SAPIENZA
UNIVERSITÀ DI ROMA

DOTTORATO DI RICERCA IN ENERGIA E AMBIENTE
PhD Program in ENERGY and ENVIRONMENT

XXXVI Cycle

Optimization and Management of Microgrids in the Built Environment based on Intelligent Digital Twins

PhD Thesis

Scuola di Dottorato in Scienze e Tecnologie per l'Innovazione Industriale
Facoltà di Ingegneria Civile e Industriale
Dipartimento di Ingegneria Astronautica, Elettrica ed Energetica

Sofia Agostinelli

Advisor

Prof. Fabrizio Cumo

A.A. 2022-2023

I would like to express my heartfelt gratitude to all the people who have supported me throughout this journey; without them, I would not be the person I am today.

I am deeply thankful to prof. Fabrizio Cumo for his belief in me from the very beginning, allowing me to have unique and irreplaceable experiences during my years as a researcher. Thank you for your endless support and guidance.

Immense gratitude goes to my mother, always supportive and encouraging, a role model of resilience and strength. I also extend my deepest thanks to my father, who has always been there for me, and to my brother, my best friend, who has been a constant source of support and care in both good and bad times.

My gratitude extends to my friends, both old and new, for their love, support, and for always encouraging me to do my best and sharing in the joy of my successes. I am also thankful to all the people I have collaborated with over the years for being a part of this journey.

Lastly, I would like to express my gratitude to Dr. Michael Grieves, who profoundly influenced my research with his fascinating and enlightening insights, and to the Digital Twin Consortium, for being an invaluable place of sharing and discovery.

Thank you to each and every one of you.

Sofia

Table of Contents

Abstract

The interplay between the built environment and energy use has profound implications for global energy consumption, emissions, and the transition towards sustainable systems. Recent analyses reveal that buildings are responsible for consuming approximately 40% of global energy, contributing to over 70% of electricity consumption and nearly a third of all carbon emissions.

As the energy paradigm shifts, so does the dynamic between infrastructure and energy sources. Distributed Energy Resources (DERs), including microgrids, nanogrids, and behind-the-meter energy storage, offer both challenges and opportunities. Additionally, innovations such as grid-interactive efficient buildings and Electric Vehicle (EV) integration pathways (like EV-to-grid and EV-to-home) are broadening the horizons of demand response, heralding an age marked by heightened demand flexibility.

The past decade's advancements on the Internet of Things (IoT) have opened an era of interconnected sensors and devices. This transition from limited to expansive data repositories facilitates advanced anomaly detection, predictive maintenance, and smart features based on Machine Learning (ML) and other Artificial Intelligence (AI)-powered methods. In this scenario, the emergence of the Digital Twin paradigm (DT) is redefining how we model the contexts in which these innovations operate, even though this concept remains under-explored in the energy and smart grid sectors.

This research was conducted with the cooperation of ENEA (Italian National Agency for New Technologies, Energy and Sustainable Economic Development) and delves into the challenges and potential of integrating AI with Digital Twins to support and enhance energy management and optimization in built environments. Using a testbed at the Faculty of Architecture of Sapienza

University of Rome known as the SmartLAB—a building provided with an IoT network, a photovoltaic system, and an EV charging station—the microgrid is monitored and regulated via an AI-driven energy management system (EMS). Through the deployment of IoT devices, vast data streams are generated enabling insights for an Intelligent Digital Twin (IDT) framework built on open-source methods, integrating Building Information Modeling (BIM), IoT, and AI within a Proof of Concept (PoC) Ecosystem. The research's core aims are to craft an autonomous system capable of discerning energy loads using Naïve Bayes Classifier with Association Rule Mining ensuring efficient energy consumption management and optimized distribution among distinct loads/users. Moreover, it seeks to enhance the well-being of SmartLAB occupants by overseeing CO_2 levels, lowering volatile organic compound (VOC) concentrations through IoT-guided, data-informed procedures. This approach demonstrates adaptability across multiple contexts, from grid interactions to smart ecosystems, improving real-time control and providing data-driven energy optimization strategies.

Moreover, the proposed approach embraces scalable and cross-disciplinary strategies, paving the way for further integrations, such as space management systems and predictive maintenance. The findings highlight the advantages of embracing DT technologies within the built sector, while ongoing advancements are also discussed towards DT-based Smart Cities and Energy Communities for a more sustainable built environment.

Chapter 1

Introduction

1.1 Background

The responsibilities of cities in energy consumption and CO_2 emissions are well understood, and it's expected that their impact on climate change will further increase (Marco Casini, 2022). Urban areas are continuously growing, both in terms of inhabitants and occupied space, at a rate of 10,000 square meters added every minute (every year, equivalent to the total size of Japan). By 2021, cities will cover 3% of the Earth's surface, but they will account for two-thirds of global energy demand and 70% of CO_2 emissions (European Commission. Joint Research Centre. Institute for Energy and Transport., 2015).

To address climate change and related environmental risks, coupled with the recent threat of the energy crisis, the European Union has set a series of goals for its member states. The Green Deal outlines the long-term goal to achieve zero CO_2 emissions by 2050, preceded by an intermediate phase of a 55% reduction by 2030 (European Commission. Directorate General for Research and Innovation., 2020). The European Commission entrusts cities with a key role towards climate neutrality. As such, they have been propelled and accelerated through the EU's mission "100 climate-neutral cities by 2030—by and for the citizens" (abbreviated as "Cities Mission"), part of the Mission Climate-Neutral and Smart Cities (European Commission. Directorate General for Research and Innovation, 2021), introduced by the Horizon Europe program. On April 28, 2022, the Commission announced the 100 chosen cities that will receive the Commission's support to achieve the goal of being climate-neutral and smart by 2030, aiming to make them innovation hubs and benchmarks for the rest of European cities.

Recognizing the massive carbon footprint of cities and the targets set to reduce it, there's a need to support cities in speeding up their green and digital transformation since climate mitigation heavily relies on urban actions. So, while climate objectives are set at national and supranational levels, urban-level governments and decision-makers are directly involved in their realization. The city community, being the end recipient of local policy strategies and measures, must adapt to the changes they bring about. Cities must grapple with this responsibility, but also with unique opportunities provided by various funding programs, with potential benefits not only for energy/climate matters but also for livability, overall urban efficiency, health, and citizens' quality of life (Vivi et al., 2019). Cities can be pioneers in reducing emissions and air pollutants by leveraging synergies across different sectors of the urban energy system and harnessing new technologies supporting the energy transition available today.

New methods and inventories for assessing CO_2 emissions at the city level have been developed to help decision-makers define actions and strategies to achieve decarbonization goals (Fuller et al., 2020). However, focusing solely on the energy sector isn't enough to ensure a carbon-neutral transition: connecting and combining the sectors of transportation, heating, cooling, water management, and waste is vital. This results in a large number of factors and variables at play and a high level of complexity.

Over the past hundred years, the energy sector has played an indispensable role in fueling both the manufacturing and service sectors of worldwide economy. Its influence has been paramount in driving economic and societal progress, consequently uplifting the living standards, despite the unpredictability of the global market. Therefore, refining the efficiency and sustainability of our energy systems has always been at the forefront of global attention. Highlighting this commitment, the sustainable development goals (SDGs) laid down a roadmap with seventeen specific objectives to enhance human development. Among these, the seventh goal prioritizes the creation of energy systems that are not only sustainable and innovative but also affordable and reliable (Imbulana Arachchi & Managi, 2021)

However, the past two decades have seen the looming shadow of climate change, primarily driven by the emissions of carbon dioxide (CO_2) from fossil fuel utilization, including coal, oil, and natural gas. This has presented a monumental challenge to the world's energy strategies (B. Chen et al., 2019). Despite the global community's earnest efforts to pivot towards sustainable energy solutions, a significant portion of the world still relies heavily on carbon-intensive and polluting energy infrastructures. An undeniable correlation exists between rising energy consumption, economic growth, and surges in CO_2 emissions. While there has been a notable reduction in CO_2 emissions with the increased adoption of renewable energy (RE) sources, the complete replacement of fossil fuel-based energy seems distant (Renewables REN21. Global status report., 2018).

This urgency to tackle climate change and the associated CO_2 emissions has necessitated a global shift in energy policies, focusing on mapping out feasible routes to a sustainable energy transition. Indeed, we are witnessing these shifts globally. Historical reviews over the past decade have highlighted the evolution of energy transitions, especially in regions like Europe and the United States (Gales et al., 2007). With these transitions comes the need for innovative energy modeling, which bridges the gap between novel energy infrastructure, policy formation, ecological impact, and energy supply security. The overarching goal remains managing the balance of costs, advantages, and challenges during this transition to ensure sustainability. In response, several governments are fortifying their commitment to this cause. For instance, the European Union (EU) has pledged to achieve a net-zero greenhouse gas (GHG) emission economy by 2050. Similarly, post the Paris Agreement, China has marked 2030 as its milestone year for peak emissions (Deloitte, 2020).

One focal point of this global shift is "decarbonization," a term emphasizing the continuous reduction of carbon intensity in energy over a period. The consensus among experts is clear: transitioning to renewable energy sources is pivotal for effective decarbonization. Numerous countries are taking leaps in this direction, with nations like Paraguay, Norway, and Costa Rica already garnering significant portions of their energy from renewable sources (Kroposki et al., 2017). While 2015's COP21[1] in Paris was a landmark event setting global decarbonization objectives, it was the subsequent COP22 in Marrakech, hailed as “the COP of action”, that paved the path for their practical implementation. Among the multiple objectives set during COP21, two were directly tied to energy decarbonization: Goal 7, ensuring universal access to sustainable energy, and Goal 9, focusing on sustainable industrialization and innovation (Di Silvestre et al., 2018).

Parallel to these energy discussions, the dawn of the digital age ushered in the fourth industrial revolution. "Digitalization," as defined by Gartner, embodies

[1] 21st Conference of the Parties to the United Nations Framework Convention on Climate Change (UNFCCC). It took place in Paris, France, from November 30 to December 12, 2015. The conference is most notably recognized for the resulting "Paris Agreement," a global pact where participating countries agreed to undertake efforts to combat climate change, primarily by aiming to limit global warming to well below 2 degrees Celsius above pre-industrial levels and to pursue efforts to limit the temperature increase to 1.5 degrees Celsius. The Paris Agreement represents a significant international commitment to address the challenges posed by climate change.

the transition to a digital-centric business paradigm, leveraging digital tools for new revenue streams and value creation. This digital shift in the energy domain is accelerating, exemplified by the adoption of innovative digital tools like smart meters, state-of-the-art control systems, artificial intelligence, and the revolutionary "Digital Twin" concept (Gonzalez et al., 2023).

1.2 Green and Digital Challenges

The built environment, which comprises the sum total of human-made structures in our landscapes—from residential and commercial edifices to infrastructure systems—plays a pivotal role in determining the environmental and climatic well-being of our planet. Historically, the construction and operation of these structures have been among the leading contributors to global carbon emissions. With buildings alone accounting for nearly 40% of annual global greenhouse gas emissions when considering their entire lifecycle, it becomes evident that the built environment remains at the heart of the climate crisis conversation (International Energy Agency, 2022).

This significant carbon footprint is the result of multiple factors (**Figure 1**) energy-intensive construction processes; the extraction, processing, and transportation of building materials; and the ongoing energy consumption during a building's operational phase. Traditional construction methods, often reliant on non-renewable resources, coupled with inefficient operational systems in older buildings, have exacerbated energy consumption patterns. As a result, we find a scenario where the growth in urban landscapes, essential for socio-economic reasons, seemingly works counterintuitively to global sustainability goals.

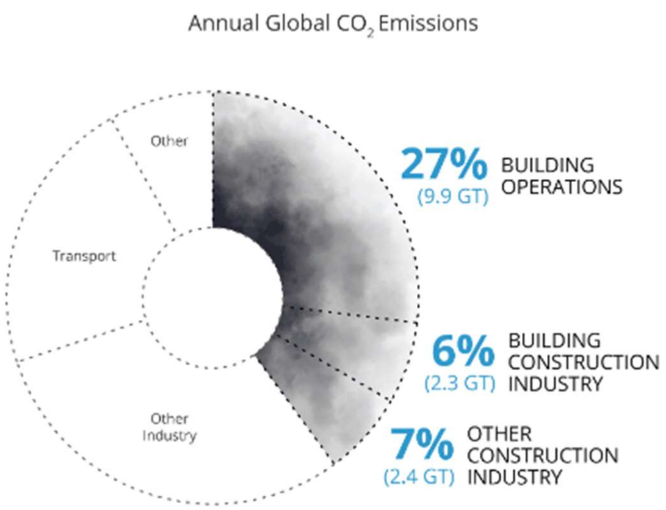


Figure 1. Building Construction Industry and Other Construction Industry represent emissions from concrete, steel and aluminium for buildings and infrastructure respectively. Architecture 2030. All Rights Reserved. Data Source: IEA (2022), Buildings, IEA

However, as concerns about the climate crisis intensify, there's been a concerted push towards reimagining the built environment. Central to this transformation is the idea of sustainable construction and operations. The overarching objective is to minimize the environmental footprint at every stage of a building's lifecycle. This approach calls for a harmonious amalgamation of advanced technologies, innovative practices, and an acute awareness of environmental needs.

As stated by the C40 network, a network of mayors from almost 100 leading global cities collaborating to undertake necessary actions to address the climate crisis, one of the main accelerators of the transformation path of cities towards greater resilience and better adaptive and mitigation capacities is "clean construction"[2]. This refers to the evolution of the global construction industry towards sustainability, efficiency, quality, and equity (Marco Casini, 2022). In fact, construction is one of the sectors most responsible for the global climate crisis, as it contributes to over 20% of global greenhouse gas emissions **Figure 2**. Construction materials and the building sector are also responsible for over 30% of global resource consumption. The impact of the construction industry is set to increase with the rising need for new buildings and infrastructures: by 2025 we will need to build 1 billion new homes. Globally, about 60% of buildings that will exist by 2050 have yet to be built. This means building a city the size of Stockholm or Milan (1.5 million people) every week until 2050 or a city the size of Singapore or New York every month until 2050 (C40 Cities, 2023).

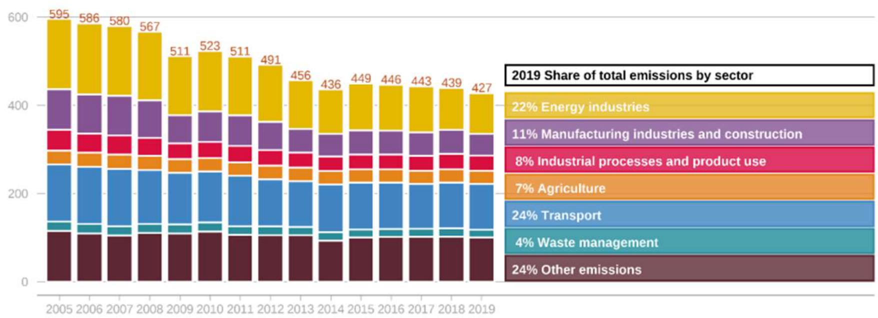


Figure 2. Total GHG emissions by sector (MtCO2e) (rounded data). Data source: EFA (GHG trends, GHG estimates, UNFCCC reporting)

The construction sector is also a key component of economic and employment growth in Europe, contributing significantly to the GDP and employing over 11 million people (European Commission. Joint Research Centre. Institute for Energy and Transport., 2015). In Europe, the challenge is focused on the rehabilitation and management of the existing building heritage, rather than on

2 The "clean construction" principle aims to prioritize the renovation of existing buildings and ensure that new buildings and infrastructures incorporate the principles of the circular economy, starting from the design phase, moving through the choice of materials, the construction phase, and up to disposal/reuse. Moreover, it promotes social equity, the reduction of air pollution, the creation of jobs by investing in sustainable local businesses and educating and retraining workers. Source: https://www.c40.org/.

the construction of new buildings and infrastructure. Buildings are long-term assets expected to remain useful for 50 or more years, and it is estimated that 75-90% of today's existing buildings will remain in use until 2050. With a low demolition rate (0.1% per year), a low renovation rate (1.2% per year), and a shift to new, highly energy-efficient constructions (1% additions per year), Europe's challenge mainly concerns energy-efficient renovations and investments in the existing building stock (Buildings Performance Institute Europe, 2016)

The European Industrial Strategy[3] highlights the need for a greener, more digital, and resilient building ecosystem. Several other European initiatives underline the role of construction in achieving goals such as sectoral renewal, process and economic circularity, adaptation to climate change and mitigation of its effects, employment. Among these are the Renovation Wave, EU Climate Adaptation Strategy, Zero Pollution Action Plan, Bioeconomy Strategy, European Skills Agenda for sustainable competitiveness, social fairness, and resilience.

The digitization of the construction sector is identified as a key element to enable development that contributes to the aspects listed above. According to the Industry 5.0 approach, digital technologies provide a new paradigm for production in general, and construction in particular (European Commission. Directorate General for Research and Innovation., 2020). It is based on four fundamental components: People, Collaboration, Sustainability, and Technology. This approach offers a vision of the industry that goes beyond efficiency and productivity as the only objectives and reinforces the role and contribution of the industry to society (European Commission. Directorate General for Research and Innovation, 2021). In this context, the potential and transformative implications of digitization for the construction ecosystem are significant:

- Digitization bridges the different scales of the built environment: starting from data on individual components, moving to the building scale, information can then be transmitted and incorporated into building production processes up to urban planning;
- Digitization bridges different professionals, as well as between workers and users: design work and construction sites can become more collaborative, while the management of a building, asset, neighborhood, or city can be optimized;
- Digitization can help create trust and transparency and improve decision-making in building construction processes and urban planning.

The deployment of these potentials will be facilitated by the full diffusion of Building Information Modeling (BIM) and tools for collaborative data-driven

[3]https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-industrial-strategy_en. European Industrial Strategy.

design, and, in the longer term, by Digital Twins (DTs), automation and robotics, and, more generally, by data-driven tools and services that exploit the latest advances in the Internet of Things (IoT) and sensors, cloud computing, massive Big Data processing, and Artificial Intelligence (AI).

In recent years, European and national initiatives have been launched, reference frameworks drafted, and roadmaps drawn up to support the digitization of the construction ecosystem. In addition, existing financing instruments have been strengthened and new ones created. However, construction remains one of the least digitized sectors in the EU, and the application of digital technologies follows a traditional value chain.

The analysis conducted by the Joint Research Centre of the European Commission confirms that architectural and engineering design activities for buildings and infrastructure increasingly use digital tools, while their adoption is still very limited in the construction and maintenance phase (European Commission, 2019).

Low digitalization rates are also found in the public sector, especially in administrative processes related to construction and medium and long-term strategic planning. For these, it is essential to rely on up-to-date databases, predictions, simulations, monitoring, and interactions with citizens and stakeholders based on transparency and clarity of communication. Therefore, in the context of the general need to digitize content, materials, and processes, DTs also represent a strategic technological innovation for cities (Federico Cinquepalmi, 2019).

Within the age of digitization and enabling technologies, which have now emerged as invaluable allies in the fight against carbon emissions in the built environment, digital tools and platforms, combined with innovative construction methodologies, offer the potential to drastically reduce energy consumption and resource wastage.

Building Information Modeling (BIM) stands out as a game-changing technology in this domain. BIM allows architects, engineers, and construction professionals to collaboratively design, visualize, and simulate a structure's performance in the digital realm long before its physical construction (Khan et al., 2021). Through such early-stage virtual analyses, inefficiencies can be identified, optimized designs can be conceptualized, and sustainability can be integrated as a core principle rather than an afterthought. Moreover, BIM's ability to integrate energy simulations ensures that energy performance is central to design considerations, leading to inherently more efficient structures (Murtagh et al., 2020).

Additionally, the deployment of sensors and Internet of Things (IoT) devices in modern structures ensures continuous monitoring of energy consumption, waste generation, and resource utilization (X. Li et al., 2022). This real-time data, when fed into advanced analytics platforms, can provide insights into areas of inefficiency, allowing for prompt remedial actions. It facilitates a feedback-driven approach to building operations, ensuring ongoing optimization and reducing the carbon footprint in the operational phase (S. Tang et al., 2019a).

The rise of Smart Buildings, empowered by Artificial Intelligence (AI) and Machine Learning (ML), heralds a future where structures are not just passive entities but are actively engaged in reducing their environmental impact (J. Zhang et al., 2015). They autonomously adjust heating, cooling, and lighting systems based on occupancy patterns and external environmental conditions, leading to considerable energy savings. Moreover, these buildings integrate renewable energy sources, like solar panels and wind turbines, into their energy matrices. Advanced algorithms ensure optimal utilization of renewables, reducing reliance on grid energy that often has a significant carbon footprint.

However, the digitization wave doesn't stop at individual structures. On a broader scale, the concepts of Smart Cities look to transform entire urban landscapes into hubs of sustainability (Deng et al., 2021). In this scenario, Digital Twins (DT) which are digital replicas of physical entities, are being created for entire cities (J. Zhang et al., 2015). These digital counterparts allow urban planners and policymakers to simulate various scenarios, analyzing potential strategies for reducing carbon emissions on a city-wide scale. It's an approach that integrates micro-level efficiencies of individual structures into a macro-level strategy for urban sustainability.

Beyond technological advancements, the shift towards sustainable built environments requires a change in mindset. Green certifications and sustainability benchmarks are driving this change, with numerous structures now aiming for certifications like LEED (Leadership in Energy and Environmental Design) or the BREEAM (Building Research Establishment Environmental Assessment Method) (Gauthier & Wooldridge, 2012). Such certifications provide tangible goals for builders, designers, and owners, ensuring that sustainability is ingrained in modern structures.

1.2.1 Key enablers and benefits of digital transformation

The ongoing digital revolution, frequently named the Fourth Industrial Revolution (Industry 4.0), is being propelled by the rapid growth of computing capacities and the influx of data (Reischauer, 2018). Digital technology integration into various sectors has been a topic of discussion in numerous industry studies for a while. Yet, the Mckinsey Global Institute recently highlighted an intensified acceleration in digitization, driven in part by the global pandemic[4]. This acceleration has seen revenues double from prior projections made before COVID-19. Mobile and cloud technologies have seen a massive upswing in their adoption, a trend that is projected to continue as economies go towards recovery. In this scenario, cyber-physical systems offer a promising avenue, with their potential to save time and costs, paving the way for a more resilient economic built environment (Sepasgozar, 2020).

However, the integration of Industry 4.0, amplified remote work capabilities, enhanced supply chain integration, and increased transparency in business processes and might be the catalysts that push resilient companies to thrive in a post-pandemic environment (Rani et al., 2022). Pioneering business paradigms are centering on digitalization, which spans across products, processes, and resources. Taking the construction sector as a case in point: there is a growing inclination towards the Digital Twin (DT), seen as a refined progression from Building Information Modeling (BIM), serving as a catalyst for industry metamorphosis. This shift towards digital solutions is projected to bolster the architecture, engineering, and construction (AEC) market's valuation from $7.188 million in 2020 to an impressive $15.842 million by 2028, with an annual growth rate of 10.7% from 2021 to 2028 (Allied Market Research., 2021).

The profound shift towards digital innovations provides a glimmer of optimism for industries like construction, suggesting that weathering the COVID-19 storm might result in amplified productivity in forthcoming years. For industries across the board, digital mechanisms are vital, from health monitoring and contact tracking to ensuring on-site safety. Concurrently, cloud technologies and mobile platforms foster remote collaborations among designers and builders and allow for swift recalibrations in supply chains.

Among the new technologies, Digital Twins (DT) are emerging as strategic tools to guide and inform data-driven decision-making processes. They enable the utilization of the vast amount of machine-readable data produced by the physical city, monitoring trends and consumption, detecting dynamics, and simulating future scenarios. At the heart of recent debates and scientific research, the concept of DT has been discussed in literature rapidly increasing in both the number of contributions and the breadth of related topics. Saeed et al. summarize the potential of DTs "in enabling a new perspective for understanding, interacting with, and responding to our living system. This opens novel dimensions in multidisciplinary research areas that can redefine the way we perceive and interact with the city. These new methods of planning, management, and operation have the potential to enhance the living experience,

[4] McKinsey & Company. How COVID 19 Has Pushed Companies over the Technology Tipping Point Final.pdf. 2020. Available online: https://www.mckinsey.com

efficiency, and performance of the city, its urban realm, and the built environment in general"(Saeed et al., 2022a).

