September 9, 2020, 11:30-13:00, the user turned off the lighting system. However, the door and motion sensor value indicated no-one in the room.
- On September 10, 2020, 11:30-13:00, the lighting system was off but the air conditioner was still on. The door and the movement sensor show that the user was absent.

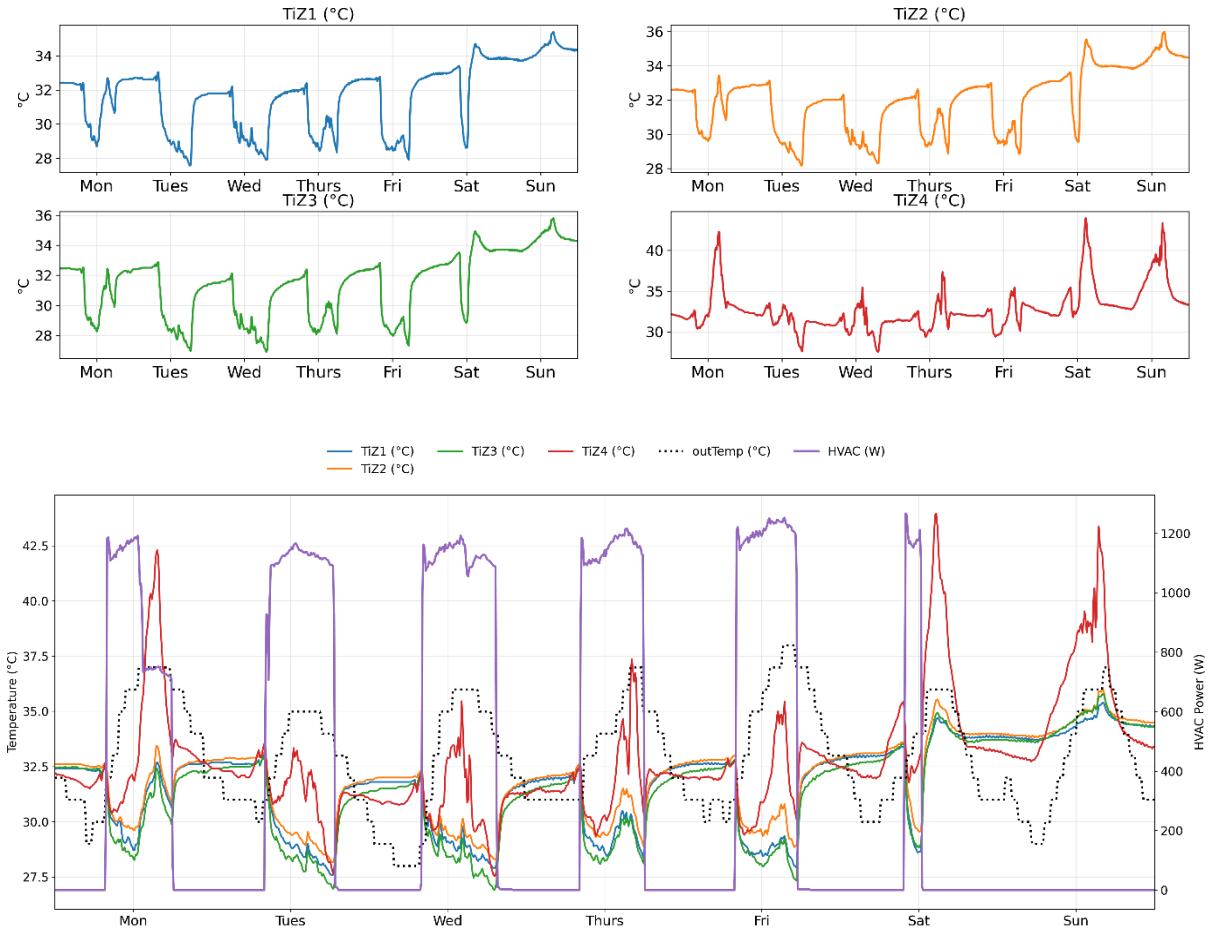


Figure 25. Correlation of HVAC operation vs environment conditions in VHH's office

Figure 25 shows that the HVAC is operating during working hours, and the heat distribution in the room. The window (where the TiZ4 sensor is located) absorbs a lot of heat, so the temperature is always higher than the other in room areas.

Experiment 7: Environment Data analysis in VHH's platform

Environment conditions effect to how users operate electrical devices in the office.

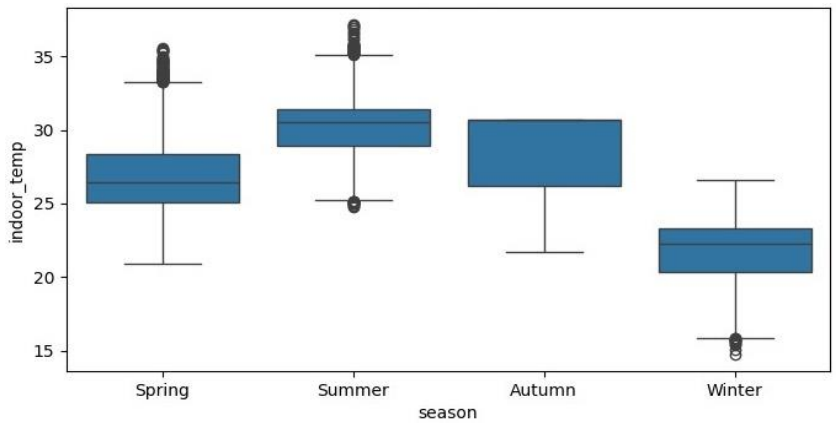


Figure 26. Description of indoor-Temperature distribution by seasons

In Figure 26, in summer, room has high temperature with high median and IQR of indoor temperature, which easily triggers HVAC to cool more. However, in winter, it has the lowest temperatures, with little variation, low median and narrow IQR. Spring/Autumn has average temperatures, moderate dispersion. Summer and spring have very high outliers' points, possibly due to door opening and closing, high occupancy or set-point temperature.

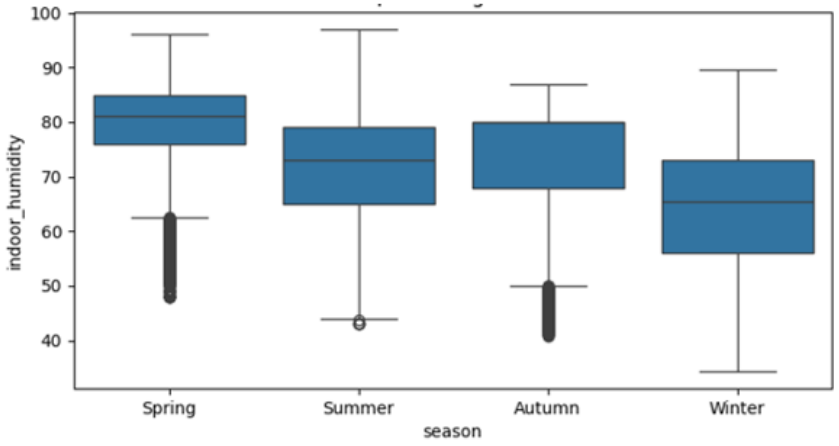


Figure 27. Description of indoor-humidity distribution by seasons

Figure 27 shows that Summer median humidity increases and has large variations, creating a feeling of heat even though the temperature is not too high. Winter humidity is lower with a low median but large variations, so there are times when the air conditioner needs to be turned on to dehumidify. Spring-Autumn session is more balanced with average humidity, but there are still outliers by fluctuating weather.

Experiment 8: Correlation data analysis in VHH's platform

Correlation between sensor data allows for the creation of virtual sensors or data compensation for missing variables. [36]. Figure 28 shows the data correlation between temperature sensors and consumption during 1 week in August 2020. The red color in the figure represents a strong correlation (correlation value>0.9) between variables. The correlation between Total power and TiZ3 and HVAC power is relatively high (with correlation values of 0.7 and 0.93, respectively). This indicates that pair [TiZ1, TiZ2] can be used to compensate for the missing data of TiZ3 and TiM2; [Total, TiZ3] to compensate for missing HVAC data.

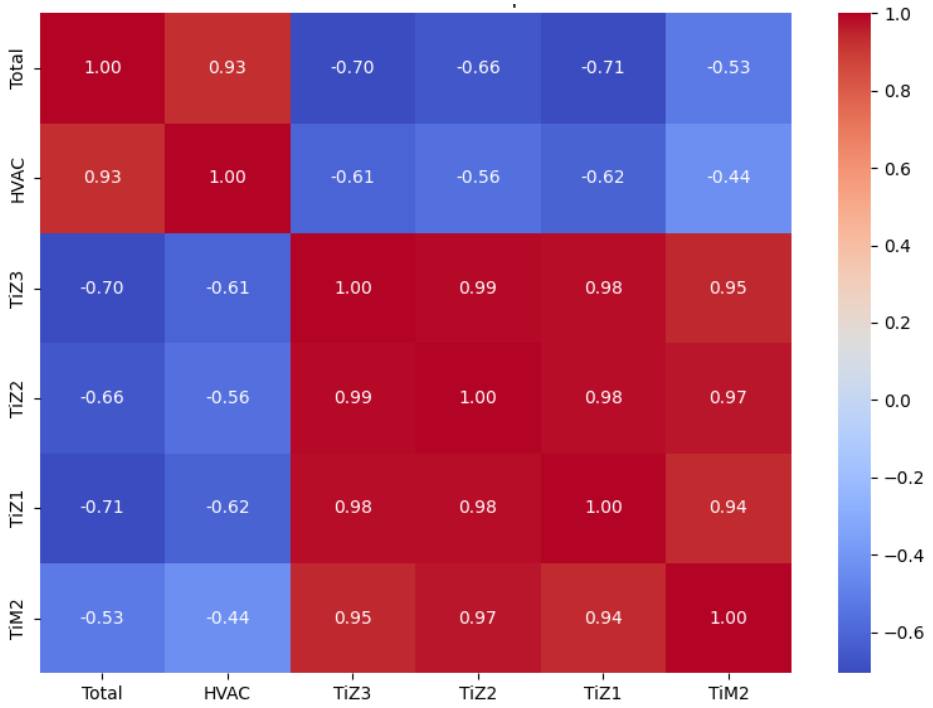


Figure 28. Correlation temperature and consumption data in VHH's platform

Experiment 9: Fault sensor detection in VHH's platform

Data quality assessment and error detection are always challenges in IoT sensor systems. Some problems affect the accuracy of error warning and system quality diagnosis. Many errors look like sensor failures, but in fact they are time shift errors, temporary connection loss, sensor stuck-at constant. Abnormal data are not caused by sensor errors but by sudden changes in the real environment. This study proposes a method for data quality analysis based on multi-sensor data correlation and system operating context. For example:

- Check for sudden large jumps in data values and compare the anomalies with real environmental data.
- Use correlation between sensors to detect sensors drifting away from the sensors in the group to diagnose faults and compare context over time.

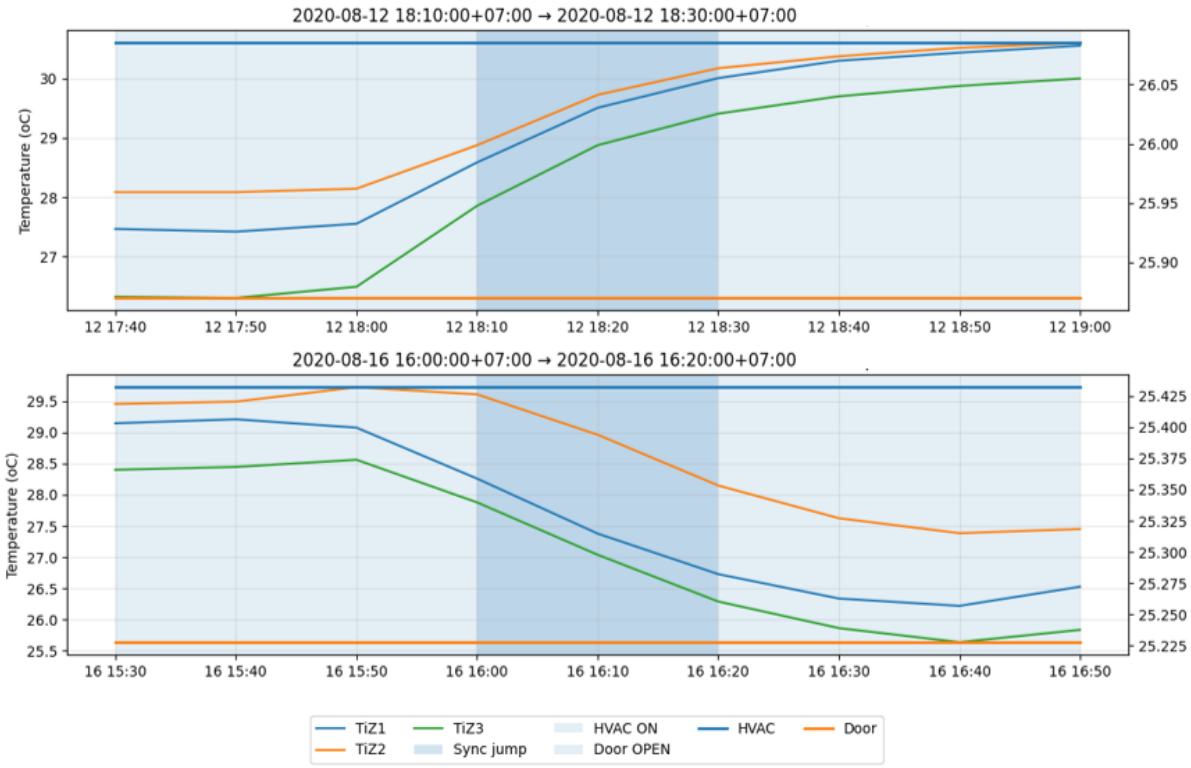


Figure 29. Jump events of temperature values during HVAC ON and Door OPEN

Figure 29 shows two unusual events of temperature value increase and decrease of sensors (TiZ1, TiZ2, and TiZ3). However, the measured data have the same up/down trend relating to status of air conditioner and door. This trend is reasonable, there is no error from the physical sensor

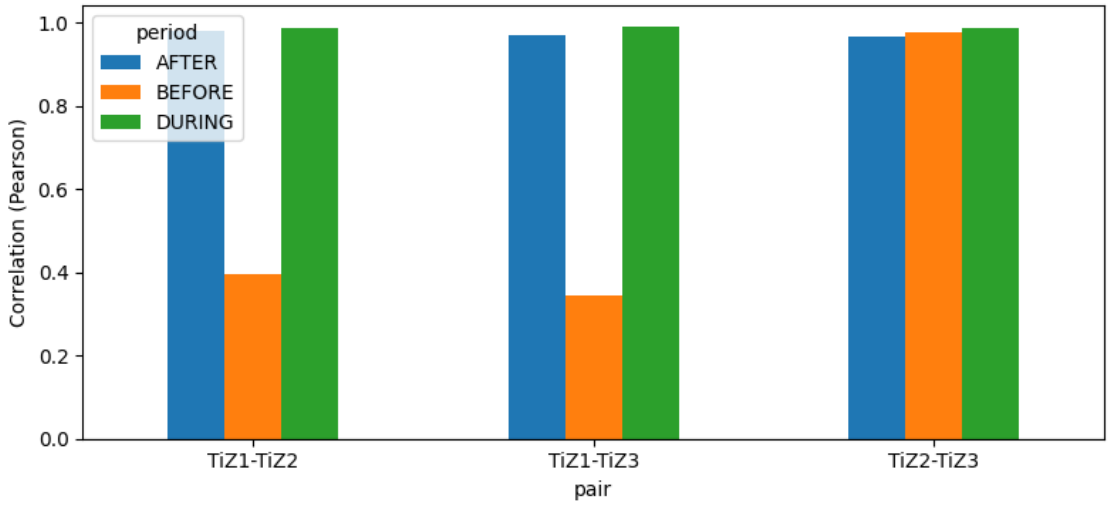


Figure 30. Correlation pair sensors (TiZ1–TiZ2, TiZ1–TiZ3, TiZ2–TiZ3) in one-week BEFORE, DURING, AFTER from (10/05/2021-30/05/2021)

Figure 30 shows that during the week (May 10, 2021-May 16, 2021) the TiZ1 sensor behaved abnormally, deviating from its correlation with TiZ2 and TiZ3.

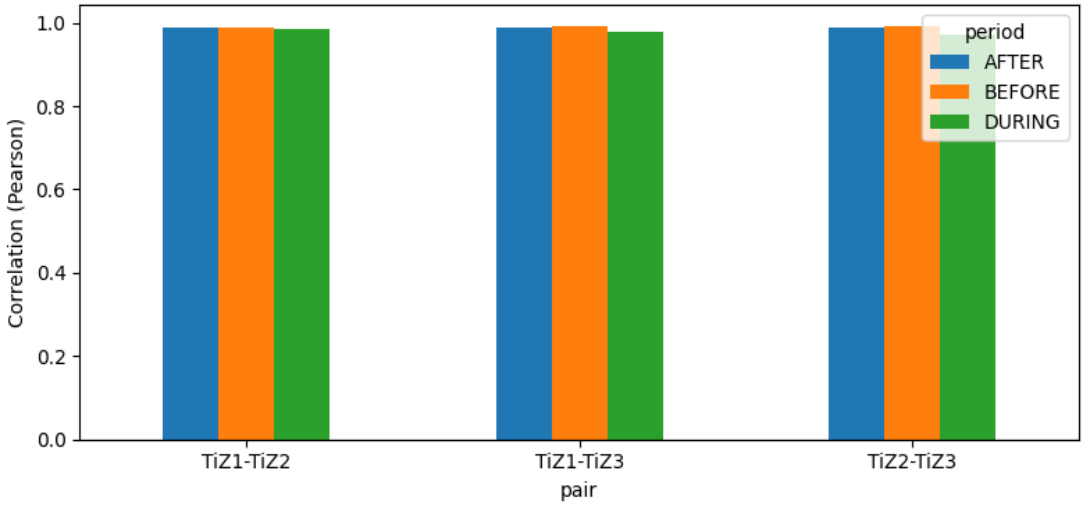


Figure 31. Correlation pair sensors (TiZ1–TiZ2, TiZ1–TiZ3, TiZ2–TiZ3) in one-week BEFORE, DURING, AFTER from (17/05/2021-06/06/2021)

Nevertheless, in Figure 31 the following weeks, this sensor was strongly correlated with the other two sensors. The TiZ1 sensor could not damage, but only temporarily faulty. The two sensors TiZ2 and TiZ3 maintained stable values.

Experiment 10: Lighting system maintenance in VHH's platform

Most electrical equipment often operates until it fails or has problems. This factor received little attention, however is prevalent in many buildings. In real Buildings, most lighting systems are old and designed over ten years, no changes until fault. The working time data of electrical devices data could support users' plans for maintenance devices (such as repairing or replacing them) on time.

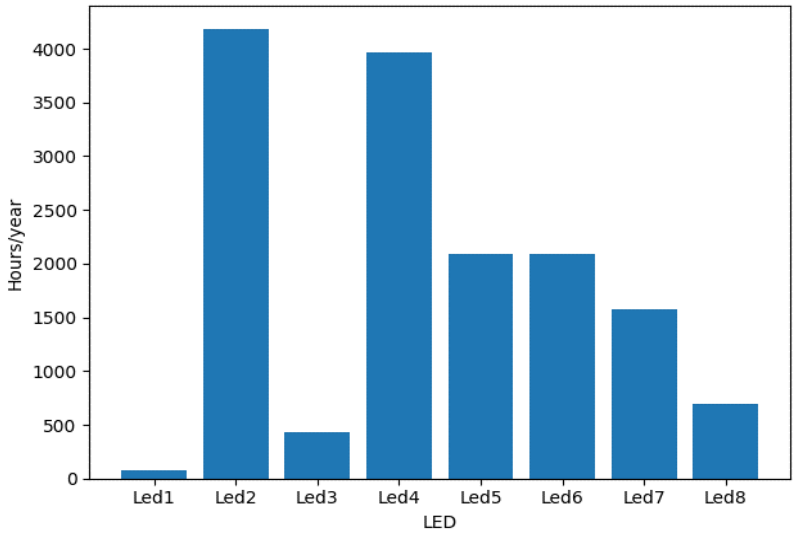


Figure 32.Working time of Led groups in a year.

In Figure 32, users could find out that group Led1 has lowest performance while groups Led2 and Led4 have the best. By monitoring lighting system, users can calculate time to replace Led.

2.2.4. Buildings energy services

The data-fusion layer described above aggregates measurements data into meaningful information about the building state. From this foundation, the IoT-BEMS can provide the following energy services:

Surveillance services. Real-time access to indoor environmental conditions, energy consumption, PV production, equipment status, sensor power status, and door/window openings. These data are presented on dashboards for display on smartphones and PCs, enabling users to respond to abnormal situations quickly.

Maintenance services. Besides fixed maintenance schedules, measurement data enables early fault detection and the issuance of warnings. This supports predictive maintenance and helps users plan repair and replacement actions more accurately.

Control services. Control strategies based on fused data (HVAC, lighting and plug loads) can improve building performance while maintaining user comfort.

Data analysis and reporting. Historical data support data-driven decision-making, improve energy management, and provide indicators for broader energy transition initiatives.

In the next section, we investigate data-quality methods to ensure that these services operate reliably on the low-cost IoT-BEMS platform.

2.3. Data quality assurance Framework

In a low-cost IoT-BEMS platform, data quality is a prerequisite for reliable operation of energy services. Data with missing samples, spike noise, drift will cause errors to propagate to visualization, data fusion, load/PV forecasting, and control, thus reducing the effectiveness of nZEBs.

The direct relationship between data errors and the reliability of control strategies has been emphasized in studies on building energy performance [30] . Therefore, this thesis proposes a framework to ensure data quality at two levels:

- ✓ Level 1: At single-sensor level [40];
- ✓ Level 2: At sensor network level [54], [39], [55].

Experiments 1–3 in section 2.3.3 validate the effectiveness of the level 1, while experiments 4–6 in section 2.3.4 validate the level 2. The latter two quality-assured datasets are used directly for building energy services and case study evaluation in Part 3.

2.3.1. The state of the art

Energy analysis, forecasting, and control models are only reliable when built on complete and accurate measurement data. In the PV systems, enhanced data processing and verification are prerequisites for improving the reliability of performance analysis and system robustness [56].

Recent literature reviews also confirm that low-quality data is the main cause of performance degradation in sensor-based monitoring systems and Machine learning models [57].

For building data, dataset characteristics such as sequence length, sampling frequency, diversity, and especially data quality need to be strictly considered when building models for energy services [58].

At sensor network level, systematic reviews of sensor data quality have pointed out typical errors in reliable monitoring including: missing values, spikes/outliers, long-term drift/bias, stuck-at, and network-level errors such as packet loss bursts, latency/jitter, and time misalignment [50], [57]. These errors break the integrity and temporal structure of the data – the foundation of energy optimization and forecasting models.

Data errors in low-cost IoT-BEMS often come from the following sources:

- ✓ Sensors' hardware: low-cost sensors are sensitive to manufacturing errors, aging, environment condition, or hardware design. LCS studies also highlight the heterogeneity of quality between manufacturers and the need for periodic

calibration/maintenance [59].

- ✓ WSN transmission network with packet loss, transmission delay and clock skew cause time series data errors [59].
- ✓ Data integration when merging multiple sources in IoT-BEMS, heterogeneity in measurement standards/semantics and representations between devices can create data inconsistencies [60]; Multi-sensor data often have different sampling frequencies, so consistent resampling–time alignment is needed [61], [62] otherwise it will degrade the consistency of the time series.

Current research is developed in two directions corresponding to two levels of quality assurance. At the single sensor level, many works proposed low-cost sensor calibration using ML and virtual sensing to reduce measurement bias and physical element dependence [40]. At the sensor network level, many studies apply error handling standards according to international standards, typically IEC 61724 for PV data to eliminate gaps, duplicates and false values [63].

The Machine Learning with data fusion for low-cost sensors is considered a promising solution for outlier detection and real-time data compensation [39], and has been extensively used for outlier detection based on forecast errors [55] as well as data compensation in computationally constrained environments [64].

In the context of PV/load forecasting, probabilistic models such as GPR and deep learning models (CNN-LSTM) have been proven effective and can be integrated as the core of data compensation/uncertainty estimation in online operations [64], [65].

Existing studies on data quality indicate that it is rare to address both levels in a unified framework. However, in low-cost IoT-BEMS deployments for buildings, these two levels need to be tightly coupled. Poor measurement integrity at the source will inevitably affect network level analysis, while errors can distort sensor readings. This motivates the development of a two-level data quality (DQ) framework in this study.

Table 3. Classification of data faults in low-cost IoT-BEMS and considered faults in the thesis.

Fault group	Fault type	Description	Example in IoT-BEMS	Considered?
Missing data	Block missing	Consecutive missing samples	RF packet loss, node freeze	✓
	Scatter missing	Randomly missing samples	Random network interference	✓
Outliers/ Spikes	Spike/ impulse	Short-duration & large-amplitude anomalies	ADC noise, transient transmission error	✓
Drift/ Bias	Linear/ nonlinear drift	Long-term gradual deviation	Sensor aging, environment effects	X (baseline handling)
Stuck-at Constant	Constant value	Sensor output stuck at a fixed value	Firmware fault, sensor failure	X (baseline handling)
Time misalignment	Delay/ shift	Timestamp delay/ time shift	Gateway clock offset, latency	X (baseline handling)
Inconsistency multi-node	Cross-sensor/ cross-system correlation	Missing data requiring correlation exploitation	Correlated PV stations/ zone sensors	✓

Table 3 summarized from systematic reviews on sensor/WSN data quality and fault types [50], [57] and from energy data quality (Loads/PV) studies [56], [63]. The (✓) mark indicates the subset of faults observed and quantitatively evaluated in the thesis.

This study focuses on errors that directly affect the data series and the forecasting model (block or scattered data loss, outliers, and inter-sensor inconsistencies). Long-term errors such as sensor drift, stuck values, or time drift are assumed to be addressed by baseline calibration and periodic maintenance, so they are not analyzed in detail. The classification is used in the subsequent step to design the data quality framework.