Recently, pilot projects, proofs-of-concept, and applications of DT in various areas have proliferated, and several research groups are working to highlight, categorize, and group the various use cases (White et al., 2021). Thanks to these studies and prototypes, the maturity level of DT is gradually evolving, demonstrating its potential and helping to gain the consensus of key stakeholders and decision-makers in the field of urban policies, planning, and the construction sector. Simultaneously, scientific research and practical experiments are also highlighting limits, risks, and open issues related to the enabling technologies of DT and the production/management of the data needed to power such technologies.

Traditionally, sectors such as consumer goods manufacturing have pioneered the path to digitization, often outpacing the construction industry. Now, leading digital advancements, namely Building Information Modeling (BIM) and Digital Twins (DT), are steering the construction domain towards a new era of data interoperability. While BIM serves as the cornerstone, encapsulating both geometric data and non-geometric attributes vital to construction processes (Sacks et al., 2018), DT extends this capability by integrating BIM's static blueprints with dynamic real-time data, offering a multifaceted view of the built environment, allowing simulations and insights (Schleich et al., 2017). BIM's pivotal influence in the construction realm is underscored by its adaptability, with applications spanning the entire building lifecycle — from the nascent design phase to the more mature operations and maintenance stages.

Building Information Modeling (BIM) serves as an innovative and cohesive platform that archives comprehensive building data. This documentation is pivotal in bolstering the planning, construction, and subsequent maintenance spanning a facility's lifespan (Volk et al., 2014). Intricately designed, BIM goes beyond the traditional confines of 3D computer-aided design (CAD) by assimilating supplementary data encompassing building specifications, chronological scheduling, financial approximations, and maintenance oversight, often referred to as 4D, 5D, and 6D dimensions respectively (Wong & Fan, 2013). The foundational objective behind is to mitigate unnecessary expenditures by precluding discrepancies and miscalculations during the design and construction phases.

Presently, the utility of BIM spans across diverse sectors, prominently in architecture, construction, engineering, and facility management (AEC/FM). It plays a crucial role in tasks such as design visualization, ensuring design coherency, detecting inconsistencies or clashes, adhering to lean construction principles, accurate cost and time prognostications, and in fostering heightened collaboration and transparency amongst stakeholders (Volk et al., 2014).

Contemporary endeavors are channeling efforts to augment BIM's capabilities by integrating real-time data streams from sensors and other IoT devices (S. Tang et al., 2019b).

Moreover, the rapid proliferation of the Internet of Things (IoT) has been significantly influenced by Moore's law, which postulated that the number of transistors on semiconductor chips would double approximately every two years, leading to enhanced computational capabilities at diminished costs (Chui et al., 2010)This evolution of IoT stands poised to reshape multifarious facets of the global economy. This transformation is evident in myriad developments ranging from connected and self-driven vehicles (E.-K. Lee et al., 2016), aerial drones or 'flying robots' (Hossein Motlagh et al., 2016), to smart home ecosystems (M. Wang et al., 2013). Fueling this transition is a symbiotic partnership between prominent appliance manufacturers and tech giants like Amazon, Google, and Microsoft.

Wireless Sensor Networks (WSN) and data analytics converge to form the foundational pillars of the digital twin creation (Tao, Cheng, et al., 2018). When envisioning a building’s digital twin, one might either extrapolate a 3D CAD model from the existing BIM or fabricate a distinct 3D representation of the structure in question. By weaving together an array of sensor networks, the digital twin offers a vivid, real-time portrayal of the asset. This dynamic perspective paves the way for instantaneous analytics, insightful decision-making, and the enhancement of building efficiency and occupant comfort.

Traditional BIM architectures do not inherently accommodate real-time data and predominantly serve design, construction, and maintenance operations, areas which do not generally demand real-time insights (Bruno et al., 2018). On the other hand, the digital twin embodies the real-time essence of a tangible asset. Its modus operandi centers around real-time data, predominantly fed by sensor systems. This data serves to chronicle and monitoring the instantaneous structural and environmental of an asset. Such a mechanism is pivotal for rendering accurate digital twin simulations and insightful data analytics (Qi & Tao, 2018a)

Over the years, the Digital Twin (DT) paradigm has been embraced across various industries as a strategy to mitigate asset risks, enhance traceability, and optimize maintenance, ultimately extending the asset's lifecycle (Hlady et al., 2018). DTs can be pivotal in numerous asset-related areas, from singular asset performance to intricate systems like manufacturing processes, wherein multiple components interact uniquely (Shubenkova et al., 2018).

For instance, a survey, as illustrated in **Figure 3**, highlights that the manufacturing sector is at the forefront in DT research and implementation.

I. Errandonea, S. Beltrán and S. Arrizabalaga / Computers in Industry 123 (2020) 103316

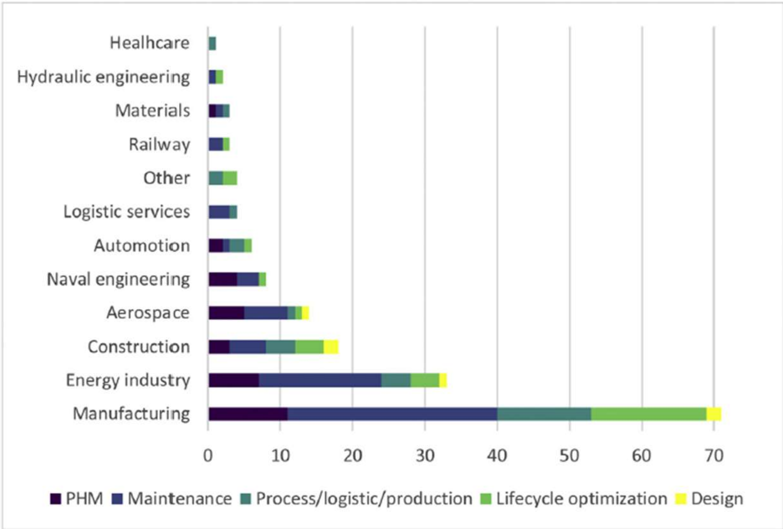


Figure 3. Digital Twins in industries

Following closely are sectors related to gas and oil extraction, wind turbine management, typically associated with offshore infrastructures. The construction and aeronautical industries are also significant adopters of DT. It's noteworthy that within this context, diverse case studies, be it buildings or bridges, are all categorized under construction. These industries often have a robust R&D investment tradition, possibly driven by stringent safety mandates, which necessitate considerable expenditure on maintenance activities.

This paradigm shift in embracing the digital era holds transformative implications for the Built Environment (BE). As we move forward, a vision of BE surrounded by digital technologies and propelled by data is not just an aspiration but a real objective. This transformation is underpinned not only by a drive to elevate the digital and physical scaffolds of BE but also by the need to contain the ecological footprint.

Taking the UK as a case in point, the nation's strategic blueprint is veering towards rejuvenating pivotal sectors by tapping into the untapped reservoirs of digital innovation and the surging streams of big data. This seismic shift isn't just a superficial overlay; it aims to enrich the very core, augmenting the value extracted from myriad services ingrained within BE.

In this scenario, schools and hospitals, not just brick-and-mortar establishments, are interconnected entities, responsive to real-time needs and adaptable to changing scenarios, as well as transportation networks, not merely routes connecting points A to B, but also dynamic systems, optimizing flows based on real-time data, ensuring efficiency and sustainability. Energy systems are not rigid frameworks but flexible matrices, responsive to fluctuating demands and adapting seamlessly to offer optimal output. This is the promise the digital era holds for the BE, where innovation doesn't just streamline processes but fundamentally reshapes the very way we conceive and interact with our environment.

The digital transformation of the built environment (BE), when synergized with a unified management strategy for the cyber and physical realms, promises to offer several benefits to society. According to demographers' predictions, the world's population is set to rise to about 9.8 billion by mid-century (European Commission. Competence Centre on Foresigh, 2022). Nearly 70 percent of this population increase is expected to be concentrated in urban centers. This change will bring complex challenges, including increasing energy consumption and intensifying vehicular traffic jams (United Nations. Department of Economic and Social Affairs, 2018).

However, the contrast between the size of the construction industry and its digital competence reveals a paradox. Despite its colossal scale, the sector's digital advances appear nascent, especially when compared with those of other industries. This digital chasm is not only emblematic of the heterogeneous technological terrains that characterize regions like Europe. It underscores the innate complexities and resistance to transformation inherent in the construction ecosystem.

The fusion of academic research and real-world observations reveals a growing awareness of the indispensability of digital integration and a comprehensive elevation that strengthens operational efficiency, supports environmental protection, and reinforces safety protocols. This vision recognizes the profound ripple effects that seemingly minor changes in industrial practices can propagate in global economic landscapes and environmental paradigms.

Table 1. Digital technologies and related construction activities. Source: Digitalising the Construction Sector. Jan 2019 Committe for European Construction Equipment

Task	Enabling technology	Result
Intelligent design and planning	BIM, Significant Data	Reduction in design errors and enhancement in the quality of the design and engineering processes through virtual and digital simulation (i.e. digital twin).
Fleet management	Internet of Things, Big Data	Remote monitoring through the use of sensors and analytics of status and location of construction equipment (i.e. work vehicles) so as to reduce costs, improve

		energy efficiency and limit idle time of machines.
Predictive maintenance	Internet of Things, Big Data, AI	Monitoring of the condition of machines through the use of sensors and analytics with the aim of conducting preventive maintenance and to reduce potential failures.
Innovative fabrication methods	Internet of Things, Big Data, AI, BIM, Drones	Rethinking construction processes and operations in a “smart perspective” leveraging the information gathered through data collection and analytics to speed up processes, reduce costs, improve energy efficiency and operators’ safety.
Monitoring and Evaluation of Resilience	BIM, Internet of Things, Big Data, AI, Drones	Better “real-time” surveying in all of the phases of the project such as providing adequate support to on-site operators as well as monitoring resilience of infrastructure following the end of the project.
Autonomous equipment and driverless vehicles	Internet of Things, Big Data, Robotics	Introduction of driver-assistance systems and autonomous driving to improve construction processes and reduce physical workload on the construction site.

Despite its cutting-edge nature, the Digital Twin (DT) model has its limitations. While it offers insights from extensive data, human intervention is still crucial in final decisions. There are challenges, like software issues that disrupt data transfer or problems syncing real-time data. The ultimate goal for DT is to perfectly blend the digital with the physical, creating a space where virtual and real worlds inform each other.

Efforts are underway to integrate DT into the future plans of cities. The development of DT-focused cities showcases this progress. Though they might not solve all urban problems, the benefits they offer across various sectors are undeniable. These benefits stem from combining Virtual Reality (VR), the

Internet of Things (IoT), and detailed 3D modeling, all enhancing the effectiveness of smart cities.

Digital Twins are now a growing feature in city planning worldwide. By 2020, an impressive 118 cities had incorporated DT into their urban plans, leading to initiatives like reducing carbon footprints and better traffic management. The potential of DT for cities is immense, and everyone eagerly awaits their broader impact (Lehtola et al., 2022).

Yet, developing a smart city is complex. Balancing people's changing expectations, the appeal of new technologies for better governance, and the need for proactive solutions is challenging. That's where DT shines as they're great at spotting potential issues and preventing them. Comparing this foresight with actual results turns data into valuable, actionable insights.

1.2.2 The journey from green production to advanced distribution

In Italy the National Recovery and Resilience Plan (NRRP)[5] encouraged esteemed research entities, chiefly ENEA, the Italian National Agency for New Technologies, Energy, and Sustainable Economic Development, the primary institution for research, technological innovation, and advanced services in the fields of energy, environment, and sustainable economic growth in taking the lead in championing advancements in the realm of renewable energy. Their efforts span across the spectrum of thermal solar to energy-efficient technologies like photovoltaic, bioenergy, and bio-refineries tailored for energy generation. There's also an emphasis on constructing smart cities, ensuring judicious energy consumption, pioneering smart grid systems, propelling sustainable transit mechanisms, and formulating cutting-edge thermal cycles and fuel cells (Ministry of Economic Development, 2019). The overarching theme remains the integration of robotics, ICT, and pivotal digital innovations.

Central to the NRRP is the emphasis on pioneering computational infrastructure and advanced data transmission frameworks. Such strides will supercharge research hubs, propelling them to lead in realms like cloud-based computing, model-driven solutions, and applications tethered to the vast

[5] The PNRR, Piano Nazionale di Ripresa e Resilienza, translates to the National Recovery and Resilience Plan in English (NRRP). Introduced in response to the economic and social challenges posed by the COVID-19 pandemic, this strategic initiative aligns with the broader European Recovery and Resilience Facility. The primary objective of the Plan is to stimulate Italy's economic recovery, focusing on several pivotal areas: digitalization, innovation, competitiveness, cohesion, education, and health.
The plan aims to modernize Italy's infrastructure, promote green initiatives, and ensure a more resilient economy against future shocks. It presents an opportunity not only to repair the damages caused by the pandemic but to reshape Italy's economy in a sustainable, inclusive, and forward-looking manner.

universe of Big Data and the Internet of Things (IoT). These are seen as essential tools in addressing multifaceted energy challenges.

Aligned with EU directives and Italy's long-term blueprint for curbing greenhouse emissions, the aspiration is to drastically slash greenhouse gas emissions, anywhere between 80% to 100% by mid-century using 1990 as the baseline. This audacious goal leans heavily on an energy composition dominated by renewables, a sharp curtailing of energy appetite, wide-scale electrification, and a significant boost in hydrogen production. The plan also underscores the criticality of digital tech, energy storage matrices, intelligent grids, and energy conservation across varied sectors.

A transformative wave is being observed in power generation facilities, largely driven by an amalgamation of sensor technology and pioneering software applications. These tools are groundbreaking in their ability to detect anomalies and flag potential threats, setting the stage for predictive interventions. The very essence of constant surveillance, enabled by these sensors, ensures real-time detection of performance gaps, leading to enhanced operational efficiency.

Machine learning and artificial intelligence (AI) algorithms are integrated with voluminous data harvested not just from isolated power plants but from the entirety of a manufacturer's portfolio. Over time, this layer of automation empowers plants to autonomously oversee and refine their operational metrics.

This digital metamorphosis isn’t confined to just electricity creation; it also reshapes distribution and commercial paradigms. It paves the way for consumers to demystify their consumption economics and equips entities like ESCOs and aggregators to curate bespoke offerings by leveraging data platforms.

The surging digital integration in metering and home automation equips enterprises to harness data for granular customer insights. Yet, it's a double-edged sword. While companies can tap into vast data pools for intricate customer profiling, consumers, armed with granular consumption insights, can engage more proactively in the marketplace.

However, the digital footprint of energy ecosystems isn't without its challenges. Cybersecurity looms large as a significant concern, with the potential to compromise individual privacy. Notably, the International Energy Agency (IEA) has embarked on a mission to digitalize energy infrastructure, with the dual objectives of fortifying resilience and championing decarbonization (International Energy Agency, 2022).

Companies are harnessing digital levers to optimize operations and cultivate a more interconnected energy ecosystem. A testament to this is the skyrocketing investment in digital tech within the electric sector, clocking in at an astounding $47 billion in 2016. By 2040, the anticipated annual savings for stakeholders are projected to be $20 billion. Ultimately, the goal is to harmoniously integrate renewable sources, ensuring networks can seamlessly calibrate their energy offerings to fluctuating demands.

The energy paradigm shift is not just about green energy production. It's a holistic transformation anchored in digitalization that encompasses all stakeholders, from energy magnates to the end consumer. The European Commission, in its proactive stance, has wrapped up a public consultation phase, laying the groundwork for a strategic digitalization action plan for the energy domain. This blueprint is geared towards ensuring that digital advancements play a pivotal role in the Green Deal and fostering a unified energy marketplace.

However, a collective strategy is needed. The digital energy marketplace must be transparent and competitive, while steadfastly adhering to core tenets:

- Ethical governance
- Robust data protection
- Unwavering commitment to privacy and cybersecurity

Intricacies unique to the energy domain cannot be overlooked. The European Commission, in its initial phases, had circulated a roadmap delineating the challenges and outlining the objectives.

The European Union's emphasis is clear in the digital energy action plan. As the EU Executive articulates, a reassessment is pivotal to ascertain if current mechanisms empower citizens to navigate a digitalized energy marketplace efficiently.

The surge in digitalization has, unfortunately, left the energy ecosystem vulnerable to cyber threats, jeopardizing supply security. The entire spectrum of the energy supply chain, right from power plant management to consumer-centric services and intelligent grid systems, is undergoing a digital overhaul.

In essence, the energy renaissance pivots on three cardinal pillars:

- Greening the electricity generation matrix,
- Innovating novel storage modalities, such as hydrogen,
- Infusing digital elements to revolutionize energy production, distribution, and consumption paradigms.

Digitalization of energy in power plants has been made possible through the integration of sensors and cutting-edge software. This combination aids in detecting anomalies and pinpointing potential risks. As a result, proactive measures can be taken before any damage transpires. This strategy is known as predictive maintenance and is most effective when implemented at times that don't disrupt regular production activities. Furthermore, the continuous monitoring enabled by these sensors can promptly highlight inefficiencies, thus enhancing the plants' overall performance and efficiency.

The most noticeable impact of digitalization is on the networks responsible for transporting and distributing generated electricity. In this realm, electronic meters, often termed "smart meters," serve as the cornerstone. These meters facilitate the creation of smart grids, which in turn ensure efficient management and equilibrium of the electrical system.

Such an approach becomes crucial when dealing with intermittent renewable energy sources like wind and solar. They can be seamlessly integrated into the network using this system. Historically, the energy distribution model was predominantly top-down, where energy moved in one direction: from the producer straight to the consumer. However, recent trends indicate a paradigm shift toward distributed generation. This trend allows a growing contingent of minor producers and consumers to contribute electricity back into the network.

Data from ENEL (Italian National Board for Electric Energy) suggests that approximately 17% of future investments in networks will prioritize innovations in transmission (i.e., by the Transmission System Operator or TSO) and, more crucially, the distribution (by the Distribution System Operator or DSO) of renewable energy sources.

As such, ICT tools employed for these endeavors embrace a data-centric methodology. They harness machine learning and artificial intelligence (AI), amalgamated with extensive data sets derived not just from one unit but from every facility under the umbrella of a given producer. Over time, such automated learning systems will empower plants to autonomously monitor their performance metrics and make necessary adjustments.

1.2.3 Digital Distributed Energy Resources (DER) Management

As the shift towards sustainable energy intensifies, distributed energy resources (DER) are expanding in scale, variety, and intricacy across the globe. Today's DER-oriented power infrastructure represents a state-of-the-art entity, encompassing a range of equipment with modifiable consumption, localized generation units, and energy storage mechanisms. While DER boasts numerous advantages, they also pose challenges concerning system consistency, dependability, adherence, cybersecurity, among others. For instance, variable outputs from renewable sources, triggered by meteorological shifts, can result in intermittent voltage spikes in the distribution grid and disruptions in relay protection mechanisms; additionally, wind energy units might instigate persistent fluctuations in the electrical grid (Ren et al., 2023). The automation of DER management necessitates the extensive incorporation of Internet of Things (IoT) instruments, whose inbuilt software is particularly vulnerable to cyber intrusions (Sharma et al., 2020).

The intricacy of adequately delineating a power system, particularly in the context of DER, is underscored by the necessity of a great array of heterogeneous models, potentially imposing hefty costs. The pivotal challenge lies in enabling the Digital Twin (DT) to swiftly and accurately mirror intricate, prolonged power distribution processes, which are dispersed over vast regions and influenced by an intricate web of latent technological, socio-cultural, and economic parameters. The models that structure a DT manifest variation both on a horizontal spectrum, referring to the specific components they characterize, and vertically, relating to the particular concerns or perspectives they epitomize.

To address this diversity, systems engineering offers an approach via viewpoint-centric DT construction, a technique deeply rooted in the guidelines of the ISO/IEC/IEEE 42010 standard (Gharaei et al., 2020). This fosters a synchronized, clear understanding of the system. Moreover, to reinforce the base semantics and sidestep potential variances in the theoretical comprehension within the DT, incorporating a sturdy semantic structure is crucial, as exemplified by the distributed energy ontology framework (Andryushkevich et al., 2019)

Delving into the sphere of standardization and aiming for automated model integration, cutting-edge mathematical methodologies based on category theory emerge as an apt solution. Within this analytical paradigm, research efforts are directed towards carving a trajectory for the development of a DT-focused DER control platform that resonates coherently with the previously mentioned benchmarks.

1.3 Research framework

Over the past few years, Digital Twins (DT) have firmly established themselves as pivotal tools across a wide spectrum of industries. In the automotive sector, they provide intricate simulations of vehicular operations, allowing for advanced diagnostics and performance enhancements. Manufacturing industries utilize DTs to replicate factory processes, optimizing production lines and predicting maintenance needs, ultimately leading to decreased downtimes and increased output. The aerospace domain leverages DT for simulating flight conditions, ensuring design robustness, and forecasting maintenance, a critical component for safety and efficiency. Moreover, in the healthcare sector, DTs have been transformative. They've enabled the creation of patient-specific digital replicas which aid in tailoring personalized treatments and ensuring the precision of diagnostics. Through these applications, it's evident that the integration of DT has resulted in enhanced operational efficiency, risk mitigation, and improved decision-making across varied sectors.

However, the integration of DT in the AECO domain holds vast potential for the energy sector, enhancing operations through precise data modeling and real-time monitoring.

The potential of Digital Twins expands even further when introduced to the built environment, particularly in the realm of energy management. Within the AECO sector, DT can model and predict how energy is consumed, stored, and distributed in real-time. This data-driven approach not only assists in optimizing existing infrastructures but also provides invaluable insights during the design phase of buildings, ensuring energy-efficient architectures. As buildings and structures become more complex, integrating renewable energy sources and adaptive systems, the role of DT becomes even more pronounced. By modeling how different energy systems interact within the built environment, stakeholders can make informed decisions, ensuring sustainability and efficiency while reducing operational costs.

At various scales of the energy ecosystem (**Figure 4**):

- *Bulk Generation* sees the harnessing of large-scale energy sources, where DT can model and predict optimal output levels and maintenance schedules.
- *Transmission Utility Operations* can benefit from DT's ability to forecast transmission bottlenecks and optimize grid flow.
- *Distribution Utility Operations* leverage DT for efficient power distribution, demand forecasting, and grid stability.
- *Behind the Meter* involves the end user's side, focusing on consumption optimization. This encompasses technologies such as:
  - *PV (Photovoltaic) Systems* which can be modeled for peak solar capture times.
  - *Wind Turbines* where DT can simulate optimal blade configurations.
  - *EV (Electric Vehicle) Charging Stations*, predicting and managing peak charging times.
  - *Batteries* for optimizing charge and discharge cycles.
  - *Heat Pumps* and Smart Thermostats for ensuring optimal heating or cooling with minimal energy wastage.
  - *Microgrids* which aggregate and manage these diverse energy sources.

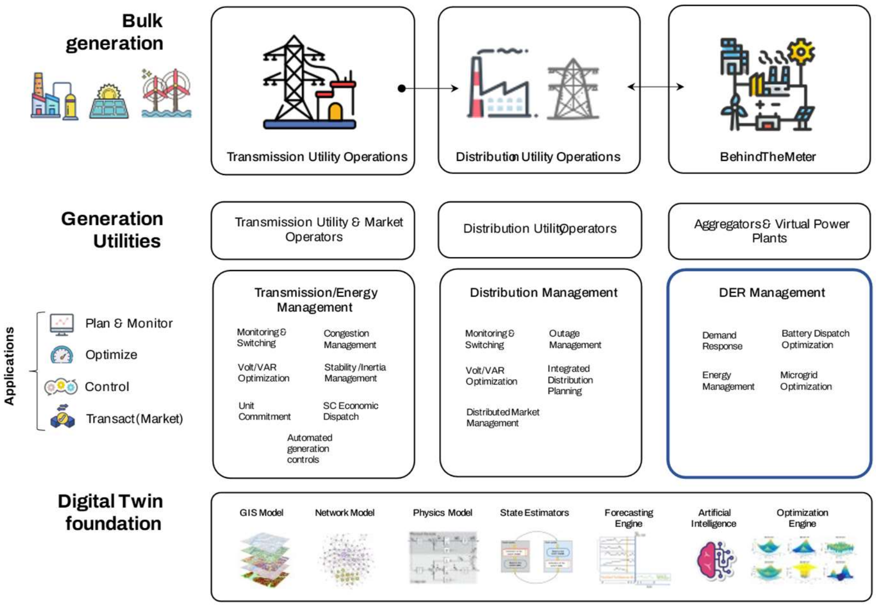


Figure 4. Digital Twins enabling Electric Grid Reliability, Resiliency and Affordability

Microgrids, with their inherent ability to operate autonomously from the main grid, stand out as a quintessential application for DT, especially in the *"Behind the Meter"* context. This term essentially refers to all energy systems and processes that end-users have direct control over, from energy generation, storage, and consumption. Digital Twins can replicate these intricate microgrid systems, providing real-time data on energy usage patterns, storage capacities, and potential grid disruptions. Such a granular level of detail allows for advanced forecasting, ensuring that during peak consumption times or grid disturbances, energy can be efficiently managed, and outages are minimized. Furthermore, DT can simulate different scenarios, aiding in strategic planning and ensuring optimal energy distribution based on varied inputs like weather forecasts, peak usage times, or even unexpected outages.

It is at the microgrid scale that the present research particularly sharpens its focus. The beauty of microgrids lies in their ability to operate autonomously, interfacing seamlessly with the main grid or functioning in isolation during disturbances.

Distributed Energy Resources (DER) Management is at the heart of modern energy systems, especially in the context of microgrids. With an array of energy sources, from photovoltaic systems and wind turbines to batteries and heat pumps, the orchestration of these resources is paramount. Digital Twins offer a holistic view of these diverse energy sources, providing a platform where each can be managed in synchrony with the others. This encompasses facilitating effective Demand Response, where energy consumption can be adapted in real-time based on market conditions, and Battery Dispatch Optimization, ensuring that energy storage systems are used efficiently. Beyond these, DTs aid in a comprehensive approach to Energy Management, balancing energy consumption, storage, and generation, and ensuring that the microgrid operates at peak efficiency. The integration of DT in DER Management represents the future of energy, where data-driven insights guide actions, leading to a sustainable and efficient energy landscape.

Through this research framework, the aim is to unearth the vast potentials of Digital Twins in transforming how energy is managed and optimized in the built environment.

This thesis is structured as follows:

Chapter 2 – Background and *Literature review:* this chapter delves deep into existing literature, offering a comprehensive review of previous research and studies related to the Digital Twin (DT) paradigm and its applications across various industries and particularly in the Built Environment frameworks. By examining established definitions, theories, practices, and findings, the chapter aims to provide a solid foundation for the subsequent sections.

Chapter 3 – Methodology and Use Case Scenario (PoC): in this chapter, the research methods employed for the thesis are detailed. It outlines the research design, data collection techniques, and analysis methods. The chapter also presents the architectural design of the framework being proposed for DT-based Microgrid Management System. The proposed system is capable of recognize and manage the main loads (HVAC for air quality control (IAQ), electrical systems, automotive, etc.) in real-time, balancing them with production systems from renewable sources and energy from the national electrical grid.

The Use Case Scenario specifically focuses on how the Digital Twin (DT) paradigm can be integrated into a Distributed Energy Resource (DER) Management System including Indoor Air Quality (IAQ) control system as a Proof of Concept (PoC). The chapter elucidates the challenges encountered, the solutions employed, and the results derived from the application of DT in the Intelligent Energy Management space.

Chapter 4 - Discussions: the final chapter synthesizes and reviews the main insights derived from the literature, the efficacy of the methodology and framework proposed, and the implications of the use case scenario towards smart district and energy communities. Additionally, it provides final remarks and recommendations for future research, potential applications in Vehicle-to-Grid (V2G) scenarios, and improvements in the field.

This page intentionally left blank.

Chapter 2

Background and Literature Review

2.1 Digital Twin paradigm

The digital twin (DT) concept is defined differently depending on the domain context such as manufacturing, automotive, aerospace etc. and it has been recently explored and introduced in the built environment sector (McKinsey Global Institute, 2022) As the DT concept rapidly gained attention and widespread interest, it became a research area for the AECO sector and contributions in literature increased over time after 2016 towards an effective conceptualization. In this regard, in the UK the National Infrastructure Commission (NIC) issued the Data for the Public Good report in 2017, and the Centre for Digital Built Britain (CDBB) introduced The Gemini Principles report in 2018, in order to define policy standards and frameworks for the implementation of digital technologies as an interconnection between society and the built environment (Batty, 2018).

Moreover, the DT concept is also closely related to the domain of smart cities and buildings as it enables predictive insights into a city management approach based on digitalization (Mohammadi & Taylor, 2017), providing big data management capabilities in urban spaces (Oliver, 2018).

2.1.1 Definitions and key enablers

Back in 2002, the concept of the Digital Twin model emerged in the world of Product Lifecycle Management (PLM) without an official name (M. W. Grieves, 2005). Shortly after, it received a name, but it went through some changes. Initially, it was called the "Mirrored Spaces Model" (MSM) in 2005 (M. W. Grieves, 2005). However, the name evolved into the "Information Mirroring Model" in 2006 (M. Grieves, 2006). Finally, in 2011, it became known as the "Digital Twin," a term coined by John Vickers of NASA (Grieves M, 2011). Throughout this evolution of names (**Figure 5**), it's important to note that the core concept and model remained consistent and unchanged.

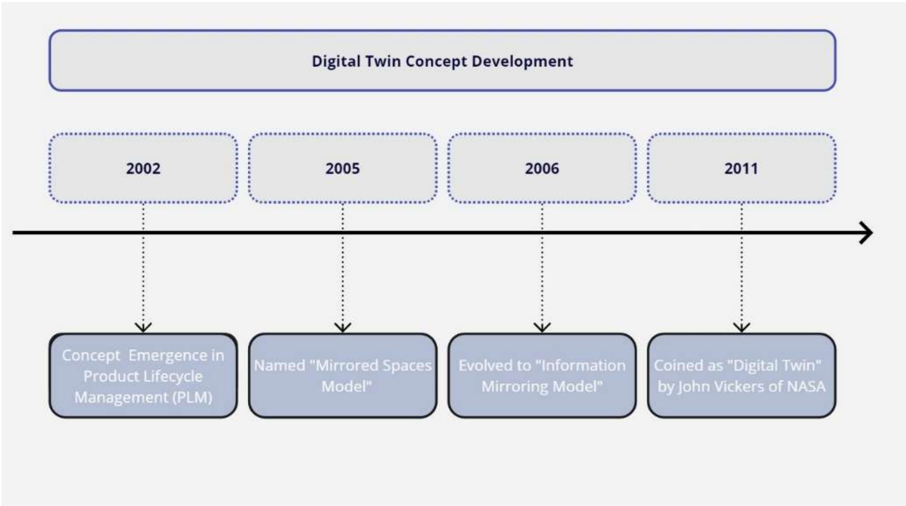


Figure 5. Evolution of Digital Twin Concept

Later in 2012 it was further evolved in the aerospace industry related to modelling and simulation as it was defined by Glaessgen and Stargel as “an integrated multiphysics, multiscale, probabilistic simulation of a vehicle or system that uses the best available physical models, sensor updates and fleet history, among others, to mirror the life of its flying twin. The digital twin is ultrarealistic and may consider one or more important and interdependent vehicle systems” ((Shafto M. et al., 2010); (Tuegel et al., 2011a)).

Examining the essence of a Digital Twin (DT) is a valuable endeavor. The Digital Twin Model, as depicted in **Figure 6** comprises three fundamental components:

- The "Physical Twin" on the left side represents an existing or intended physical entity in the real world.
- The "Digital Twin" on the right side is its virtual or digital counterpart residing in the digital realm.

- Facilitating the seamless exchange of data and information between these two counterparts is the "Digital Thread".

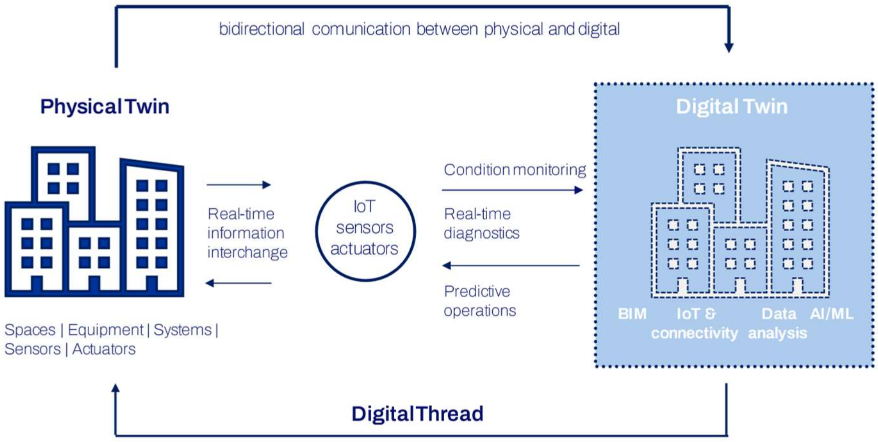


Figure 6. Digital Twin concept's fundamental components

The concept of the Digital Twin came into being roughly twenty years ago and has since been a recurring theme in the literature. It's important to note that there's not universally accepted and standardized definition at present. The definition tends to vary based on factors such as the specific application area of the Digital Twin, the discipline involved, the scale of the physical element being replicated, and other contextual variables.

Regarding this conceptual ambiguity, several interdisciplinary and multidisciplinary investigations have recently commenced taxonomic analyses tracing the gradual development of the Digital Twin concept. Some of these studies are referenced to trace the historical trajectory of the concept.

DTs are widely used by NASA for spacecraft, as well as the U.S. Air Force uses it for jet fighters (Tuegel et al., 2011a). It's also being considered for aircraft health in general. In addition, it's being explored for use in IoT deployment and factory production.
Even the oil industry is looking at using Digital Twins for ocean-based production platforms. And in the field of medicine, there's talk of using Digital Twins of humans to improve patient health.

Moreover, many software providers that deal with product development and product lifecycles are adopting the Digital Twin terminology. Big players like Dassault Systems, PTC, Siemens, and General Electric have all embraced the concept and are actively using it in various ways (Ghenai et al., 2022).
In Grieves' own words, he affirms that despite the evolving terminology, the fundamental concept of the Digital Twin model has remained remarkably consistent since its inception in 2002. The core idea revolves around the creation

of a digital information structure representing a physical system as an independent entity. This digital information serves as a 'twin' of the data inherent to the physical system, remaining linked with the physical system throughout its entire lifecycle (M. Grieves, 2016).

Many DT definitions are wide spreading over time and several research studies introduced the concept of cyber-physical connection (Haag & Anderl, 2018) (Tomko & Winter, 2019)) and real-world mirroring at different levels from products, assets, buildings (Buckman et al., 2014), districts to national ecosystems (Bolton et al., 2018).

While looking at specific definitions of the concept related to process management in the AECO sector, many conceptualizations from different points of view can be observed. Control and monitoring capabilities combined with intelligence skills are some of the main goals for digital twins in the built environment throughout its lifecycle.

The Digital Twin Consortium defines it as “a virtual representation of real-world entities and processes, synchronized at a specified frequency and fidelity” introducing the synchronization concept as a fundamental element.

The Gemini Principles report provided by the Centre for Digital Built Britain defines the DT as a dynamic model of the physical asset, controlled and operated through real-time sensor data measuring performance. As a recently born concept, DT is often resulting in confusion about its intended uses, performances and outcomes, so identifying the main levels of DT in the built environment could be valuable in obtaining clearer definitions.

Digital twinning is generally conceived as virtually simulating the relevant behaviour of physical objects in real-world environments (Hochhalter J.D. et al., 2014) as the main components of a typical digital twin are the physical entities in the physical world, the digital models in the virtual world and the data that tie the two worlds together (Michael Grieves, 2014).

Moreover, many definitions emphasize the concept of synchronisation between the digital and virtual world, such as (Garetti et al., 2012) define DTs as “the virtual representation of a production system that is able to run on different simulation disciplines that is characterized by the synchronisation between the virtual and real system”.

Hereafter, a selection of 31 definitions from research literature are included in the following Table 2 and grouped according to six different key layers which specifically identify the main key points according to the field of interest.

Table 2. Digital Twin definitions and key layers

Key layers	Definition	Field	Ref.
Integration [K.1]	[D.1] Integrated multiphysics, multiscale, probabilistic simulation composed of physical product, virtual product, data, services and connections between them.	Aerospace	(Glaessgen & Stargel, 2012)
	[D.2] Ultrarealistic integrated multiphysics, multiscale, probabilistic simulation of a system.	Complex equipment	(Tao F, 2017)
	[D.3] Comprehensive physical and functional description of a component, product or system, together with all available operation data	Product lifecycle	(Boschert & Rosen, 2016a)
	[D.4] A means to link digital models and simulations with real-world data, creates new possibilities for improved creativity, competitive advantage and human-centred design.	Built environment	(Arup, 2019)
	[D.5] Big collection of digital artefacts that has a structure, all elements are connected; there exists metainformation as well as semantics.	Manufacturing	(Rosen et al., 2015)
Connection [K.2]	[D.6] New mechanisms to manage IoT devices and IoT systems-of-systems.	Industrial IoT	(Canedo, 2016)
	[D.7] One where the virtual object exchanges data flows with the physical one in both directions.	Manufacturing	(Menegon & Isatto, 2023)
Information [K.3]	[D.8] The notion where the data of each stage of a product lifecycle is transformed into information.	Product lifecycle	(Abramovici et al., 2017)
	[D.9] Comprehensive physical and functional description of a component, product or systems.	Smart manufacturing	(Shao et al., 2019a)
	[D.10] Digital information construct about a physical system.	Aerospace	(Tuegel et al., 2011b)
Simulation model [K.4]	[D.11] Simulation based on expert knowledge and real data collected from existing systems.	Machine engineering	(Gabor et al., 2016)

	[D.12] Reengineering computational model of structural life prediction and management.	Product lifecycle	(M. Grieves & Vickers, 2017)
	[D.13] Virtual models for physical objects to simulate their behaviours.	Smart manufacturing	(Qi & Tao, 2018b)
Virtual replica [K.5]	[D.14] Computerized clones of physical assets.	Industrial production	(Banerjee et al., 2017)
	[D.15] Virtual and computerized counterpart of a physical system.	Production systems	(Negri et al., 2017)
	[D.16] Functional system formed by the cooperation of physical production lines with a digital copy.	Industrial production	(Vachalek et al., 2017)
	[D.17] Cyber copy of a physical system.	System of systems	(Alam & El Saddik, 2017)
	[D.18] Digital model that dynamically reflects the status of an artefact	Healthcare	(Bruynseels et al., 2018)
	[D.19] Digital replica of physical entity with two-way dynamic mapping	Manufacturing	(El Saddik, 2018)
	[D.20] Virtual representation of production system that is able to run on different simulation disciplines.	Product lifecycle	(Garetti et al., 2012)
	[D.21] Digital mirror of physical world.	Smart manufacturing	(Guo et al., 2019)
	[D.22] Virtual model of physical object.	Smart manufacturing	(Mabkhot et al., 2018)
	[D.23] Dynamic digital representation of a physical system.	System engineering	(Madni et al., 2019)
	[D.24] Virtual representations of physical manufacturing elements, such as personnel, products, assets and process definitions.	Manufacturing	(Shao et al., 2019b)
	[D.25] Virtual representation of real product.	Industrial component	(Schroeder et al., 2016)
	[D.26] Virtual model of physical asset.	Manufacturing	(Talkhestani et al., 2018)

	[D.27] Digital copy of a physical system.	Industrial production	(Wärmefjord et al., 2017)
	[D.28] A realistic digital representation of assets, processes or systems in the built or natural environment.	Built environment	(Bolton et al., 2018)
	[D.29] Dynamic digital replica of physical assets, processes and systems, involving internet of things (IoT) devices and information feedback from citizens.	Built environment	(Q. Lu et al., 2020)
Capabilities [K.6]	[D.30] Dynamic virtual representation of a physical object or system across its lifecycle, using real-time data to enable understanding, learning and reasoning.	Built environment	(National Infrastructure Commission (UK). United Kingdom, 2017)
	[D.31] Systematic approach consisting of sensing, storage, synchronisation, synthesis and service.	Manufacturing	(J. Lee et al., 2013)

Moreover, it can be noted that over the past decade, the concept of Digital Twin (DT) has undergone a transformation driven by several factors such as the "explosion of data" in terms of the exponential surge in data generation from diverse sources such as IoT devices, sensors, mobile phones, social media, and network infrastructure:

- Advanced sensing technologies: significant advancements in the capability to detect and digitally replicate physical "objects" on a large scale. This is made possible through technologies like Light Detection and Ranging (LIDAR), photogrammetry applications, and Unmanned Aerial Vehicles (UAVs).

- Software advancements: the continuous development of software tools and platforms that facilitate the creation and management of Digital Twins.

- High-Performance computing (HPC): the availability of high-performance computing resources, enabling complex simulations and real-time data processing.

- Machine Learning and AI: the maturation of Machine Learning (ML) and Artificial Intelligence (AI) algorithms, which enhance the analytical capabilities of Digital Twins.

In this scenario, the Digital Twin Consortium took a significant step towards establishing a unified foundation for future advancements by releasing the following definitions:

A digital twin is a virtual representation of real-world entities and processes, synchronized at a specified frequency and fidelity.

Digital twin systems revolutionize businesses by expediting comprehensive comprehension, optimal decision-making, and effective action.

Digital twins harness real-time and historical data to encapsulate the digital twin past and present while simulating potential future scenarios.

Digital twins are driven by desired outcomes, tailored to specific use cases, fortified through seamless integration, constructed upon a foundation of data, guided by domain expertise, and implemented within the realm of IT/OT systems (Digital Twin Consortium, 2020).

Such definitions provide a clear and encompassing understanding of what a Digital Twin entails, highlighting its transformative potential and its reliance on data, expertise, and integration for effective application across various domains.

In 2021, (Deren et al., 2021) traced the origins and evolution of the Digital Twin (DT) concept, starting from its inception in industrial design and manufacturing. They also explored how the concept expanded in scale and intersected with the notion of smart cities, thanks to advancements in related technologies.

Several systematic literature reviews have delved into the DT paradigm. (Semeraro et al., 2021), (Trauer et al., 2020), and (Barricelli et al., 2019) have contributed to understanding the multifaceted nature of DT.

Caprari et al. in 2022, emphasized the multidisciplinary nature of the DT concept, showcasing its intersections with various disciplines and its application in diverse sectors. In his comprehensive overview, Caprari presented a collection of transdisciplinary studies (Caprari et al., 2022), offering a taxonomic analysis of the historical development of the DT concept and its connections to various fields of development and application (Valk, 2020).

Lastly, Shahzad et al. (Saeed et al., 2022b) explored the characteristics, challenges, and applications of Digital Twins in the built environment. They compiled a selection of the most notable and widely cited definitions of "digital twin" in the literature (Brilakis et al., 2019), (David et al., 2018), (Madni et al., 2019), (B. N. Silva et al., 2020), (Shahat et al., 2021), (Centre for Digital Built

Britain, 2019), (Glaessgen & Stargel, 2012), (Y. Chen, 2017), (Arup, 2019), (Boschert & Rosen, 2016b)], providing a synthesis of each to enhance our understanding of this concept's diverse interpretations and applications.

2.1.2 Primary objectives

The primary goal of the Digital Twin (DT) is to facilitate remote and real-time monitoring and control of physical assets. The data integrated into the digital representation of the asset serves several critical functions:

- Anomaly detection: it helps in identifying irregularities or deviations from normal operation.
- Simulation: the data enables the running of simulations to predict how the asset will behave under various conditions.
- Failure prediction: it assists in forecasting potential failures or issues before they occur, enabling proactive maintenance.

This wealth of information is then harnessed to optimize the operations of the asset. It's worth noting that while the initial focus of DT was primarily on real-time monitoring, it has since expanded in scope. Now, DT is also recognized as a valuable tool for building and testing products in virtual environments and supporting design and manufacturing processes.

As illustrated by Dr. Grieves, the concept of Digital Twins encompasses different types, each tailored to specific phases in the system's lifecycle (M. Grieves & Vickers, 2017). Here are the key types (**Figure 7**):

- Digital Twin Prototype (DTP): the DTP serves as the design version and includes all its variations. DTPs are crucial for complex manufactured products.
- Digital Twin Instance (DTI): the DTI represents the Digital Twin of each individual produced artifact. DTIs are created for products where it's essential to maintain comprehensive information about the product throughout its lifespan. Examples include airplanes, rockets, manufacturing floor equipment, and automobiles.
- Digital Twin Aggregates (DTAs): DTAs are composite representations that combine multiple DTIs. They provide both longitudinal and latitudinal insights into behavior. Longitudinally, they correlate past state changes with subsequent behavioral outcomes. For instance, this can help predict component failures when specific sensor data patterns emerge. Latitudinally, DTAs enable a learning process, where a small

group of DTIs can acquire knowledge from actions and share this learning with the broader set of DTIs.

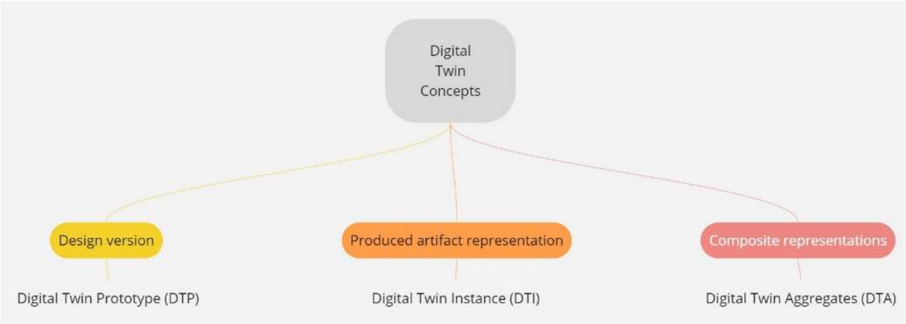


Figure 7. Key Digital Twin Concepts (M. Grieves & Vickers, 2017)

The review of existing literature presented herein introduces the Digital Twins paradigm (**Figure 8**) focusing on conceptual and process models pertinent to various domains, each tailored to specific use-cases. Furthermore, this review delves into the integration of cyber-physical processes and Internet of Things (IoT)-based technologies, alongside artificial intelligence and machine learnins as essential structural components for the development of Intelligent Digital Twins (IDTs).

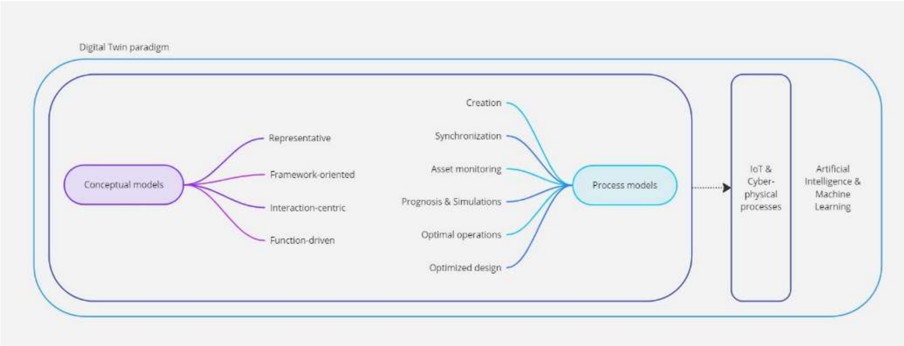


Figure 8. Digital Twin paradigm's main components from literature review

2.1.3 Conceptual models

The concept of the Digital Twin (DT) originally had specific applications, such as assessing structural integrity and aiding in monitoring and maintaining physical assets. However, the DT's broad potential, involving the creation of a precise digital replica of physical assets, has garnered interest across multiple sectors and applications. This has led to diverse definitions as researchers explore various possibilities.