2.3.2. Methodology

In this study, we focus on tools that support non-expert users, who are essential in ensuring BEMS data quality. The data quality assurance framework is built based on the characteristics of low-cost IoT-BEMS systems and data errors observed in real deployments. The proposed Data Quality Assurance pipeline with two level of quality assurance is described as follows:

(Level 1) At the single-sensor level: focuses on reducing errors at the

measurement origin, including:

(1) Sensor calibration during development to achieve target errors;

(2) Improving box design to maintain stability in harsh environments;

(3) Adding virtual/soft sensors for reducing missing and drift due to physical components.

(Level 2) At the sensor network level: handles errors arising at the network and post-collection data levels. This level implements the three-step pipeline: Detection, Correction/Imputation, and Validation in an online solution. Machine learning models (MLR, GPR online) are used to detect missing, spike, drift, and inconsistency between nodes. With online models, the computational resources (CPU/RAM) are reported to demonstrate real-time feasibility.

Overall, (Level 1) ensures the measurement accuracy of input data, while (Level 2) ensures the integrity, synchronization, and consistency of data in real-time operations.

Estimating errors factors are often used including RMSE, MAE, and R2 [66][67].

2.3.3. Single Sensor level- Experiments

2.3.3.1.Sensor calibration & validation

An emerging trend among users is the self-development sensors using Arduino platforms combined with various open modules [45]. The study aims to provide a structured process with detailed steps to assist users in developing their customized sensor solutions.

Experiment 1: Testing a self-develop DC meter vs reference instruments

In this experiment, we presented sensor estimation process in range 0-5A DC current and 0-60V DC voltage. Materials for experiment shown in Figure 33 including:

- ✓ Standard signal generator are a Generator DC Current/Voltage

(Alimentation continue 0-100V 0-60A 6000W) and Load 5A.

- ✓ Reference devices: Keysight Digital Multi-meter 34450A, 5 ½ Digit, OLED Display; Agilent 34450A Digital Multi-meter.
- ✓ Power supply (15V) was from DC power supply device of ELC - AL936N

Figure 33. Devices for calibration experiment

The estimating sensor error process applied for testing in two parts:

(1) Testing with DC current part:

- Current input LTS-25P was supplied by Generator DC Current/Voltage; Value of generator current (I_{in}) was measured by Agilent 34450A Digital Multi-meter;
- Value of current sensor (I_{out}) was recorded on raspberry Pi. Result is RMSE of Iout (1-5A) = 0.7%;

(2) Testing with DC voltage part:

- Voltage input LV 25P was supplied by Generator DC Current/Voltage;
- Value of generator voltage (Vin) was measured by Keysight Digital Multi-meter 34450A, 5 ½ Digit, OLED Display;
- Value of voltage sensor (V_{out}) was recorded on raspberry Pi. Result is RMSE of Vout (0-63V) = 0.97%.

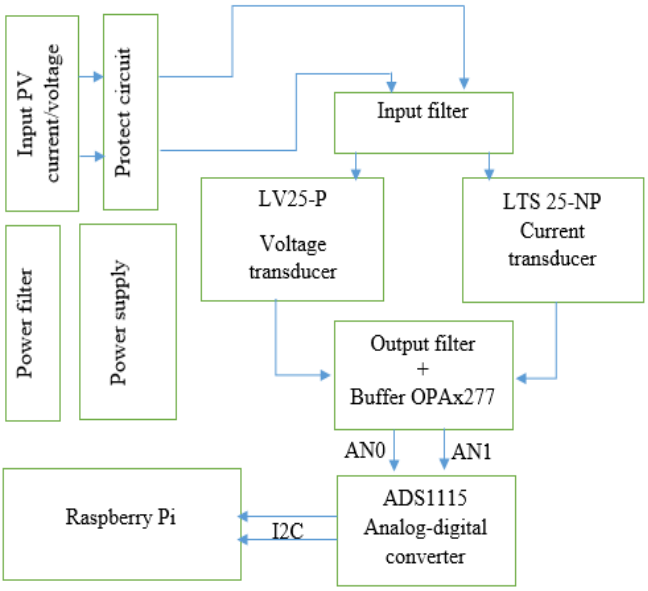


Figure 34. Block diagram of DC meter module

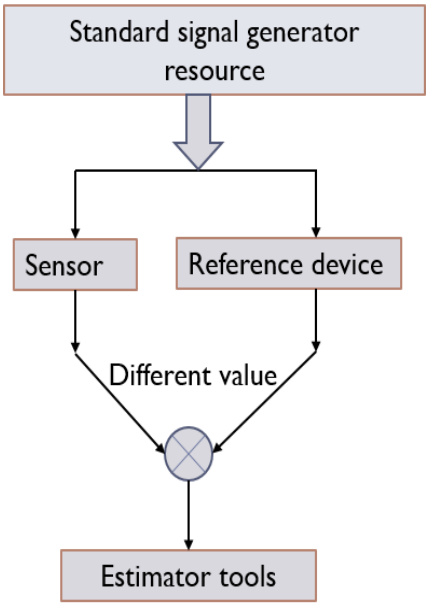


Figure 35. Processing of DC meter calibration

The diagram of the DC meter illustrated in Figure 34 and the processing of DC meter calibration in Figure 35. In the experiment, we used high-precision devices to calibrate the DC meter module.

2.3.3.2. Robust design against environment (packaging)

Experiment 2: Testing a self-develop sensor in high humidity environment condition and improving the protective box design

In this experiment, the BME380 multi-parameter sensor, developed by the team, was tested by comparing its measurement data with that of the AEOTEC MultiSensor 6 sensor. As presented in [36], the self-developed BME280 sensor showed a 15% error in humidity values when the room humidity exceeded 80%. The first version (V1) of humidity sensor met the target error under RH<80%. In a higher RH environment, the sensor error increased and a persistent bias appeared due to a humid micro-climate created in sensor's protective box.

To reduce moisture condensation, the box was redesigned with the following changes: (1) increasing the air gap and area to increase ventilation; (2) changing the position of the sensing head to avoid condensate from directly contacting the sensing surface (see Figure 36).

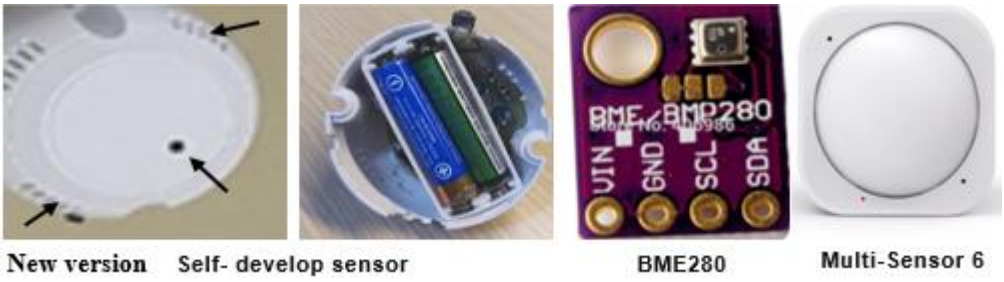


Figure 36. Description Self-develop sensor & commercial sensor

In one day experiment data at a high humid (>80%), the new version achieves better accuracy with error RH under 5% deviation comparing with Multi-sensor.

2.3.3.3. Virtual-sensor to reduce technical issues

By combining data, it is possible to infer critical information without an increase in physical sensors. Virtual-sensors are developed basing on mathematical equations and measure data to reduce number of physical sensors. In this study [68], making a charging battery virtual-sensor to mitigate technical risks.

Experiment 3: Virtual sensing (State of Charge without charging sensor) [68]

The experiment on greenhouse platform presented how calculate State of Charge (SoC) while lacking measured data in charging battery mode. Our study in [68] identified SoC based on relationship of charge/discharge mode with battery voltage changes. Therefore, the study proposed a function D(t) to observe the duration time of charging/discharging battery modes:

$$D(t) = \begin{cases} n, & V_B(t) - V_B(t-1) < 0 \text{ with } n > 0 \\ -n, & V_B(t) - V_B(t-1) > 0, \end{cases} \quad (2.1)$$

In discharging mode, we have Pbat=$P_{discharge}$

$$P_{discharge} = P_{loads} + P_{loss-discharge} - P_{PV} - P_{AC\_in} \quad (2.2)$$

$$E_{loss} = \sum_{n=start-discharge}^{end-discharge} P_{loss-discharge}(t_n) \times (t_n - t_{n-1}) \quad (2.3)$$

In charging mode, we have P_{bat}=P_{charge}

$$P_{charge} = P_{loads} + P_{loss-charge} - P_{PV} - P_{AC\_in} \quad (2.4)$$

The experiment data includes start-time, stop-time, C_i , C_{end} - Corresponding with battery capacity at initial, endpoint with (C_i = C_{end}).

For three-day data, assume that the losses were constant, $P_{loads}(t)$ is power of total loads at time t.

$$E_{loss} = E_{PV} + E_{Grid} - E_{loads} \quad (2.5)$$

Based on equations [1–6], the state of charge were identified.

$$\rightarrow SoC(t_n)=SoC(t_{(n-1)})+P_{bat}(\Delta t) \quad (2.6)$$

In this section, the experiment demonstrates a virtual SoC sensor based on Eq. (2.6) to replace the physical sensor.

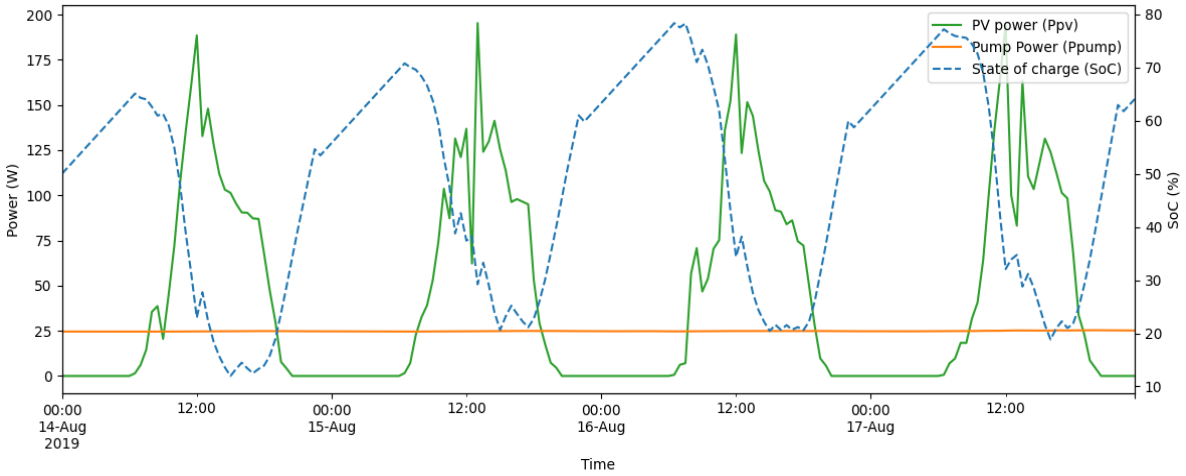


Figure 37. Description relationship of State of charge, Pump, and PV generation

Figure 37 illustrates the three-day trend of power flows (PV generation, Pump power and State of Charge) without physical charging battery sensor installation. In the next section, we will present a promising approach sensors network for data quality.

2.3.4. Sensor Network level - Experiments

2.3.4.1 Introduction

In sensor network, the evaluation and calibration of individual sensors become complexity and impractical. Prediction Models based machine learning techniques are increasingly valuable tools for enhancing quality building management [39]. Practical experiences highlight that datasets in sensor networks are crucial for ensuring system reliability. However, selecting input features and dataset sizes will effect to prediction errors if not properly managed.

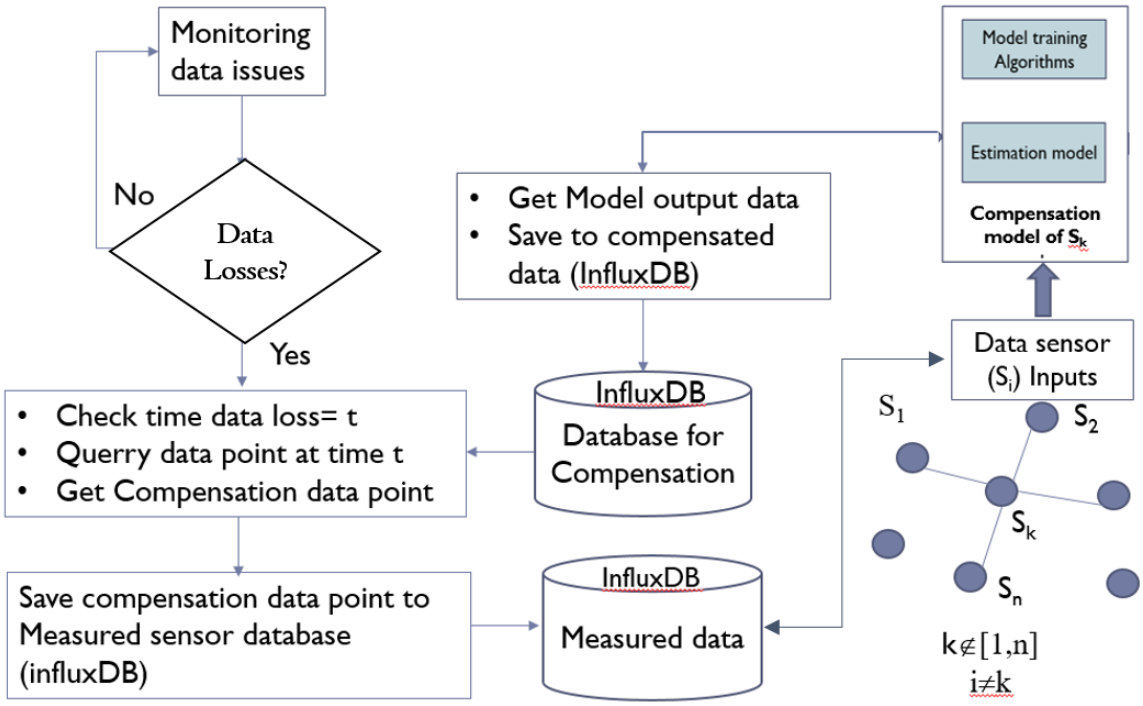


Figure 38. Deployment workflow of an online data compensating model

Figure 38 shows the second level of the data-quality framework, which handles missing or faulty measurements at the network level. When a data-loss event occurs, the system detects missing timestamps and sends a query to the compensation database. A compensation model, trained on neighboring sensors and historical data, then estimates the missing values. The estimated points are stored in an InfluxDB database and later merged with the measured data for visualization, analysis and control.

This workflow illustrates how data quality is maintained in practice on a low-cost platform. Data losses and sensor faults are processed automatically by the IoT-BEMS backend. Low-skill users do not need to intervene, while the system still preserves a consistent time series for analytics and control.

2.3.4.2 Experimental setup in VHH's platform

The test case is a low-cost IoT platform in the Vietnam-Korea Vocational college of Hanoi (VHH). The computer used for experiments is a general one in the offices with configured CPU of Intel Core i7-8550U 1.80 GHz. Two months of measured data (July and August 2020) with a 10-minute timestamp were used in real experiments. Table 4 shows measured data symbols in the experiments.

Table 4. Symbol of variables used in Experiments

No	Variables	Definition
1	TiM2	Temperature at node Mi02
2	TiZ1	Temperature at node Z01
3	TiZ2	Temperature at node Z02
4	TiZ3	Temperature at node Z03
5	Tout	Outdoor temperature from a web service.
6	HVAC	Value at point k of HVAC power in the office
7	P_{Total}	the total power consumption in the office
8	PVi with (i=1, 2, 3,4)	The power of the i^{th} PV system nearby.

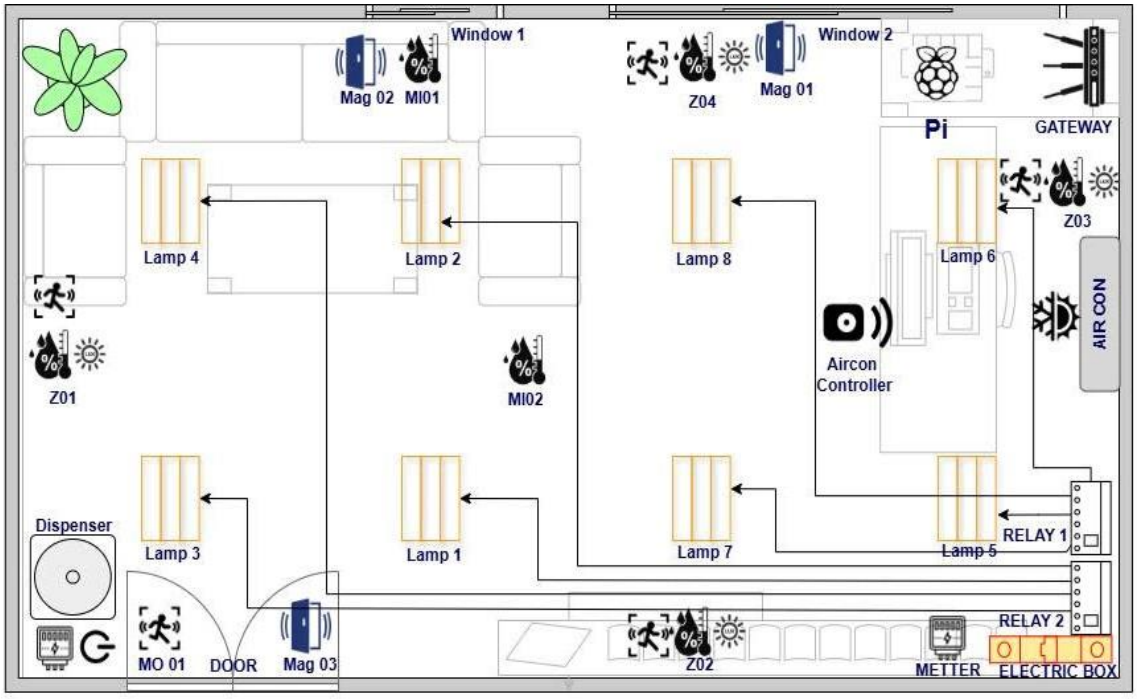


Figure 39. Description diagram sensor network of testbed in VHH

Figure 39 shows sensor types and locations in real platform. There are 2 energy meters (measuring sub-loads: HVAC and Total). Among some multi-sensors, four multi-sensors (Z01, Z02, Z03, and Mi02) placed on wall surfaces of building.

2.3.4.3 Machine learning for abnormal data detection & compensation

Nevertheless, practical challenges, such as computational latency and data availability, must be carefully considered under source-constrained conditions.

The literature [69], [70] indicated that Gaussian Process Regression (GPR) is a Machine Learning model for effectively managing data uncertainty. However, practical research on this model still has gaps in terms of limited-source computation and diversity data.

Main objectives:

- Propose efficient online machine learning models to detect different data faults and compensate for data in the IoT-BEMS platform.
- Focus on computational performance and available data.

a. Methodology

a.1 Multiple linear regression (MLR) [71]

"*Multiple linear regression is a generalization of simple linear regression to the case of more than one independent variable, and a special case of general linear models, restricted to one dependent variable*".

The basic model for multiple linear regression is:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} + \varepsilon_i \qquad (2.7)$$

Where: Each observation i = 1, ... , n.

In formula above, we consider *n* observations of one dependent variable and p independent variables. Y_i is the i^{th} observation of the dependent variable, X_{ij} is i^{th} observation of the j^{th} independent variable, j=1, 2, …, p ; β_j represent parameters to be estimated, and ε_i is the i^{th} independent identically distributed normal error.

a.2 Gaussian process regression (GPR) – Online model [72]

The GPR technique has several significant advantages over other methods due to providing a precise measurement with quantified reliability. However, the GPR is quite complex compared with other regression models, so its computational speed is limited. We proposed experiments to estimate the suitability of the GPR technique for the data compensation models, which could run online.

GPR is a statistical approach approximating input-output mappings from empirical data in prediction models. A Gaussian process regression equation:

$$y = f(x) + \varepsilon \qquad (2.8)$$

Where: $y \in \mathbb{R}$: observed scalar output.

$f : \mathbb{R}^{D\times 1} \to \mathbb{R}$: Latent Regression function.

$x \in \mathbb{R}^{D\times 1}$: Input feature.

$\varepsilon \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right)$: Gaussian measurement noise with zero mean and standard deviation σ_ε.

A Gaussian Process function is:

$$f(x) \sim \mathcal{GP}\left(\mu(x), k\left(x, x'\right)\right) \text{ , and } y \sim \mathcal{N}\left(\mu(X), K + \sigma_\varepsilon^2 I\right) \quad (2.9)$$

Where: $\mu(x)$: Mean function (often set to 0).

$k\left(x, x'\right)$: Covariance (kernel) function with $x, x' \in \mathbb{R}^D$ (input vectors)

$X \in \mathbb{R}^{n\times D}$: Training - input matrix whose i^{th} row is $x_i^\top$.

$y = \left[y_1, \ldots, y_n\right]^\top \in \mathbb{R}^{n\times 1}$: Training outputs stacked as a column vector.

$I \in \mathbb{R}^{n\times n}$: Identity matrix.

$$K = k(X, X) \in \mathbb{R}^{n\times n}, K_{ij} = k\left(x_i, x_j\right) \quad (2.10)$$

Because updating a batch GPR requires re-inverting $K + \sigma_\varepsilon^2 I$ with $\mathcal{O}\left(n^3\right)$ cost per update, an Online GPR is preferred.

RBF kernel function use in Experiments:

$$k(x, x') = \sigma_f^2 \exp\left(-\frac{\| x - x' \|^2}{2\ell^2}\right) \quad (2.11)$$

Where: $x, x' \in \mathbb{R}^D$: Input vectors

$\sigma_f^2 > 0$: Signal variance (amplitude of the latent function); $k(x, x) = \sigma_f^2$

$\ell > 0$: Length-scale (smoothness) – (If Length-scale value is small, the function varies rapidly; If Length-scale value is large, the function varies slowly).

By Equ. (2.8) ÷ Equ. (2.11), the Gaussian process model is computed and fitted to the training set using maximum log likelihood method to tune the hyper-parameters **(**Signal variance and Length-scale).

b. Experimental results

Experiment 4: Evaluating MLR and GPR techniques for compensating temperature data

In this experiment, two models, GPR and MLR, are for room temperature compensation. The purpose is to compare performance of the GPR and MLR technique for online data compensation models.

Data collection: Ten day data in Aug, 2020 is separated in (Train: 8 days)/ Validate: 1day/ Test: 1 day).

Data input: TiZ1, TiZ2, and TiM2lag1, time features (sin/cos of hour and day).

Data output: compensated PV1 power.

In this experiment, Temperature data at Z01 and Z02 node is model's input. The compensation temperature point of node M02 is the model's data output. In data preprocessing, time features (Trend features) were added in dataset to improve model quality.

In this experiment, GPR with an RBF kernel vs trend and seasonal components, trained and test, then producing sequential forecasts on validation and periodically refitting every 1 hour using a 72 hour window. An alpha of 3e-5 improves numerical stability. Estimators (RMSE, MAE) are used to estimate on fault region.

This work injects data faults into the validation set to simulate fault data and evaluate model robustness**.** Fault data imputation simulation based on injecting missing data in validation dataset (in Table 5):

- **Block missing:**

(1) NaN: Set NaN for TiM2 validation faulty in specific range (with k points). Number of points in block equal k points. In this experiment, k equal approximating 30% total number of samples in validation dataset.

(2) Spike: create noise in the block. With scale of 3, each point in the block is randomly "mutated" by 3 times the standard deviation.

➢ **Scatter missing:** Set NaN for k discrete points in TiM2 validation data.

Table 5. The results of data compensation error on validation TiM2 dataset

Type of missing data of sensor TiM2			**Estimators (°C)**	**GPR**	**MLR**
Block missing	30%)	NaN	RMSE	0.48	1.4
			MAE	0.41	1.38
		Spike	RMSE	0.35	1.39
			MAE	0.14	1.38
	15%	NaN	RMSE	0.05	1.11
			MAE	0.04	1.1
		Spike	**RMSE**	0.022	1.15
			MAE	0.018	1.16
Scatter missing	30%		RMSE	0.32	1.35
			MAE	0.17	1.27
	15%		RMSE	0.36	1.63
			MAE	0.19	1.47

The table above shows that across all data-loss scenarios (30% block and 30% scatter, and 15% block and 15% scatter), the GPR model consistently has lower RMSE and MAE than the MLR model.

- ✓ For the case of 30% block missing and NaN data, the GPR error (RMSE = 0.48; MAE = 0.41 °C) is still much lower than MLR (RMSE = 1.4 °C; MAE =1.38 °C). When the reduced loss rate data are up to 15%, GPR model error drops rapidly (RMSE = 0.05 °C; MAE = 0.04 °C). In contrast, MLR model error remains above 1 °C.
- ✓ For Spike block missing, GPR still gives very small errors (RMSE is only about 0.35 - 0.022 °C, MAE = 0.14 - 0.018 °C), while MLR fluctuates around 1.15-1.39 °C. This shows that GPR not only compensates well when the data are NaN but also “flattens” spikes, keeping the error at a very low level even when large-amplitude noise appears.
- ✓ For scatter missing, GPR continues to maintain lower error (RMSE = 0.32-0.36 °C compared to 1.35-1.63 °C for MLR).