Authors have expanded their definitions to include additional elements, such as specific hardware or software components, expected behaviors, and the context of DT operation. Some have integrated the DT into specific application frameworks, encompassing potential use-cases and sophistication levels.

Several adaptations of the initial DT conceptual model have emerged, better aligning with diverse use-case needs. This review of existing literature suggests the introduction of four primary structural model categories: representative, framework-oriented, interaction-centric and function-driven.

- Representative: these models serve as exemplary representations, offering a clear and fundamental understanding of the Digital Twin's structure and functionality.
- Framework-oriented: these models are centered around a structured framework, emphasizing the systematic arrangement of components and their relationships within the Digital Twin.
- Interaction-centric: these models prioritize the portrayal of interactions and interfaces within the Digital Twin, highlighting how various elements communicate and operate together.
- Function-driven: these models revolve around the functionalities and services provided by the Digital Twin, emphasizing its practical utility and application in specific contexts.

Within the **Representative** category, numerous conceptual models closely resemble the original one outlined in the literature (M. W. Grieves, 2019), characterized by minor formal adjustments, additions of components, and more detailed specifications.

For instance, one conceptual model, proposed by (Madni et al., 2019) expands beyond the traditional framework. In this model, the physical asset not only conveys performance, health, and maintenance data but also includes event occurrences and actions.

In another approach, (Liu et al., 2021) introduced a conceptual model featuring separate representations for an information model and a decision-making model within the digital asset.

Furthermore, communication between the physical asset and the digital asset should encompass data related to geometry, performance, and context. (P. Wang & Luo, 2021) present a representative DT conceptual model that considers the asset's life cycle and use-case descriptions. In this context, the physical asset compiles raw and processed data, process descriptions, and use-case descriptions, while the digital asset incorporates a variety of data models, algorithms, and digital processes.

Other examples in this category underscore remote control capabilities and the interaction of users with the DT, where DT capabilities are considered as services. For instance, (Y. Lu & Xu, 2019) propose a conceptual model addressing on-demand manufacturing services, wherein the DT manages a series of physical controllers and interfaces with remote users through cloud services.

(Damjanovic-Behrendt & Behrendt, 2019) adopt an alternative approach by defining three interoperability "managers" responsible for data exchange and asset control, without specifying the exchanged data. These managers facilitate monitoring, decision-making, and simulations, respectively.

Additionally, some conceptual models consider communication among diverse DTs and the ability to simulate these interactions.

In summary, this category encompasses a range of conceptual models that build upon and refine the foundational Digital Twin concept to suit various applications and requirements within the built environment. In the context of Architecture, Engineering, Construction, and Operations (AECO), it contemplates integration with Building Management Systems (BMS), identifies various users interfacing with the DT (operators, owners, authorities, and end-users), and underscores the relevance of contextual visualizations due to the asset's location in the built environment.

Framework-oriented conceptual models fall into a category that elucidates the structure of Digital Twins (DT) by portraying them as an assembly of diverse models operating within a communication framework. Their focus lies in delineating the assortment of models essential for crafting a precise digital counterpart of a physical asset, which the original DT conceptual model does not explicitly articulate.

For instance, (Z. Zhang et al., 2020) present a comprehensive DT comprising models for product definition, geometric shape, manufacturing attribution, behavior rules, and data fusion. Similarly, (Vrabič et al., 2018) propose distinct models encompassing geometry, environment, dynamics, control, sensors, and machine learning.

In a broader context, (Tao et al., 2019) introduce a four-model conceptual framework, embracing geometry, physics, behavior, and rule models. This framework also integrates user services and access to domain knowledge.

(Zheng & Sivabalan, 2020a) present a DT concept model featuring three core components: a digital model integrating assets and environment, a computational model for simulations, and a graph-based model organizing interactions among datasets. Notably, this model distinguishes between digital replicas of physical assets and the real environment.

(Stark et al., 2019) adopt an approach that explicitly segregates hardware and software components, considering various model types, data, processes, and inter-component communication.

In all these models, the effective interaction among diverse models is pivotal for accurately mirroring changes in the physical asset's condition. (Terkaj et al., 2019) introduce a network-configured DT model with nodes for configurator, ontologies, simulation, and visualization. This configuration, enhanced with elements from other models, holds promise for applications in the built environment.

For AECO sectors, determining the relevance of models for specific use cases is crucial. The configurator can then select the pertinent models and offer requested services via the visualization node, offering flexibility and adaptability. This modular conceptual model facilitates integration with existing capabilities in AECO, encompassing ontologies, simulations, and contextual visualizations.

Interaction-centric DTs are presented as an intermediary that bridges the "physical space" and "digital space," uniting digital and physical entities. Unlike the previous models that emphasize DT as a digital replica of physical assets, interaction-centric models focus on its role in enhancing processes rather than assets.

For instance, (Zheng & Sivabalan, 2020b) propose a model where the DT acts as an interface linking manufacturers, service providers, users in the physical space, and cloud-based services in the digital realm. Similarly, (Y. Wang et al., 2020) describe a DT that facilitates connections between manufacturing, cloud processing, and physical product inspections during operations.

In these models, the specific linkage between physical and digital assets remains implicit. The DT serves solely as a conduit for digital processes to augment physical ones. Unlike the previous models, which prioritize assets, these conceptual models concentrate on processes. This difference poses limitations for AECO sectors, where physical assets hold paramount importance. In contrast, in manufacturing, the focus is often on the manufacturing processes rather than the products themselves.

In summary, the higher level of abstraction and the absence of a clear physical-digital correspondence among assets in interface-oriented models may restrict their suitability for AECO use cases. Nevertheless, these models excel in accommodating digital components and services that lack a direct physical counterpart.

According to the **function-driven** conceptual model the user is placed at the core of their structural framework. These models prioritize complex workflows by considering the interactions among processes and people. Services are explicitly modeled, with DTs serving as wrappers for these services. The primary goal is to automate intricate workflows through task orchestration, distinguishing between automation (for single tasks) and orchestration (for multi-step processes across various systems).

For instance, (Aheleroff et al., 2021) used a service-based DT conceptual model to orchestrate diverse processes, leveraging various technologies for wetland maintenance scheduling. This involved real-time monitoring, control, and prioritization of maintenance activities. In this model, a supervisor sets goals for an orchestrator, a machine-learning-enabled entity managing multiple DTs. These DTs can represent physical assets or human operators who interface with the orchestrator through their service capabilities. The orchestrator utilizes historical data, synthesizes supervisor-defined goals, proposes solutions, and executes the selected solution.

This conceptual model deviates significantly from the traditional DT paradigm, positioning the DT as a component within a larger system where orchestrators and supervisors play crucial roles. It focuses on managing multiple DTs in complex workflows, expanding the DT's use-case beyond its original intent. In the context of Architecture, Engineering, Construction, and Operations (AECO), this model could be applied to manage intricate construction or maintenance operations.

Process models provide high-level descriptions of sequential and parallel activities, rules, guidelines, and behavior patterns that lead to desired outcomes.

In summary, service-based conceptual models prioritize user-centric, complex workflows with explicit modeling of services. They extend the traditional DT paradigm to accommodate intricate task orchestration. Additionally, process models encompass a variety of activities in DT creation, synchronization, and operationalization, with resource-centric approaches aligning closely with AECO practices.

2.1.4 Process models

The DT paradigm encompasses various process models that serve different purposes within the domains of manufacturing and the built environment. These models can be grouped into several categories, including creation, synchronization, asset monitoring, prognosis and simulations, optimal operations, and optimized design.

The creation of process models involve the automated generation of DT instances as a DT model is initially defined, and specific DT instances are created for each product to be manufactured. For instance, (K. T. Park et al., 2019) proposed a four-step process for DT creation, involving combining digital components, incorporating manufacturing equipment, importing functional definitions, and visualizing the DT instance in a 3D environment. While these processes are common in manufacturing, they differ significantly from creating Building Information Models (BIM) or geometrical DT models for built assets.

DT creation processes focus on functional attributes and relationships among components, which are not typically captured in BIM models. Additionally, the small variations between built assets and the relatively low number of instances make direct translation of manufacturing approaches to the built environment challenging. However, these approaches can support BIM model federation for interdisciplinary assessments in AECO.

Synchronization is a fundamental requirement in the DT paradigm, involving two-way data exchange between physical assets and their digital counterparts. Event-based synchronization, proposed by (K. T. Park et al., 2019), aligns different data types with specific events and timestamps, enabling diverse synchronization processes for tracking historical performance, real-time monitoring, and future schedules. In the AECO sectors, such synchronization processes could correspond to obtaining data related to construction progress, current tasks, and future construction plans. The rate of synchronization, known as the "twinning rate," is crucial, with real-time synchronization being vital for Industry 4.0 manufacturing. However, the specific use-case should determine latency requirements, which may not be as stringent in AECO compared to manufacturing.

Asset monitoring in the DT paradigm often focuses on performance monitoring to detect faults promptly and facilitate maintenance.

In AECO, Structural Health Monitoring (SHM) and building services monitoring are predominant. For example, (Davila Delgado & Hofmeyer, 2013) demonstrated the integration of physical and digital infrastructure assets for structural condition data and long-term asset management. In building services, DT process models, such as the one by (Q. Lu et al., 2020), are used for anomaly detection in heating and ventilation systems. These models can be adapted for gradual and abrupt fault detection in various infrastructure and built assets.

Prognosis and simulations are essential aspects of the DT paradigm, enabling predictive maintenance and condition-based assessments. Prognostic simulations predict future trends or state changes based on current asset states, while reactive simulations predict future states due to unexpected disturbances. In the AECO context, prognosis and simulations are often carried out at the design stage, focusing on future behaviors rather than current operations. However, DT-based approaches aim to simulate asset operations based on current performance data. Validation of simulations is crucial, though relatively unexplored in the DT context.

Optimal operations leverage DTs to optimize complex processes by adjusting parameters within limits to minimize costs or maximize efficiency. There are two main types of process models: those focused on optimizing manufacturing processes and those targeting equipment operation. (Schluse et al., 2017) presented a DT process model for simulation-based optimal operations of complex equipment tasks, which integrates real-world and digital interactions and simulates user interactions. In the AECO context, this approach can optimize the operation of various plant equipment or coordinate large-scale construction activities efficiently.

Optimized design using DTs involves two primary approaches. One uses DTs to simulate the performance of physical assets and refine designs based on simulation results. The other leverages historical performance and condition data from physical assets to improve future designs. In AECO, these approaches can be applied to optimize designs based on real-world data and enhance simulations to better align with actual asset performance. The use of evolutionary processes and benchmarking further supports the optimization of designs in the built environment.

In summary, the DT paradigm offers a versatile set of process models applicable to various industries, including manufacturing and AECO, with the potential to enhance asset management, performance optimization, and design processes. These models bridge the gap between physical assets and their digital counterparts, offering opportunities for improved efficiency and decision-making.

2.1.5 Cyber-physical processes

Cyber-physical security also stands out as a paramount concern as its significance arises from the fact that if the information flowing between the Digital Twin and its Physical Twin is not adequately safeguarded or if the system remains vulnerable to unauthorized access, the Digital Twin becomes a potential source of harm. Even in scenarios where the Digital Twin is primarily responsible for monitoring, the incapacity of the system to shield its data from external breaches poses not only problems but also endangers the well-being of the system's owner or user.

To delve deeper into this matter, cyber-physical security entails the protection of interconnected digital and physical systems. In the context of Digital Twins, this involves securing the data exchange and communication between the virtual representation (Digital Twin) and its real-world counterpart (Physical Twin).

Ensuring the security of a Digital Twin ecosystem encompasses several crucial aspects:

- Data Encryption: employing advanced encryption techniques to scramble data during transmission, making it incomprehensible to unauthorized parties.

- Access Control: implementing strict access controls and authentication mechanisms to ensure that only authorized personnel can interact with the Digital Twin.
- Intrusion Detection: employing intrusion detection systems to swiftly identify and respond to any suspicious activities or attempts to breach the system.
- Regular Updates: keeping the Digital Twin's software and security protocols up-to-date to patch any vulnerabilities that may emerge over time.
- Redundancy: establishing backup systems and data storage to prevent data loss or system failure in case of a cyberattack.

In essence, cyber-physical security is akin to safeguarding the Digital Twin's system, ensuring its integrity and reliability. It's a critical concern not only for protecting valuable data but also for maintaining the safety and functionality of the systems that rely on Digital Twins, be it smart cities, industrial facilities, or healthcare systems. Therefore, addressing cyber-physical security comprehensively is critical to harness the full potential of Digital Twins while mitigating risks to individuals and society as a whole.

The concept of Digital Twins (DT) and Cyber-Physical Systems (CPS) share many similarities, as both describe the integration of digital entities with physical entities. While these terms emerged around the same time, DT in 2005 and CPS in 2006, there is still no universally accepted definition for either. Multiple definitions exist, reflecting the evolving nature of these concepts.

For instance, CPS can be defined as systems "with integrated computational and physical capabilities that can interact with humans through many new modalities" (Baheti, 2011). Alur describes CPS as "a collection of computing devices communicating with one another and interacting with the physical world via sensors and actuators in a feedback loop." Meanwhile, (Tao, Zhang, et al., 2018) provide a broader definition of CPS as "multidimensional and complex systems that integrate the cyber world and the dynamic physical world," emphasizing the integration of computation, communication, and control of physical processes.

The relationship between DT and CPS is still a subject of debate, with three major trends emerging:

- DT as an aggregation of CPS concepts: some argue that DT is essentially a repackaging of existing CPS concepts, specifically developed for aerospace and later applied to manufacturing.

- DT as a subset of CPS: others view CPS as the overarching term, with DT serving as a subsidiary concept used for specific use cases, such as asset monitoring and maintenance.
- DT and CPS as slightly different aspects of the same paradigm: this perspective suggests that CPS focuses more on fundamental scientific aspects, while DT concentrates on practical implementations.

DTs emphasize digital models, while CPS emphasizes computation, communication, and control. Additionally, CPS research emphasizes sensors, actuators, and control, while DT research centers on models and data.

To further clarify the comparison between CPS and DT, it's essential to consider their attributes, functions, main use cases, and key differences. Notably, one fundamental distinction is that DT refers to an information construct describing a digital replica of a physical asset and its data connections, whereas CPS refers to a system integrating digital and physical components. This nuanced difference has significant implications.

In a DT solution, a physical asset has a corresponding digital replica, allowing for behavior simulation, condition monitoring, and predictive analysis. In contrast, a CPS solution primarily focuses on enhancing control and optimization of physical processes through the integration of digital components without the need for a direct correspondence between the digital and physical elements.

Making a comprehensive comparison between Cyber-Physical Systems (CPS) and Digital Twins (DT) in terms of their defining attributes, functions, primary use cases, and key distinctions, CPS can be classified as a scientific category, revolving around integrating physical processes with computer systems, emphasizing communication, computation, and control of sensors and actuators in physical systems. It operates on a one-to-many correspondence basis, allowing one digital system to correspond to multiple physical assets, and it plays a pivotal role in enhancing physical processes through real-time monitoring and control. In contrast, Digital Twins are categorized as engineering tools, focusing on data and digital models, and operate on a one-to-one correspondence basis, where one physical asset corresponds to one digital asset. Digital Twins serve as near-real-time digital replicas of physical products or processes, encapsulating all relevant information throughout their lifecycle phases.

The proliferation of Internet of Things (IoT) technology has ushered in a significant surge in interconnected smart devices (L. Da Xu et al., 2014), (Stankovic, 2014) To ensure the meaningful utilization of these devices, they must possess the capability to capture data relevant to their intended functions. In the context of the built environment, these smart devices may collect a wide array of information, including but not limited to traffic patterns, temperature, humidity, and energy consumption. Depending on their specific design, smart devices can capture essential attributes of objects, enabling the realization of enhanced efficiency and intelligence within cyber-physical systems (CPS). These CPS encompass critical infrastructure systems such as energy, transportation, manufacturing, agriculture, and healthcare (G. Xu et al., 2016), (Mahmud et al., 2017), (Bartolini et al., 2020). Given the vast volume of data generated within these CPS, it becomes imperative to embrace advanced networking, data analysis techniques (e.g., deep learning), and cloud/edge computing technologies within smart systems (Hatcher & Yu, 2018), (D. Wu et al., 2019a), (W. Shi et al., 2016). This strategic approach facilitates the efficient collection, transmission, analysis, and sharing of pertinent data, thereby empowering physical systems with the intelligence required for improved monitoring and control capabilities (Cai et al., 2021).

The concept of a Digital Twin (DT) encompasses a data-driven combination of software and hardware that intricately portrays a real physical system, encompassing all of its functionalities, use cases, statuses, and information across various life-cycle phases (Z. Cai et al., 2021b). While integrating DT into Cyber-Physical Systems (CPS) offers evident advantages, it introduces a host of challenges in the domains of modeling, computation, networking, and data analysis. Moreover, CPS imposes exceptional requirements related to latency, reliability, safety, scalability, security, and privacy, among others. While advanced networking, computing, and data analysis technologies can contribute to realizing DT, several critical issues need attention. These include defining the theoretical foundations and modeling techniques to ensure the accurate and dependable representation of real-world systems by DT, designing Machine Learning (ML) and Deep Learning (DL) models to enable real-time processing of vast datasets, and addressing the security and privacy concerns associated with DT and the collection and dissemination of privacy-sensitive information. To achieve a precise representation of physical systems, a multi-domain and multilevel design approach must be integrated into the lifecycle of these systems. Additionally, comprehensive investigations are required to extend DTs to IoT-based smart systems driven by data science and engineering, given the widespread application of IoT across diverse physical systems.

Existing surveys have examined DT frameworks in industrial systems (Kritzinger et al., 2018), (Danilczyk et al., 2019)and some studies have categorized DTs in the power grid (Tzanis et al., 2020a), (Tzanis et al., 2020b) However, given the extensive use of IoT in various physical systems, a comprehensive examination of extending DTs to IoT-based smart systems is imperative.

As a DT system serves as a replicated version of a target physical system, utilizing a model to continuously emulate various functions of the physical counterpart, the DT must establish a connection with the corresponding physical entity, allowing it to gather and update the state of physical components. Consequently, a DT model can not only simulate but also predict, control, optimize, and learn from the real-world entities it represents.

2.1.6 Cybersecurity

The integration of Information Technology (IT) and Operational Technology (OT) systems, in conjunction with the inherent data exchange involved in the creation and operation of digital twins, reveals a complex and evolving landscape of cybersecurity threats. The failure to adequately mitigate these risks poses significant threats to the integrity, availability, and confidentiality of both the digital twin and its associated data (Stouffer, 2023). This necessitates a comprehensive understanding of the various types of threats and the implementation of tailored cybersecurity strategies.

Data Integrity Threats: the efficacy and reliability of a digital twin are fundamentally dependent on the accuracy and integrity of the data it receives from its physical counterpart. Any form of data manipulation or corruption can lead to erroneous modeling and analysis, potentially culminating in suboptimal or detrimental decisions. Cyber-attacks that specifically target data integrity, such as Man-in-the-Middle attacks or data tampering, pose a significant threat to the security of digital twins (Alcaraz & Zeadally, 2015). These threats underscore the necessity for robust data validation and verification mechanisms to ensure the fidelity of the digital twin's data.

Unauthorized Access: Digital Twins, which often process and store sensitive and proprietary data, are prime targets for cybercriminals seeking unauthorized access. This unauthorized access could be aimed at data theft for industrial espionage purposes or gaining control over the physical systems mirrored by the digital twin (Roman et al., 2013). Implementing stringent access control measures and continuous monitoring for unauthorized access attempts is crucial in safeguarding these digital assets.

Malware and Ransomware: as interconnected systems, digital twins are vulnerable to malware or ransomware attacks. Such an attack has the potential to disrupt the operation of the digital twin, or even result in a shutdown of the corresponding physical system (Collier et al., 2014).

IT-related Threats: in the IT domain, vulnerabilities can arise from various sources, including network vulnerabilities, inadequate access controls, or the use of outdated systems. Given IT's critical role in transmitting and processing the data utilized by digital twins, any compromise in IT systems can have far-reaching effects on the digital twin's functionality and reliability. Therefore, maintaining up-to-date IT infrastructure and employing comprehensive network security strategies is imperative.

OT-related Threats: Operational Technology (OT) encompasses the hardware and software responsible for monitoring and controlling physical devices, processes, and infrastructure, particularly in industrial settings (Boyes et al., 2018). The increasing convergence of OT with IT systems, especially with the adoption of IoT devices, has led to enhanced connectivity but also exposed OT systems to new cybersecurity threats. Compromising OT systems can have direct and severe impacts on the physical systems they control, potentially leading to physical damage and safety hazards.

Privacy Concerns: The nature of data processed and stored by digital twins, especially in sensitive sectors like healthcare, raises significant privacy concerns. Ensuring compliance with data protection regulations and implementing robust encryption and anonymization techniques are critical in addressing these privacy issues.

Given the diverse applications of digital twins across various sectors, a one-size-fits-all approach to cybersecurity is not feasible. Customized cybersecurity strategies tailored to specific applications of digital twins are essential. However, certain general strategies, such as regular system updates, robust access controls, and advanced threat detection mechanisms, can provide a solid foundation for cybersecurity efforts across different implementations of digital twins.

In the rapidly evolving digital landscape, the security of Information Technology (IT) and Operational Technology (OT) systems has become a critical concern, especially in the context of digital twins. Digital twins, as virtual replicas of physical systems, rely heavily on the seamless and secure exchange of data between the IT and OT domains. This introductory section delves into the various strategies and measures necessary to fortify the cybersecurity of both IT and OT technologies. exploring the implementation of network security measures, data security protocols, and regular updates in IT systems, while also addressing the unique challenges posed by OT systems, including legacy systems and network segmentation. Additionally, it is highlighted the importance of addressing privacy concerns introducing Italian Cybersecurity Framework, which provides a comprehensive approach to managing cybersecurity risks in the context of digital twins. This framework, developed by Sapienza University's Research Center of Cyber Intelligence and Information Security (CIS), offers a tailored solution for organizations, particularly small and medium enterprises, to navigate the complexities of cybersecurity in the digital twin environment.

Network Security Measures: the integrity of IT networks is critical in managing the flow of data to and from digital twins. To protect these networks, the implementation of comprehensive security measures is essential. This includes deploying advanced firewalls, sophisticated intrusion detection systems, and designing secure network architectures. These measures are pivotal in safeguarding against unauthorized access and potential data breaches. Furthermore, the utilization of encryption and secure communication protocols is imperative to maintain the confidentiality and integrity of data, both in storage and during transmission (Keoh et al., 2014). This dual approach ensures a fortified barrier against cyber threats.

Data Security Measures: the protection of data associated with digital twins involves securing it during transit and while at rest. This security is achieved through the application of robust data encryption techniques, secure data storage solutions, and regular data backups. Additionally, stringent access control measures are crucial in preventing unauthorized data access (Z. Xiao & Xiao, 2013). These measures collectively form a comprehensive shield, safeguarding critical data against various cyber threats.

Regular Updates and Patch Management: the landscape of cyber threats is continually evolving, making regular updates and patch management of IT systems a necessity. This proactive approach is vital in addressing emerging security vulnerabilities and reducing the risk of cyber-attacks. A structured patch management process ensures timely application of updates, thereby enhancing the resilience of the IT infrastructure (Othmane et al., 2019).