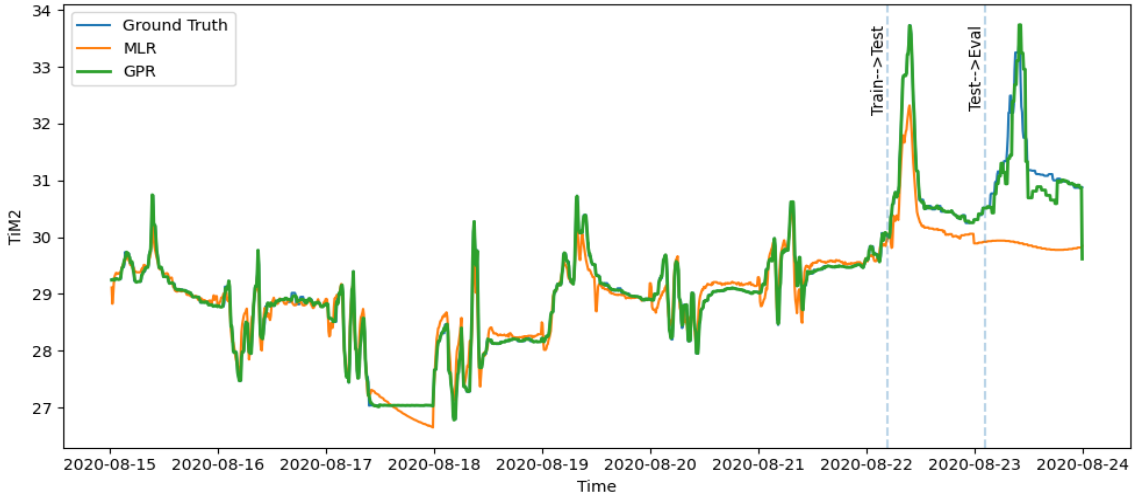


Figure 40. Comparison performance of MLR vs GPR on ten days dataset with 30% Block-missing data (NaN)

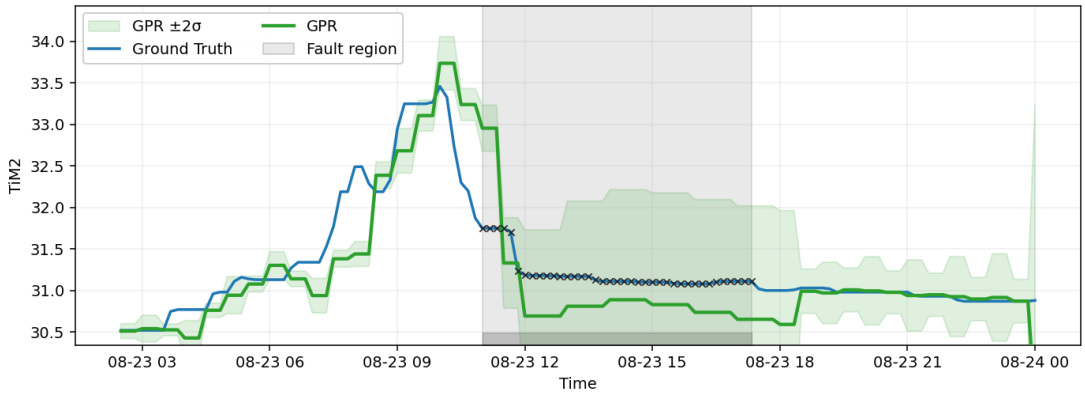


Figure 41. Evaluation imputation performance with 30% Block-missing data (NaN)

Figure 40 shows performance of MLR and GPR model for data compensation on ten days. Figure 41 presents clearly by indicating missing points, missing region and confident. GPR present a better performance than MLR in test and validation steps. In summary, GPR provides better, more stable data compensation than MLR in both block and scatter data losses, especially when the data loss ratio is high.

Experiment 5: Evaluating the GPR model's computational performance.

In this experiment, we propose a method for compensating indoor temperature data using online models. The primary objective is to balance computational performance (on training time) with effective data compensation.

The study in [73] indicated that training data comprising a half-week to one-week data are for feasible online data compensation models. In the study, we investigate the models' data compensation horizons. Two models (Model 1 and Model 2) were trained

using different dataset sizes collected in July 2020, respectively. Then. a 2.5-day dataset comprising 350 samples collected was used to evaluate performance of models.

Table 6. Estimation factors of the compensation of TiM2 data model

	Model with 2.5-day test dataset	Model 1	Model 2
TiM2	RMSE (°C)	0.18	0.22
	MAE (°C)	0.49	0.5
	R2	0.984	0.986
	Training Time (second)	3.25	19

In Table 6, the estimated indoor temperature values were compared against actual data and both results are summarize with both RMSE, MAE and R2 values.

- ✓ With a 2.5-day test dataset (350 samples), both models achieve good accuracy: RMSE =0.18÷0.22 °C, MAE = 0.49÷0.50 °C and R^2 = 0.984÷0.986, which is acceptable for building energy management.
- ✓ For Model 1, training time is about 3.25 s and 19 s, confirming that both can be retrained online, but Model 1 is better suited for near-real-time updates on resource-constrained devices.

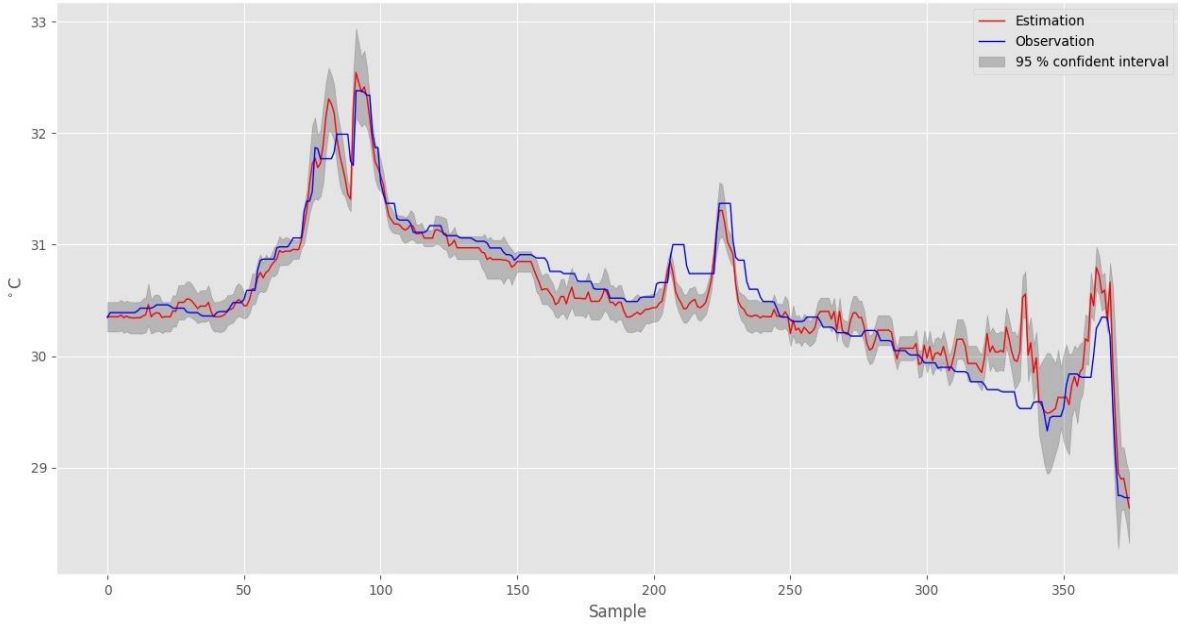


Figure 42. Estimation and observation of indoor temperature of model 1

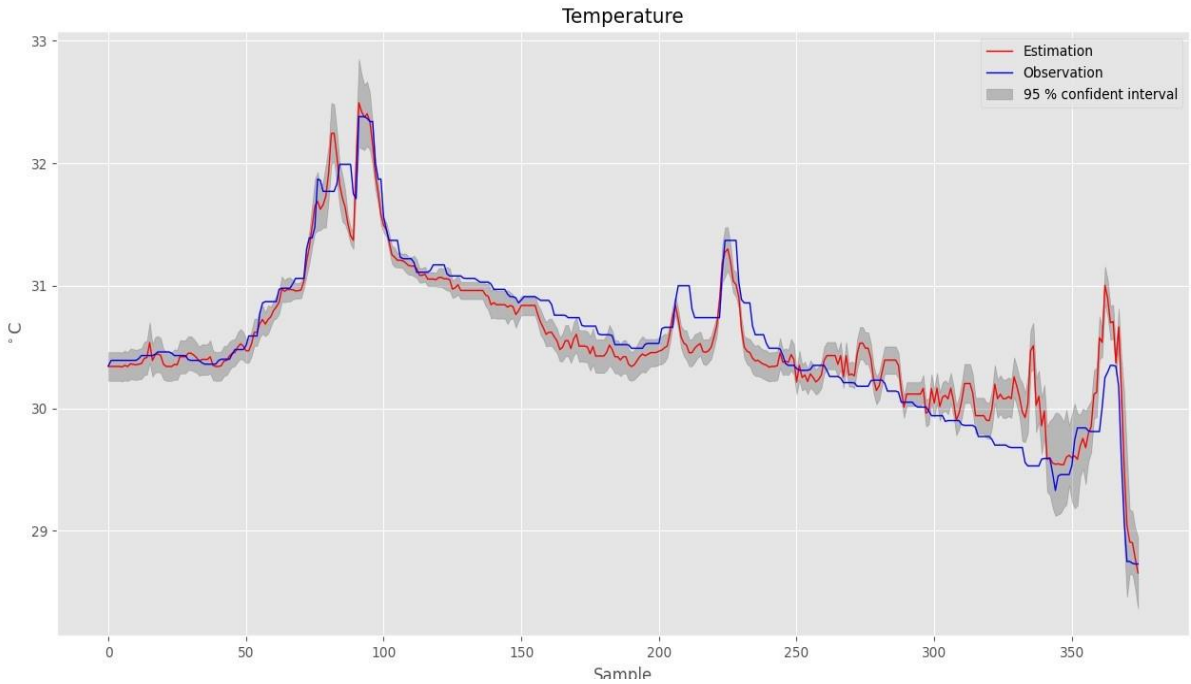


Figure 43. Estimation and observation of indoor temperature of model 2

Figures 43 and Figure 44 further show that, over a 2.5-day loss period, the compensated temperatures closely follow the measurements, indicating that the proposed compensation approach remains reliable over several days of missing data.

Experiment 6: The GPR data compensation models of local PV systems.

The Gaussian Process Regression is for developing online photovoltaic (PV) power compensating models [73]. The study investigated the development of a compensated data model for PV generation based on the correlation of multiple PV systems.

Data collection: The 1-week dataset is split into Train: 4 days, Validation: 1 day, and Test: 2 days. PV data are time-aligned and resampled at 10-minute intervals. To simulate data loss, missing blocks are created in the daytime interval [06:00–18:00], with block lengths of 6–12 samples.

Data input: PV2, PV3, two time-lagged (1hour) values of PV1 (lag1, lag2), and time features (sin/cos of hour and day).

Data output: compensated PV1 power.

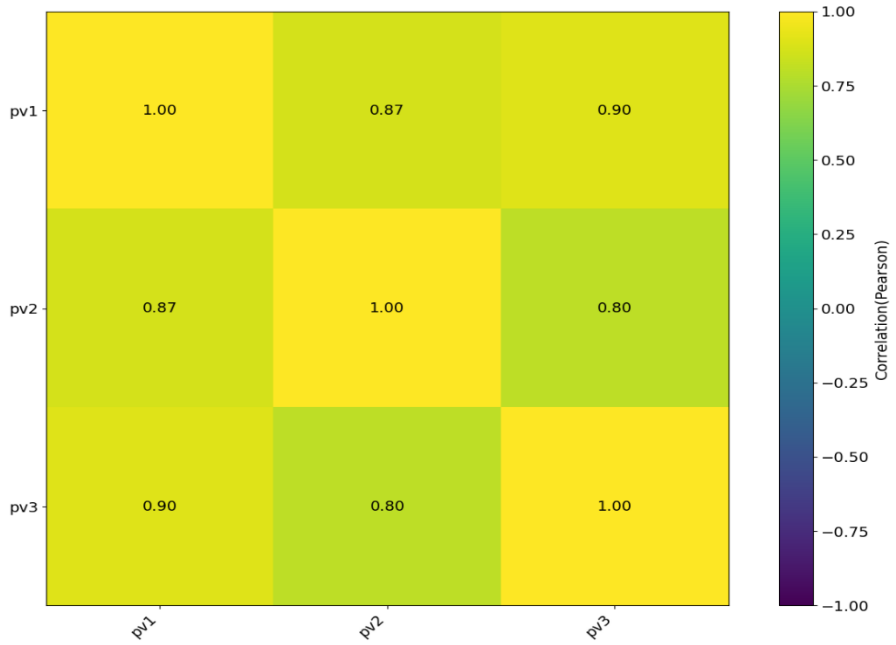


Figure 44. Correlation generation data of four PV systems

Figure 44 shows the strong correlation between PV1 and the PV2–PV3 stations over the selected 1-week period (data collected in January 2022). The final model is refitted over the entire Train before the Validation/Test prediction. When compensating for each missing block in Validation/Test, the model only uses data before the start of the block (history) to learn. This reflects the real-world model deployment (no looking into the future).

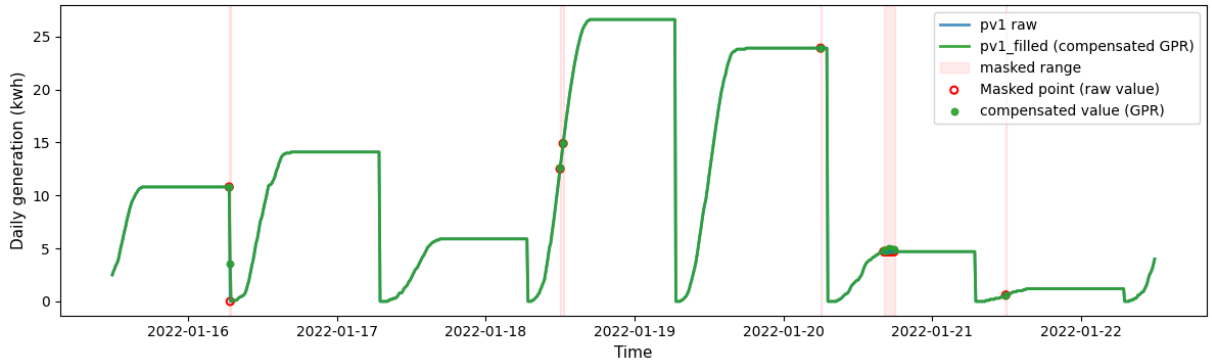


Figure 45. GPR compensation data PV1 (sample rate 10min)

Figure 45 shows the data compensation results at some missing data points in a week (Train/Validation/Test). The result on Test reported at the masked points has a Coverage value of 95.1%, showing that the model is reliable. The error on Test MAE is 0.347, RMSE is 0.878, and R2 is 0.976, showing that the model

matches the points accurately. This confirms that the model reproduces the true PV1 power and is sufficiently robust for practical data-compensation applications.

2.4. Adaptation of the IoT-BEMS platform to Local Conditions

In developing the IoT-BEMS system, the conceptual foundations and architectural framework were initially derived from a large-scale IoT platform in France that relied mainly on commercial sensors. To deploy in Vietnam, several adaptations were required to account for local conditions. From the research experiences in France, projects in Vietnam were designed to priorities affordability, flexibility, and practical applicability. Instead of deploying high-cost commercial sensors, the systems are implemented by integrating low-cost commercial and self-developed devices. This integration was validated through experimental activities to ensure the platform's interoperability, reliability, and scalability under local constraints. The adaptation process focused on reducing hardware costs and maintaining data quality, system robustness, and user engagement. Therefore, this study provides valuable insights to scale the IoT-BEMS infrastructure in Vietnam.

2.5. Conclusions of part 2

In this part, the work proposed a highly interactive, modular, heterogeneous IoT architecture (RF24/ZigBee/Z-Wave/Wi-Fi, Raspberry Pi, time-series database, and Real-time monitoring). Based on the implementation of real projects using low-cost sensors and open-source platforms:

- Cost initial infrastructure cost ≤ a few thousand USD.
- With no-code, pre-configured templates and “step-by-step” instructions, basic skill users can deploy in one day for a standard room such as VHH's testbed, schedule maintenance once a month, and the system automatically warns when data is missing or incorrect.

This confirms the low-cost, low-skill friendly platform, suitable for deployment conditions and easy to expand (many rooms/hundreds of measuring points).

The proposed pipeline (including system design, data management and configuration, and a two-level DQ insurance framework) enables reproducibility and maintenance in real-life operations.

Overall, part 2 has provided a practical roadmap for designing an IoT-BEMS platform with reliable data. In the following part, we will focus on a case study to apply the platform to evaluate users' interventions and feedback, as well as energy-efficiency strategies.

Part 3. The implementation of IoT-BEMS platform – A case study

3.1. Overview

In this part, we present the application of IoT techniques for managing various subsystems in an actual building. First, we introduce the experimental platform integrated with IoT technologies. Subsequently, we provide insights into the electrical devices within the testbed environment.

The study explores how technology and data can be leveraged to implement energy efficiency solutions. The correlation analysis of environmental conditions, user behaviors, and energy provides a foundation for developing energy services to reduce overall energy use and electricity costs [74].

Although the study was deployed on a small-scale platform, the findings from this case study can serve as a basis for scaling up to larger applications through the widespread use of IoT devices [75].

In addition, it is important to note that scientific studies in Energy Systems have widely adopted a small-scale validation strategy. Thus, this research provides a scalable foundation for broader real applications.

3.2. Introduction VHH' project

This is a project under the cooperation between the Vietnam-Korea Vocational College of Hanoi City (VHH) and the University of Science and Technology to develop an experimental platform for research purposes.

In the platform, building services included the following:

(1) Monitoring building status;

(2) Maintenance devices (replace old-date lighting devices, sensor outage battery);

(3) Automatic control (schedule and energy-saving scenarios) for Water dispenser, Lighting system (with 24 units of Lamp) and Air conditioner (12,000 Btu);
(4) Access database and analysis.

Figure 46. Description location PV system in VHH [36]

Figure 46 is an overview of the VHH building, which is responsible for the energy strategy experiments with a PV system a Hybrid Inverter 5 kW 1 phase with 04 lead-acid battery work in series [36].

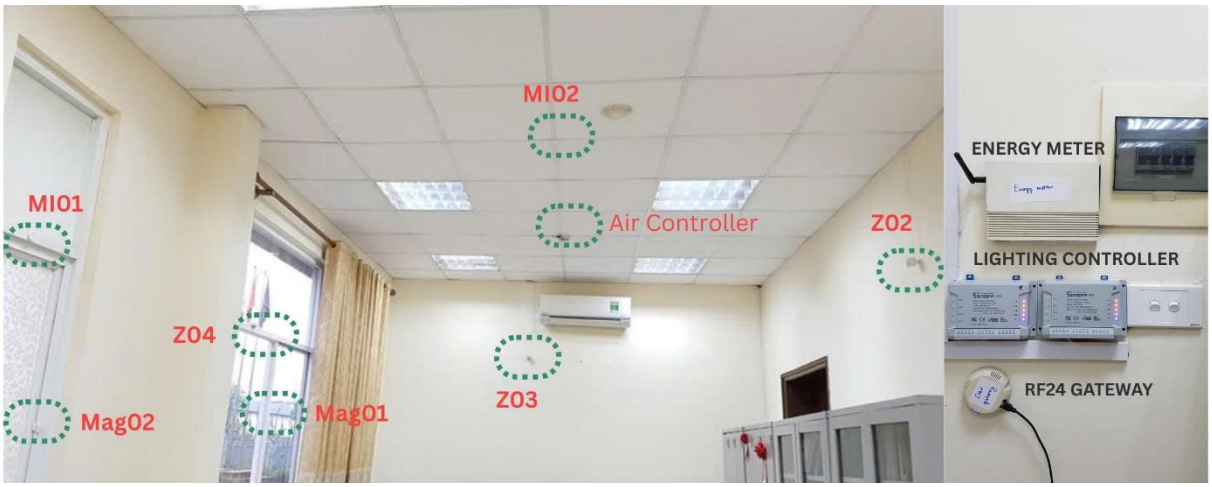


Figure 47. Setting up an IoT experimental platform including location of sensor nodes, air conditioner and lighting controllers, and energy meters [36]

Figure 47 shows an IoT experimental room layout in VHH for real-time monitoring and control using wireless sensors and controllers network. The network integrates multi-communication protocols (RF24, Z-Wave, ZigBee, and Wi-Fi) via a Raspberry Pi-based local server, enabling real-time data collection, energy monitoring, and HVAC control within testbed.

Table 7. IoT Network Configuration for the testbed

Component	Function / Description
01 - Central Server Raspberry Pi 3	Acts as a local server and data hub; manages three gateways (RF24/Wi-Fi, Z-Wave, ZigBee) and stores real-time sensor data.
01 - RF24 / Wi-Fi Gateway	Transmits measured data from custom energy meters to the server and sends control commands back to actuators.
01 - AEOTEC Z-Wave Stick	Interfaces with Z-Wave multi-sensor nodes for indoor environmental data acquisition.
01 - Xiaomi ZigBee / Wi-Fi Hub	Connects ZigBee sensors for temperature, humidity, and door status monitoring.
02- ZigBee Sensors Xiaomi Multi sensor (T, RH)	Measure indoor temperature and relative humidity.
03- Xiaomi Door Sensor	Detect open/close state of doors and windows.
04- Z-Wave Sensors AEOTEC Multi sensor 6	Monitor multi-environmental parameters ((T, RH, Light, Motion,…).
01- RF24 one phase Energy Meters	Monitor power consumption of the water dispenser.
01- RF24 four-phase Energy Meters	Monitor multi-channel electrical consumption
02- Wi-Fi Sonoff 4CH Pro	Control 8 LED lighting channels
01- Broad-Link SP3 Mini	Control the water dispenser power supply.
01- Wi-Fi Controller	Enables remote and automated HVAC control.

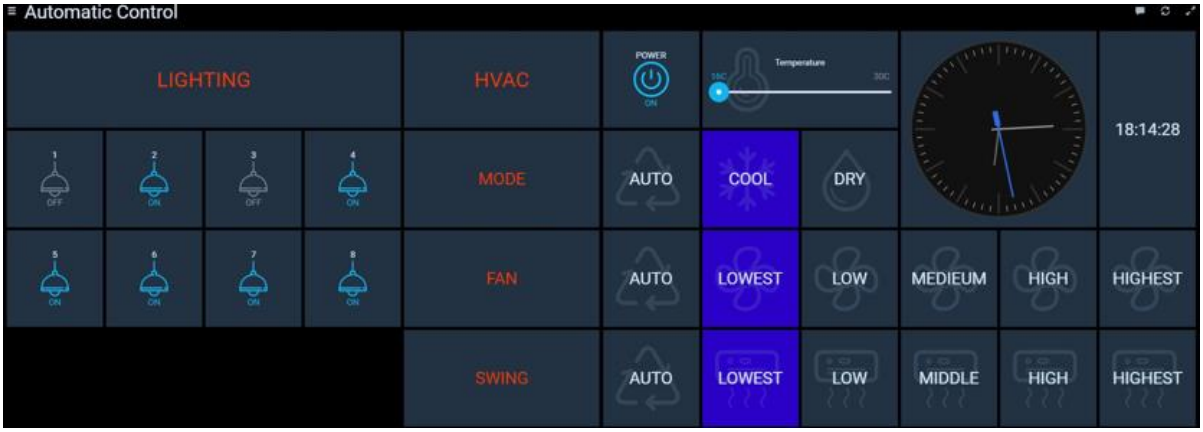


Figure 48. Description a Control Users Interface in VHH platform

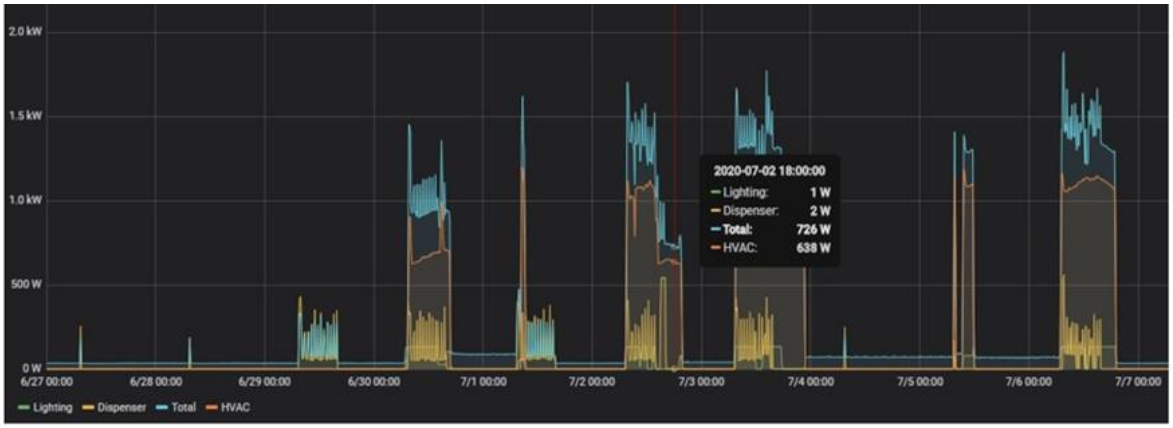


Figure 49. Description a Dashboard Users Interface in VHH platform

Figure 48 and Figure 49 present Users Interface for control lighting and Air conditioning, and monitoring energy consumption. Users could interact with IoT platform using laptop or smartphone.

3.3. Deployment of IoT-BEMS platform for nZEBs – A case study

This section highlights the practices for monitoring and controlling energy targets. Data collection and processing can determine the actual building performance and the effects of user behavior on energy demand. Feedback from measurement systems is an effective means of influencing and changing behavior [34], [36].

This work focuses on strategies to influence behaviors and operate the building more efficiently. Measurement data could provide insights into the relationships among environmental conditions, behaviors, and energy [76].

Monitoring plans identified significant energy savings, behaviors that require change, interventions to address the problem, and user benefits.

3.3.1. Environmental conditions vs Energy Correlation analysis

The purpose of this section is to demonstrate the relationship between environmental conditions (such as temperature and humidity) and energy consumption, providing a baseline for understanding user-driven behavioral effects in the next section.

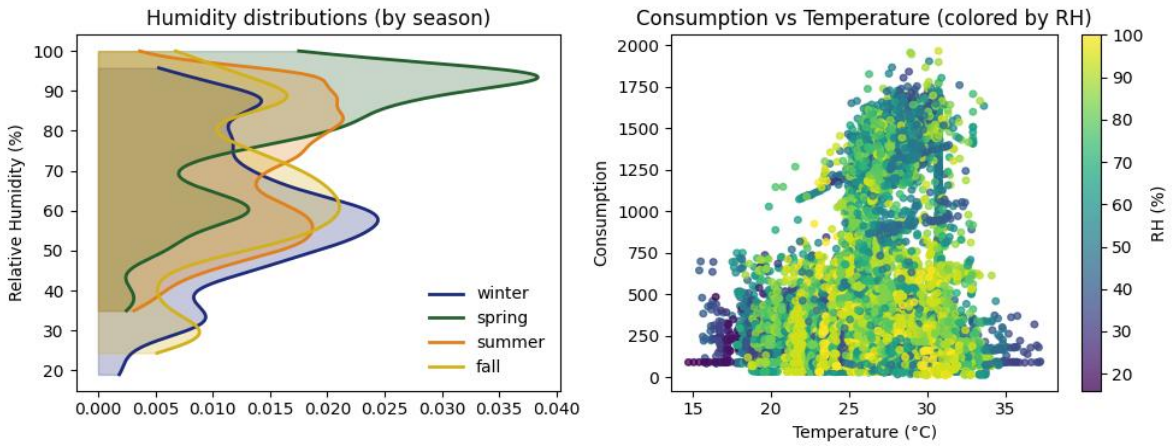


Figure 50. Description correlation of Temperature, Humidity vs Consumption by season in VHH's office

Figure 50 shows the seasonal distribution of humidity and the relationships among temperature, humidity, and consumption. Summer has high moisture (relative humidity [RH]> 60%), which significantly increases the cooling load. The scatter plot shows that, at the same outdoor temperature, higher RH values are associated with higher consumption. At high temperatures (over 30°C), a saturation zone appears, in which decreasing the set-point no longer increases cooling efficiency, indicating the capacity limit.

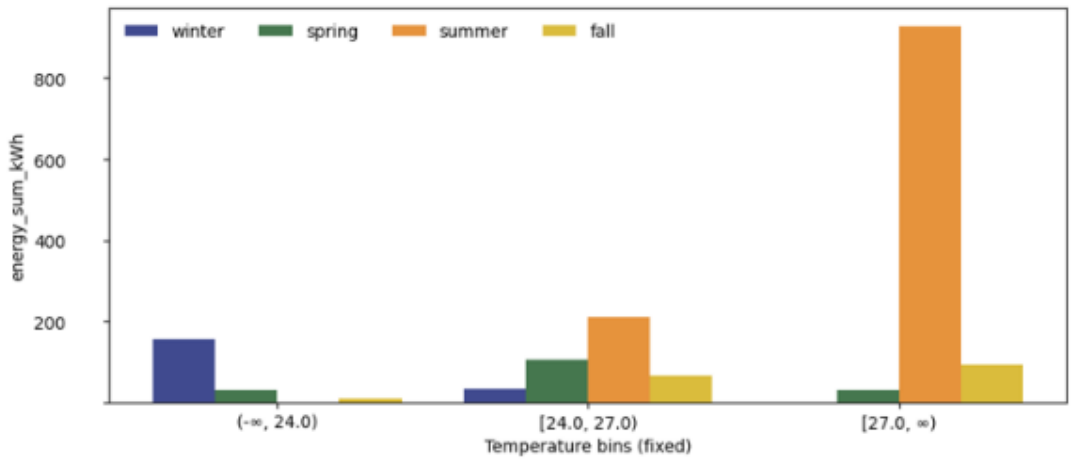


Figure 51. Energy by bins vs season in VHH's platform

Figure 51 shows HVAC consumption by temperature and season. Consumption increases when T_{out} exceeds 27 °C, and summer has the highest values. For the same temperature range, summer always uses more energy than other periods because both heat and humidity rise together. In spring and autumn, consumption is much lower thanks to moderate weather.

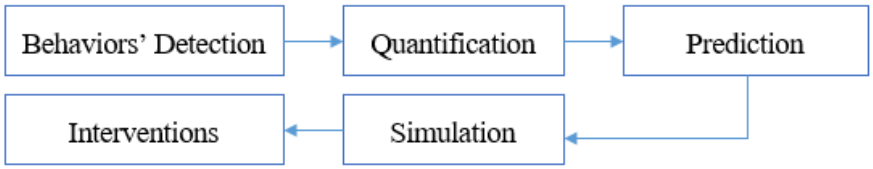
Although weather strongly influences HVAC demand, daily consumption is not fully explained by environmental factors. Behavioral factors such as set-point adjustments and openings also have a significant impact. Therefore, the following section focuses on quantifying the effects of behavior under different environmental conditions.

3.3.2. Data-driven behavior change analysis

Aim of this section:

- Analyze HVAC control behaviors to reveal patterns in energy use.
- Measure HVAC consumption associated with key behaviors (habit/time, set-point, door status, occupancy).
- Propose behavior-change measures that improve efficiency.

Proposed a data-driven cycle for efficient, and sustainable operations:



Evaluation:

- Assess actual operation performance and user behavior.
- Quantify energy-saving potential achievable through behavioral interventions.

3.3.2.1. Behaviors' HVAC on and Opening vs Energy Correlation

The relationship between behaviors and energy could reveal wasteful factors. This study analyses the relationships among opening, HVAC operation, and consumption over five months to identify the behavior of wasted energy in an office.

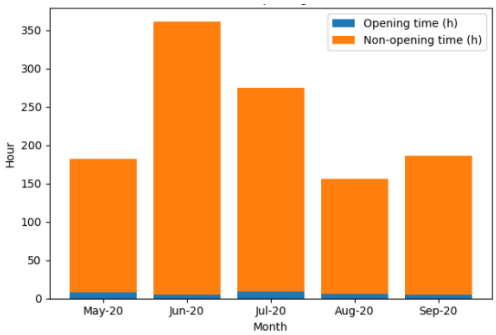


Figure 52. HVAC On (h) Total=1159.6h, Opening =2,9%

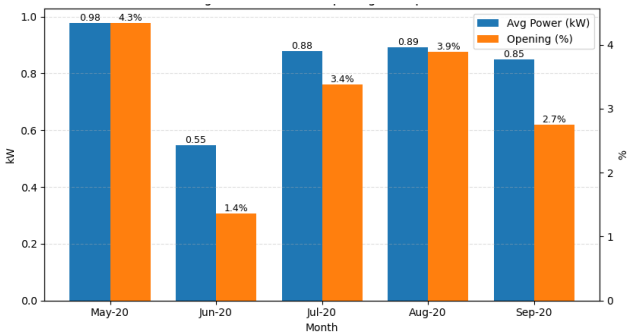


Figure 53. Average HVAC power and Opening Portion

Observation of waste energy behavior in Figure 52 shows an opening when

HVAC is on. In Figure 53, the significant increase in average HVAC power is attributable to the opening behavior. In May, this value (0.978 kW) was more than twice that of June (0.547 kW), corresponding to the opening (4.4% and 1.4%). The information could help managers require staff to engage in energy-saving efforts. The analyzed changes in energy behavior and the potential energy savings could inform HVAC operational scenarios.

3.3.2.2. Set-point temperature and Energy correlation

According to results in the previous study [36], this work further analyzes the relationship between set-point vs indoor temperature, and HVAC consumption across operating sessions.

Data processing: (1) Consider HVAC is on; working hours (06AM–19PM); door/window are closed. (2) Grouping by outdoor conditions: Outdoor temperature ≥ 30°C and divide RH into 2 groups (RH≤60%, RH in the range of 60–70%) for fair environmental comparison. (3) Analyze by the median value of the session and use IQR (Q25–Q75) to represent the variation (this will reduce noise compared to coarse dispersion).

Explanation of symbols:

- Average indoor temperature: Tin, compared to set-point: $T_{setpoint}$: The line represents the average indoor temperature when on (Tin). The shaded area represents the range (Q25–Q75).
- Tracking error = $T_{in} - T_{setpoint}$: Equal to 0 when the set-point temperature equals the room temperature.
- HVAC energy used per on session (kWh/session): The line shows the average energy consumption and the shaded area where the interquartile range (IQR) can be analyzed. The label n at each point indicates the version currently on.

Results and discussions:

In Figures 54–55, T_{out} is over 30°C, T_{in} stays around 30°C, and the tracking error ($T_{in} - T_{setpoint}$) remains positive (about 5–6°C). This means the system does not reach the set-point, which is consistent with a limited-power condition under high heat load. When the set-point is changed from 24°C to 25°C and then 26°C, Tin changes very little and the error does not shrink. Thus, lowering the set-point does not improve the indoor temperature.

Figure 54 (zoomed to 21–23°C) shows sessions with very low set-points (21–23°C), the error is still clearly positive. Even setting an extremely low set-point does not help the room reach the target temperature.

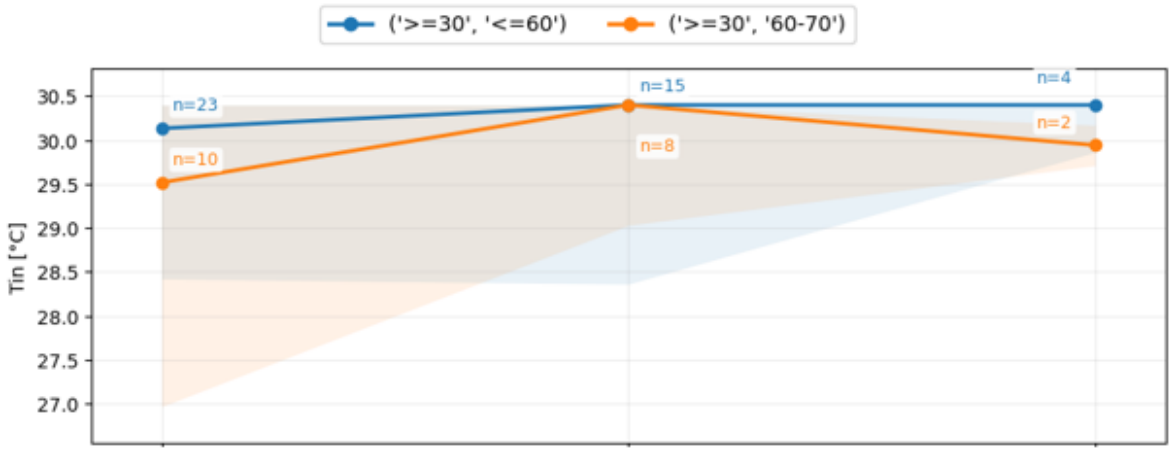


Figure 54. Indoor temperature response to set-point (median/IQR), 24–26°C

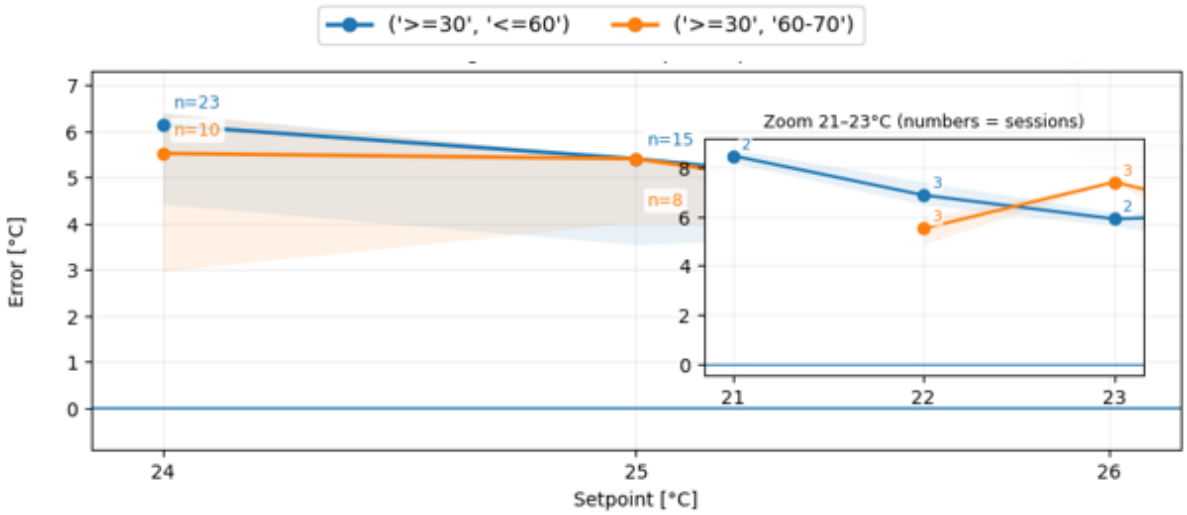


Figure 55. Tracking error ($T_{in} - T_{setpoint}$), session median ± IQR, 24–26°C

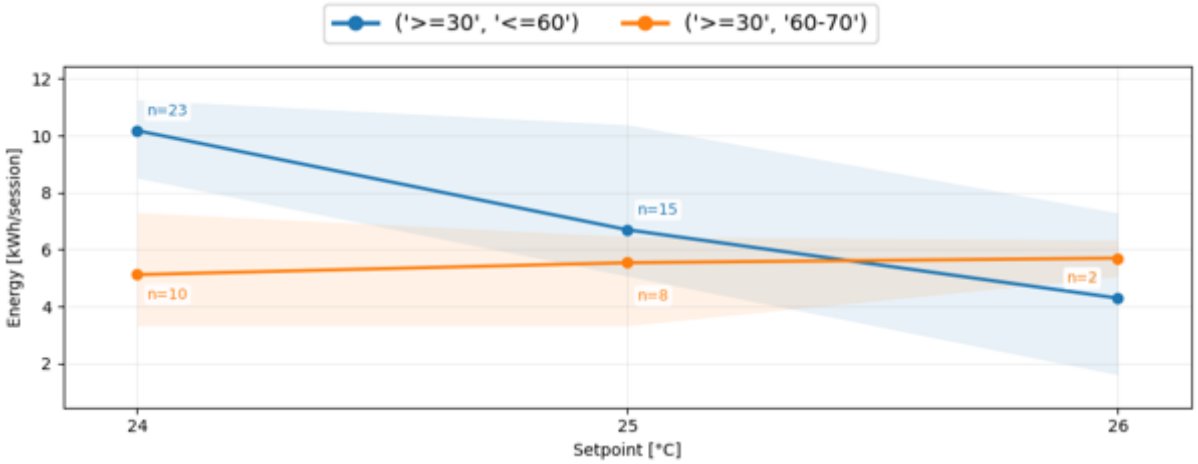


Figure 56. HVAC energy per ON-session (kWh/session) in the 24–26°C set-point range

In Figure 56, energy consumption per session increases as the temperature set-point is lowered (most noticeable at 24°C), but the Tin index does not decrease proportionally. This indicates an "oversetting" phenomenon: users set the temperature too low, increasing energy consumption while providing only minimal comfort benefits. Figures 54-56 also show the differences between humidity groups. In high humidity conditions, users can lower the temperature set-point, but the Tin index only reacts slightly. This indicates weak cooling performance in hot and humid conditions.

Interventions: Re-examine HVAC system size and consider performance upgrades when high temperature saturation is observed. Higher temperature set-points or adaptive temperature set-points are recommended. In saturation conditions, lowering the temperature set-point often increases energy consumption without a commensurate improvement in comfort.

3.3.2.3. HVAC consumption modeling - Experiments

Using historical data to develop a model and then assume a change in $T_{setpoint}$ and estimate what will happen to energy & thermal conditions if the policy is applied.

The experiment purpose:

- Simulate policy based on the condition if the $T_{setpoint}$ is increased/decreased (1°C), identify how expected energy consumption changes.

- Using the Pareto curve to analyze Energy use (kWh) and Comfort penalty ΔT to optimal savings with the minimal thermal comfort impact.
- Understand the mechanism by using a Heat map HVAC ON probability against T_{out} vs $T_{setpoint}$ to see when raising $T_{setpoint}$ is most effective.

a. Methodology

(1) Gaussian Process-based HVAC consumption prediction model

$$P_{HVAC} = f(T_{setpoint}, \Delta T, \text{occupancy}, T_{out}, T_{in}, \text{hour}, \text{weekday}) \quad (3.1)$$

- Change simulated $T_{setpoint}$ on entire sample (step 1°C).
- Expected energy:

$$\sum P_{on} \times P_{HVAC} \times \Delta t/1000 \quad (3.2)$$

(P_{on}=1 when HVAC is ON and P_{on}=0 when HVAC is OFF)

- Comfort ΔT: penalty if the set-point is outside the zone

(2) Description of measured Data

Measured data are collected from the building IoT system over 7 days with a sampling period of 10 minutes.

Main variable groups:

- System status: HVAC_ON, $T_{setpoint}$
- Environmental conditions: indoor vs outdoor temperature (T_{in}, T_{out})
- Human factors: occupancy
- Energy: P_{HVAC}
- Variables: hour, weekday, temperature difference $\Delta T = T_{in} - T_{setpoint}$.
- Behaviors: HVAC ON/OFF, Change $T_{setpoint}$

Assumption: increase $T_{setpoint}$ by +1°C, other factors remain no change.

Estimated Data: One day data: 2020-09-10;

b. Results and analysis

The policy is to reduce energy use (kWh) while increasing ΔT slightly.

Table 8. Simulation results of constant set-point policies: expected daily HVAC energy and comfort penalty (10 September 2020)

Set point	Expected energy use (kWh/day)	ΔT (policy)
22	3.39	8.5
23	3.60	7.5
24	4.03	6.5
25	4.76	5.5
26 (baseline)	**5.44**	**4.5**
27	4.76	3.5
28	4.03	2.5

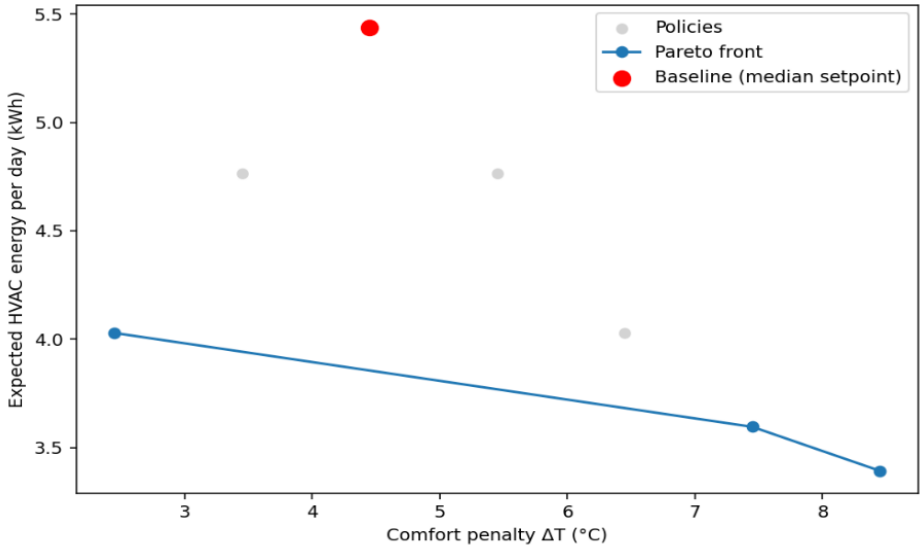


Figure 57. Pareto curve for energy and comfort analysis (10/09/2020)

On the summer day, Table 08 and Figure 57 report the expected daily HVAC consumption and the comfort penalty ΔT under different constant-cooling set-point policies. The baseline operation (at 26 °C) consumes about 5.44 kWh per day with comfort penalty ΔT (4.5°C).

From three policies (at 22 °C, 23 °C, and 28 °C), the 28 °C set-point is attractive for reducing the expected daily HVAC consumption to 4.03 kWh (about 26% energy savings). So, the comfort penalty must decrease by about 2 °C compared to the baseline. In contrast, set-points at 22–23 °C could achieve larger energy savings (around 34–38%), but at a higher comfort penalty (+3–4 °C) (not suitable for general office use).

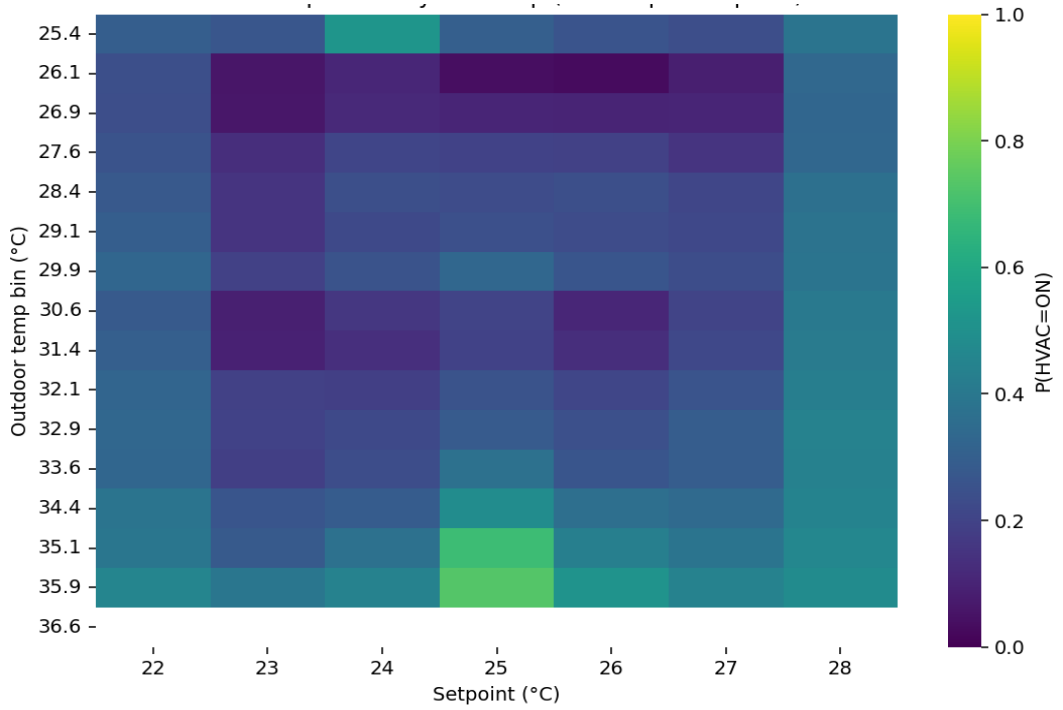


Figure 58. HVAC ON probability heat map (T_{out} vs $T_{setpoint}$)

A Heatmap (Figure 58) is used for adaptive control suggestions. The 25–27°C set-point column for the high T_{out} zone is lighter in colour (lower HVAC on probability). In the lower T_{out}, the effect of increasing $T_{setpoint}$ is negligible. This suggests that the set-point can be flexible by season. These results illustrate set-point recommendations for both energy efficiency and comfort, and quantify the trade-offs associated with more energy-saving strategies.

Interventions:

- On one summer day, increasing the set-point from 26 °C to 28 °C reduces the expected HVAC consumption from 5.44 to 4.03 kWh/day (26% savings) with a lower comfort penalty (about 2 °C).
- At high outdoor temperatures (above 30 °C), Figure 58 shows that increasing the set-point by +1 °C reduces the probability that the HVAC will turn on. Therefore, a simple control rule is increasing $T_{setpoint}$ by +1 °C whenever T_{out} > 30 °C, to limit peak cooling demand while keeping ΔT within the comfort limits of the Pareto analysis.

3.3.2.4. Behavioral change strategies

- **Real-time feedback:** display warnings when the room is empty but the Air conditioning is still on or the door is open when cooling.
- **Target setting:** sets a target to reduce P_{HVAC}/week.
- **Automation:** turn off Air conditioning after 15 minutes of no occupancy, synchronize with working schedules or occupancy sensors.

3.3.3. Rooftop Building - Energy strategies analysis

Maximizing self-consumption is essential for balancing demand and local power supply. This section evaluates self-consumption capability by comparing load indices across different time scales.

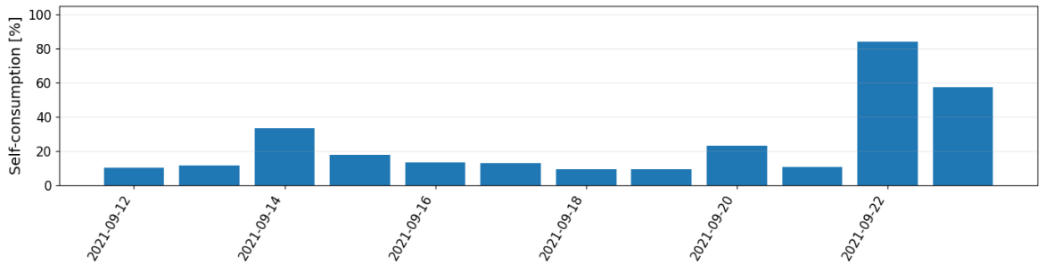


Figure 59. Daily self-consumption from 12/9/2021-23/9/2021

On the daily scale, the self-consumption rate fluctuates widely from 10% to over 82% (see Figure 59). The results in different near days indicates that the variation is not only from weather, it could be together with both storage plan and user behaviors at smaller scale [36]. It is necessary to investigate in smaller scale to improve self-consumption rate.

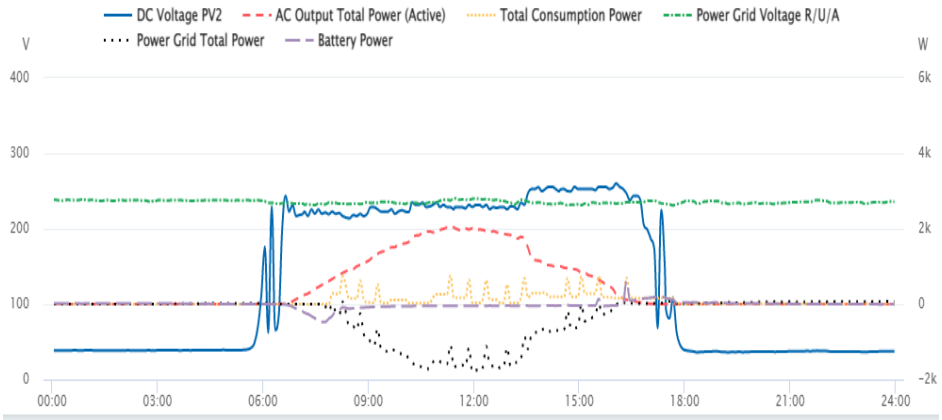


Figure 60. Time-series of inverter operating variables and power-flow components over a 24-h period (3 December 2021).

Figure 60 shows that although the PV system generated 8.6 kWh, the consumption only 2.2 kWh, and the battery discharged from 8AM to 4PM because it's SOC was almost complete full before the PV peak. This behavior led to inefficient storage operation and a low self-consumption rate (25%).