Securing OT Technologies: OT systems, particularly legacy systems, were often not designed with modern cybersecurity considerations. Their integration with IT systems exposes them to new cyber threats. Developing secure solutions for OT systems, often from the ground up, is a complex yet critical task (Colombo et al., 2016).

Network Segmentation: network segmentation is a proven strategy in network security and is particularly effective in securing OT systems. It involves isolating the OT network from other networks, including IT networks, to prevent cross-network breaches. This strategy not only limits the spread of malware but also provides enhanced control over network traffic (Colombo et al., 2016).

Regular Security Assessments: conducting regular security assessments is crucial in identifying and addressing vulnerabilities in OT systems. These assessments should cover technical, operational, and procedural aspects, allowing for comprehensive risk management (Humayed et al., 2017).

Addressing Privacy Concerns: In the realm of digital twins, privacy concerns are paramount, especially given the sensitivity of the data involved. A holistic approach that combines technical measures like data anonymization and encryption with adherence to data protection regulations is essential. Raising user awareness about privacy implications is also crucial in ensuring data privacy.

In 2015 Sapienza University (at the Research Center of Cyber Intelligence and Information Security – CIS) produced the "National Framework for cyber security" aimed at providing to organizations a homogeneous and volunteer approach to face up cyber security in order to reduce the risk linked to the cyber threat.

The approach of this Framework is strinctly linked to a risk analysis and not to technology standards and therefore can be applied to digital twins.

The Italian Cybersecurity Framework derives much from the Framework for Improving Critical Infrastructure Cybersecurity (NIST) but has been tailored according to the Italian production context with a specific focus on small and medium enterprises. The National Framework derives from the NIST Framework the basics of Framework Core, Profile and Implementation Tier, adding the priority and maturity levels to the 98 Subcategories of the Framework Core.

The NIST Framework offers a highly flexible framework, which is mostly targeted at crucial facilities; at CIS they developed it according to the characteristics of the social and economic system of our country, reaching a cross-sector framework that can be contextualized in specific production sectors or in company types with specific characteristics, therefore realizing a suitable framework also for DT approaches.

In the Cybersecurity Framework, three important concepts have been introduced:

1. Priority Levels: These define the priority associated with each Subcategory of the Framework Core. It is important to note that organizations are at liberty to adapt their priority levels based on their type of business, size, and individual risk profile.
2. Maturity Levels: These describe the various ways in which each Subcategory of the Framework Core can be implemented. Each enterprise must carefully evaluate the selected maturity level in accordance with its business size and risk profile. Generally, higher maturity levels necessitate more significant investment in terms of financial and management resources. It is also worth noting that for some Subcategories, establishing maturity levels may not be feasible.
3. Framework Contextualization: This involves tailoring the Framework to a specific productive sector, business type, or individual business. This process entails selecting the Function, Category, and Subcategory from the relevant Framework Cores and specifying the priority and maturity levels that are most suitable for the specific implementation context. (Roberto Baldoni, 2016).

2.1.7 IoT and networking technologies

The effectiveness of DT technology hinges on its capacity to ingest vast datasets and derive meaningful insights through advanced Machine Learning (ML) and Deep Learning (DL) techniques. This enables real-time forecasting and prediction. For instance, when DT is integrated into manufacturing processes, it can facilitate proactive planning for repair and maintenance activities, thereby mitigating potential failures during manufacturing. To ensure seamless integration between DT and physical systems, it is imperative for the DT to access real-time state information from the physical systems, a feat achieved through the utilization of Internet of Things (IoT) sensors and networking technologies.

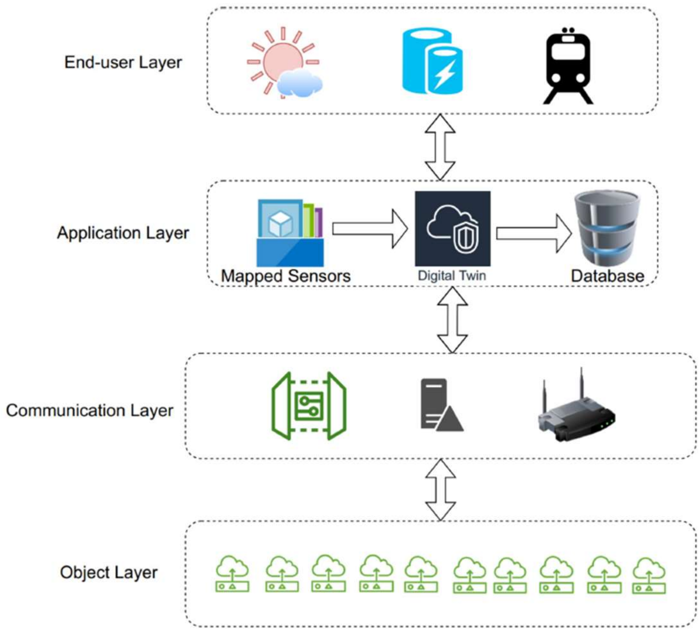


Figure 9. DT architecture for IoT (Qian et al., 2022)

The proliferation of Internet of Things (IoT) devices across diverse real-world applications and physical environments has introduced a great challenge associated with the collection, aggregation, storage, and analysis of the data they generate. **Figure 9** presents a generic architecture for IoT systems comprising four distinct layers: the object layer, communication layer, application layer, and end-user layer.

The object layer encompasses all IoT sensors, serving as the source of data and information for various IoT-driven applications. This layer encompasses all the components that constitute the physical system. Meanwhile, the communication layer furnishes the necessary communication network infrastructure to connect IoT devices and gather data for DTs.

For instance, multiple edge gateways can be strategically deployed to collect and consolidate data from sensors, subsequently transmitting this data to the application layer. These IoT sensors and gateways employ distinct communication protocols for data collection and transmission. Upon acquisition of information at the gateway within the object layer, it is forwarded to the application layer, which encompasses the digital systems of the Digital Twins (DT).

Within the application layer, the DT system first aligns all sensor and gateway information with the digital model. Simultaneously, an IoT naming service is employed to assign names to the mapped IoT sensors and local gateways. This naming convention enables the digital system to locate various resources (Hatcher et al., 2021). Subsequently, the digital system harnesses Machine Learning (ML) and Deep Learning (DL) models for predictive analysis. This allows the digital system to regulate actuators in the target layer based on the outcomes of these predictions. Additionally, the results of prediction and analysis can offer diverse services within the end-user layer.

The end-user layer serves as a gateway for user services, allowing users to submit requests to the application layer. Subsequently, the digital system processes the received requests and provides responses accordingly. Within the architectural framework of the Digital Twin (DT) system, data representation and communication protocols assume critical roles in facilitating data sharing within the DT and delivering services to end users.

Data representation is fundamental for enabling system components to comprehend data from diverse domains (Jaloudi, 2019a) (Al-Sarawi et al., 2017a). In the realm of DT, several widely utilized data representation protocols include the DT Definition Language (DTDL) (GitHub, n.d.), FIWARE (Conde et al., 2022), OPC Unified Architecture (OPC UA) (Ala-Laurinaho et al., 2020), and the Feature-Based DT Framework (FDTF) (Autiosalo et al., 2020). Notably, DTDL, an open-standard platform introduced by Microsoft (GitHub, n.d.), plays a significant role in enabling data transmission within the system and between different systems. The communication protocol plays a crucial role in facilitating information exchange among IoT devices and IoT systems (Jaloudi, 2019b), (Al-Sarawi et al., 2017b). In the realm of IoT, several prominent communication protocols are commonly employed, such as the Constrained Application Protocol (CoAP) (Kome et al., 2018), the OASIS Standard Message Passing Protocol (MQTT) (D. Silva et al., 2021a), the Modbus TCP/IP Protocol (Cagnano et al., 2020a), and Ultra Reliable Low Latency Communication (URLLC) (Tan et al., 2020). A summary of these communication protocols is provided in Table 3.

Table 3. IoT communication protocols

Protocol Name	Protocol Type	Protocol Characteristics
DTDL	Data Representation	As an open-standard platform, it defines six characteristics of IoT components and enables seamless data transmission between different DTs.
FIWARE	Data Representation	It supports DT data transmission and the processing of contextual information received from various IoT components.
OPC UA	Data Representation	As a modeling framework, it can retrieve information from raw data, support data manipulation, and provide monitoring capabilities.
FDTF	Data Representation	As a DT structure, it enables the DT system to share information based on the data link between DT components.
CoAP	Communication	As a specialized web communication protocol based on the User Datagram Protocol (UDP), it is tailored for resource restricted devices, supports the transmission of data via Hypertext Transfer Protocol (HTTP) and provides a publish and subscribe mechanism to simplify the process of obtaining continuous data from the sensor.
MQTT	Communication	As a communication protocol based on Transmission Control Protocol (TCP), it enables lightweight way for IoT devices to communicate, provides reliable data transfer, and can establish a long-existing outgoing TCP protocol to enable transmission.
Modbus TCP/IP	Communication	As a communication protocol based on Transmission Control Protocol (TCP), it realizes the connection between industrial devices, provides reliable data transfer, and contains built-in checksum protection.
URLLC	Communication	As a communication protocol, it tends to achieve low latency and reliability in the transmission process between IoT devices.

2.1.8 Towards Intelligent Digital Twins (IDT)

The synergy between Artificial Intelligence (AI) and Digital Twins (DTs) represents a pivotal advancement in various industries, notably manufacturing. Researchers, like (Rathore et al., 2021), have underscored the transformative potential of AI techniques when applied to DTs. By harnessing advanced AI capabilities, DTs can evolve into intelligent systems capable of not only replicating physical processes but also making critical decisions. These AI-powered DTs can excel in tasks such as process optimization, resource allocation, safety and fault detection, predictive maintenance, and real-time decision-making. However, to fulfill these roles effectively, DTs must exhibit qualities like accuracy, robustness, and autonomy. Real-world systems are dynamic, influenced by ever-changing operational and environmental factors. This dynamism necessitates that DT models remain in sync with the current state of physical systems. Challenges arise in maintaining this synchronization due to inherent limitations in modeling and data collection, leading to residual errors. Paradoxically, these errors present opportunities for further progress.

The conventional view of DTs portrays them as passive repositories of product information, requiring user-initiated queries for accessing data and predictions. However, a fundamental shift is occurring towards Intelligent Digital Twins (IDTs), marked by their proactive and anticipatory nature. IDTs constantly monitor physical systems, learning from real-time sensor data and optimizing their behavior to align with evolving conditions. This proactive approach relies heavily on machine learning and AI techniques, as noted in (Jaensch et al., 2018) and (J. Wang et al., 2019) These methods empower IDTs to adapt to machinery process degradation, fine-tune parameters, and continuously improve their models. (Sapronov et al., 2018) even employ machine learning for refining DT parameters, while (Maschler et al., 2021) explore cross-phase transfer learning to reduce discrepancies when using real data. (Cronrath et al., 2019) delves into reinforcement learning to correct model and data errors.

As described by Grieves, Artificial Intelligence (AI) and Modeling & Simulation (M&S) stand as two complementary facets of the digital realm, offering distinct capabilities that, when combined, yield a powerful intelligence for Digital Twins within a virtual realm.

AI is the effort of computers emulating human intelligence, focusing on tasks traditionally associated with human thinking. Conversely, M&S pertains to computers replicating the dynamics of the physical universe, excluding human intelligence. These parallel tracks converge to furnish digital twins with a unique form of intelligence. The bedrock of both AI and M&S lies in the relentless evolution of computing power. Since the early 1970s, the exponential growth of computing performance, epitomized by Moore's Law, has been an instrumental driving force. This steady march forward, as depicted in the logarithmic graph (Figure 6), has witnessed computing performance doubling approximately every 18 months, culminating in the awe-inspiring 54 billion transistor density in 2020. Extrapolating this trajectory, we anticipate a staggering six trillion transistors or equivalent capabilities by 2030. The implications of such computational prowess are profound, ensuring that the concepts elucidated herein will remain unconstrained by computational limitations.

Historically, the Digital Twin realm has been one of passivity—a reservoir of product information awaiting user inquiries. This conventional Digital Twin, described as an information repository, functions with users extracting data from it as needed. It serves as a resource for users seeking to minimize physical resource wastage by making informed decisions. The Intelligent Digital Twin (IDT) (M. Grieves, 2022), on the other hand, embodies a paradigm shift by embracing activity. The traditional Digital Twin has often been depicted as a static receptacle of product information, with users initiating requests for specific predictions when necessary. Contrarily, the IDT operates as a dynamic and engaged entity, actively participating in the decision-making process. The essence of this transformation lies in the establishment of bidirectional communication links between the Digital Twin and its virtual surroundings and the Physical Twin and its real-world environment.

For the IDT to be active, it must remain in an online state, perpetually scanning the Physical Twin (PT). The IDT's activation pivots from the physical aspect of the digital twin model to its active counterpart. Depending on the lifecycle phase, the physical side encompasses the environment, the physical twin, and the human stakeholders involved. Consequently, the IDT metamorphoses from a passive repository into a proactive one, offering continuous agent assistance. Unlike the conventional digital twin, which primarily relies on human-driven goal seeking, the IDT shares this responsibility with its human users.

A defining attribute of the IDT is its anticipation, mirroring a human trait. The Future-Running Simulation (FRS) (M. Grieves, 2022) is a concept defined by Grieves to explain this anticipation capability, constantly simulating complex products within the digital twin. Leveraging the exponential advancements in computing power, the IDT can orchestrate multiple scenarios, harnessing Bayesian probabilities to estimate the likelihood of adverse events. This empowers humans to assess and mitigate risks proactively, substantially enhancing decision-making.

As we navigate this realm of AI-empowered DTs, the key challenge lies in ensuring the fidelity of the virtual model to the physical system it represents. Imperfections or discrepancies between the two realms can be problematic, necessitating continuous updates and validation before deployment. Assessing the virtual model's confidence and its alignment with real-time data is key. Since physical systems evolve and are influenced by various factors, DT models must mirror this dynamism and adapt to changing conditions.

Beyond manufacturing, the combination of Digital Twins and AI holds immense promise in energy management. This synergy can optimize grid operations, reduce human intervention, facilitate intelligent decision-making, and enable forecast-based management models, as shown by (Y. Li & Shen, 2022) Achieving power balance, efficient load management, and microgrid integration are primary goals in power grid optimization, enhancing overall resilience (Bazmohammadi et al., 2022a).

Nevertheless, several challenges loom over the successful implementation of AI-empowered DTs. These include the availability of historical data for training machine learning algorithms, data quality and resolution, communication latency, model interpretability and repeatability, and the adequacy of computational resources (J. Guan et al., 2021).

In the context of the built environment, the application of Digital Twins has seen substantial interest, particularly in maintenance management and energy performance simulations. However, research in the realm of digital twins for building energy efficiency is relatively nascent, with a limited number of scientific articles and recent publication dates indicating its emerging nature.

This nascent field reveals gaps in areas like data integration systems, complex autonomous decision-making, and data visualization—key aspects to facilitate understanding and interpretation by non-technical stakeholders (Bortolini et al., 2022)

Examining the Intelligent Digital Twin (IDT) across its lifecycle reveals two key phases: DT design and DT utilization.

Design Phase: during this stage, models are created, and data is gathered and stored.

Utilization Phase: this phase consists of three sub-stages:

- Setup: parameters are adjusted based on system observations.
- Run: the DT performs its functions (e.g., monitoring, decision support) once parameters are validated.
- Maintenance: updates and maintenance are conducted to account for deviations.

Focusing on the utilization phase, it can be highlighted how AI techniques enhance DT accuracy.

In the setup phase, the critical task involves validating and refining the virtual model's parameters to ensure an accurate representation of the physical system. This meticulous process is initiated by human operators due to the initial uncertainty regarding the DT's parameters. The aim is to minimize any disparities between the physical and virtual twins, and to achieve this, artificial intelligence techniques, such as Reinforcement Learning, are employed.

In this context, Reinforcement Learning acts as a crucial tool, aligning the two twins by fine-tuning the DT's parameters. This alignment process plays a pivotal role in establishing the foundation for accurate real-time reflections of the physical system within the digital realm, ultimately enhancing the DT's overall efficacy.

The Run phase is characterized by the dynamic nature of the DT model, which operates with predefined parameters. During this phase, the DT must continuously ensure that its behavior effectively mirrors the current state of the physical system. To achieve this, it engages in vigilant monitoring and compares the nominal behaviors calculated by the models with observations from the physical twin. Artificial intelligence techniques, particularly Supervised or Unsupervised Learning, are indispensable at this stage. These AI methods have the critical role of identifying unexpected patterns within the dataset.

Algorithms like Isolation Forest, Local Outlier Factor, Principal Components, or DBSCAN are frequently employed to detect anomalies in the data distribution. It is worth noting that human expertise is not directly involved in this phase, but it may be called upon to analyze anomalies identified by the AI tools.

When unsupervised learning algorithms detect deviations in the DT's behavior, it becomes crucial to assess the severity of these deviations. This assessment is the responsibility of the human operator, who, if necessary, triggers the maintenance phase. Here, the evaluation of DT parameters differs from the setup phase, as parameters have evolved, requiring updates to maintain accuracy. Despite this difference, the functionalities of this phase closely resemble those of the setup phase, with many of the same algorithms potentially being applicable. The primary goal remains to ensure the DT's ongoing precision and reliability as it continues to operate with the physical system.

The true value of a Digital Twin perpetually hinges upon its application in real-world contexts. The pivotal question remains: can the information derived from the Digital Twin replace wasteful expenditure of physical resources, such as time, energy, and materials? The answer lies in the cost-effectiveness of collecting, processing, storing, and retrieving this information in comparison to the expense of squandering tangible resources. As the cost of virtual bits continues its exponential decline, physical atoms concurrently become scarcer, even at the rate of inflation, underscoring the necessity and urgency of this paradigm shift.

2.2 Digital Twins in the Energy sector and Distributed Energy Resources (DER)

The energy sector is currently undergoing a significant transformation towards digitalization, driven by the need of decarbonization and sustainable development (Kueppers et al., 2021), (Tsoutsanis & Meskin, 2019), (Farhana et al., 2021), (Huang et al., 2022). For organizations in this sector to thrive in the digital age, it's crucial for their leaders to anticipate market changes and implement flexible procedures while seamlessly integrating cutting-edge technologies.

Digitalization plays a pivotal role in enhancing the security, efficiency, and sustainability of energy systems. By harnessing data collection and analysis technologies, digitization can significantly boost energy efficiency. These data are processed using technologies like artificial intelligence and then transmitted to devices capable of effecting physical changes to optimize energy utilization (Borowski, 2021).

However, there is currently a limited body of research addressing the applications of digital twins within the energy industry. In this study, we aim to explore the main potential uses of digital twins in the energy sector exploring instances where digital twins are already being utilized in the energy industry in order to establish a framework of criteria for evaluating this technology.

Identifying the diverse applications and use cases of digital twin technology within the energy sector is key (**Figure 10**) and will be further discussed in terms of (1) applications of digital twins in energy production, encompassing various forms of energy generation, from fossil fuels to renewables such as photovoltaics and wind turbines, and even nuclear; (2) the utilization of digital twins in the domain of energy consumption, spanning buildings, transportation, and industrial applications; and (3) the deployment of digital twins in the energy storage industry, covering mechanical, thermal, battery, and hydrogen storage devices.

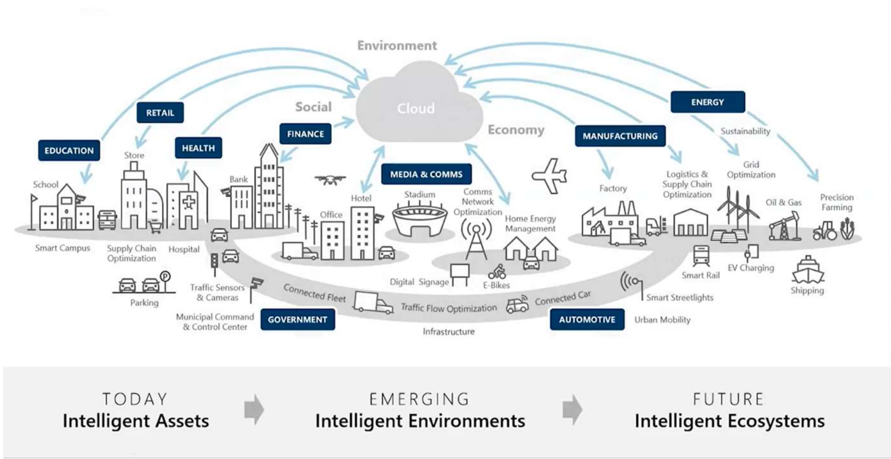


Figure 10. Energy and the Built World

As modern cities continue to evolve, they give rise to intricate energy layers like transportation systems, smart grids, and microgrids, each confronting a multitude of challenges in multi-dimensional energy management. Within power grids, challenges encompass remote data transfer within the grid and various analyses reliant on real data. These complexities can be effectively addressed through the implementation and analysis of a comprehensive digital twin framework within each segment.

Emerging smart technologies bring additional capabilities, including advanced computing and communication, which are challenging to represent in the physical layer alone. These enhanced capacities elevate the sophistication of power systems, necessitating novel approaches to tackle these challenges by providing enhanced insights into the physical infrastructure.

By establishing a robust DT power system, real-time and historical data can be effectively managed in a secure and efficient manner, contributing significantly to system operations by supporting maintenance, design, and operational management.

Despite being a relatively new and complex concept, the exploration of DT-based power systems is still in its preliminary stages. DTs hold promise for diverse applications within power systems, including fault detection, load prediction, operator behavior analysis, power system control and analysis, health condition assessment of power tools, and more.

Given that real-time data on energy infrastructure is often not readily available, DT architectures can play a pivotal role in sustainable dispatching system operation and management. Consequently, DTs facilitate expedited decision-making through comprehensive and optimal management, bridging the gap in energy infrastructure.

The concept of the smart city's DT is rooted in its ability to analyze the dynamics resulting from the intricate interdependencies between infrastructures, including energy systems and various technologies. To realize the smart city's DT, it becomes imperative to deploy DTs across all energy systems within the smart city framework (**Figure 11**). Indeed, energy systems are intricately woven into the fabric of the smart city, playing a relevant role in its functioning and sustainability.

The smart grid, a vital energy infrastructure, utilizes information and communication technology (cyber systems) to facilitate two-way communication for both energy and information transfer. This enhances grid monitoring, control, and consumer engagement. The physical components of the power grid include power generation facilities responsible for electricity generation, transmission facilities for power delivery, and power distribution facilities supplying electricity to consumers.

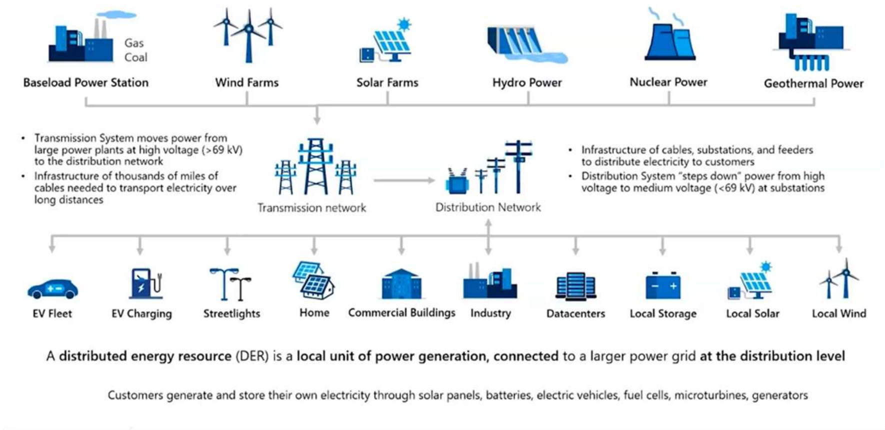


Figure 11. Electric grid assets and infrastructure

The integration of cyber systems has given rise to the smart grid, the next-generation power grid designed to offer more efficient energy services. It encompasses reliable and intelligent distribution management, renewable energy integration, energy storage, grid monitoring, control, and the incorporation of electric vehicles.