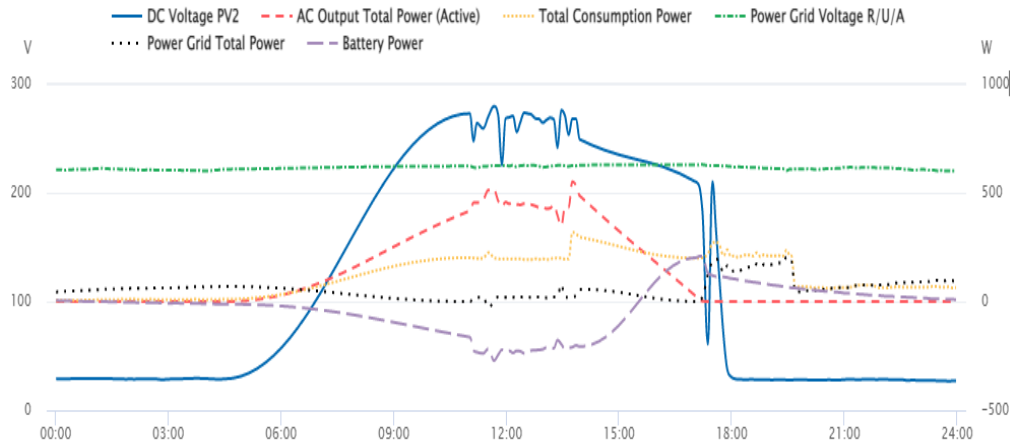


Figure 61. Time-series of inverter operating variables and power-flow components over a 24-h period (4 December 2021)

In Figure 61, when the user actively planned the charge–discharge schedule to keep the SOC lower before the PV peak, the self-consumption rate increased to 75% and grid export dropped markedly. A local control strategy charges the battery during overvoltage periods and discharges it when grid demand is high. This helps stabilize the voltage and improves on-site PV utilization.

The minute-resolution experiments also show that user decisions when scheduling battery charge and discharge directly affect storage performance and the building's self-consumption level.

3.3.4. Energy efficiency project analysis

After installing and operating the monitoring platform at VHH, we estimated the energy efficiency after one year. Table 9 shows that after 1 year of operating the IoT-BEMS platform, there is a significant reduction in electricity consumption: from 5,095 to 1,941 kWh/year. In contrast, the monitoring system consumed only 69 kWh/year. The electricity saved reached 3,154 kWh/year (around of 62%). Based on an average electricity price of 2,000 VND/kWh, the cost saving is about 6.31 million VND/year, the payback period is approximately 2.38 years.

Table 9. Summary of Energy Efficiency Calculation in VHH platform

Item	Value	Notes
Energy consumption before the project **(Ei)**	5,095 kWh/year	Baseline operation
Energy consumption after implementation **(Ee)**	1,941.05 kWh/year	Consumption of electrical devices after applying IoT-BEMS platform
Energy consumption of the monitoring system	69.33 kWh/year	Additional load of the IoT-BEMS
Energy saved **Es = Ei − Ee**	3,153.95 kWh/year	Reduction achieved through monitoring and control
Electricity tariff **(Pe)**	2,000 VND/kWh	Public college tariff
Annual electricity cost savings $\mathbf{R_b = Es \times Pe}$	6,307,900 VND/year	Direct financial benefit
Energy saving proportion **Ps = Es / Ei × 100%**	62%	Percentage reduction relative to baseline
IoT-BEMS investment cost **(Iv)**	15,000,000 VND	Hardware and installation costs
Payback period **(Iv /Rb)**	2.38 years	Excluding maintenance and depreciation

These results show that the IoT-BEMS platform delivers practical efficiency: continuous monitoring, detailed load analysis, optimal self-consumption, and improved electricity usage behavior. Combining with a rooftop PV system, the platform helps increase energy efficiency and reduce operating costs with a short payback period.

3.4. Conclusions of Part 3

Part 3 applies the proposed IoT-BEMS platform to a real office building in Vietnam to analyze how environmental conditions and user behavior jointly affect HVAC energy use and rooftop PV self-consumption. Using one year of monitored data, the analysis quantifies how outdoor temperature and humidity affect HVAC loads.

Choices of set-points, window opening and operating schedules, together with behavior-driven demand, can be translated into simple control rules. Raising the cooling set-point (about 2°C) can reduce daily HVAC energy use by 25–30% while maintaining comfort within acceptable limits. In hot periods, these simple

dynamic set-point rules work better than very low fixed set-points in the saturation zone.

On the rooftop PV–Battery system, IoT-BEMS data show that monthly self-consumption is around 40%. At the daily scale, values vary strongly (10–82%), depending on how the storage system is operated. Minute-level experiments confirm that user actions in scheduling battery charge and discharge can raise self-consumption from 25 to 75% on a given day. Simple local rules charging the battery during overvoltage periods and discharging it when demand is high, help stabilize the connection point voltage and increase on-site PV utilization.

The energy-efficiency VHH's project analysis shows an annual electricity reduction of about 62% in the monitored system. The simple payback time for the IoT-BEMS investment is estimated at roughly 2.4 years.

Overall, Part 3 shows that a low-cost IoT-BEMS can deliver actionable, high-resolution feedback on environment, behavior and storage, and support effective user-oriented nZEB operation in real Vietnamese buildings.

Part 4. Optimal energy management strategies toward nZEBs

4.1. Overview

4.1.1. Optimal energy management strategies context

Solar energy has high energy potential and low environmental impact. However, its weather dependent, variable output can cause grid fluctuations, additional losses, and power quality issues if it is not well controlled.

To address these challenges, many authors have introduced the concept of **energy autonomy** in PV systems, defined as "*the ability of the energy system to function fully, without the need of external support in the form of energy imports, through its local energy generation, storage and distribution systems*" [77]. In practice, this autonomy is achieved by combining on-site PV, battery storage and controllable loads within a suitable control architecture.

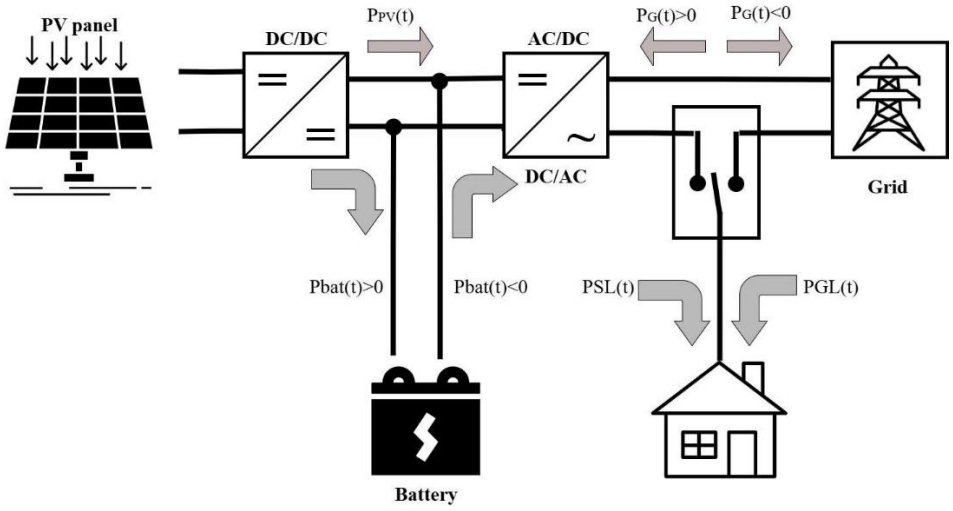


Figure 62. Diagram of energy flows in the Battery-PV system

Figure 62 illustrates the main energy flows in a grid-connected PV–battery system. PV panels supply DC power that can (i) feed AC loads through an inverter, (ii) charge the battery when surplus is available, and (iii) inject excess energy into the grid. When PV and battery power are insufficient, the grid supplies the loads and can also recharge the battery. During grid outages, the PV–battery subsystem can still support critical loads, provided that the state of charge (SoC) remains within safe limits.

In this thesis, PV–battery architectures are considered for optimal energy management toward nZEB targets. The proposed IoT-BEMS platform is to monitor energy flows in real time and to implement control strategies to improve energy autonomy level.

4.1.2. Key aspects of building energy management strategy design

In the context above, this work focuses on designing optimal operating strategies for small and medium-sized existing buildings equipped with low-cost monitoring and control systems. The goal is to obtain strategies that can run on an IoT platform and also provide energy saving and cost benefits.

The following key aspects are considered in the strategy design:

- Model accuracy and computational performance, to ensure seamless integration into a low-cost IoT platform.
- Flexibility of the optimization algorithms, so that different objectives and constraints can be addressed.
- Practical feasibility of the energy management strategies, ensuring they can be implemented and operated in real buildings.

The following sections build on these principles: Section 4.2 introduces the energy models used in the optimization (PV, battery and load models), Section 4.3 presents the optimization problems and algorithms, and Section 4.4 applies them to real case studies in France and Vietnam.

4.2. Energy Modeling

Buildings are complex systems, so developing full dynamic models is often costly and data-demanding. In this work, instead of detailed numerical models, we adopt simplified model approaches to IoT-based building applications. Following [2], building models are grouped into empirical, analytical and numerical types. Numerical models require many physical parameters and high computing resource, which are rarely available in small and medium-sized buildings.

Empirical and analytical models are more suitable for low-cost IoT platforms [30].

Therefore, this thesis focuses on simple empirical and analytical models that are:

- Accurate enough for control and energy management,
- Lightweight model for embedded or low-cost PCs,
- Easy to embed into optimization algorithms.

Concretely:

- Analytical models are used for PV production and battery (SoC).
- Empirical (data-driven) models are used for load and demand forecasting based on historical data, without detailed construction parameters.

These models form the basis for the energy management strategies developed in Sections 4.3 and 4.4.

4.2.1. PV production model

4.2.1.1. General context

The development of PV systems recently has increased uncertainty variables in the electricity systems and is a significant challenge in balancing energy demand and power supply. Many studies show that PV forecasting is essential for both grid operations and local energy management in buildings [78]. PV forecasts are typically categorized by time horizon, ranging from very short-term (seconds to minutes) and short-term (hours to days) to medium and long-term [78].

Based on the input data sources, forecasting PV production models are classified into two approaches: direct and indirect [79]. The direct method predicts PV production using historical PV power output data. This approach requires historical power data, which is not always available and accessible. The rest method uses weather-forecasting data obtained from a meteorological station or a web service as input to the PV power-forecasting model. However, the model developed by this method depends on the accuracy of weather forecast data. This study uses weather forecasting data as the input of the PV forecasting model.

For online applications, PV forecasts are continuously updated as weather data arrive. The PV production prediction model can be derived from solar radiation

[29]. The power produced by the photovoltaic panels (PV) is given by the equation below:

$$P_{PV}(t) = \eta_{PV}.S_{PV}.I_{PV}(t) \qquad (4.1)$$

Where: S_{PV} –PV panels area (m2); η_{PV} – Efficiency of PV system; I_{PV} – Total radiation on the plane of PV (W/m2), as following:

$$I_{PV}(t) = DNI(t).\cos\theta(t) + DHI(t).\frac{1+\cos\beta}{2} + GHI(t).\rho.\frac{1-\cos\beta}{2} \qquad (4.2)$$

$$I_{PV}(t) = DNI(t).\cos\theta_z(t) + DHI(t) \qquad (4.3)$$

DNI - Direct irradiation (W/m2); DHI - Diffuse horizontal radiation (W/m2); GHI - Global horizontal radiation consists of (DNI, DHI, and Reflection) (W/m2); β - Tilt angle of PV panels (rad); θz – Zenith angle (rad); θ – Theta angle (rad); ρ - Albedo coefficient; $n_{ref} = 0.2$ for polycrystalline modules; $n_{real} = 0.8$ and $n_{inverter} = 0.95$ as a standard value it leads to $n_{PV} = 0.15$

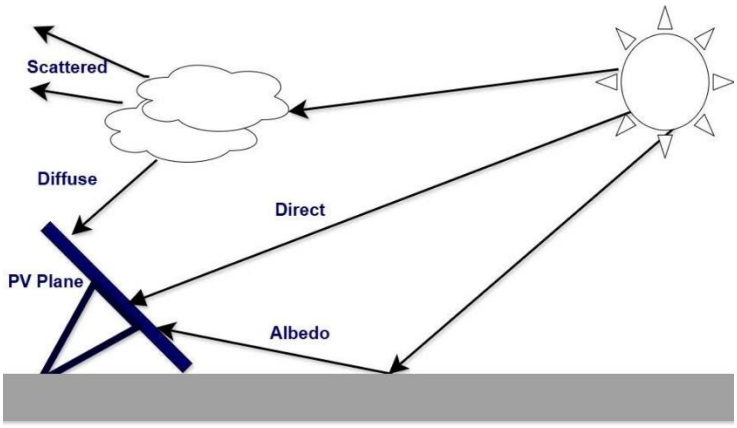


Figure 63. Components of solar radiation (direct, diffuse and reflection) to a PV plane

In Figure 63, I_{PV} represents the global solar irradiance on the panel plane (W/m^2), which is composed of 3 components (direct, diffuse and reflection) and is derived from the weather data of Direct Normal Irradiance (DNI) ID (W/m^2) and Diffuse Horizontal Irradiance (DHI) Id (W/m2).

$$I_{PV}(t) = I_D(t).\cos(\theta(t)) + I_d(t).\frac{1+\cos(\beta)}{2} + [I_D(t).\cos(\theta_z(t)) + I_d(t)].\rho.\frac{1-\cos(\beta)}{2} \qquad (4.4)$$

- ✓ β is the tilted angle of panel plane (radian). Horizontal panel correspond to $\beta = 0$;

- ✓ θ (radian) is the angle between direct irradiance and the normal of the panel (Beware: If θ>pi/2 then $I_D(t).cos(\theta(t)) = 0$)
- ✓ θ_z (Radian) is the zenith angle (It is noted that $\theta_z + \alpha = \pi/2$ with α is sun's altitude angle).
- ✓ ρ is the reflection coefficient of ground (also called albedo and is considered equal to 0.2 as a standard value)
- ✓ γ (Radian) is the PV surface azimuth angle.
    - It is the panel's orientation compared to the South Pole.
    - Its value range is between $-\pi$ and π. γ =0 for facing to South; $\gamma = \pi/2$ for facing to West, $\gamma = -\pi/2$ for facing to East.

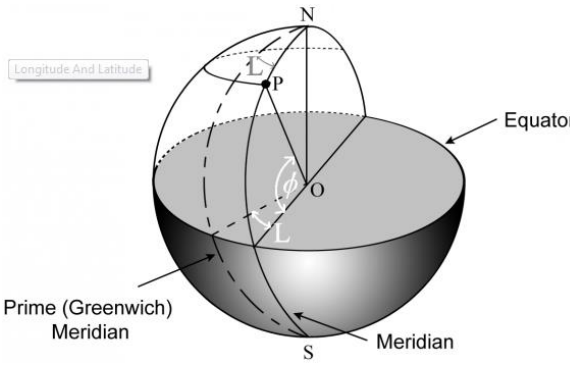


Figure 64. Geometric definition of site coordinates on Earth: latitude (φ) and longitude (λ) [29]

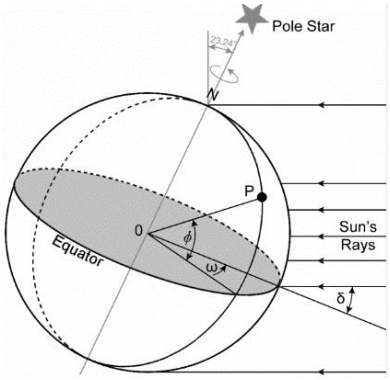


Figure 65. Definition of solar angles at a site: declination, hour angle, and local latitude [29]

- δ (Radian) is the angle between the equatorial plane and the direct irradiance, also called declination angle :

$$\delta(t) = \frac{\pi}{180} \cdot 23.45 \cdot \sin\left(\frac{2\pi.(284 + n_{jour})}{365}\right) \qquad (4.5)$$

- n_{jour} is the day of the year (n=1 on the 1st January).
- ω is the hour angle (radian). The hour angle is dependent on the true solar time (*TSV* in hour):

$$\omega(t) = 15.(TSV - 12).\frac{\pi}{180} \qquad (4.6)$$

$$TSV = t - \tau - \frac{L}{15} + \frac{ET}{60}$$
$$ET = 9.87 \cdot sin(2B) - 7.53 \cdot cos B - 1.5 \cdot sin B$$
$$B = 2\pi \cdot \frac{n_{jour} - 81}{365}$$

Where: *t* (hour) is the time indicated on our watch. ↿=1 in winter days and ↿=2 in summer days. L (in degree) is the longitude of location regarding prime (Greenwich) meridian. Its value is negative in East and positive in West. *ET* (in minute) is the correction of the time equation (See Figure 64, Figure 65).

4.2.1.2. Developing an online PV production prediction model

In this work, PV forecasting is used for control-oriented energy management on a low-cost platform. Recent studies show that lightweight online models, updated on sliding windows of short- horizon data, can provide sufficiently accurate forecasts while remaining computationally efficient for hardware [64].

In this context, the proposed PV model is designed to (i) run on a Raspberry-Pi-class device, (ii) use only a short recent history of local measurements, and (iii) exploit exogenous weather forecasts (cloud cover / nebulosity) that are updated online. This design explains why the training dataset is short, as the model is continuously updated with the most recent data.

(a) Methodology – Clear Sky Model

The PV power output can be characterized using measured irradiance (W/m²), but the measurements are not always available in many locations [80]. Clear-sky models therefore provide a useful alternative by estimating solar irradiance without requiring on-site radiometric sensors.

Several families of clear-sky models come from straightforward geometric or more complex models [80], [81]. Although complex models can be more accurate in highly variable aerosol location, models often achieve comparable accuracy with much lower computational cost and fewer input variables [83], [84].

In this work, the online PV prediction model adopts a clear-sky–based indirect approach:

- ✓ A clear-sky irradiance curve is computed for the site, then corrected using nebulosity forecasts obtained automatically from a web service.
- ✓ The irradiance is finally converted into expected PV power using nominal system parameters and empirical efficiency factors [77], [84], [85].
- ✓ The model is embedded in a low-cost Raspberry Pi within the PV monitoring system: Nebulosity forecasts are updated every 3 hours; the model parameters are periodically re-fitted on a short recent history of measured PV data; approach real-time operation on low-computing hardware and accurate for short-term energy management.

In previous work, the forecasting irradiance model is developed by approaching the clear sky model and the Nelder-Mead optimization algorithm [86].

(b) Data collection

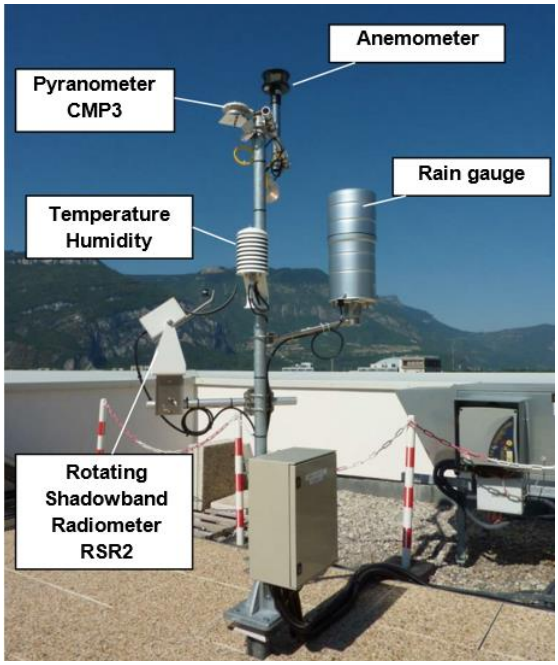


Figure 66. Description of the weather station on GreEn-ER rooftop

The measured GHI/DNI dataset from Campbell Scientific weather station (including Rotating Shadow band Radiometer) in Figure 66.

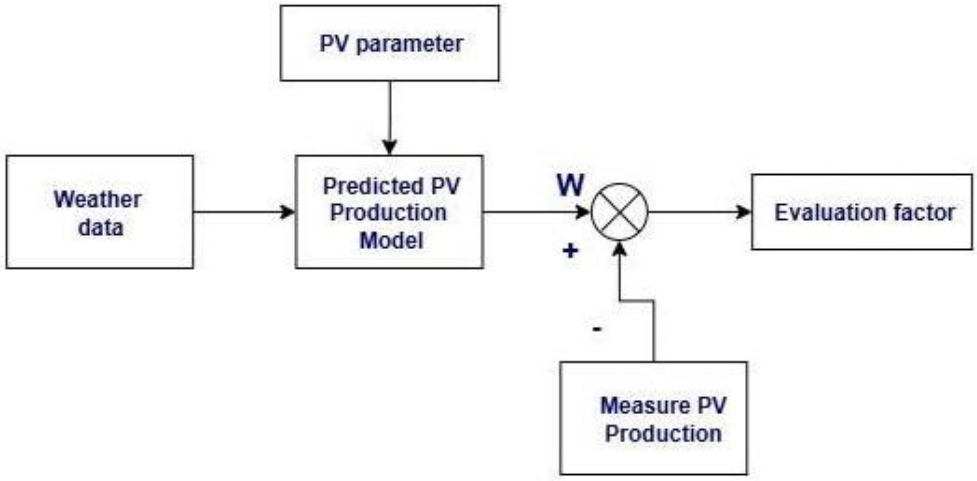


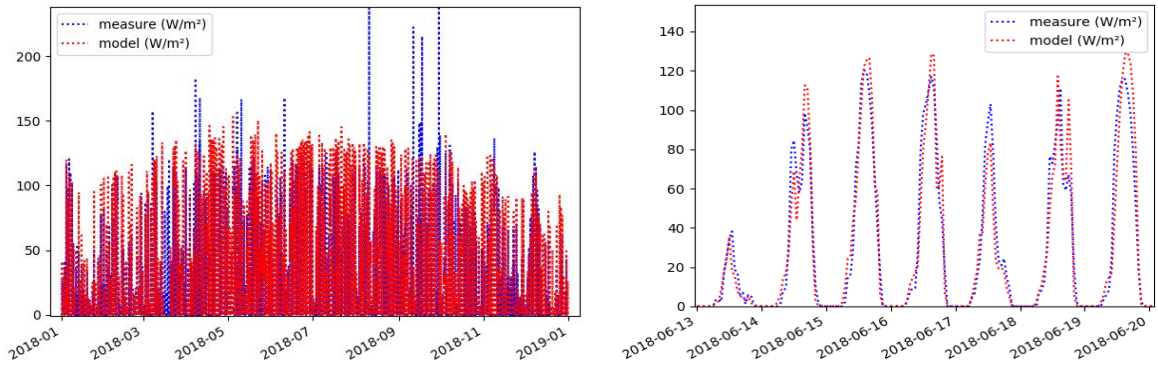
Figure 67. Description of validation model diagram

(c) Evaluation factor:

$$\text{RMSE} = \sqrt{\text{mean}(\text{model value} - \text{measured value})^2} \qquad (4.7)$$

(d) Results and Validation

In Figure 67, validation model solution is described with RMSE. Measured PV power data from a real PV's system in 2018. PV parameters include (location (longitude and latitude), area of PV, tilted angle, and azimuth angle of PV plane).



(a) PV prediction in one year (b) PV prediction in one week

Figure 68. Comparison of PV power measure (W/m²), and PV model output

In Figure 68.a. model data (the red dot line) are sometimes much lower than measured data (the blue dot line). For example, in Sep 2018, there was sometimes a blue dot point's value approximating higher twice the red dot point's value. In Figure 68.b (for a one-week prediction horizon), the model could achieve a better prediction result (average of RMSE=5.2W). Therefore, the model is suitable for short-term

prediction. The model quality depends much on the accuracy of weather forecast data. In this work, the online predicted PV production model is available on a raspberry Pi 3 on Greenhouse.

4.2.2. Battery model [29]

In battery model, we are interested of linking storage capacity and charge/discharge process at any given time. This model is described by [29].

Battery charging process:

$$C_{bat}(t) = C_{bat}(t - \Delta t).(1 - \sigma) + \eta_c . P^c_{bat}(t).\Delta t \quad (4.8)$$

Battery discharging process:

$$C_{bat}(t) = C_{bat}(t - \Delta t).(1 - \sigma) - \frac{P^c_{bat}(t)}{\eta_{dc}} \quad (4.9)$$

Where: $C_{bat}(t)$ - The battery's available capacity (Wh) at time t and t-Δt.

σ is the self-discharge rate of the battery.

η_c and η_{dc} are the charging and discharging efficiency of the battery.

$P^c_{bat}(t)$, $P^{dc}_{bat}(t)$ are the charging/discharging power at time t, respectively.

If we consider an ideal battery with perfect charging and discharging performance and no self-discharge, it can be simplified by:

$$C_{bat}(t + 1) = C_{bat}(t) + P_{bat} * \Delta t \quad (4.10)$$

Where: P_{bat} (t) (W) is the capacity of the charge if its value is positive, of the discharge if it is negative. The battery's charging and discharging capacity will be optimized according to energy production and consumption.

Energies balance: $P_G(t) = P_{loads}(t) + \eta*(P_{bat}(t)-P_{PV}(t))$ (4.11)

Where: PG (t) is the electrical power exchanged with the grid (W); P_{PV} (t) is the power from the PV panels supplied to the system; P_{loads} (t) is the total load power (W).

Electricity will be imported from the grid if its value is positive otherwise fed to the grid if its value is negative. The inverter's efficiency η, in this study is assumed to be 1.

4.2.3. Load model

Forecasting a building's energy use is a key input to any optimization strategy.

However, it is difficult to accurately predict demand because it depends on factors such as weather and occupant behavior. In this work, the load model is developed using historical data collected from the building. Data input includes TiZ3, T_{our}, $P_{Total,t-1}$; Data output is $P_{Total,t}$, with $P_{Total,t}$ is prediction total power consumption.

The study developed total consumption model based on GPR technique and the input data includes the indoor temperature (TiZ3) and the historical total power consumption ($P_{Total,t-1}$). The training dataset is one week data in July 2020, the test dataset is one week data in August 2020**.** In this work, the root mean square error is 25.6W, and the mean absolute error is 51.5W, and 95% confidence interval. Since, the load model could be good to predict total power consumption in one week horizon prediction.

4.2.4. A model-based energy management workflow in IoT-BEMS platform

To link the energy models of Section 4.2 with the optimization problems discussed in Section 4.3, Figure 70 summarizes the overall workflow of the proposed model-based energy management strategy implemented in the IoT-BEMS.