The smart grid is essentially a highly distributed grid system, integrating information communication technologies like sensing, networking, data analytics, and machine learning. These technologies enhance grid reliability, efficiency, and security.

The smart grid, being a large-scale distributed system, requires comprehensive knowledge of the physical objects' states, such as voltage and current. To achieve this, a multitude of sensors are deployed in the grid to measure critical object states, providing valuable insights. Static and dynamic state estimation methods, along with bad data detection algorithms, are employed to handle invalid or corrupted measurement data, which can result from measurement errors, sensor failures, or cyber threats. Notably, cyber threats can directly or indirectly impact grid operations by injecting malicious measurement data or manipulating control information.

Addressing cyber threats necessitates monitoring not only the physical objects but also the state of cyber components. Here, Digital Twins offer a promising solution by bridging the physical and cyber worlds and creating a linkage between them.

2.2.1 Data-driven modeling strategies

DT employs a data-driven approach to map the physical grid onto the digital grid, facilitating comprehensive monitoring and management. In the DT architecture discussed earlier, the smart grid system is divided into distinct layers: the object layer, communication layer, application layer, and end-user layer.

In the object layer, grid components encompass power generation, transmission, and distribution facilities. The communication layer connects these objects, facilitating data exchange with the application layer and intra-object layer communication. For instance, the communication layer's gateway stores device information from the object layer.

When data is required by the application layer, the gateway retrieves it from specific devices as needed. Moreover, the application layer can construct a digital model based on device distribution information from the gateway.

The application layer comprises a digital representation of gateways from the physical system, incorporating device information from the object layer. It acts as an intermediary between the application and object layers, with AI models trained using data collected from the object layer and stored in data storage for future utilization. If there are changes in the physical object layer, the application layer can send commands to actuators to adjust grid operations based on real-time data and trained models. Additionally, the AI model can update itself to reflect any changes in the object layer.

The end-user layer leverages information processed by the application layer to offer various services, including smart grid management systems, autonomous vehicle smart charging systems, and hybrid energy management systems.

Numerous research initiatives have explored DT's application in the smart grid (Tzanis et al., 2020c), (Saad et al., 2020a), (Danilczyk et al., 2021a). For example, General Electric (GE) introduced two DT models for wind farms. In the object layer, a communication network enables wind turbines to communicate, with middleware involving cloud-based infrastructure and digital models to collect data and remotely control turbines. A graphical user interface

(GUI) in the application layer provides visualization and control capabilities for wind farm management.

(Saad et al., 2020b) proposed a microgrid DT model employing IoT to counter cyber-attacks. It includes microgrids, local controllers, and area controllers, featuring real-time balancing algorithms to mitigate the impact of attacks.

Similarly, (Tzanis et al., 2020d) applied DT to manage a large number of smart grid devices. They used spike neural networks (SNN) in smart meters to detect fault nodes and a transient state estimator (TSE) to monitor the grid's dynamic state. (Danilczyk et al., 2021b) employed deep learning algorithms to detect physical faults in the smart grid system by analyzing data from the supervisory control and data acquisition system (SCADA).

(Baboli et al., 2020) proposed a DT framework with artificial neural networks (ANN) for distributed smart grids, ensuring real-time model generation, verification, and identification.

2.2.2 Microgrid and complex systems

A microgrid is a self-contained energy system comprised of distributed energy resources and interconnected loads. This versatile asset can operate in either island mode or grid-connected mode, offering flexibility and control (Liang et al., 2019). Microgrids are designed with the overarching goal of enhancing the overall performance of energy systems, focusing on sustainability, security, efficiency, economics, and effective energy management. Key attributes influencing microgrid performance include reliability, security, flexibility, self-sufficiency, and optimality.

While extensive research has been dedicated to improving microgrid performance, the exploration of Digital Twins in the analysis, design, control, and development of microgrids is a relatively novel research area. This section delves into the application of DT services in microgrid contexts and highlights recent studies in this emerging field.

The smart grid, functioning as an energy-based Cyber-Physical System (CPS), represents a complex system characterized by significant uncertainties affecting both its cyber and physical components (Dileep, 2020), (Z. Guan et al., 2015), (H. Xu et al., 2020). These uncertainties can originate from various sources, including grid outages. In essence, the smart grid remains susceptible and vulnerable to random events such as power load imbalances, outages, and even external disruptions from its surroundings. It is paramount to acknowledge the critical significance and inherent fragility of the smart grid, given the unprecedented challenges it confronts.

The evolution of big data and the Internet of Things (IoT) opens up opportunities to address these challenges. The utilization of advanced big data analysis tools such as Machine Learning (ML), Deep Learning (DL), data mining, and statistics holds the potential to predict and anticipate potential risks within the smart grid (Ponnusamy et al., 2021). This predictive capability enables the implementation of proactive measures to mitigate accidents and disruptions effectively.

Moreover, harnessing the power of big data enables the creation of a virtual smart grid environment. This virtual simulation environment can replicate real-world accidents and disturbances, allowing for comprehensive investigation and the development of robust mitigation strategies. In essence, the synergy between big data analytics and the smart grid offers a path towards enhancing resilience and reliability in the face of dynamic and uncertain challenges.

DTs play a pivotal role in microgrid environments, offering diverse capabilities for scheduling, optimization, and planning, especially in scenarios where conducting real-time experiments is impractical. Microgrids encompass renewable energy resources, demand components, and communication network segments, each of which can benefit from dedicated DTs.

To establish a microgrid DT, individual DTs are constructed for each segment. Subsequently, an aggregate DT is formed, comprising DTs from all three segments. This holistic approach enables comprehensive monitoring and control of the entire microgrid.

Notably, DTs are increasingly finding applications across a wide spectrum of industries, including discrete manufacturing, process manufacturing, energy (power), oil and gas, mining and metals, automotive, life sciences and medical, aerospace, infrastructure, and defense. In the energy sector, DTs offer solutions to various challenges, including:

Predicting energy demand for individual consumers using machine learning (ML) approaches within planning and operational DTs.

- Improving grid management and distribution by leveraging real-time data-driven simulation models for distributed energy sources.
- Enhancing maintenance of solar arrays by identifying abnormal behavior for timely repairs.
- Anticipating maintenance needs for wind farms to support service teams effectively.

The integration of DTs into microgrid applications and broader energy systems underscores their potential to drive innovation, efficiency, and resilience in the evolving energy landscape.

The concept of Digital Twins (DT) offers practical and efficient possibilities in real-time data presentation and forecast analysis. Instead of relying solely on historical data, utilizing microgrid digital data outputs can lead to more accurate predictions for energy management and future system responses.

Accurate load growth prediction is crucial for optimizing and making microgrid development programs more efficient, particularly as the development of microgrids depends on future load expansion. The integration of DT can significantly impact long-term microgrid development planning.

Machine Learning (ML) algorithms have demonstrated their prowess in solving prediction and diagnosis problems. Load forecasting, which depends on factors that influence load consumption patterns, includes Short-term forecasting (STLF), Mid-term forecasting (MTLF), and Long-term forecasting (LTLF). Deep Learning (DL) methods have been applied in microgrids for load forecasting, enhancing prediction accuracy.

Network studies encompass a wide array of analyses, including restoration, reliability, prediction, energy hub, uncertainty, and physical and cyber security. Implementing DT-based power grids can enhance network behavior under various conditions, reducing response times from minutes to seconds. Recent studies on power grids utilizing DT techniques are summarized in Table 3.

DT's predictive capabilities are invaluable in planning maintenance processes, enabling simulations for optimal solutions and protection strategies, ultimately streamlining system maintenance.

DT technology facilitates remote distance control and monitoring, enhancing reliability and reducing costs. Distance protection relays and fault location algorithms have been tested using DT, and DT-based distributed networks have been used for troubleshooting in transformer performance evaluation.

Energy hubs, which convert different types of energy in urban systems, require efficient energy consumption management. DT's real-time data control, monitoring, and analysis capabilities make it a promising tool for improving the performance of energy hub systems. Integrating DTs from energy carriers such as water, electricity, and gas into the energy hub's DT aggregation can address energy management challenges effectively.

Successful implementation of DT relies on robust IT infrastructure, as it requires interconnected and powerful systems. While cloud services like Amazon, Microsoft, and Google offer significant benefits, they also present security challenges in data analysis. Edge computing can mitigate data transmission delays, increase bandwidth, and play a pivotal role in preprocessing, storage, and analysis for DT applications.

2.2.3 Energy supply, consumption and storage

Uncertainty is a significant factor in power systems due to unpredictable elements and their random behaviours. Addressing uncertainty is crucial for solving power system problems and achieving effective system control. Real-time monitoring and analysis are essential to manage the uncertain behaviour of random factors in any system.

In power systems, the integration of random entities such as wind turbines and solar photovoltaics, coupled with the unpredictability of consumer consumption, contributes to the inherent uncertainty in power networks and energy systems. Numerous techniques have been proposed to model the random behaviour of uncertainty parameters (D. Silva et al., 2021b). However, accurate modelling and real-time analysis are essential to understand and control the system effectively.

Digital Twins (DT) technology offers valuable support for understanding and managing the random behaviour of physical entities through real-time monitoring, control, and online analysis. With DT technology, the real-time physical behaviour of an agent can be observed continuously, effectively reducing uncertainty. It enables the thorough investigation of real-time physical entities, providing high fidelity and accuracy, which is a desired feature in DT technology to support the uncertainty of physical systems.

The concept of an energy hub involves the simultaneous communication and exchange of energy carriers within a network (K. Yang et al., 2021). Energy hubs serve as converters for various types of energy in different urban system components. These hubs extend beyond the electricity sector, encompassing gas and urban water, and play a pivotal role in managing energy consumption across multiple sectors (Cagnano et al., 2020b). Integrating and efficiently managing energy consumption within these hubs requires systems and approaches that analyze behavioral changes and monitor energy consumption. Real-time access to accurate data is critical for this purpose.

Digital Twins, with their capabilities in real-time data control, monitoring, and analysis, promise enhanced performance for energy hub systems. Integrating initial DTs from energy carriers such as water, electricity, and gas into the DT aggregation of energy hubs can effectively address the challenges of energy management, integration, and exchange.

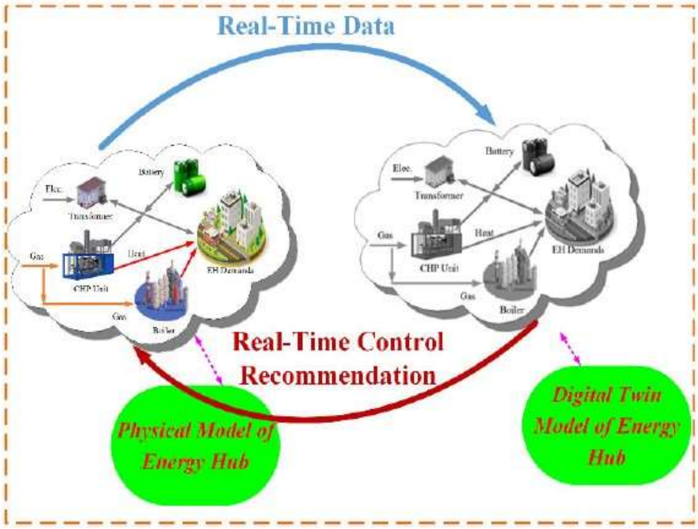


Figure 12. An example of DT concept's Energy Hub

Figure 12 provides an illustrative example of a DT concept applied to an energy hub, highlighting its potential in enhancing the performance and efficiency of energy systems in urban environments.

2.2.4 Load prediction and forecasting

Real-time data availability is a fundamental requirement for various forecast analyses in areas such as load forecasting and system response prediction. The accuracy of forecast outputs is directly linked to the accuracy and real-time nature of the input physical data. Extensive literature has established a direct correlation between prediction accuracy and the use of Digital Twins (DT) services (D. Wu et al., 2019b).

The specificity of DT in presenting real-time data makes it a practical and efficient tool for forecast analysis. Instead of relying solely on historical data, microgrid digital data outputs can be utilized to achieve more precise predictions for energy management and future system responses. Accurate load growth predictions are particularly vital for optimizing microgrid development programs, as the expansion and growth of load in the future are key factors in microgrid planning and efficiency.

Machine Learning (ML) algorithms have demonstrated their robust predictive and diagnostic capabilities. Load forecasting, which depends on various factors that can influence consumption patterns, includes Short-term forecasting (STLF), Mid-term forecasting (MTLF), and Long-term forecasting (LTLF). Deep Learning (DL) methods have been extensively applied in microgrid load forecasting, improving prediction accuracy.

For example, in STLF, Deep Neural Networks (DNN) have been employed (Din & Marnerides, 2017) Feed-forward DNN and Recurrent-DNN models have been compared Fare clic o toccare qui per immettere il testo.(He et al., 2017), and Deep Belief Networks (DBN) combined with parametric Copula models have been suggested for hourly load forecasting (Dedinec et al., 2016). Additionally, DBN composed of multiple layers of Restricted Boltzmann Machines (RBMs) has been used for STLF, fine-tuning parameters through supervised back-propagation training (Wen et al., 2020). DRNN-GRU models have been presented for STLF and MTLF using consumption data (Estebsari & Rajabi, 2020) and CNN has been used for residential load forecasting Fare clic o toccare qui per immettere il testo.(Tong et al., 2018). Stacked Denoising Auto-Encoders (SDAs) have been proposed for electricity load forecasting, with the output data from SDAs used as input for Support Vector Regression (SVR) models (Marino et al., 2016). LSTM methods have been introduced for load demand prediction at hourly and minute ahead levels (Khodayar et al., 2019), and DBN has been used for wind and PV power prediction (W. Wu et al., 2016). A DNN method, including Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN), has been proposed for deterministic short-term wind power forecasting (Francisco et al., 2020).

Integrating these high-capability ML algorithms into a digital environment based on DT enhances the efficiency and accuracy of real-time forecasting and analysis. Adapting prediction models with real-time data ensures the development of accurate prediction models, enabling effective planning and operation of power systems.

2.2.5 Plan and monitoring

Effective energy management within a microgrid necessitates the presence of control systems in each system component. These control systems facilitate the collection and access of data to the central decision-making system, enabling optimal energy management. The installation of measurement systems at various points in the network is essential for data quantification. Control systems and energy management systems rely on measurement data for informed decision-making. However, challenges arise in achieving optimal decisions due to errors in data estimation from measurement devices.

In (R. S. Srinivasan et al., 2020), a framework for energy management based on Digital Twins (DTs) is proposed, where smart meters are employed to collect entity data. This approach leverages DT technology to enable real-time monitoring of renewable energy resources, addressing real-time operational and monitoring challenges (Sivalingam et al., 2018). Similarly, (Ebrahimi, 2019) suggests a DT-based platform for energy demand management.

The utilization of DTs extends to optimizing the management of wind farms, as demonstrated in (Lamagna et al., 2021), where a DT for wind farms is employed for performance evaluation and health management. Additionally, (Steindl & Kastner, 2021) applies DT technology to display, process, and evaluate the performance of energy storage systems within a virtual environment. This technology plays a pivotal role in determining scheduling programs for the operation process.

Ontologies and semantic web methods are also introduced as common techniques for DT modeling (Glatt et al., 2021)]. These methods enable the creation of layered structures for specific smart grid architectures and various energy systems. Ontologies can be used to model customer load profiles, power generation systems, and substructures of generation through organized constructions. Semantic web technologies offer a level of abstraction for energy system simulation and modeling, adaptable to different energy systems (Y. Xu et al., 2019).

Edge-based DTs have emerged to create virtual simulation environments for performance assessment, enhancing the resilience of microgrids (H.-A. Park et al., 2020). These approaches collectively enhance the efficiency and effectiveness of energy management within microgrids, addressing the challenges posed by real-time data processing and decision-making.

2.2.6 Real-time analysis and optimization

To establish a Digital Twin (DT)-driven model, it's essential to consider three key components: physical space, virtual space, and the data interaction between these spaces. This comprehensive model requires the identification of nodes within a specific system, which are then meticulously modelled. The integrated model incorporates various data sources, including historical data and sensor data, to create an accurate representation of the system under diverse states (H. Yang et al., 2019).

Evaluating the states of the DT model over a specified time period allows for the resolution and analysis of the model. This perspective provides a comprehensive understanding of the system's state, facilitating analysis and decision-making. As most models exhibit continuous changes over time, dynamic simulation of the model is imperative. Designing a DT involves various phases, including modeling physical systems, connecting real-time data, and evaluating and adapting the model. A virtual model can be constructed using historical data from different system states, integrated with information based on system dynamics. In this context, three main approaches are employed: physics-based modeling, data-driven modeling, and a combination of both. Physics-based models are constructed using mathematical and physical models, with machine learning techniques such as AI used for parameter estimation when there is insufficient information (Ahmadian et al., 2018), (Danilczyk et al., 2021c). For instance, (Khaled N., 2020) employs an artificial neural network (ANN) to adjust inverter model parameters.

Data-driven models are utilized when it's challenging to formulate mathematical representations of system elements. These models can incorporate extensive historical data, making them suitable for scenarios where a physics-based model may be inadequate. Data-driven models are continuously improved with real-time data to enhance accuracy and alignment with the actual system. It is essential to note that secure and accurate data transmission is a fundamental capability of the DT.

Managing large volumes of data from multiple sources is a challenging task, necessitating advanced data analysis approaches for preprocessing. Data is shared through secure and reliable communication systems, selected based on specific requirements, including quantitative and qualitative aspects. Communication options encompass wired, wireless, and hybrid systems.

Preprocessing and standardizing the raw data are crucial steps to prepare it for specific purposes such as management and control. Big data analytics, often implemented in cloud-based environments, is instrumental in processing large datasets to extract meaningful insights. Cloud computing's significant storage and computational capabilities facilitate various applications, including data mining, dimension reduction, filtering, and processing. Big data analysis plays a pivotal role in DT modeling, particularly in the smart grid and smart city domains (Song, Cai, et al., 2020), (Song, Jiang, et al., 2020). Additionally, it helps integrate continuous and discrete information flows within microgrids. Fog and edge computing can also be deployed to reduce data transfer latency and integrate various services. A case study of artificial intelligence approaches to edge computing is discussed in (Reka & Dragicevic, 2018).

Maintaining the consistency and accuracy of a DT model is a significant challenge, given the continuous shifts in operating states and environmental conditions of physical systems over time. Continuous model updates are necessary, as real-time data is collected through control and monitoring applications and processed using data analysis techniques throughout the system's lifecycle. Evaluating and adapting the model is a critical step in establishing a DT. Parameters can be adjusted, physical configurations optimized, and various stimulus procedures tracked for model adaptation. Changes in observed data over a specific time period can be used to update the model. Advanced machine learning algorithms, including deep learning and reinforcement learning, are effective strategies for these purposes. A survey of machine learning techniques is discussed in (Y. Zhang et al., 2018)

In practice, a real-time DT model can be developed using tools such as Simulink, Simscape, MATLAB, and Raspberry Pi hardware (Carvalho et al., 2020). The desired model is deployed on Raspberry Pi hardware, which acts as a real-world representation of the entity being modelled. Raspberry Pi hardware communicates the system's state and input/output data to cloud services like Amazon AWS, enabling the creation of a parallel DT model in the cloud. This cloud-based DT model, utilizing the same inputs as the physical model, conducts real-time, accurate, and efficient system diagnosis (Carvalho et al., 2020). The hardware's Wi-Fi capability facilitates seamless communication with the cloud to share the physical states of entities implemented in the hardware environment. This synchronized approach ensures that the DT model remains consistent and up-to-date with the physical system it represents.

2.2.7 Diagnostic systems and fault detection

Faults in the microgrid environment are inevitable and can pose significant challenges to the grid's operators and control systems. These faults, if left undetected or unaddressed, can lead to widespread issues within the entire system. Therefore, the timely detection of faults is crucial to prevent irreparable damage. A fault detection system must possess the capability to detect faults promptly, determine their time and location, and initiate fault interruption procedures promptly. To achieve this, access to real-time data is essential. Real-time analysis plays a relevant role in addressing these challenges by enabling the detection of faults as they occur. Implementing a Digital Twin (DT) of the microgrid allows for real-time analysis of the entire network, contributing to the stability, reliability, and flexibility of the microgrid.

In (Jain et al., 2020), researchers investigate the development of DTs to facilitate fault detection. A fault detection system for distributed energy resources, based on the DT concept, is developed in (Tzanis et al., 2020e). (Palensky et al., 2022) proposes a monitoring system that relies on DT technology to predict faults in a physical power converter, integrating it with a digital model controlled by a controller. Additionally, (Joseph et al., 2018) presents a digital model of a power system to track dynamic voltage faults and forecast post-fault dynamic behavior.

The concept of DT is harnessed in (Bazmohammadi et al., 2022b) for controller design and distributed energy resources. In (Goia et al., 2022), DT technology is developed and deployed to evaluate the operation of microgrid controllers. These initiatives leverage DTs to enhance fault detection and overall grid performance, ensuring the microgrid's reliability and stability.

2.3 Digital Twins and Indoor Air Quality (IAQ)

Enhancing the comprehension of indoor air quality (IAQ) and the factors that influence it can lead to improved management of indoor environmental quality (IEQ), resulting in decreased health hazards and enhanced well-being for occupants. The domains of energy, health, and economy are closely interconnected with the IAQ concept, particularly concerning aspects such as air ventilation, public health, and productivity.

Based on extensive reports from scientific sources, the proportion of time during which individuals are exposed to indoor environments (such as residential buildings, workplaces, vehicles, public transportation, and public facilities like schools, hospitals, museums, theatres, and libraries) exceeds 80% in developed nations, over 87% in the United States, and around 85-90% in Europe. These estimates could be even higher due to the emerging COVID-19 pandemic and the resulting changes in lifestyle, such as lockdowns and remote work. Consequently, it is of utmost importance to thoroughly consider the implications of indoor air quality (IAQ) on human health and overall well-being.

The study and characterization of IAQ, coupled with the growing attention and practical initiatives to establish healthy and comfortable indoor environments, have the potential to significantly enhance the quality of life and productivity for occupants. Notably, exposure to indoor air pollution can surpass outdoor air pollution exposure by more than twofold. Shockingly, it's estimated that a global population of approximately 3 billion individuals experiences inadequate IAQ levels on a daily basis.

Addressing IAQ concerns is crucial for promoting public health and well-being Exposure to poor indoor air quality (IAQ) can lead to a range of health issues, including irritation, allergic symptoms, impaired cognitive abilities, reduced productivity, dizziness, headaches, restlessness, asphyxia, coma, cancer, and even death. Various scientific studies have linked inadequate IAQ to these adverse health outcomes (World Health Organization. Regional Office for Europe, 2010).

2.3.1 IAQ and External Ventilation Flow Rate evaluation

Indoor Air Quality (IAQ) refers to the quality of the air within buildings and structures, specifically how it affects the health and comfort of the occupants. External air flow rate, often referred to as ventilation, is a critical factor that directly influences IAQ.

In summary, the connection between IAQ and external air flow rate is crucial. Proper ventilation and fresh air intake directly impact the quality of the air indoors by diluting pollutants, controlling CO_2 levels, preventing moisturerelated issues, removing odors, and enhancing overall comfort. To maintain a healthy indoor environment, it's essential to design and operate ventilation systems that provide an appropriate external air flow rate based on the specific requirements of the space and its occupants.

The absence of explicit and detailed legal directives pertaining to Indoor Air Quality (IAQ) can largely be attributed to the inherent variability, diversity, and challenges associated with gathering consistent analytical data for all the factors and sources involved. IAQ is a complex concept that encompasses a wide range of factors rather than being confined to a single measurement element. Furthermore, there are numerous architectural and design variables that are interconnected with IAQ. Complicating matters further, the conflicting energy efficiency goals across various parameters, as discussed earlier, contribute to the intricate, intertwined, and adaptable nature of IAQ on both spatial and temporal scales.

To address this complexity, the development of IAQ indexes becomes crucial. These indexes serve to explain, categorize, and enhance the quality of indoor air by providing user-friendly and comprehensive scores (rankings) of IAQ levels within indoor environments. Despite the global development of various IAQ indexes in recent years, their specific relevance to assessing IAQ levels hasn't been fully explored. Aiming for a comprehensive and expedient assessment of IAQ through an index can streamline the creation of effective measurement, qualification, and maintenance protocols for control purposes. However, choosing an appropriate evaluation measure for this purpose presents a significant challenge due to the indexes present in the scientific literature. Moreover, the distinction between health risk-based and comfort-based indexes remains unclear.

2.3.2 IAQ measurement and protocols

The primary objective of this review is to comprehensively identify and categorize the existing IAQ indexes on a global scale (World Health Organization. Regional Office for Europe, 2010) as a specific regard to buildings energy efficiency strategies.

Contemporary comprehensive methodologies and standards, exemplified by Standard NBN EN 16798–1:2019, are being progressively implemented to integrate building energy efficiency with indoor environmental conditions (Olesen B. W., 2012). The growing trend of constructing tightly sealed dwellings has incentivized architects and building companies to enhance the performance of ventilation systems and bolster overall energy efficiency. However, the tension between achieving "energy efficiency improvements" and adhering to "IAQ guidelines" necessitates the creation, advancement, and optimization of multifaceted strategies for indoor air purification.

The pursuit of increased air exchange rates (AERs) to enhance indoor air quality leads to higher energy consumption by ventilation systems, thereby potentially diminishing building energy efficiency. This presents a challenge that requires innovative solutions for providing adequate ventilation while minimizing energy consumption. Several factors come into play, including IAQ standards, the type of ventilation systems, and occupants' activities. This duality of interest arises between policies aiming for high IAQ and those emphasizing reduced building energy use. The design of modern indoor environments must address these conflicting demands to effectively tackle emerging challenges and requirements (Šujanová et al., 2019)

Notably, the European Union's policy-making body, through the revised Energy Performance of Buildings Directive (EPBD, 2018/844), has mandated that energy performance criteria set by EU member states' executive administrations should optimize health, indoor air quality, and comfort measures. To facilitate the transition towards an energy-efficient and decarbonized building stock by 2050, EU governments must adopt a comprehensive approach encompassing these pivotal factors (Buildings Performance Institute Europe (BPIE), 2019): incorporating indoor environmental quality (IEQ) measures into long-term renovation strategies; integrating IEQ considerations with Energy Performance Certificates (EPCs); formulating strategies that are cost-optimal and evaluating the various factors influencing IEQ; ensuring certification, agreement, and quality control measures to actively support the provision of acceptable IEQ.

Numerous indoor environments rely on mechanical ventilation systems that introduce a limited amount of outdoor air, which can lead to the accumulation of indoor pollutants. According to the ASHRAE standard (ANSI/ASHRAE Standard 62.1-2022, 2022), three distinct approaches—namely, Ventilation Rate, IAQ, and/or Natural Ventilation are employed to meet ventilation criteria.

In the ventilation rate procedure with prescriptive method, ventilation rates are predetermined based on building usage, the number of occupants, and floor area. On the contrary the IAQ procedure in a performance-based approach defines outdoor air intake rates and other parameters considering pollution sources, Exposure Limit Values (ELVs), and the perceived acceptability of indoor air quality. A procedure involving natural ventilation is based on the inflow of outdoor air through openings into indoor spaces and can be used in conjunction with mechanical ventilation systems. However, contemporary energy-efficient designs, characterized by reduced air leakage and tightly sealed constructions, can substantially curtail natural ventilation. This shift towards high-sealed indoor environments can lead to the accumulation of indoor pollutants due to inadequate air exchange rates (Godish & Spengler, 1996)

It's important to note that the recommended minimum levels for indoor air quality indicators may not always correspond to optimal levels. For example, the (EN 16798-1:2018, 2018) standard specifies an absolute minimum ventilation value of 4 liters per second per person, while scientific research often proposes values of 6 to 7 liters per second per person or even higher when additional parameters such as productivity and learning are taken into account.

As policymakers increasingly prioritize concepts like "energy transition," "carbon neutrality," and "net-zero emissions," regulations and policies addressing outdoor air pollution are being developed and implemented more rapidly than those focused on indoor environments. Additionally, changes in inhabitants' behaviours and lifestyles due to climate change can also influence indoor pollutant concentrations (Vardoulakis et al., 2015). Consequently, there is a growing recognition of the need to address the impact of indoor pollution sources on indoor air quality (IAQ).

While controlling the emission sources to reduce indoor contaminants is a viable approach when sources are well-known, it's important to note that new substances are continuously being identified as harmful to health. Therefore, the prevention or reduction of indoor contaminant emissions often has limitations, may be technically challenging, and might not always be cost-effective (Kwok et al., 2022). In this context, the role of ventilation emerges as critical in effectively controlling and maintaining good IAQ. Ventilation plays a pivotal role in ensuring that indoor air remains free from pollutants and meets acceptable quality standards.

The current standards and guidelines for Indoor Air Quality (IAQ) aim to establish recommended concentration levels, values for indoor climate parameters, and appropriate air ventilation requirements. These guidelines serve as the scientific foundation for legally enforceable standards. The recommended levels of pollutant concentrations are described using various terms, such as, Exposure Limit Values (ELV)[6], Threshold Limit Value (TLV)[7], Lowest

[6] Exposure Limit Values (ELV): These are the levels of pollutants considered acceptable over a specified period of time. ELVs also serve as benchmarks to assess whether IAQ is improving or deteriorating.

[7] Threshold Limit Value (TLV) the maximum average airborne concentration of a hazardous material to which healthy adult workers can be exposed during an 8-hour workday and 40-hour workweek.

Concentration of Interest (LCI)[8], Toxicity Reference Value (TRV)[9], and Occupational Exposure Limit (OEL)[10].

The definitions below provide a clear understanding of the terminology used to describe pollutant concentrations and their associated regulatory measures in the context of IAQ standards and guidelines (UK Department for Environment Food & Rural Affairs, 2010).

The primary goal of this chapter is to provide a comprehensive overview of various IAQ indexes with different applications. This compilation and classification aim to be a valuable resource for professionals working in diverse fields, including Building Energy Management Systems (BEMS), HVAC design, architecture, indoor air monitoring device manufacturing, indoor air purifier production, healthcare, and various scientific and engineering disciplines.

The literature presents several perspectives for classifying IAQ indexes. Although researchers agree that IAQ indexes can be categorized based on both subjective and objective principles, they can be further classified based on their intended application.

A classification is focused on specific application: IAQ indexes can be categorized based on their intended application. This includes health-related indexes that focus on pollutants and factors affecting human health, comfort-related indexes that assess indoor conditions for occupant comfort, and energy-related indexes that consider the energy efficiency aspects of ventilation and air exchange.

The main classification is Subjective vs. Objective. Subjective evaluation of indoor air quality involves assessing occupants' perceptions and experiences related to the quality of the air they breathe. This assessment is often gathered through surveys, questionnaires, and feedback from building occupants. The focus is on how individuals feel about the indoor environment, whether they are satisfied or dissatisfied, and whether they find the air comfortable to breathe. Subjective evaluation takes into consideration factors such as odours, stuffiness, humidity, and overall comfort. This approach provides valuable insights into occupants' experiences but may not provide quantitative data about pollutant concentrations.

---

[8] Lowest Concentration of Interest (LCI): The maximum acceptable pollutant levels that limit emissions from construction products before they are used by end-users. This is meant to control emissions from building materials and products.

[9] Toxicity Reference Value (TRV): A toxicological indicator used to assess or quantify the risk to human health posed by exposure to pollutants.

[10] Occupational Exposure Limits (OELs): Directing values for chemical substances present in workplace air. These values are designed to ensure the health and safety of workers with regard to their exposure.

Objective evaluation of indoor air quality is centred around quantifiable measurements of pollutants and parameters that directly influence air quality. This approach involves using scientific instruments and measurements to directly assess the concentration levels of various pollutants present in the indoor environment. These pollutants may include volatile organic compounds (VOCs), particulate matter (PM), carbon dioxide (CO_2), formaldehyde, and more. The objective evaluation also involves comparing pollutant concentrations to established Exposure Limit Values (ELVs) or guidelines to determine if the indoor environment meets recommended safety standards. This approach provides concrete and measurable data about the actual state of indoor air quality.

In summary, subjective evaluation focuses on occupants' feelings and experiences, while objective evaluation involves concrete measurements of pollutants and adherence to health-related standards. Both approaches contribute to a comprehensive understanding of indoor air quality and guide efforts to improve the indoor environment for the well-being and comfort of occupants.

The IAQ indices highlighted in this analysis are those primarily reliant on measurements of concentrations of three parameters (airborne pollutants). The following two tables present, in chronological order, the literature referenced IAQ indexes related to residential and tertiary environments and industrial buildings, respectively.

Table 4. Residential IAQ indexes

Author, Year	Index	IAQ indicators* (parameters)
(Fanger, 1988)	The perceived air quality is measured in decipol (dp) One dp is the perceived air quality (PAQ) in a space with a sensory load of one olf (one standard person) ventilated by 10 L/s. Percentage of dissatisfied: PD 395 $\times$ exp($-3.25C^{0.25}$) for C $\leq$ 3.13 dp PD = 100% for C > 3.13dp	VOCs, CO_2
(M. H., Shi & Tao, 2000)	Air Quality caused Percentage Dissatisfied index (QPD) = exp($5.98 - \frac{112}{c}$). C=decipol	CO (mgm^{-3}), CO_2 (%), Bacteria number ($1m^{-3}$)
(Jokl, 2000)	$L_{odour(CO2)} = 90 \times \log\left(\frac{\rho iCO2}{485}\right),\ \Delta\rho CO2 = 167350\ (\ln(PD) - 5.98)^{-4}$	CO_2, TVOC

	$L_{odour(CO2)} = 50 \times \log(\frac{\rho iTVOC}{50})$, $\Delta\rho TVOC = 46000\ (\ln(PD) - 5.98)^{-4} - 10$	
(Ribéron, n.d.)	Air stuffiness index (ICONE) = $(\frac{2.5}{log(2)})$ $log(1+f_1+f_2)$ $f_1 = \frac{n1}{n0+n1+n2}$ $f_2 = \frac{n2}{n0+n1+n2}$ n_0 between 0 and 1700 ppm, n_1 between 1000 and 1700 ppm, n_2 greater than 1700 ppm	CO_2
(Balocco et al., 2014)	Air change efficiency (ACE) $ACE = \frac{Vtv/Vvent}{Tzj}\ x\ 100,$ V_{tv} is the total volume of the room, V_{vent} is the mass flow rate of incoming ventilating air, *tzj* is the average value of *mean age of air* in different zones, Local Air Change Efficiency (LACE) = $\frac{Vtv/Vvent}{T}$ x 100, τ is the *mean age of air,* Ventilation Effectiveness (VE) = $\frac{c_e-c_s}{Cz_j-c_s}$, c_e is the concentration of contaminants at the exhaust point, Cz_j is the mean value of contaminant concentration within a specific zone, C_S is the contaminant concentration at the air inlet Contaminant Removal Effectiveness (CRE) = $\frac{c_e}{Cz_j}$	CO_2, PM
(M. Wang et al., 2013)	Pi = $\frac{c_i}{S}$ where Pi was the pollution index for the *ith* source location, Ci is the value of pollutant concentration for the *ith* source location, and S is the standard for indoor air quality.	HCHO, $PM_{2.5}$
(Koufi et al., 2017)	IAQ = $\frac{C - C_{out}}{Cth - C_{out}}$ Average concentration in the interior (C), the concentration of extracted air (C_{out}), the concentration "threshold" (C_{th}).	CO_2
(Gugliermetti & Astiaso Garcia, 2018)	Air Quality Index (AQI) = (1- $\frac{\sum_{i=0}^{N}(a_i \times c_{gas})}{N}$ x 100, $\alpha_{i=} \frac{c_i, limit}{c_{max}}$ Where $c_i, limit$ is the regulatory limit concentration for the ith substance, Cmax is the higher regulatory limit concentration among the analyzed gases, α_i is the weight coefficient for the ith substance.	H_2S, CO

Table 5. Tertiary and industrial buildings IAQ indexes

Author, Year	Index	IAQ indicators* (parameters)

(Cariou et al., 2005)	GAPI (Global Airborne Pollutant Indicator) = Σi Wi Ci (mean pollution) GAPIn = $\sum_{t=1}^{n} \frac{GAPI_t}{n}$ $W_{i=\sqrt[2]{X_i Y_i}}$; $X_i = \frac{V_i}{V_{ref}}$, V_i is the volume of molecule i	VOC_s
(Sarbu & Sebarchievici, 2013)	The olfactory pollution degree of a room: $C_i = C_P + 10 \frac{G}{L_P}$; C_i is the indoor air quality in decipol (dp), C_P is the outdoor air quality in dp, G is the contaminant concentration of the room air in olf, L_p is the outside airflow rate, in l/s. S = kCβ S is odorant intensity (magnitude); C is the odorant concentration in ppm, β is the exponent (0.2 – 0.7) of psychophysical function, k is the constant characteristic of material. (Fanger)PPD = 395 exp (-3.66Lp0.36) for Lp ≥ 0.332 l/s, PPD = 100, for Lp < 0.332 l/s.	CO_2, odor
(Rojas, 2016)	Integral evaluation method $I_i = \frac{\int_0^T (C_i(t) - EL_{\ \ lower})dt}{ELV_{upper} - ELV_{lower}}$ where $C_i(t)$ is the concentration of pollutant i, ELV_{lower} and ELV_{upper} are the lower and upper exposure limit values, t is the time and T is the occupancy period	CO_2, TVOC and RH
(Piasecki & Kostyrko, 2020)	$IAQ_{index} = (100 – PD_{IAQ(CO2)})$; $PD_{IAQ(CO2)}$ =395. exp (-15.15 $C_{co2}^{-0.25}$), $PD_{IAQ(CO2)}$ = 407. Exp (-15.05 (-15.15 $C_{co2}^{-0.25}$) $IAQ_{(OI)index}$= (100 – $PD_{IAQ(OI)}$); $PD_{IAQ(OI)}$= $\frac{1}{1+\frac{1}{\exp(2.14 . OI - 3.81)}}$, OI: Odour Intensity When the indoor environment is hot and humid (value of air enthalpy h>55 kJ/kg) and IEQ sub-component of $IAQ_{(h)}$ is introduced in addition to the sub-component $IAQ_{(CO2)index}$ $IAQ_{(h)index}$ = (100 – $PD_{IAQ(h)}$), $PD_{IAQ(h)}$ = $\frac{100}{1+exp(-3.58+0.18(30-t_a)+0.14(42.5-0.01\ \ _v)}$ where t_a is the air temperature within the tested range from 20 to 29°C and p_v is the partial pressure of water vapour within the tested range from 1000 to 3000 Pa.	Odour, RH, CO_2
(Y. Chen et al., 2018)	DALY index and the $PWE_{p,y,h,f} = \frac{1}{P_{p,y}}$ Σ_j ($EXP_{j,f,h}$ x $P_{p,y,j}$). Population-weighted annual mean exposure to $PM_{2.5}$ (PWE): Where $P_{p,y,j}$ is the size of subpopulation j in the province (or country), p and year y.	$PM_{2.5}$

	$EXP_{j,f,h} = \Sigma_k (t_{j,k,h} \times C_{f,k,h})$. $EXP_{j,f,h}$ is the daily exposure of subpopulation *j* using fuel *f* for heating or non-heating season, $t_{j,k,h}$ is the proportion of time a subpopulation *j* spent in microenvironment *k* in a heating or non-heating season (h), $c_{f,k,h}$ is the area concentration of $PM_{2.5}$ in microenvironment *k* in a heating or non-heating season *h* in a household using fuel type *f*.	
(T. Zhang et al., 2019)	$P_i = \frac{C_i}{S}$ Where P_i was the pollution index for the *i*th source location, C_i is the value of pollutant concentration for the *i*th source location, and S is the standard for indoor air quality	HCHO, $PM_{2.5}$

2.3.3 Demanded Controlled Ventilation (DCV)

Demanded Controlled Ventilation is an advanced system for mechanical controlled ventilation, which refers to a system in which airflow within a building is actively managed and controlled by mechanical equipment such as fans, blowers, and air handling units (AHU). This type of ventilation is distinct from natural ventilation, which relies on natural air movement driven by wind and temperature differences.

In mechanical controlled ventilation, the introduction of fresh outdoor air and the removal of indoor air pollutants are regulated by mechanical systems. These systems are designed to ensure that a sufficient amount of outdoor air is brought into the building while simultaneously expelling indoor pollutants, excess humidity, and odours. Mechanical ventilation can be essential for maintaining good indoor air quality (IAQ) in spaces that are well-sealed, heavily insulated, or located in areas with adverse outdoor air conditions.

Mechanical controlled ventilation systems can be designed with various levels of complexity, incorporating features like variable airflow rates, humidity control, and pollutant filtration. The goal of mechanical ventilation is to create a healthier and more comfortable indoor environment by ensuring a steady supply of fresh air and effective removal of indoor pollutants.

There are various models of ventilation (**Figure 13**), including:

- Supply Ventilation: fresh outdoor air is supplied to the indoor spaces through mechanical systems. This method helps ensure a constant supply of fresh air, but it may not effectively expel indoor pollutants.
- Exhaust Ventilation: indoor air is expelled from the building using mechanical systems, creating negative pressure that draws in outdoor

air through natural leakage points or intentionally provided intake vents.

- Balanced Ventilation: this approach combines both supply and exhaust systems to maintain a balance between indoor and outdoor air. it seeks to provide controlled and consistent ventilation without creating significant pressure imbalances.
- Heat Recovery Ventilation (HRV) and Energy Recovery Ventilation (ERV): these systems capture the heat or energy from the exhausted indoor air and transfer it to the incoming outdoor air. This helps improve energy efficiency while maintaining ventilation.

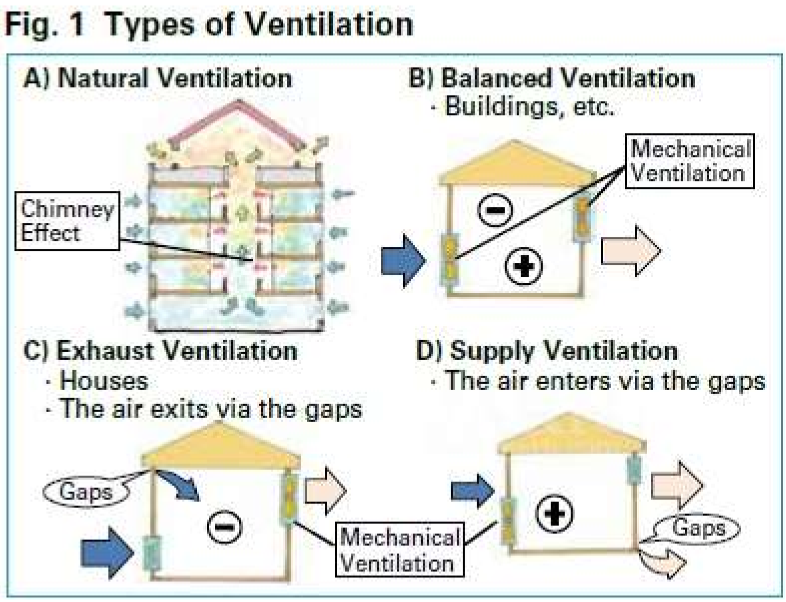


Figure 13. Types of ventilation

The two foremost priorities for the future of the construction industry are undoubtedly ensuring the healthiness and comfort of indoor spaces while simultaneously contributing to sustainability. It is for this reason that the combination of ventilation and energy efficiency emerges as the central emphasis with which new constructions should be designed, as well as the renovation of pre-existing structures.

Ventilation refers to the process by which indoor air is renewed within a building. Traditionally, ventilation was manual and achieved by opening doors or windows to allow outdoor air to enter. While this approach might be effective, it contradicts the goal of energy efficiency, as experience has shown that the periods of open windows can be lengthy, resulting in significant energy consumption. As energy conservation is no longer an option but a mandate embedded in European regulations concerning energy efficiency in the construction sector, the EU's objective is to achieve complete decarbonization of the building stock by 2050.

It is precisely in the need to ensure sustainability and health that controlled mechanical ventilation becomes crucial in buildings. This entails the installation of systems that regulate the inflow of external air without causing substantial thermal imbalances and without impacting energy consumption. In essence, energy efficiency is maintained without compromising air quality and thermal comfort.

There are two ways to implement a mechanical ventilation system: dingle-flow where the fan is used to expel stale air, while fresh outdoor air enters naturally through grilles; dual-flow system which involves two fans, one for extracting stale air and another for introducing clean air. The dual-flow system allows for precise control of air flow and is therefore the most efficient in terms of energy efficiency.

The reference parameter within which the architectural project operates is that of green buildings or passive constructions, so in this new ecological model of sustainable buildings, where airtightness is a priority, natural ventilation is no longer a viable option.

The difficulty in controlling airflows in this type of solution, with the risk of high energy consumption and poor air quality, has led to the recognition that natural ventilation is no longer sufficient or suitable.

Conversely, controlled mechanical ventilation systems provide a triple guarantee: proper air exchange, maintenance of thermal comfort, and energy efficiency. Without a doubt, the dual-flow system is the most comprehensive option, as it offers a significant enhancement: the possibility of incorporating a heat recovery unit, both in domestic environments and industrial buildings, as well as the filtration of outdoor air and improved sound insulation for the building.

DCV management strategies are commonly founded on occupancy-related control markers. Nevertheless, approaches encompassing the regulation of pollutants not contingent upon occupancy are demanded and more viable with the emergence of economical sensors (ES).

In the following, some techniques for the enhancement of demand-controlled ventilation (DCV) through the utilization of measurements of IAQ parameters using some correlation assessments are introduced. The objective is the study of the effectiveness of DCV in curtailment of yearly energy consumption and the proportion of instances when indoor air quality (IAQ) parameter concentrations exceeded specified limits (Justo Alonso et al., 2023).

The pursuit of reducing uncontrolled air leakage has highlighted the necessity for ventilation systems employing demand control to efficiently ensure satisfactory indoor air quality (IAQ) (Mysen et al., 2005). Notably, in countries like China, Canada, and the USA, the recirculation of return air is a cutting-edge practice. Here, the minimum outdoor air (OA) fraction is influenced by both IAQ standards requirements and the aim to decrease heating, cooling, and dehumidification loads on air handling unit (AHU) coils. However, insufficient OA fractions in conjunction with tightly sealed buildings could potentially compromise IAQ (Nazaroff, 2013). Airborne pollutants that would otherwise be ventilated might be recirculated within the space (**Figure 14**), and environmental parameters such as temperature and humidity could escalate to undesirable levels. Simulations have indicated that well-regulated recirculation of a portion of return air can provide protection against outdoor pollutants and lead to reduced annual energy consumption (Justo Alonso, Dols, et al., 2022). This is primarily due to the relationship between indoor and outdoor pollutant concentrations, which is contingent on OA supply and filter efficiency (Ekberg, 1994).