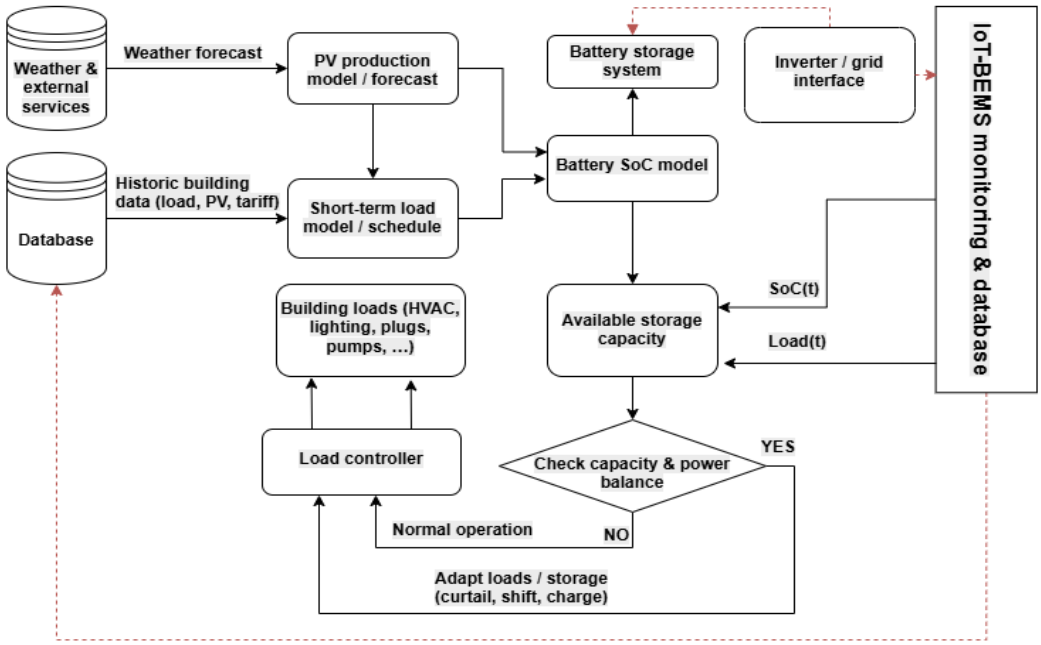


Figure 69.General workflow of model-based optimal energy management in IoT-BEMS.

Figure 69 summarizes how the proposed IoT-BEMS uses models to manage energy flows. Historical building data (load, PV and tariffs) and weather forecasts are first

collected and stored in the database. The PV production, short-term load and battery SoC models work together to predict generation, demand and available storage capacity. These predictions feed the optimization block, which checks the power balance and storage limits and then computes control actions, such as charging or discharging the battery, shifting loads or curtailing surplus PV.

The resulting setting points are sent to the load controllers and inverter, while the monitoring system records SoC and load values for further analysis. The same workflow is used in the Greenhouse and VHH case studies, with different objectives (self-consumption and system sizing in the Greenhouse, electricity-bill minimization in VHH).

4.3. Optimization Problem & Algorithm- Programming languages

4.3.1 Problems & Algorithms

Building energy management leads to constrained optimization problems, where the objective is to minimize an energy-related cost (electricity bill) or maximize on-site self-consumption, subject to power balance and battery SoC constraints. Because of non-linear components (battery, tariff structure, operating modes, etc.), these problems are non-linear. A wide range of optimization techniques apply in building, including Linear Programming (LP), Quadratic Programming (QP), Mixed-Integer Linear Programming (MILP), Sequential Quadratic Programming (SQP), as well as meta-heuristics such as Genetic Algorithms (GA) and Particle Swarm Optimization (PSO) [87]. This approach enables rapid convergence for moderate-sized problems and is well-suited to low-cost hardware.

In this thesis, SQP is adopted as a good compromise between robustness, computational efficiency and ease of integration into an automatic optimization tool [87]. SQP iteratively solves a sequence of quadratic sub-problems that approximate the original non-linear objective and constraints, using first-order derivatives to update the search direction. This approach provides fast convergence for the moderate-size problems and is well-suited for a low-cost hardware.

Principle of Sequential quadratic programming (SQP) [29]: algorithm solves

a sequence of optimization sub problems, each of which optimizes a quadratic model of the objective subject to constraints.

By the direction searching process, d_k, from initial possible point, x^k, and approximating nonlinear problem, Quadratic programming formula is:

$$\min_{d_k} \frac{1}{2}.\, d_k^T\,.\, H_k.\, d_k + \nabla f\left(x^k\right)^T.\, d_k \qquad (4.12)$$

$$\nabla h\left(x^k\right)^T.\, d_k + h\left(x^k\right) = 0$$

$$\nabla g\left(x^k\right)^T.\, d_k + g\left(x^k\right) \leq 0$$

$\nabla f(x^k)$ Gradient of the scalar objective function $f(x^k)$, $\nabla h(x^k)$ and $\nabla g(x^k)$ is the Jacobian of the equality constraints and inequality constraints of $h(x^k)$ and $g(x^k)$ respectively.

$$\nabla f(x^k) = \begin{pmatrix} \frac{\partial f}{\partial x_1^k} \\ \cdots \\ \frac{\partial f}{\partial x_n^k} \end{pmatrix}; \nabla h(x^k) = \begin{pmatrix} \frac{\partial h_1}{\partial x_1^k} & \cdots & \frac{\partial h_1}{\partial x_n^k} \\ \cdots & \cdots & \cdots \\ \frac{\partial h_p}{\partial x_1^k} & \cdots & \frac{\partial h_p}{\partial x_n^k} \end{pmatrix}; \nabla g(x^k) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1^k} & \cdots & \frac{\partial g_1}{\partial x_n^k} \\ \cdots & \cdots & \cdots \\ \frac{\partial g_q}{\partial x_1^k} & \cdots & \frac{\partial g_q}{\partial x_n^k} \end{pmatrix}$$

$H_k = \nabla_x(\nabla_x L(x^k, \lambda^k, \mu^k))$ - The Hessian of the Lagrangian function

$$L(x, \lambda, \mu) = f(x) + \lambda^T.\, h(x) + \mu^T.\, g(x) \qquad (4.13)$$

$\lambda \in \Re^p$; $\mu \in \Re^q$ - The Lagrange multipliers.

$x^{k+1}=x^k+\alpha d_k$ - An approximating x^{k+1} is defined by searching direction of the initial variable x^k, $\alpha \in [0,1]$ - the step space, determined by to satisfy the objective function and constraints [29][88].

In the case studies of Sections 4.4.2 and 4.4.3, this SQP-based solver is used to compute daily optimal schedules for PV–battery operation and controllable loads under different objectives (self-consumption maximization and electricity-bill minimization).

4.3.2 Programing languages - Software

The Grenoble Electrical Engineering Laboratory (G2Elab) developed a new open-source NoLOAD using Python, which is easy to understand for designers [89].

NoLoad is a lightweight library for non-linear optimization that relies on

Automatic Differentiation (AD) to compute derivatives. Role of Library:

- It uses Sequential Least Squares Quadratic Programming algorithm, providing a stable and low-cost computing solution, suitable even for embedded hardware.
- The tool allows the designer to define constraints on both input and output parameters of the model and to specify one or multiple objective functions to be minimized.
- NoLoad automatically analyses inputs and outputs of a given model and selects the appropriate forward or reverse AD mode to improve computational performance.
- It also supports problems with vector-valued constraints, which is convenient for complex engineering models.

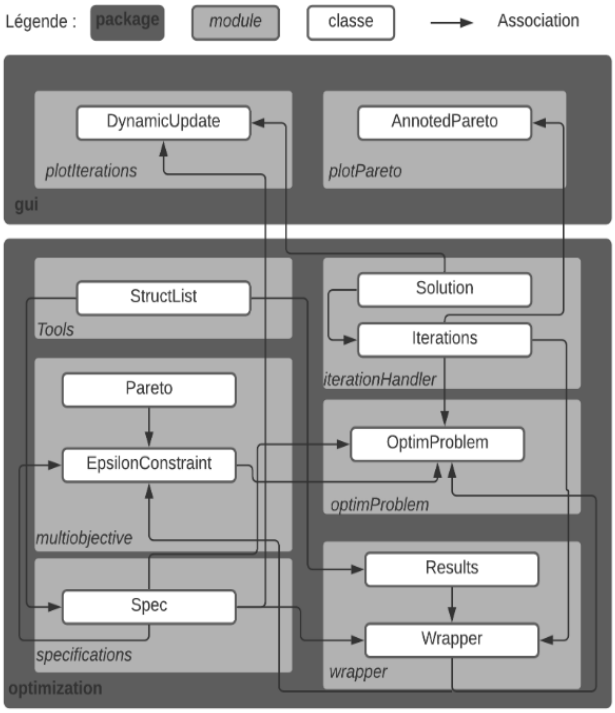


Figure 70. Overview of the NoLoad library architecture [89]

Figure 70 describes structure of the NoLoad library. It solves two kinds of optimization problems for non-linear systems: system sizing and computing optimal controller. In the next part, the study will apply these methods to control strategies on two case in France and Vietnam.

4.4. Optimal energy management applications – Study cases

In this section, the modelling and optimization framework introduced in Sections 4.2 and 4.3 is applied to case studies targeting nZEBs.

Two experimental platforms are considered:

- An aquaponics Greenhouse installed on the roof of a building in France, used as a self-consumption and self-supply testbed.
- The VHH building in Vietnam is a grid-connected office/educational facility operated under Vietnamese electricity tariffs.

From real platforms, three studies are analyzed:

- Section 4.4.1 – Energy balance analysis in nZEBs-Greenhouse testbed: evaluation of annual and seasonal energy balance using load-matching indicators (γ_{supply}, γ_{load}) based on monitored PV production and load data from Greenhouse;
- Section 4.4.2 – Energy management strategies for PV–Battery system sizing: formulation of a data-driven sizing problem on Greenhouse platform, combining monitored data, simplified models and optimization to explore trade-offs between autonomy, curtailment and storage capacity;
- Section 4.4.3 – Optimal control strategies for minimal electricity bill: application of the same modelling and optimization framework to VHH Platform. The objective is to minimize the electricity bill while improving self-consumption with operational constraints.

4.4.1. Energy balance analysis in a nZEB [68]

4.4.1.1. Case description - Problem statement

This section uses an aquaponics greenhouse as an experimental nZEB case study to assess the energy balance. The goal is to quantify how well on-site PV generation matches energy demand across different time scales and to identify periods of deficit and surplus. The results provide a basis for the subsequent system sizing strategy developed in Section 4.4.2.

Following the methodology adopted in European solar projects [28], the relationship between on-site supply and electrical demand is characterized by two load-matching indicators**.** The supply cover factor express fraction of PV production is self-consumed in the building. The load cover factor express fraction of the building demand is covered by on-site PV generation.

Over a time horizon [t_1-t_2] , the cover factor (γ_{supply}) and load cover factor (γ_{load}) are defined in [28] with equations below :

$$\gamma_{\text{supply}} = \frac{\int_{t1}^{t2} \min[g(t) - S(t) - \zeta(t) * load(t)]dt}{\int_{t1}^{t2} g(t)dt} \quad (4.14)$$

$$\gamma_{\text{load}} = \frac{\int_{t1}^{t2} \min[g(t) - S(t) - \zeta(t) * load(t)]dt}{\int_{t1}^{t2} load(t)dt} \quad (4.15)$$

$$S(t) = s_c(t) - s_{dc}(t) \quad (4.16)$$

Where:

t_1/t_2 are start/end of the evaluation period;

$g(t)$ is energy production; ζ(t) is energy losses;

load(t) is the power of loads;

S(t) is the storage energy balance defined by Equation 3. In which, s_c/s_{dc} are charging/discharging storage energy.

In this case study, γ_{supply} and γ_{load} are computed at daily and monthly scales; PV production and load data come from the Greenhouse platform. The formulation provides a quantitative basis to evaluate how close the monitored building operates to nZEB targets, and to benchmark future scenarios with improved control strategies and PV–battery configurations.

4.4.1.2. Results and discussion

Energy balance in the nZEB Greenhouse is evaluated using monitoring data for load consumption, PV production and system losses, combined with PV production simulations from PVSyst.

The loads have a nominal power of 42 W, and the system losses are approximated as

10% of the nominal load (4.2 W). Assuming pump operates 24 h/day. Over one year, the loads will consume 368.4 kWh, while the PV system produces 406.4 kWh, giving a yearly supply cover factor of $\gamma_{supply,year}$ = 0.91 and a yearly load cover factor of $\gamma_{load,year}$ = 1.0 [68]. This indicates that, on an annual basis, local PV generation is sufficient to cover the total demand, with only a small energy surplus.

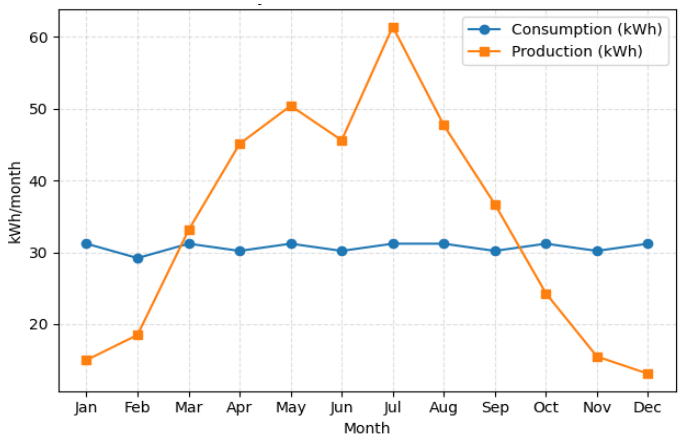


Figure 71. Consumption vs Production on Greenhouse project

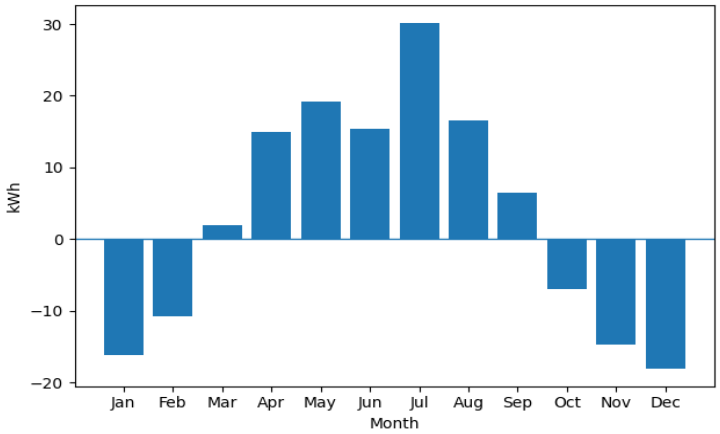


Figure 72. Monthly Energy Balance (>0: Export vs <0: Import) on Greenhouse project

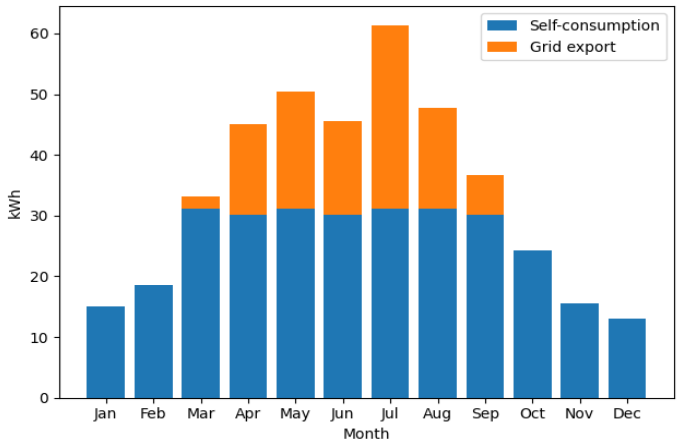


Figure 73. Distribution of production on Greenhouse Project

Figure 71 compares monthly consumption and production. Load demand is almost flat over the year (around 30 kWh/month), whereas PV generation shows strong seasonality, with high production in summer (May–August) and deficits in winter (January–February and October–December). This pattern is reflected in Figure 72, where positive values correspond to monthly surplus (export) and negative values to monthly deficit (import). Figure 73 decomposes PV production into self-consumed energy and grid export. During summer months, most of the surplus is exported, as the battery capacity is not sufficient for seasonal storage. In contrast, winter operation relies partly on grid import despite platform annual balance.

These results highlight that, for a small platform, PV sizing alone is not enough to guarantee self-sufficiency at all times. Complementary demand-side management (load shifting or curtailment in winter) and storage/control strategies are required to improve seasonal matching between supply and demand.

4.4.2. Energy management strategies approach system sizing [68]

Based on the energy balance analysis in Section 4.4.1, this subsection formulates a system sizing and operation problem for the autonomous Greenhouse.

4.4.2.1. Greenhouse's platform description

In the SERRE project, the greenhouse is installed on the roof of the building in Grenoble. It's a closed-loop, recirculating hydroponic system combining fish farming and plant cultivation, aiming for energy self-sufficiency and sustainable food production. Therefore, the greenhouse requires an energy monitoring and control system to ensure continuous operation when solar power is insufficient.

Figure 74. Overview of the Greenhouse testbed in Grenoble [68]

Figure 74 shows the aquaponics Greenhouse with PV system and an electrical box for setting up the monitoring and control system. The electrical box stores the inverter, protection devices, batteries, and a low-cost monitoring and control board (based on an Arduino and a Raspberry Pi). The Raspberry Pi acquires power data from the energy meters and sends control commands to the pump via the Arduino.

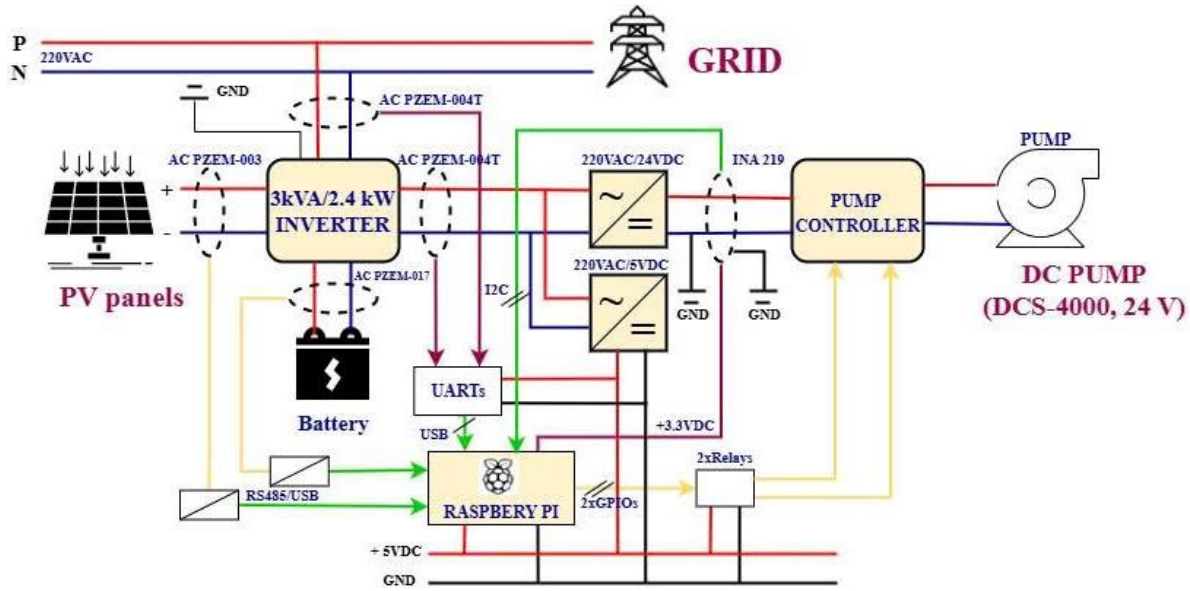


Figure 75. The Greenhouse Low-cost monitoring and control hardware [68]

A diagram of the PV–battery–load architecture and the main component ratings is shown in Figure 75. The system consists of six PV panels (64 W/module); two 12 V&22 Ah batteries; a 2.4 kW inverter charger, and a 24VDC pump. This configuration shows that the Greenhouse testbed is fully monitored and controlled using inexpensive hardware, consistent with the low-cost IoT-BEMS approach.

4.4.2.2. Problem formulations

The greenhouse is operated as an islanded DC system. The sizing problem is to select PV and battery capacities, together with a simple pump control rule, such that:

- The IT monitoring load is always supplied;
- The aquaponics pump operates as long as possible over 24 h;
- Curtailment and unnecessary oversizing of PV and battery are limited.

To evaluate performance, the load-matching indicators introduced in Section

4.4.1 are used at daily scale: γl_{oad} close to 1 indicates that the local production is sufficient to cover the daily demand, whereas γ_{supply} measures self-consumption.

The objective of the sizing problem is to maximize the self-consumption factor (γ_{supply}) under constraints on the load cover factor (γ_{load}) and State of charge (SoC).

Objectives function:

$$J = \text{abs}\left(\sum_{t=t0}^{t=T} P_{\text{Loads}}(t) * \Delta t - P_{\text{pv}}(t) * \Delta t\right) \qquad (4.17)$$

Input data: GreenHousePV-Prediction.csv; GreenHouseLoad-History.csv

Decision variables: $P_{pump}(t)$, $P_{bat}(t)$

Power balance: $P_{grid}(t) = P_{loads}(t) + P_{losses}(t) + P_{bat}(t) - P_{pv}(t)$

Boundary conditions: $0 < P_{pump}(t) < P_{max}$ and $-(C_{max}-C_{min})/dt < P_{bat} < (C_{max}-C_{min})/dt$

For finding control step: a test time response of actuator in one minute was carried out. Pump could be controlled to change from min power to max power and reversely.

For the optimization algorithm, assume that:

$$-P_{max} < dP_{pump}/dt < P_{max}$$

Initial conditions:

C_{nom}=27.2V*22Ah=600Wh; C_{min}=20%*C_{norm}=120Wh;

C_{max}=80%*C_{bnorm}=480Wh; C_{init}=50%*C_{norm}= 300Wh;

P_{max}=25W; $P_{criLoad}$=10W

All symbols of variables in Greenhouse test case are presented and explained in Table 10.

Table 10. Symbols of variables in Greenhouse test case

Parameter	Unit	Description
J	Wh	Objective function: Total surplus PV energy
T	hour	Optimal time
P_{bat}	W	Capacity of the battery
C_{norm}	Wh	Rate capacity of the battery (600Wh)
C_{min}	Wh	20% Cbat_norm
C_{max}	Wh	80% Cbat_norm
C_{init}	Wh	Initial capacity of battery = 50% C_{norm}
$C_{bat}(t)$	Wh	Capacity of battery at time t
$P_{pv}(t)$	W	Predicted Power produced by the PV system at time t
$P_{Load}(t)$	W	Toal consumption of Loads at time t
$P_{pump}(t)$	W	Consumption of Pump at time t

$P_{losses}(t)$	W	Power losses at time t
$P_{criLoad}$	W	Power consumption of IT components (10W)
Pmax	W	Highest consumption of Pump when it works
Pmin	W	Lowest consumption of Pump when it works
Pzero_pump	W	Consumption of Pump when it is not work (0W)
Total_load	W	Total power consumption of loads (IT and Pump)
$P_{pump\_optimal}(t)$	W	Optimal consumption of Pump at time t
$P_{bat\_optimal}(t)$	W	Optimal battery capacity at time t
SoC (t)	%	State of charge of battery at time t
$SoC_{optimal}(t)$	%	Optimal state of charge of battery at time t

4.4.2.3. Optimization strategy

Building on the modelling and optimization framework introduced in Sections 4.2 and 4.3, this subsection defines the specific control strategy for the Greenhouse PV–Battery system.

Autonomous PV–Battery systems with similar objectives have been widely analyzed in the literature [90], [91], [92], often using probabilistic indicators such as Loss of Load/Power Loss Probability. Other studies have highlighted the impact of inverter operation and conversion losses on the energy balance [93], [94]. Although these losses are sometimes neglected in control design. In this work, measured data from the Greenhouse are used to build empirical models of PV production, load profiles and conversion losses.

The objective of the experiment is to ensure a continuous power supply to the IT system and to minimize pump shutdown time to preserve the aquaponics ecosystem. When PV power is available, it is used to supply the IT and pump loads and to charge the battery; when PV power is insufficient or unavailable, the battery discharges to support the loads.

The optimization strategy adjusts the pump power and battery charge/discharge schedule based on stage of charge (SoC) and power balance constraints to maximize self-consumption while avoiding storage oversizing.

4.4.2.4. Results and discussion

On a typical summer day, the PV array produces about 1059 Wh while the total daily consumption is around 730 Wh. At daily scale, the load cover factor is γ_{load} = 0.9, indicating that local production almost meets the demand, and the supply cover factor is γ_{supply} = 0.7, meaning that about 70% of PV energy is self-consumed and 30% is surplus.

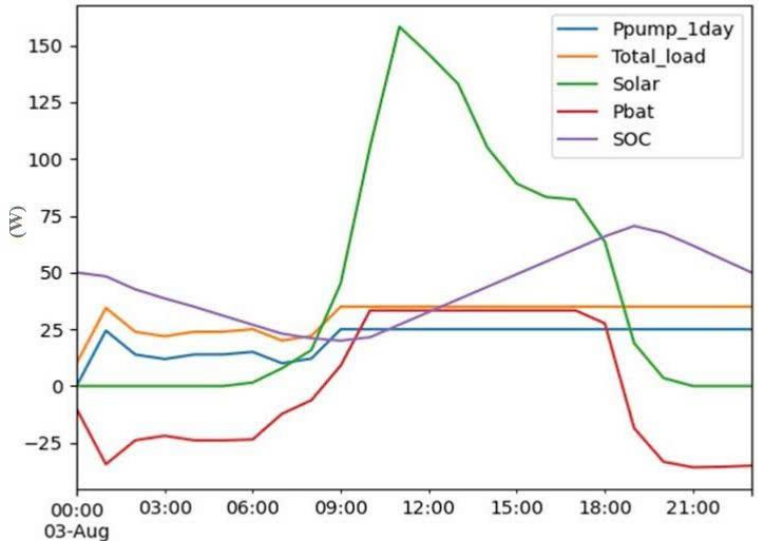


Figure 76. Results of optimal management in Greenhouse

Figure 76 illustrates the optimal schedules of pump power, battery power and battery state of charge.

During daytime (9:00–18:00), PV production supplies both the loads and battery charging. At night, the battery discharges to support the loads as long as the SoC remains above a minimum threshold. In the optimized strategy. The pump operates at reduced power when SoC drops below a given level, and is temporarily stopped when SoC becomes critically low, while the IT load remains supplied.

Compared to a naive strategy where the pump always runs at nominal power, the proposed control extends the pump operating time by more than two hours under the same PV–battery sizes, and avoids deep battery depletion that could interrupt the IT system. This shows that even a simple optimization-based strategy can significantly improve the autonomy and robustness of a small nZEB-like system using measured data and a low-cost control platform.

4.4.3. Energy management strategy approach minimal electricity bill

In this section, the optimization framework is applied to a typical Vietnamese office building, using the VHH platform as a case study. The goal is to design energy management strategies that minimize the electricity bill under a time of use tariff, while maintaining safe operation of PV–Battery system and loads. The PV and battery models are reused from the Greenhouse project. While In the load profile and tariff structure reflect Vietnamese conditions, characterized by high cooling demand and strong PV production in summer, but limited storage capacity due to investment constraints.

4.4.3.1. VHH's platform description

The VHH platform represents an office building in Vietnam equipped with:

- On-site grid-connected PV system;
- A battery storage system with limited capacity;
- Monitored building loads with significant HVAC contribution;
- A time-of-use electricity tariff with off-peak, normal and peak periods, where peak prices are substantially higher than the feed-in tariff.

Measured data from VHH (building consumption, PV production, and battery operation) are used with the PV and battery models for testing energy management strategies. The NoLoad optimization tool, developed by G2Elab (University Grenoble Alpes), is used to solve optimal control problems.

4.4.3.2. Implement energy management strategies

a. Methodology Approach

Energy management is formulated as an optimization problem that coordinates PV, battery and grid exchanges in response to the tariff profile.

Two stakeholder perspectives are considered:

- *Grid operators:* reduce peak demand and power fluctuations by the grid;
- *Consumers:* maximize on-site use of PV energy and minimize the electricity bill.

The control variables include the battery charging/discharging power and the fraction of building load supplied by PV, battery and grid over time.

b. Optimization formulation

The optimization problem is defined over a 24-hour horizon with a time discretization consistent with the monitoring data. Objective is minimize the daily electricity cost, such as sum of energy imported from the grid weighted by time of use prices, while encouraging self-consumption of solar energy.

Objective function:

$$J = \left(\sum_{t=t0}^{t=T} P_{im}(t) * \Delta t * \ Price_{import}(t) \ - \ P_{ex}(t) * \Delta t * Price_{export}(t)\right) \quad (4.18)$$

Power balance: $P_{grid}(t) = P_{loads}(t) + P_{bat}(t) - P_{pv}(t)$

Decision variables: battery charge/discharge power, grid import/export, and operational status of flexible loads (when applicable).

Constraints: Power balance between PV, battery, grid and loads at each time step; Battery state-of-charge (SoC) limits and maximum charge/discharge power; Operational constraints on loads; Grid exchange limits if applicable.

Contraints include:

- $C_{min} \leq C_{bat}(t) \leq C_{max}$;
- $C_{init} = C_{bat}(T)$;
- $P_{dis\_max} \leq P_{bat}(t) \leq P_{ch\_max}$;
  If $P_{bat}(t) \leq 0$: charging power of battery: $P_{ch}(t) = P_{bat}(t)$;
  If $P_{bat}(t) > 0$: discharging power of battery: $P_{dis}(t) = P_{bat}(t)$;
- $P_{grid}(t) > 0$: $P_{im}(t)$;
- $P_{grid}(t) < 0$; $P_{ex}(t)$;
- $P_{pv}(t) >= P_{ex}(t)$;
- Assumption: $P_{ex\_max} = 2000W$; $P_{im\_max} = 2000W$;

This optimization is implemented and solved in NoLoad library, using a Sequential Least Squares Quadratic Programming (SLSQP) solver.

The symbols of parameters values for operation of PV system described in Table 11.

Table 11. Description of parameters and symbols used in simulation

Parameter	Unit	Description

J	VNĐ	Objective function is total energy costs to be paid by users
$P_{im}(t)$	W	Power consumption taken from the grid at time t: $P_{im}(t)$
$P_{ex}(t)$	W	Power pumped into the grid at time t: $P_{ex}(t)$
$P_{grid}(t)$	W	Power exchanged with the grid at time t
T	hour	Optimal time
ΔT	hour	Time step of data
P_{bat}	W	Power of the battery
$P_{dis\ max}$	W	Battery discharge capacity limit
P_{ch_max}	W	Battery charging capacity limit
C_{norm}	Wh	Rated capacity of the lead acid battery
C_{min}	Wh	20% C_{norm}
C_{max}	Wh	80% C_{norm}
C_{init}	Wh	Initial capacity of battery = 50% C_{norm}
$C_{bat}(t)$	Wh	Capacity of battery at time t
$P_{pv}(t)$	Wh	Power produced by the PV system at time t
$Price_{import}$	VNĐ	Electricity price purchased of EVN
$Price_{export}$	VNĐ	Selling price of solar power of EVN
P_{loads} (t)	W	Total consumption of loads at time t
$P_{ch\text{-}pv}$ (t)	W	Power Battery charging capacity from solar energy at time t
$P_{ch\text{-}grid}$ (t)	W	Power Battery charging capacity from the grid at time t
$P_{pv}(t)$	W	Installed PV Power at time t
P_{ex_max}	W	Maximum Power export to grid
P_{im_max}	W	Maximum Power import from grid

c. Data collection

(1) Energy consumption data: Measured load data from the smart meters (Figure 77). The VHH test model has 3 groups of loads powered by 3 different power lines: Plugs load; Lighting load; Air conditioning load (HVAC - highest priority).

(2) PV system data: one-day history of measured PV power data;

(3) Electricity tariff of EVN: The selling price of solar power to the grid (FIT2) $Price_{export}$ =1.943 (VNĐ/kWh); In Figure 78, electricity purchase price from the grid $Price_{import}$ (VNĐ/kWh).

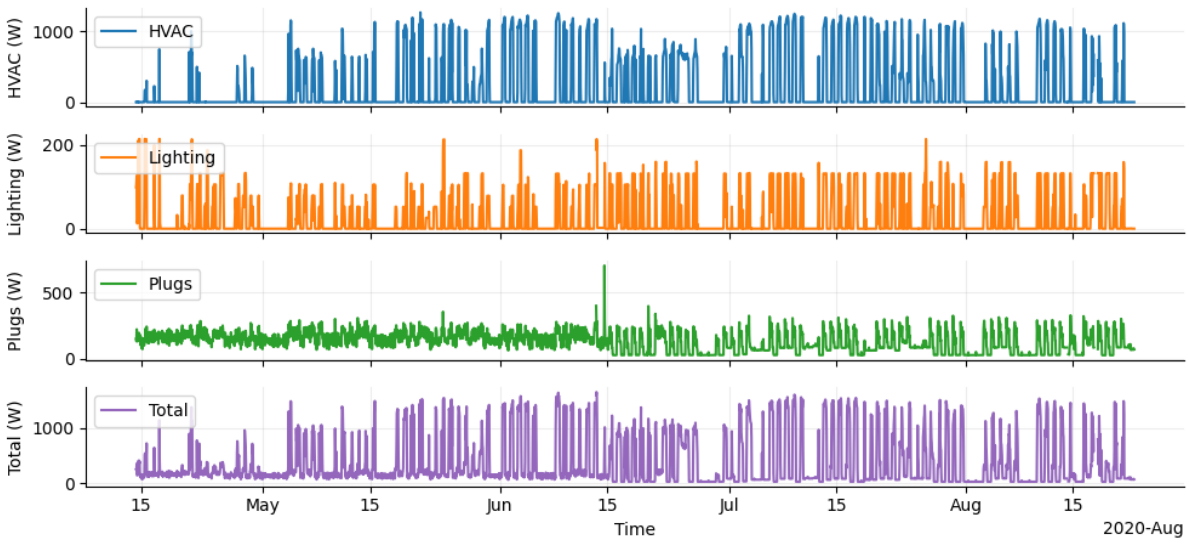


Figure 77. Detail consumption data from the smart meters in VHH's Platform

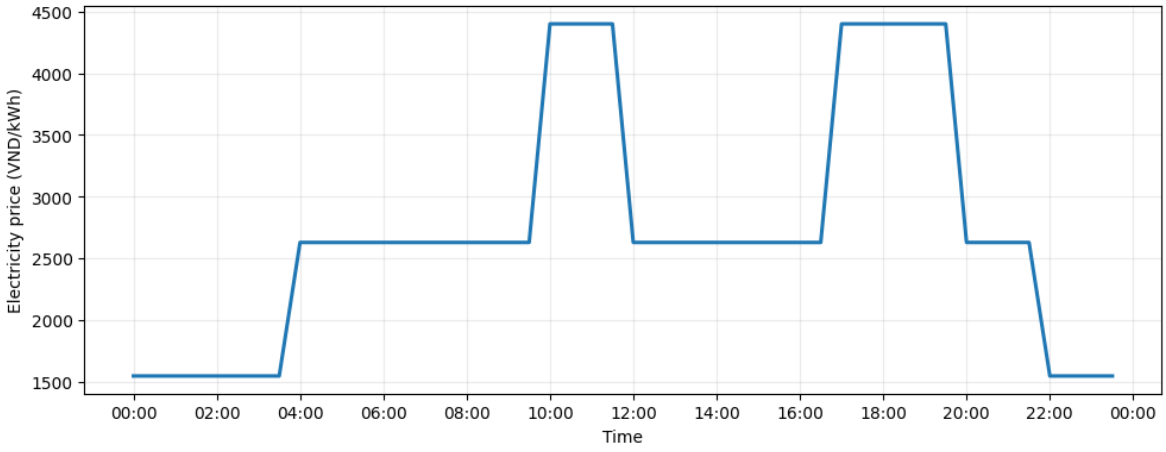


Figure 78. Electricity tariff profile applies to enterprises

d. Scenarios

The different simulation scenarios to compare benefits of the PV system with batteries.

Scenario 1 (Based-management) is presented a PV system without battery – Solar energy prioritized for on-site consumption. If the excess energy produced by PV panels will be transferred to the grid, if not enough solar energy will be taken from the grid (Figure 79).

Scenario 2 is presented A grid connected PV system with battery (Minimize electricity-bill) (Figure 80).

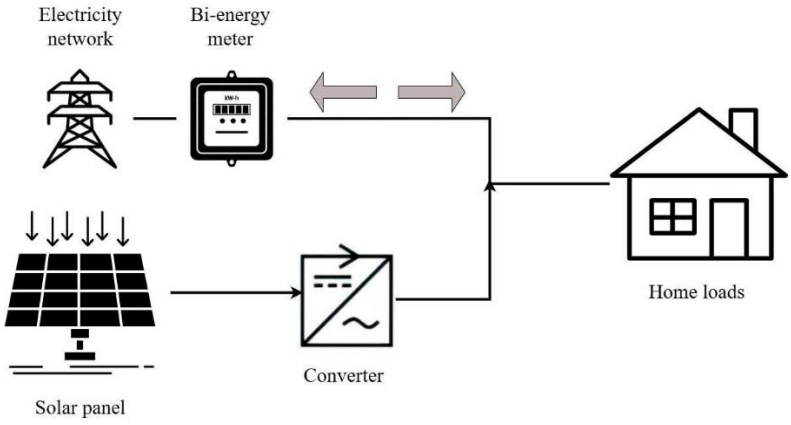


Figure 79. PV system without batteries

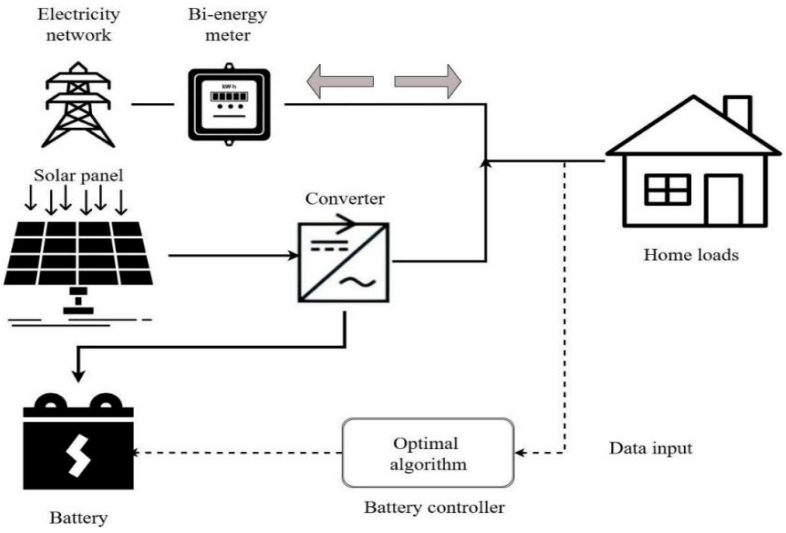


Figure 80. Grid-connected PV system built-in batteries

e. Evaluation of energy and cost benefits

This section compares the obtained optimization results with scenarios on a typical day in summer when buildings are at a time of high energy demand and PV data. The electricity tariff profile is shown in Figure 82. According to the tariff, there is a big difference between peak and off-peak hours. The electricity price in Peak hours is two times higher than the PV selling price, and 1.3 times higher than the PV selling price in normal hours. However, electricity prices at off-peak hours are lower than PV electricity prices. The objective is to evaluate the benefits of energy management solutions with/without energy storage systems.

- **Scenario 1** (Based-management): Operating PV system without battery, calculating total energy pumped into the grid and energy cost in 24h.
- **Scenario 2** (Optimization SQP method): Operating a grid connected PV system integrated batteries.

In scenario 1, power consumption from the grid is required when solar power is unavailable or insufficient to meet demand. Grid power consumption gradually decreases as PV power increases during the day.

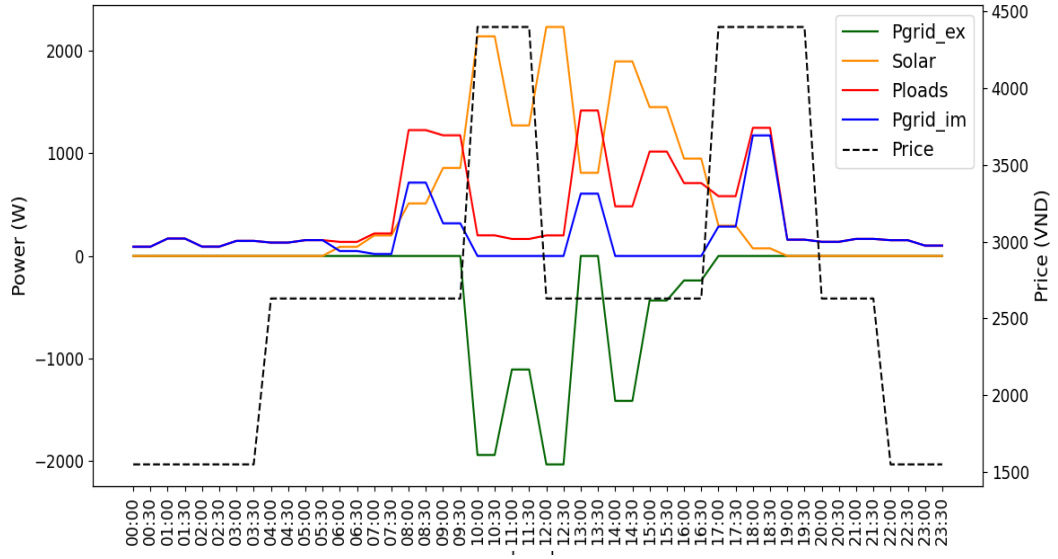


Figure 81. Simulation of operation of PV solar power system without battery

In Figure 81, the amount of PV generated is for export and self-consumption. During certain daytime periods (7:30–10:00, 12:30–14:00, and 16:30–19:00), PV generation is insufficient to meet the load demand. Therefore, the building must import from the grid at a high tariff, which increases total electricity costs.

In scenario 2, when the battery is available, and the algorithm optimizes the battery charge/discharge capacity, observing the dark blue line in Figure 81.

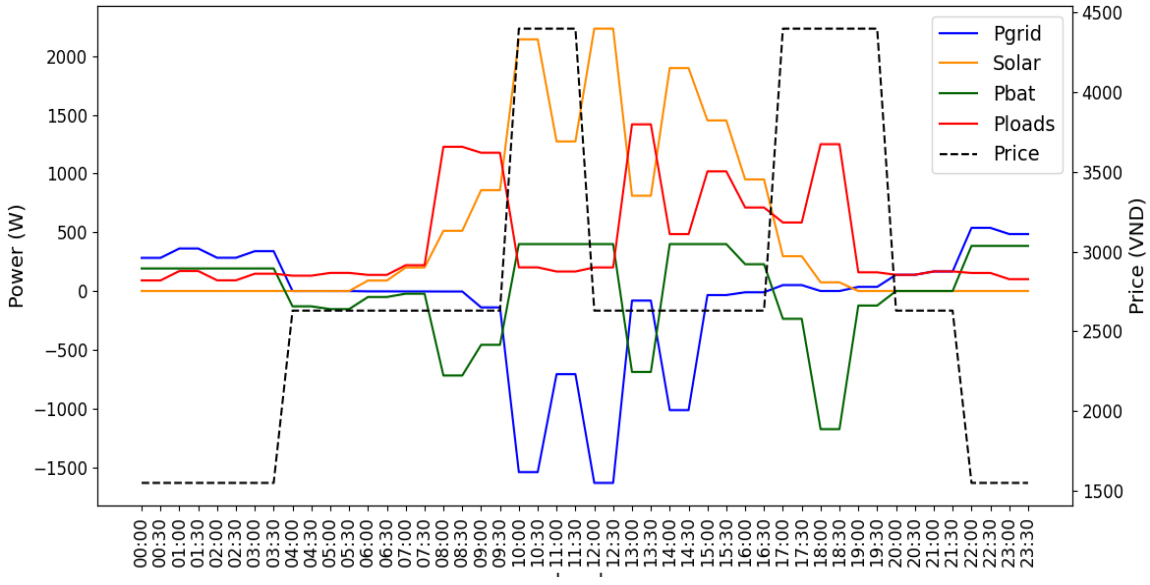


Figure 82. Simulation of PV system operation with built-in battery

Figure 82 shows that the amount of imported electricity at the time of high electricity purchase price decreases during 16h30 to 19h (when the highest purchase price of electricity). During the daytime, excess solar energy charges the battery and maximizes benefits for building owners based on the difference between the selling price of solar power and the electricity price.

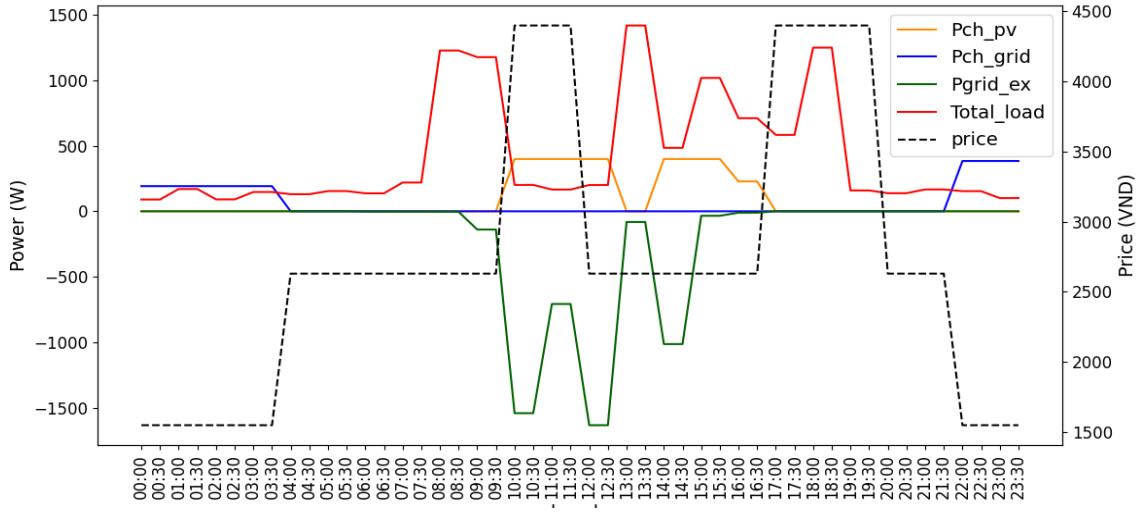


Figure 83. The relationship of energy flows and the purchase price of electricity

Figure 83 shows that battery prioritizes charging from the grid at night when electricity purchase price is lowest (lower than the selling price of solar power). The curves could give users an energy storage plan to optimize their electricity bills.

Table 12. Calculation results of cost and peak load demand in building

Scenarios	**Cost export (10^3 VND/day)**	**Cost import (10^3 VND/day)**	**Electricity Bill (10^3 VND/day)**	**Peak demand (W)**
Baseline: Building without PV system	0	30.508	30.508	1418.5
Scenario 1: Building integrated PV system without battery	13.915	14.419	0.504	1175.76
Scenario 2: Building integrated PV system with battery	10.106	4.804	-5.202	537.64

Comparing the scenarios described in Table 12, we see that the storage system reduces the peak load demand, and the electricity bill is the best obtained in the PV system with the battery. This building has around 27% of the exported energy reduction and a decrease of more than half the energy imported from the grid compared with case 1.

f. Evaluating the environmental impact

The optimization framework is used to compare the CO_2 emissions of two configurations:

- **Case 1:** Building without PV system
- **Case 2:** Building integrated PV system with battery (the best case in evaluation of energy and cost benefits)

In order to calculate global warming potential index in this study, we used parameters in Table 13 below:

Table 13. Parameters for calculation the environmental impact assessment

Parameters	Note
S_{PV} = 17.6 m2	Lifetime of monocrystalline silicon in 30 years[95]
PVGWP = 4740 ($kgCO_2e$/year/m2)	
Study time is N =30 (years)	
Bat_{GWP}=351($kgCO_2eq$ / Total whole life cycle) Cbat=800Ah	Lead-acid battery 200Ah, Total whole life cycle is 3 years [96]
VN_powergridCO_2 =815.4 ($kgCO_2eq$ / kWh)	CO_2 emission average index of Vietnam's power grid = 0.8154 (tCO_2eq/MWh) [97]
Total consumption =5095 (kWh/year)	Data in project
$Grid_{import}$ = 2498 (kWh/year)	Data in project

Case 1: Building without PV system

Environmental impact assessment index (Global warming potential (GWP)) is presented in the equation below:

$$GWP\_1\ (kgCO_2e) = VN\_powergridCO_2 * Grid_{import}\ (kWh/year) * N = 124633000.89\ (kgCO_2eq)$$

Note: In this case, the Total consumption equals $Grid_{import.}$

Case 2: Building integrated PV system with battery

Environmental impact assessment index (Global warming potential (GWP)) is presented in the equation below:

$$GWP\_2\ (kgCO_2e) = PV_{GWP} * S_{PV} + Bat_{GWP} * N/3 * C_{bat} + VN\_powergridCO_2 * Grid_{import}\ (kWh/year) * N = 62593500\ (kgCO_2eq)$$

In two cases, the environmental impact assessment index in Case 2 is lower than 2 times the index in Case 1. This confirms that the proposed methods not only lower electricity bills but also support environmental objectives for low-carbon buildings.

4.5. Conclusions of Part 4

Part 4 applied the proposed modelling and optimization framework to two case studies in France (Greenhouse testbed) and Vietnam (VHH office building) to design low-cost, control-oriented energy management strategies for nZEBs.

On the Greenhouse platform, an online PV production model was developed and validated against monitored data. Annual PV generation reaches about 406 kWh, corresponding to about 10% surplus. The applied control strategy improves on-site self-consumption, avoids battery oversizing. The pump operating time extends (2 hours) and local production almost meets the demand ($\gamma_{load} = 0.9$).

In the VHH office building, the same modelling approach is combined with a time of use (TOU) tariff to minimize the electricity bill and assess environmental impacts. The baseline building without PV, then adding PV and battery storage, turns the building into a net exporter, with the greatest reduction in daily electricity costs and a significant decrease in the annual GWP index.

The proposed low-cost IoT-BEMS, combined with simple yet powerful optimization tools, can deliver fast computations and quantitative guidance for PV–battery sizing, load control, and tariff-aware operation.

Part 5. Conclusions of Thesis

The thesis presents a practical low-cost IoT platform for monitoring, improving data quality, and optimizing energy management towards nZEBs. Low-skilled users can deploy the platform, and it is suitable for conditions in Vietnam.

Contributions of the Thesis:

- Proposed implement roadmap a Low-cost IoT-BEMS platform with RF24/ZigBee/Z-Wave/Wi-Fi integration, time series database and real-time dashboard; ready to expand and interact with existing systems.
- Proposed a data quality assurance framework based on data fusion process and data quality monitoring with online GPR for real-time data compensation.
- Propose energy models, optimization algorithms and optimal control strategies focusing on increasing self-consumption through energy storage system (ESS) and controllable load.
- Fast calculation model suitable for low-cost hardware; simulation scenarios support user planning and operational decision-making.
- Validation through case studies: (1) VHH testbed (in Vietnam): comparison of three building configurations; real PV with ESS options optimal results when combined with contextual load control. (2) Greenhouse testbed (in France): development of online PV forecasting model; at multiple resolutions to assess nZEB energy balance; load matching analysis shows demand of ESS to compensate for in-day fluctuations; maximizes self-consumption strategy.

Limitations of the Current Work

- The empirical evaluation is conducted on a limited number of buildings, and over relatively short monitoring periods. For multi-building, multi-climate and multi-season variability have not yet captured. The behavioral dataset is modest, which constrains the analysis of occupant-related effects.
- In the study, the integration of control strategies and energy storage

scheduling is only explored in a simplified way. It does not yet address complete multi-objective optimization or detailed security privacy constraints.

- The dataset and implementation are primarily used internally for the case studies; open data/code releases and systematic benchmarking against other approaches are still limited.

Future work:

- Expand multi-building/multi-climate/multi-season scale; increase behavioral sample size.
- Integrated predictive control, optimize multi-objective ESS, and enhance security–privacy.
- Promote open source/data for community reproducibility, comparison and improvement.

The thesis connects low-cost IoT devices, high-quality data, and fast model–optimization–control algorithms, and also adopts a user-oriented design. Together, these elements enable smart, efficient, and sustainable buildings that advance nZEB goals in Vietnam's energy context.

This work is supported by HaUI, VHH, IES, and G2ELab during my internship there. The data in this report has been reviewed and permitted by Dr Dang Hoang Anh - PI of HaUI's projects, Mr Bui Van Cong - PI of VHH's projects, and Prof. Benoit Delinchant – Greenhouse project, France.

The Thesis has been granted permission by VHH to use the data, system architecture and experimental results obtained from the VHH platform.

References

[1] I. Energy Agency, “World Energy Outlook 2024,” 2024. [Online]. Available: www.iea.org/terms

[2] H. A. Dang, “Modeling for building energy simulation : Application to virtual prototyping and optimal management of PREDIS MHI,” Nov. 2013.

[3] “Tracking SDG 7: The Energy Progress Report 2024, Washington D.C., USA: World Bank, Jun. 2024,” 2024. [Online]. Available: https://trackingsdg7.esmap.org/downloads.

[4] “Annual CO_2 emissions.” https://ourworldindata.org/grapher/annual-co2-emissions-per-country (accessed Aug. 31, 2024).

[5] “TRACKING SDG 7,” 2024. [Online]. Available: https://trackingsdg7.esmap.org/downloads.

[6] “Viet Nam: Efficiency and demand, IEA, 2024.” https://www.iea.org/countries/viet-nam/efficiency-demand (accessed Aug. 31, 2024).

[7] “Viet Nam: Electricity, IEA, 2024.” https://www.iea.org/countries/viet-nam/electricity (accessed Aug. 31, 2024).

[8] I. - International Energy Agency, “Electricity 2024 - Analysis and forecast to 2026,” 2024. [Online]. Available: www.iea.org

[9] “Viet Nam - Countries & Regions - IEA.” https://www.iea.org/countries/viet-nam/energy-mix (accessed Dec. 11, 2024).

[10] “Viet Nam - Countries & Regions - IEA.” https://www.iea.org/countries/viet-nam/emissions (accessed Aug. 31, 2024).

[11] “Quyết định số 13/2020/QĐ-TTg của Thủ tướng Chính phủ: ‘Về cơ chế khuyến khích phát triển điện mặt trời tại Việt Nam.,’” 2020, [Online]. Available: https://thuvienphapluat.vn/van-ban/Tai-nguyen-Moi-truong/Quyet-dinh-13-2020-QD-TTg-co-che-khuyen-khich-phat-trien-dien-mat-troi-tai-Viet-Nam-439160.aspx

[12] Iea, “Achieving a Net Zero Electricity Sector in Viet Nam.” [Online]. Available: www.iea.org

[13] S. Durovic, X. Ma, H. Chen, and A. Urakami, “Are the Barriers to Private Solar/Wind Investment in Vietnam Mainly Those That Limit Network Capacity Expansion?,” *Sustain. 2023, Vol. 15, Page 10734*, vol. 15, no. 13, p. 10734, Jul. 2023, doi: 10.3390/SU151310734.

[14] “Government of Vietnam. (2024, October 22). Decree No. 135/2024/ND-CP on mechanisms to promote rooftop solar power for self-production and self-consumption. Office of the Government. - Google Search.”

[15] “Decision of the Prime Minister No. 500/QĐ-TTg, 2023 Approving the national electricity development plan for the period of 2021 - 2030, with a vision to 2050,” 2023.

[16] “Resolution No. 55-NQ/TW dated February 11, 2020 on orientations of strategy for national energy development by 2030 with a vision towards 2045.”, [Online]. Available: https://tulieuvankien.dangcongsan.vn/he-thong-van-ban/van-ban-cua-dang/nghi-quyet-so-55-nqtw-ngay-11022020-cua-bo-chinh-tri-ve-dinh-huong-chien-luoc-phat-trien-nang-luong-quoc-gia-cua-viet-nam-den-6096

[17] “Decision No. 896-QD-TTg on approving the National Strategy for Climate Change until 2050”, [Online]. Available: https://thuvienphapluat.vn/van-ban/EN/Tai-nguyen-Moi-truong/Decision-896-QD-TTg-2022-approving-the-National-strategy-for-climate-change-until-2050/525126/tieng-anh.aspx

[18] Decision No. 280/QĐ-TTg, “National Program for Thrifty and Efficient Use of Energy for the Period of 2019 - 2030 (VNEEP3),” 2019. https://climate-laws.org/document/decision-no-280-qd-ttg-on-approval-for-national-program-for-thrifty-and-efficient-use-of-energy-for-the-period-of-2019-2030_7c0f?

[19] “Buildings - Energy System - IEA.” https://www.iea.org/energy-system/buildings#tracking (accessed Aug. 31, 2024).

[20] “2023 Global Status Report for Buildings and Construction Beyond foundations Global Status,” 2024, doi: 10.59117/20.500.11822/45095.

[21] “Diagnostic Analyses Report of Energy Efficiency Development in Three Countries: Indonesia, the Philippines and Vietnam Final Report,” Jan. 2022. [Online]. Available: https://www.energytransitionpartnership.org/wp-content/uploads/2024/04/Diagnostic-Analyses-Report-of-Energy-Efficiency-Development-in-Three-Countries-Indonesia-the-Philippines-and-Vietnam.pdf

[22] “Quy chuẩn QCVN 09:2017/BXD Công trình xây dựng sử dụng năng lượng hiệu quả.” https://luatvietnam.vn/xay-dung/quy-chuan-viet-nam-qcvn-09-2017-bxd-bo-xay-dung-158656-d3.html (accessed Nov. 28, 2023).

[23] H. C. Minh, “Energy Efficiency Improvement in Commercial and High-Rise Residential Buildings in Vietnam EXECUTIVE REPORT ON THE ESTABLISHMENT OF SEC PROFILES, NATIONAL ENERGY BENCHMARKS, EE LABELLING SYSTEM AND M&V PROTOCOL FOR HIGH-RISE BUILDINGS IN VIETNAM,” 2021.

[24] European Parliament and of the Council., “Directive 2010/31/EU of the Euroepan Parliament and of the Council of 19 May,” 2018, [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2018:156:TOC

[25] G. Paoletti, R. P. Pascuas, R. Pernetti, and R. Lollini, “Nearly Zero Energy Buildings: An overview of the main construction features across Europe,” *Buildings*, vol. 7, no. 2, May 2017, doi: 10.3390/buildings7020043.

[26] E. R. Sanseverino, H. L. T. Thuy, M. H. Pham, M. L. Di Silvestre, N. N. Quang, and S. Favuzza, “Review of potential and actual penetration of solar power in Vietnam,” *Energies*, vol. 13, no. 10, May 2020, doi: 10.3390/EN13102529.

[27] D. D’Agostino, P. Zangheri, and L. Castellazzi, “Towards nearly zero energy buildings in Europe: A focus on retrofit in non-residential buildings,” *Energies*, vol. 10, no. 1, 2017, doi: 10.3390/en10010117.

[28] Jaume Salom; Anna Joanna Marszal; José Candanedo; Joakim Widén; Karen Byskov Lindberg; Igor Sartori, “ANALYSIS OF LOAD MATCH AND GRID INTERACTION INDICATORS IN NET ZERO ENERGY BUILDINGS WITH HIGH-RESOLUTION DATA A report of Subtask A, IEA Task 40/Annex 52 Towards Net Zero Energy Solar Buildings,” 2013.

[29] Van Binh Dinh, “Méthodes et outils pour le dimensionnement des bâtiments et des systèmes énergétiques en phase d’esquisse intégrant la gestion optimale’,” University of Grenoble, 2017.

[30] N. Artiges, “De l’instrumentation au contrôle optimal prédictif pour la performance énergétique du bâtiment,” 2016.

[31] A. M. Vega, F. Santamaria, and E. Rivas, “Modeling for home electric energy management: A review,” *Renew. Sustain. Energy Rev.*, vol. 52, pp. 948–959, Dec. 2015, doi: 10.1016/J.RSER.2015.07.023.

[32] B. Seabra, P. F. Pereira, H. Corvacho, C. Pires, and N. M. M. Ramos, “Low Energy Renovation of Social Housing: Recommendations on Monitoring and Renewable Energies Use,” *Sustain. 2021, Vol. 13, Page 2718*, vol. 13, no. 5, p. 2718, Mar. 2021, doi: 10.3390/SU13052718.

[33] P. F. Pereira, N. M. M. Ramos, and M. L. Simões, “Data-driven occupant actions prediction to achieve an intelligent building,” *https://doi.org/10.1080/09613218.2019.1692648*, vol. 48, no. 5, pp. 485–500, Jul. 2019, doi: 10.1080/09613218.2019.1692648.

[34] P. F. Pereira and N. M. M. Ramos, “Influence of Occupant Behaviour on the State of Charge of a Storage Battery in a nearly-Zero Energy Building,” *E3S Web Conf.*, vol. 172, Jun. 2020, doi: 10.1051/E3SCONF/202017216010.

[35] M. Bourdeau, J. Waeytens, N. Aouani, P. Basset, and E. Nefzaoui, “A Wireless Sensor Network for Residential Building Energy and Indoor Environmental Quality Monitoring: Design, Instrumentation, Data Analysis and Feedback,” *Sensors*, vol. 23, no. 12, 2023, doi: 10.3390/s23125580.

[36] T. T. H. Vu, B. Delinchant, A. T. Phan, V. C. Bui, and D. Q. Nguyen, “A

Practical Approach to Launch the Low-Cost Monitoring Platforms for Nearly Net-Zero Energy Buildings in Vietnam," *Energies 2022, Vol. 15, Page 4924*, vol. 15, no. 13, p. 4924, Jul. 2022, doi: 10.3390/EN15134924.

[37] I. Energy Agency, "Efficient Grid-Interactive Buildings." [Online]. Available: www.iea.org

[38] "ESP32 Series Datasheet Version 5.0 2.4 GHz Wi-Fi + Bluetooth ® + Bluetooth LE SoC Including." [Online]. Available: www.espressif.com

[39] N. U. Okafor, Y. Alghorani, and D. T. Delaney, "Improving Data Quality of Low-cost IoT Sensors in Environmental Monitoring Networks Using Data Fusion and Machine Learning Approach," *ICT Express*, vol. 6, no. 3, pp. 220–228, Sep. 2020, doi: 10.1016/J.ICTE.2020.06.004.

[40] M. A. Zaidan *et al.*, "Intelligent Calibration and Virtual Sensing for Integrated Low-Cost Air Quality Sensors," *IEEE Sens. J.*, vol. 20, no. 22, pp. 13638–13652, Nov. 2020, doi: 10.1109/JSEN.2020.3010316.

[41] A. Paone and J. P. Bacher, "The impact of building occupant behavior on energy efficiency and methods to influence it: A review of the state of the art," *Energies*, vol. 11, no. 4, 2018, doi: 10.3390/en11040953.

[42] T. M. Cristino, A. Faria Neto, F. Wurtz, and B. Delinchant, "Barriers to the adoption of energy-efficient technologies in the building sector: A survey of Brazil," *Energy Build.*, vol. 252, p. 111452, Dec. 2021, doi: 10.1016/J.ENBUILD.2021.111452.

[43] "Vietnam - National Energy Efficiency Program 2019 – 2030 | VNEEP - Chương trình mục tiêu quốc gia về sử dụng năng lượng tiết kiệm và hiệu quả."
https://documents1.worldbank.org/curated/en/598851561961183317/pdf/Vietnam-National-Energy-Efficiency-Program-2019-2030.pdf

[44] A. D. B. S. Asia, "THE INTERNET OF THINGS IN THE POWER SECTOR OPPORTUNITIES IN ASIA ADB SUSTAINABLE DEVELOPMENT WORKING PAPER SERIES Study of Three," no. 48, 2017.

[45] B. Delinchant and J. Ferrari, “Standards and Technologies from Building Sector, IoT, and Open-Source Trends,” *Towar. Energy Smart Homes*, pp. 49–111, 2021, doi: 10.1007/978-3-030-76477-7_3.

[46] Y. Sun, F. Haghighat, and B. C. M. Fung, “A review of the-state-of-the-art in data-driven approaches for building energy prediction,” *Energy Build.*, vol. 221, p. 110022, Aug. 2020, doi: 10.1016/J.ENBUILD.2020.110022.

[47] “• Services to monetize the IoT infrastructure Data and device management from things to cloud”.

[48] “Microsoft Azure IoT Reference Architecture,” 2018. [Online]. Available: https://azure.microsoft.com/en-us/services/iot-central/

[49] “Internet of Things: Wireless Sensor Networks White Paper.”

[50] G. Jesus, A. Casimiro, and A. Oliveira, “A survey on data quality for dependable monitoring in wireless sensor networks,” *Sensors (Switzerland)*, vol. 17, no. 9, Sep. 2017, doi: 10.3390/S17092010.

[51] “Home | MySensors - Create your own Connected Home Experience.” https://www.mysensors.org (accessed Nov. 26, 2023).

[52] “Newly Optimized RF24Network Layer: Addressing Format: Understanding Addressing and Topology.”

[53] B. Delinchant, H. A. Dang, H. T. T. Vu, and D. Q. Nguyen, “Massive arrival of low-cost and low-consuming sensors in buildings: Towards new building energy services,” in *IOP Conference Series: Earth and Environmental Science*, Aug. 2019, vol. 307, no. 1. doi: 10.1088/1755-1315/307/1/012006.

[54] D. Yang, C. Gu, Z. Dong, P. Jirutitijaroen, N. Chen, and W. M. Walsh, “Solar irradiance forecasting using spatial-temporal covariance structures and time-forward kriging,” *Renew. Energy*, vol. 60, pp. 235–245, 2013, doi: https://doi.org/10.1016/j.renene.2013.05.030.

[55] G. F. M. Nascimento, F. Wurtz, P. Kuo-Peng, B. Delinchant, and N. J. Batistela, “Outlier Detection in Buildings’ Power Consumption Data Using Forecast Error,” *Energies 2021, Vol. 14, Page 8325*, vol. 14, no. 24, p. 8325, Dec. 2021, doi: 10.3390/EN14248325.

[56] A. Livera *et al.*, “Data processing and quality verification for improved photovoltaic performance and reliability analytics,” *Prog. Photovoltaics Res. Appl.*, vol. 29, no. 2, pp. 143–158, Feb. 2021, doi: 10.1002/pip.3349.

[57] H. Y. Teh, A. W. Kempa-Liehr, and K. I. K. Wang, “Sensor data quality: a systematic review,” *J. Big Data 2020 71*, vol. 7, no. 1, pp. 11-, Feb. 2020, doi: 10.1186/S40537-020-0285-1.

[58] B. Delinchant *et al.*, “Machine Learning on Buildings Data for Future Energy Community Services,” Jul. 2021.

[59] J. Lloret, M. Garcia, D. Bri, and S. Sendra, “A wireless sensor network deployment for rural and forest fire detection and verification,” *Sensors*, vol. 9, no. 11, pp. 8722–8747, Nov. 2009, doi: 10.3390/s91108722.

[60] S. Rinaldi, A. Flammini, M. Pasetti, L. C. Tagliabue, A. C. Ciribini, and S. Zanoni, “Metrological Issues in the Integration of Heterogeneous Iot Devices for Energy Efficiency in Cognitive Buildings,” *I2MTC 2018 - 2018 IEEE Int. Instrum. Meas. Technol. Conf. Discov. New Horizons Instrum. Meas. Proc.*, pp. 1–6, Jul. 2018, doi: 10.1109/I2MTC.2018.8409740.

[61] A. A. Aguileta, R. F. Brena, O. Mayora, E. Molino-Minero-re, and L. A. Trejo, “Multi-Sensor Fusion for Activity Recognition—A Survey,” *Sensors 2019, Vol. 19, Page 3808*, vol. 19, no. 17, p. 3808, Sep. 2019, doi: 10.3390/S19173808.

[62] A. Tawakuli, B. Havers, V. Gulisano, D. Kaiser, and T. Engel, “Survey:Time-series data preprocessing: A survey and an empirical analysis,” *J. Eng. Res. early access*, Mar. 2024, doi: 10.1016/J.JER.2024.02.018.

[63] K. A. Klise, J. S. Stein, and J. Cunningham, “Application of IEC 61724 Standards to Analyze PV System Performance in Different Climates,” in *2017 IEEE 44th Photovoltaic Specialist Conference (PVSC)*, 2017, pp. 3161–3166. doi: 10.1109/PVSC.2017.8366666.

[64] F. Najibi, E. Alonso, D. Apostolopoulou, and E. Alonso, “Gaussian Process Regression for Probabilistic Short-term Solar Output Forecast Enhanced

performance Gaussian process regression for probabilistic short-term solar output forecast," 2021, doi: 10.48550/arXiv.2002.10878.

[65] S. C. Lim, J. H. Huh, S. H. Hong, C. Y. Park, and J. C. Kim, "Solar Power Forecasting Using CNN-LSTM Hybrid Model," *Energies*, vol. 15, no. 21, Nov. 2022, doi: 10.3390/en15218233.

[66] F. Sanchez-Sutil, A. Cano-Ortega, J. C. Hernandez, and C. Rus-Casas, "Development and calibration of an open source, low-cost power smart meter prototype for PV household-prosumers," *Electron.*, vol. 8, no. 8, Aug. 2019, doi: 10.3390/ELECTRONICS8080878.

[67] S. Bell, "A Beginner's Guide to Uncertainty of Measurement Measurement Good Practice Guide," 1999.

[68] T. T. V. Hong, B. Delinchant, J. Ferrari, and Q. D. Nguyen, "Autonomous Electrical System Monitoring and Control Strategies to Avoid Oversized Storage Capacity," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 505, no. 1, Jul. 2020, doi: 10.1088/1755-1315/505/1/012045.

[69] H. Carstens, X. Xia, and S. Yadavalli, "Bayesian energy measurement and verification analysis," *Energies*, vol. 11, no. 2, Feb. 2018, doi: 10.3390/en11020380.

[70] M. C. Burkhart, Y. Heo, and V. M. Zavala, "Measurement and verification of building systems under uncertain data: A Gaussian process modeling approach," *Energy Build.*, vol. 75, pp. 189–198, Jun. 2014, doi: 10.1016/J.ENBUILD.2014.01.048.

[71] "General linear model - Wikipedia." https://en.wikipedia.org/wiki/General_linear_model (accessed Aug. 31, 2023).

[72] C. E. Rasmussen and C. K. I. . Williams, *Gaussian processes for machine learning*. MIT Press, 2006.

[73] A. T. Phan, T. T. H. Vu, D. Q. Nguyen, E. R. Sanseverino, H. T. T. Le, and V. C. Bui, "Data Compensation with Gaussian Processes Regression: Application in Smart Building's Sensor Network," *Energies*, vol. 15, no. 23,

Dec. 2022, doi: 10.3390/en15239190.

[74] D. Yan *et al.*, "Occupant behavior modeling for building performance simulation: Current state and future challenges," *Energy Build.*, vol. 107, pp. 264–278, Nov. 2015, doi: 10.1016/J.ENBUILD.2015.08.032.

[75] A. H. GhaffarianHoseini, N. D. Dahlan, U. Berardi, A. GhaffarianHoseini, and N. Makaremi, "The essence of future smart houses: From embedding ICT to adapting to sustainability principles," *Renew. Sustain. Energy Rev.*, vol. 24, pp. 593–607, Aug. 2013, doi: 10.1016/J.RSER.2013.02.032.

[76] M. Amayri *et al.*, "Estimating Occupancy In Heterogeneous Sensor Environment," *Estim. Occup. Heterog. Sens. Environ. Energy Build.*, vol. 129, pp. 46–58, 2016, doi: 10.1016/j.enbuild.2016.07.026ï.

[77] M. H. Rahman and S. Yamashiro, "Novel distributed power generating system of PV-ECaSS using solar energy estimation," *IEEE Trans. Energy Convers.*, vol. 22, no. 2, pp. 358–367, Jun. 2007, doi: 10.1109/TEC.2006.870832.

[78] K. J. Iheanetu, "Solar Photovoltaic Power Forecasting: A Review," *Sustain.*, vol. 14, no. 24, 2022, doi: 10.3390/su142417005.

[79] U. K. Das *et al.*, "Forecasting of photovoltaic power generation and model optimization: A review," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 912–928, 2018, doi: 10.1016/J.RSER.2017.08.017.

[80] M. J. Reno, C. W. Hansen, and J. S. Stein, "SANDIA REPORT Global Horizontal Irradiance Clear Sky Models: Implementation and Analysis".

[81] C. A. Gueymard, "Clear-sky irradiance predictions for solar resource mapping and large-scale applications: Improved validation methodology and detailed performance analysis of 18 broadband radiative models," *Sol. Energy*, vol. 86, no. 8, pp. 2145–2169, Aug. 2012, doi: 10.1016/J.SOLENER.2011.11.011.

[82] V. Badescu, "Verification of some very simple clear and cloudy sky models to evaluate global solar irradiance," *Sol. Energy*, vol. 61, no. 4, pp. 251–264, Oct. 1997, doi: 10.1016/S0038-092X(97)00057-1.

[83] P. Ineichen, “Comparison of eight clear sky broadband models against 16 independent data banks,” *Sol. Energy*, vol. 80, no. 4, pp. 468–478, Apr. 2006, doi: 10.1016/J.SOLENER.2005.04.018.

[84] A. Kornelakis and E. Koutroulis, “Methodology for the design optimisation and the economic analysis of grid-connected photovoltaic systems,” *IET Renew. Power Gener.*, vol. 3, no. 4, pp. 476–492, 2009, doi: 10.1049/IET-RPG.2008.0069.

[85] W. T. Jewell and T. D. Unruh, “Limits on cloud-induced fluctuation i photovoltaic generation,” *IEEE Trans. Energy Convers.*, vol. 5, no. 1, pp. 8–14, 1990, doi: 10.1109/60.50805.

[86] D. V. Nguyen, B. Delinchant, B. V. Dinh, and T. X. Nguyen, “Irradiance forecast model for PV generation based on cloudiness web service,” in *IOP Conference Series: Earth and Environmental Science*, Aug. 2019, vol. 307, no. 1. doi: 10.1088/1755-1315/307/1/012008.

[87] G. Serale, M. Fiorentini, A. Capozzoli, D. Bernardini, and A. Bemporad, “Model Predictive Control (MPC) for enhancing building and HVAC system energy efficiency: Problem formulation, applications and opportunities,” *Energies*, vol. 11, no. 3, 2018, doi: 10.3390/en11030631.

[88] P. T. B. and J. W. Tolle, “Sequential Quadratic Programming,” *Acta Numer.*, vol. 4, pp. 1–51, 1995.

[89] L. Agobert, S. Hodencq, B. Delinchant, L. Gerbaud, and W. Frederic, “NoLOAD, Open Software for Optimal Design and Operation using Automatic Differentiation,” Sep. 2021.

[90] J. G. De Matos, F. S. F. E Silva, and L. A. S. Ribeiro, “Power control in AC isolated microgrids with renewable energy sources and energy storage systems,” *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3490–3498, Jun. 2015, doi: 10.1109/TIE.2014.2367463.

[91] S. A. Klein and W. A. Beckman, “Loss-of-load probabilities for stand-alone photovoltaic systems,” *Sol. Energy*, vol. 39, no. 6, pp. 499–512, Jan. 1987, doi: 10.1016/0038-092X(87)90057-0.

[92] E. Koutroulis, D. Kolokotsa, A. Potirakis, and K. Kalaitzakis, “Methodology for optimal sizing of stand-alone photovoltaic/wind-generator systems using genetic algorithms,” *Sol. Energy*, vol. 80, no. 9, pp. 1072–1088, Sep. 2006, doi: 10.1016/j.solener.2005.11.002.

[93] D. Michaelson, H. Mahmood, and J. Jiang, “A Predictive Energy Management System Using Pre-Emptive Load Shedding for Islanded Photovoltaic Microgrids,” *IEEE Trans. Ind. Electron.*, vol. 64, no. 7, pp. 5440–5448, Jul. 2017, doi: 10.1109/TIE.2017.2677317.

[94] E. S. Gavanidou and A. G. Bakirtzis, “Design of a stand alone system with renewable energy sources using trade off methods,” *IEEE Trans. Energy Convers.*, vol. 7, no. 1, pp. 42–48, 1992, doi: 10.1109/60.124540.

[95] “AppInies.” https://base-inies.fr/tableau-de-bord (accessed Aug. 11, 2023).

[96] “Batterie Plomb Acide [capacité=200Ah] - DONNEE ENVIRONNEMENTALE PAR DEFAUT (v.1.2).” https://base-inies.fr/infos-produit/31390 (accessed Aug. 11, 2023).

[97] K. H. Tran, “Clean Production & Energy Efficiency (P116846),” Washington, D.C, Jun. 2017. [Online]. Available: http://documents.worldbank.org/curated/en/845851498493250863

Appendix A: List of Publications

No.	Publication titles	Time (to be) published	ISI-Scopus
	International publications		
1	**Data Compensation with Gaussian Processes Regression: Application in Smart Building's Sensor Network**	*Phan A.T., **Vu T.T.H**, Nguyen D.Q., Sanseverino R.E., Le T.T.H., Bui V.C. Energies 2022, 15, 9190. https://doi.org/10.3390/en15239190*	**SCIE**
2	**A Practical Approach to Launch the Low-Cost Monitoring Platforms for Nearly Net-Zero Energy Buildings in Vietnam**	***Vu T.T.H**, Benoit DELINCHANT, Anh Tuan PHAN, Van Cong BUI and Dinh Quang NGUYEN, Energies 2022, 15(13), 4924; https://doi.org/10.3390/en15134924, 05 Jul 2022*	**SCIE**
3	**Machine Learning on Buildings Data for Future Energy Community Services**	*Benoit Delinchant, Gustavo Martin, Tiansi Laranjeira, **Thi-Tuyet-Hong Vu**, Muhammad Salman Shahid, et al... , SGE 2021 - Symposium de Génie Electrique, Jul 2021, Nantes, France. ⟨hal-03638394⟩*	
4	**Autonomous electrical system monitoring and control strategies to avoid oversized storage capacity**	***Hong T T Vu** , Benoit Delinchant, Jérôme Ferrari, and Quang D Nguyen, IOP Conference Series: Earth and environmental Science (EES) (ISSN: 1755-1315),* 2020	**Scopus**
5	**Massive arrival of low-cost and low-consuming sensors in buildings: towards new building energy services**	*B.Delinchant, H.A. Dang, **H.T.T Vu**, and D.Q. Nguyen, IOP Conference Series: Earth and Environmental Science (EES) (ISSN: 1755-1315) 2019*	**Scopus**
	National publications		
1	**A New State-Space Model of Three-Phase Power Systems: Application in Positive and Negative Sequence Estimation**	*Phan Anh Tuấn, **Vũ Thị Tuyết Hồng**, Nguyễn Đình Quang, Nguyễn Thị Diệu Linh Journal of science and technology of HaUI- ISSN 1859-3585, No 50 - 2/2019*	
2	**Overview of automation in building energy management** -	***Vũ Thị Tuyết Hồng**, Đặng Hoàng Anh, Nguyễn Đình Quang Journal of science and technology of HaUI- ISSN 1859-3585, No 48 (10/2018)*	