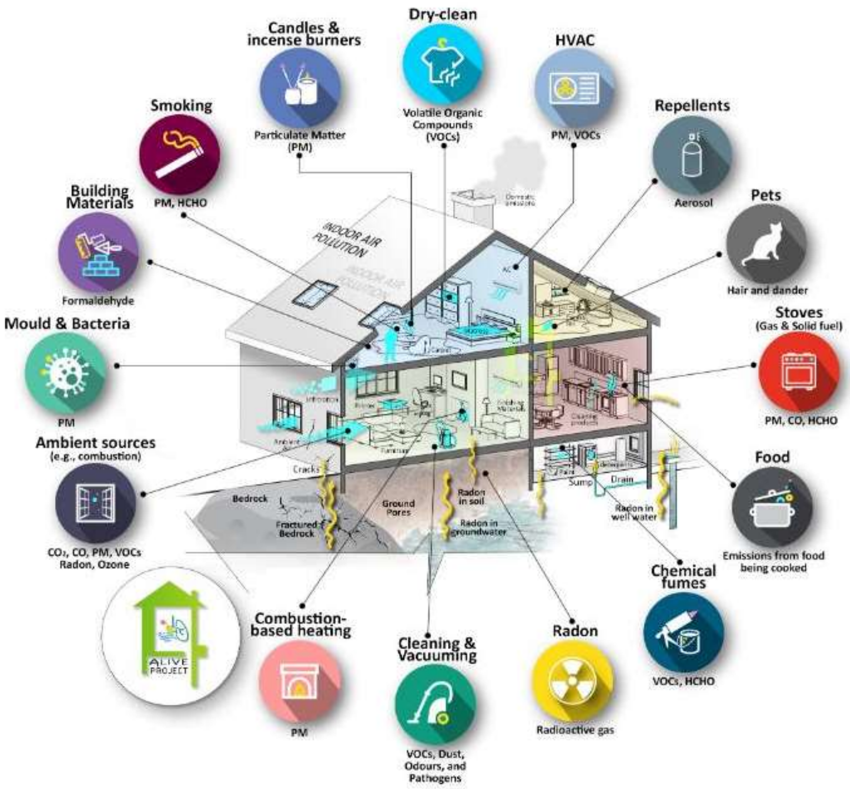


Figure 14. Source of indoor air pollutans

IAQ criteria in most countries are rooted in health impacts and perceptions of IAQ quality, (Scutaru & Witterseh, 2020). However, some pollutants like nitrogen oxides, sulphur oxides, ozone, particulate matter, and formaldehyde, are infrequently measured due to the expense of traditional measurement equipment. As a result, up until recently (Ma et al., 2021), the literature has predominantly referred to IAQ measurements as CO_2 levels, temperature, and occasionally relative humidity (RH). Manufacturers have endeavoured to bridge this gap by enabling the measurement of health-related pollutants alongside standard measurements through the integration of low-cost sensors (ES), defined in this context as sensors costing less than € 50,00.

While these sensors provide cost-effective measurement options, their accuracy often falls below that of reference equipment. Standardized communication between ES and conventional ventilation control systems is not yet established. While this may currently contribute to more complex systems, future enhancements are anticipated (Chiesa et al., 2019). Additionally, ES could contribute to reducing embodied CO_2 emissions, as many of these sensors can communicate wirelessly, facilitating emission reduction.

CO_2 serves as a reliable indicator for bio effluents (Morawska et al., 2021), making it commonly used alongside temperature to regulate ventilation through demand control (DCV). However, over 50% of pollutants in office settings are not emitted by occupants, and in residential spaces, NO2, CO, PM10, and PM2.5 may hold significant relevance. Furthermore, CO_2 does not necessarily correlate with other prevalent IAQ pollutants (Gonzalo et al., 2022), leading to potential misrepresentation if used as a proxy. Ratios of indoor to outdoor pollutant concentrations (I/O ratios) are influenced by factors like the season, building tightness, installed filters, and the specific pollutants being considered (Majd et al., 2019). Several studies have concurrently measured elevated levels of "other" pollutants alongside low CO_2 concentrations during occupancy (Choe et al., 2022). As a result, some researchers advocate using CO_2 to indicate occupant-related pollutants (Maddalena et al., 2015) while others recommend monitoring additional parameters (Society's Indoor Carbon Dioxide Position Document Committee, 2022). This is particularly important since CO_2 and temperature might not adequately detect airborne pollutants with more significant health, comfort, and productivity implications. Thus, establishing a selection protocol for essential airborne pollutants becomes imperative due to limited knowledge regarding ventilation control based on multiple IAQ parameters.

While demand control ventilation (DCV) and economizers have found common applications in office settings, literature concerning the utilization of economizers and heat recovery systems was notably absent. Similarly, limited research focused on the optimization of DCV or the recirculation of return air, considering broader perspectives encompassing energy consumption and IAQ beyond the scope of CO_2 and temperature. HVAC control sequences specified in ASHRAE Guideline 36, primarily centered around occupancy, CO_2, and temperature. The utilization of these sequences resulted in energy savings of 31% (K. Zhang et al., 2022). However, these sequences are deterministic and overlook the impact of modulating outdoor air (OA) on airborne pollutants beyond the preselected parameters. Other studies have tackled different facets of control. For example, research has assessed real-world performance (Afroz et al., 2020a), forecasted pollutants (Kallio et al., 2021), optimized ventilation control and evaluated simulation strategies (W. Wang et al., 2020).

A recent position paper by ASHRAE (Society's Indoor Carbon Dioxide Position Document Committee, 2022) urged the exploration of "Strategies for DCV using CO_2 and other indicators of occupancy that overcome limitations of current approaches and control contaminants that are not linked to occupancy." Nevertheless, the paper did not provide guidance on the selection and utilization of these additional non-occupancy-related indicators in DCV sequences.

To determine which measurable contaminants not tied to occupancy are essential for controlling ventilation, correlation analysis can be employed. Correlation analyses are pragmatic, as utilizing one of the correlated parameters in the control logic would be sufficient to represent the entire set of correlated parameters and regulate the supplied airflow rate (Justo Alonso, Wolf, et al., 2022). In the literature, Pearson and Spearman correlation coefficients are commonly adopted for correlation analyses [H.W. Davies et al]. However, these analyses focus on concurrent correlations and do not delve into the cause-and-effect relationship between variables over time. Cross-correlation functions (CCFs) offer a solution to this limitation [M. Justo et al], revealing correlations even when they exhibit temporal shifts. CCFs calculate Pearson correlations across simultaneous and time-shifted lags.

The incorporation of new or additional parameters into ventilation control strategies can be intricate and challenging. Due to their distinct origins, these parameters exhibit diverse emission profiles and strengths, which might yield conflicting control feedback. Validated simulations can streamline the trial-and-error process for ventilation controls. Numerous simulation programs are available for assessing DCV strategies, such as EnergyPlus, TRNSYS, CONTAM, and Modelica. However, most of these programs cannot concurrently simulate all aspects encompassing energy usage, airflow, IAQ, pollutant sources, and HVAC controls. Consequently, the existing simulation literature has largely focused on energy or IAQ in isolation, but none have comprehensively explored the simultaneous effects of extended IAQ (including CO_2, temperature, and several other airborne pollutants) and energy usage when utilizing DCV and recirculating return air. This weakness can be addressed through co-simulation. Co-simulation between EnergyPlus and CONTAM or CONTAM and TRNSYS can be employed when evaluating all the parameters mentioned above. Such co-simulation enables the evaluation of overall energy consumption, airflow, and IAQ in a building, and both tools are freely accessible.

The research gap that this review addresses is the objective of refining ventilation control logic to curtail annual energy consumption and minimize the instances in which CO_2, temperature, several other airborne pollutants, and RH exceed the thresholds defined in existing literature. This work serves as a pivotal step towards redefining ventilation control paradigms and accommodates the anticipated integration of various pollutant measurements facilitated by the proliferation of ES in the market.

2.3.4 IoT sensors and connectivity

In accordance with ASHRAE Standard 62.1–2019, satisfactory indoor air quality entails "air containing no recognized impurities at detrimental levels, as determined by knowledgeable authorities, and to which a substantial majority (80% or more) of exposed individuals react without discontent". Nevertheless, several real-world investigations have revealed that structures frequently fail to fulfil even the minimal standardized requisites.

The European Respiratory Society (ERS) has pinpointed particulate matter (PM2.5 and PM10), volatile organic compounds (VOCs), and carbon dioxide (CO_2) as pivotal air pollutants.

To ensure adequate management oh HVAC systems, the monitoring of IAQ holds a crucial role, which can trigger a suitable sequence of actions through real-time feedback to encourage human interventions or by directly activating automated control systems. To address these challenges, forecasting the Air Quality Index (AQI) assumes a pivotal role in managing and curtailing air pollution levels.

IoT represents a wireless network of intelligent sensors linked to the Internet, capable of gathering and transmitting data through embedded devices. Typically, IoT devices consist of five fundamental components: sensors, a processor, a memory module, a communication module, and a power source. The sensors are usually connected via a gateway, which establishes communication with the sensors while also providing storage and processing capabilities. This gateway may be hosted either in the cloud or on the edge.

The generation of time-series data from IoT devices, including sensors, machinery, and robotics, is gaining popularity. These data are swiftly generated by real-world applications such as monitoring air pollution (Joshi, 2008) Subsequently, they are transmitted to a cloud or an edge processing center for further analysis.

Traditionally, the monitoring of indoor air quality has been carried out by experts employing certified reference instruments [N. Castell et al]. However, the high initial cost and large size of such equipment make them unsuitable for widespread and continuous IAQ monitoring in buildings (Castell et al., 2017).

Recent technological progress in metal oxide semiconductors (MOS) for gas detection light scattering for particles, and non-dispersive infrared (NDIR) spectroscopy for carbon dioxide measurement has facilitated the creation of affordable sensors and monitors intended for consumer use. These monitors are typically designed to provide real-time data on air temperature, relative humidity, and various IAQ parameters, often including PM2.5, PM10, CO_2, and total VOCs (TVOCs).

Certain consumer-grade monitors are equipped with sensors for additional gases like carbon monoxide, nitrogen dioxide, ozone, or other metrics such as air pressure and noise level. The readily available consumer-grade monitors commonly store data on Internet of Things (IoT) servers, and the measurements can be accessed through web or mobile applications. The increased market availability of such consumer-grade monitors and single, affordable sensors (devices that measure individual IAQ parameters and transmit data to a logging system) has captured the interest of numerous researchers.

Up until now, numerous studies have investigated the effectiveness of affordable sensors and monitors in detecting indoor and outdoor (Z. Wang et al., 2020) particulate matter (PM). (Singer & Delp, 2018) evaluated the capabilities of budget-friendly air quality monitors in identifying fine particles originating from residential sources. They identified a reasonable correlation within a twofold factor for most sources, except for particles with an optical diameter below 0.3 µm, where there was minimal response. (Singer & Delp, 2018) recently reaffirmed these findings. Other research indicated that the accuracy of integrated PM sensors within consumer-grade monitors can be influenced by air temperature and relative humidity (Bai et al., 2020).

Similarly, the precision of CO_2 measurement using low-cost non-dispersive infrared (NDIR) sensors, commonly found in consumer-grade monitors, was shown to depend on air temperature and relative humidity (Marinov et al., 2018).

Besides direct measurements, certain devices estimate CO_2 levels based on total volatile organic compound (TVOC) measurements, resulting in notable inaccuracies (Demanega et al., 2021a). The TVOC measurement itself, conducted with metal oxide semiconductor or photoionization detector (PID) sensors, is prone to cross-sensitivity with interfering compounds. The realm of volatile organic compounds encompasses a wide range of substances, from innocuous cooking scents to hazardous chemicals like aromatics (e.g., benzene, toluene, xylene) and aldehydes (e.g., formaldehyde and acetaldehyde), which renders detecting and monitoring VOCs challenging, along with quantifying exposure.

Numerous studies have evaluated sensor performance not only concerning air quality but also encompassing other aspects of the indoor environment, such as thermal comfort. Moreno-Rangel et al. assessed five "Foobot" monitors in measuring residential air temperature, relative humidity, PM2.5, CO_2, and TVOC. The study found satisfactory accuracy for all sensors except CO_2, which wasn't measured by a dedicated sensor but rather derived through an algorithm using TVOC data.

Beyond this research, our understanding of the comprehensive performance of low-cost consumer-grade monitors and sensors remains limited.
The findings obtained from (Demanega et al., 2021b) activities underscore the limitation of using optical light scattering technology in low-cost PM sensors, as it is unable to cover the entire range of particle sizes commonly emitted from indoor sources. (Singer & Delp, 2018) assessed 2 research-grade and 7 consumer-grade monitors, concluding that consumer-grade monitors exhibit semi-quantitative responses (50–200%) to most tested pollutants, while all devices exhibited minimal or no response to events involving particles with an optical threshold of 0.3 µm. (Z. Wang et al., 2020) study confirmed this, establishing the particle detection limit at around 0.25 µm. As per specifications, the majority of consumer-grade monitors are designed to detect particles with optical diameters ranging from 0.3 µm to 2.5 µm. Depending on the pollutant source and its associated particle size distribution, closer alignment with the reference was observed when the optical particle diameter ranged from 1 µm to 2.5 µm, with most tested devices reporting around 50% of the reference concentration at worst.
Agreement diminished when sources were dominated by submicron particles (<1 µm) and during activities generating coarse particles (e.g., vacuuming). Studies also reveal that optical monitors (consumer, professional, and research-grade) might underestimate the mass concentration of larger particles produced by vacuuming, particularly if those particles have higher density.

However, due to the polydisperse nature of indoor particle sources, the response of the majority of sensors exhibited temporal correlation. A robust correlation with reference data was also identified by (J. Li et al., 2020) for the evaluated consumer-grade monitors. This implies that these devices dynamically track concentration changes and can detect events even when quantitative agreement is poor. The analyzed data suggests no consistent bias for PM2.5 sensors. It's important for end-users to recognize that PM data from current low-cost sensors should be interpreted as an indication of a shift in state or a rough estimate, rather than an exact concentration indoor area.

TVOCs comprise a wide range of volatile organic compounds, and each pollutant source generates different types of VOCs. A comprehensive study highlighted that TVOC sensors exhibit varying sensitivity to different VOC sources, depending on their working principles (Militello-Hourigan & Miller, 2018). This variability is especially notable in PID sensors, which yield accurate results in laboratory air sampling when measuring specific groups of compounds, they are calibrated for. This variability explains the different responses of monitors and the inconsistent seasonal replication in various experiments. While consumer-grade monitors managed to capture changes in TVOC concentrations over time and could be employed to detect events, users should be informed about the inaccuracies in absolute values. (Nirlo et al., 2015).

In all the tested monitors and single sensors, air temperature and relative humidity were found to impact measurement performance. These factors should be considered when interpreting the data. Additionally, variations in particle density from different sources can influence the accuracy of mass concentration measurements. To address this, using source-dependent particle densities and understanding their effects on measurement accuracy is crucial.
If sensor was positioned in close proximity to a microcontroller with a power converter within a housing accommodating multiple single acquisition set, the heat generated by the microcontroller likely interfered with the air temperature measurement, leading to overestimated temperature values.

The literature analysis performed provides a comprehensive evaluation of the performance of low-cost consumer-grade monitors and single-parameter sensors in detecting five indoor environmental parameters: particulate matter, carbon dioxide, total volatile organic compounds, dry-bulb air temperature, and relative humidity. Eight experiments were conducted to simulate indoor air pollutant sources under two different climatic conditions: cool and dry, and warm and humid.

Regarding PM measurements, even though some devices exhibited Mean Relative Errors (MRE) exceeding 100%, the dynamic responses were time-correlated for most tested devices. This implies that low-cost units could effectively detect changes in particulate matter concentrations ranging from 0.3 to 2.5 µm. On average, the best-performing monitor deviated by a factor of two from the reference.
In terms of CO_2 concentration detection, the majority of devices displayed errors within 25% from the reference for concentrations up to 3,500 ppm. Some of the best monitors achieved a deviation of just 3% from the reference.
For total volatile organic compounds (TVOC), low-cost monitors showed strong correlations with professional-grade monitors, despite a lack of precise quantitative agreement.

Comparing the two seasons, most consumer-grade monitors exhibited similar performance in both conditions, with some devices slightly closer to the reference in cool and dry conditions for PM and CO_2, and in warm and humid conditions for TVOC.

The study suggests that recent technological advancements provide an opportunity for more effective indoor air quality control. The majority of tested low-cost consumer-grade monitors could potentially trigger appropriate actions to ensure satisfactory indoor environments. This could be achieved through a feedback loop to encourage human actions or integration into building management systems with automated controllers and devices. Continuous improvement of low-cost monitoring technology is vital to enhance indoor air quality management.

As the field of environmental sensing technology continues to evolve, future research should prioritize several key areas. Firstly, there is a need to delve into the longitudinal performance of these low-cost monitoring units. This involves studying their consistency and accuracy over extended periods of time, which is essential for assessing their reliability in real-world scenarios.
Additionally, the development of robust quality control algorithms is crucial. These algorithms should be designed to minimize measurement errors and eliminate biases that might arise from various sources, such as sensor drift, calibration discrepancies, or environmental fluctuations. Such algorithms would contribute to enhancing the accuracy and consistency of the collected data.
Furthermore, the establishment of standards and guidelines for the testing of these low-cost monitoring devices is essential. A standardized framework would ensure that these devices undergo thorough and consistent evaluation processes, enabling direct comparisons between different models. This, in turn, would aid consumers, researchers, and decision-makers in making informed choices based on reliable data.
In conclusion, the future direction of research in this field should focus on the long-term performance assessment of low-cost sensing units, the development of effective quality control algorithms, and the establishment of standardized testing protocols to ensure the accuracy and reliability of these devices for indoor environmental monitoring.

2.3.5 DT-based IAQ management

Digital twins can play a crucial role in addressing these challenges and ensuring the optimal indoor air quality (IAQ) in smart buildings. By creating a digital replica of a building's HVAC system, ventilation, and other environmental parameters, digital twins can offer real-time insights into IAQ performance. This empowers building managers to detect potential problems, such as insufficient ventilation or elevated pollutant levels, in advance of critical situations and optimize the performance of HVAC systems, including temperature and humidity control, to sustain a healthy indoor environment.

The system is able to support in making informed decisions regarding maintenance, upgrades, and investments in IAQ enhancements simulating various scenarios and assess their impact on IAQ, aiding in the identification of potential challenges and the development of strategies to address them. This ensures that air quality standards are upheld even in changing conditions.

Digital Twins provide a potent solution for achieving and upholding indoor air quality standards in smart buildings, especially concerning SmartScore certification. By offering real-time insights into IAQ performance, digital twins empower building managers to proactively tackle issues, enhance system efficiency, and base decisions on data, ultimately benefiting the health and well-being of occupants. As digital twin technology advances, we can anticipate further innovative applications for enhancing IAQ and overall building performance.

Several key factors contribute to the complexity of managing indoor environments in the context of pathogen transmission and incorporate the choice towards a DT system capable of handling complex problems:

- Heterogeneity: indoor environments are highly diverse, with various functionalities and characteristics that create distinct habitats for different pathogens. Microbiomes can significantly differ even within the same building, depending on the room's usage. This variability highlights the necessity for tailored, location-specific risk management practices.
- Diverse Pathogen Habitats: pathogens can be found in various indoor habitats, such as surfaces, air, and water. These different locations offer unique transmission and exposure pathways. For instance, strategies effective at preventing surface transmission might not be as effective for airborne pathogens.
- Occupant Behavior: building occupants play a pivotal role in the transmission and exposure pathways of pathogens, especially in environments with multiple users. Their behaviors, interactions, and movements can influence the risk of infection and pathogen spread.
- Disruption vs. Functionality: traditional approaches like prolonged access control and extensive cleaning may reduce pathogen transmission but can also disrupt the standard operations and functionalities of indoor environments.

Given the dynamic and interconnected nature of human behavior, pathogen transmission, and building operations in indoor environments, there's a growing need for adaptive and advanced systems. These systems should enable real-time monitoring and control responses to the ever-changing dynamics of the building-human-pathogen system. Digital Twin emerges as a promising approach to address these challenges by providing a comprehensive and adaptable framework for managing indoor environments and mitigating pathogen risks.

According to (J. Cai et al., 2023), the DT of the indoor environment can be perceived as the interaction between the buildings' indoor environment and its digital representation model to provide real-time monitoring, analytics, and control for the environments. Therefore, three literature research clusters that develop key technologies to enable the establishment of DT have been identified, i.e., BIM, IoT sensing, and smart building control through data analytics.

Research in the field of Building Information Modeling (BIM) applications for promoting healthy indoor environments has been relatively limited. Typically, this area of study is considered a subset of design performance or facility management, with a focus on managing factors like indoor air quality, daylighting, and ventilation. However, the emergence of COVID-19 has led to increased attention on leveraging BIM to support healthier indoor environments.

In a study by (Rice, 2020), 14 building health indicators were identified, including aspects like thermal comfort, volatile organic compounds (VOCs), and sound insulation. The potential of BIM models to measure and evaluate these indicators through surveys and assessments was discussed. Most participants in the study agreed that BIM models could be used to measure these indicators. However, there were debates regarding the complexity of measurement for different indicators. Three complexity levels were identified: direct measurement based on BIM models, the need for additional plug-ins, and the requirement for extra sensors and detectors.

Some research has delved into specific indicators, such as indoor air quality (Utkucu & Sözer, 2020) and lighting (Montiel-Santiago et al., 2020), and explored how BIM can facilitate their assessment and monitoring. Beyond indicators perceptible to humans, other studies (Adams et al., 2014) (S. Li et al., 2021)) have taken a building microbiology perspective. They have extracted building information, such as layout, and correlated it with microbial communities identified through advanced sequencing techniques.

Overall, while the application of BIM in promoting healthy indoor environments is gaining attention, it remains an evolving field with various complexities related to measuring and assessing different health indicators.

The impact of design on microbial risks has been explored in research efforts. For example, (Adams et al., 2014) conducted a study that revealed variations in microbial communities between different types of rooms. Additionally, a pilot study conducted by (S. Li et al., 2021) achieved a groundbreaking integration of Building Information Modeling (BIM) with a microbial transmission model, allowing for the assessment of infection risks at the room level. This work has laid the foundation for the integration of BIM with microbial risk assessment as a means to promote health in indoor environments.

In essence, the current use of BIM for enhancing healthy indoor environments involves two primary processes.

- Integration of BIM data: this process entails extracting information from BIM and combining it with other relevant data to assess health or health-related factors using simulation or data-driven methods. Various studies have employed this approach (e.g., (Sporr et al., 2019) (Kwok et al., 2020), (Gan et al., 2021).
- Utilization of BIM Visualization: BIM's visualization capabilities are harnessed to represent and communicate health or health-related factors to users. Researchers and practitioners have employed BIM for this purpose (e.g., (Alavi et al., 2021); (Valinejadshoubi et al., 2021).

These processes collectively contribute to leveraging BIM as a tool for enhancing indoor environmental health and microbial risk management.

In the context of fostering a healthy indoor environment, Building Performance Simulation (BPS) models are commonly employed, often in conjunction with Computational Fluid Dynamics (CFD) techniques. These models serve the purpose of modeling, assessing, and controlling various aspects of indoor environments, including indoor air quality ((Y. Li et al., 2021) (Stephen Lopez et al., 2021), pathogen transmission (Peng et al., 2020) (Motamedi et al., 2022)), as well as thermal and visual comfort (Cheong et al., 2020).

However, a notable limitation of BPS models lies in their representation of occupant behavior, which can significantly impact the accuracy and reliability of simulation results. To address this limitation, Agent-Based Simulation (ABS) has been employed to create more sophisticated models that better capture the dynamic interaction between occupants and indoor environments. ABS can be integrated with BPS models to explore how such interactions influence building performance metrics, including energy consumption and disease transmission (X. Zhou et al., 2020), (Islam et al., 2022). This integration allows for more informed decisions related to building operations and control.

Furthermore, Building Information Modeling (BIM), which provides rich spatial and semantic information about building design and operation, has become an indispensable tool in data analytics. BIM is often combined with BPS to enhance the accuracy and reliability of simulation results concerning building performance (Y. Zhou et al., 2021).

In the context of smart building control, Digital Twins (DTs) equipped with advanced dynamic modeling capabilities act as the intelligent core.

They take inputs and outputs from digital building models and formulate building control as an optimization problem. Various optimization techniques, such as mixed-integer problems and nonlinear programming, can be effectively employed to solve these problems. Model Predictive Control (MPC) techniques have been implemented in real buildings, demonstrating their ability to enhance building operational practices (Ruusu et al., 2019).

The primary focus of DT-enabled smart building controls is to enhance indoor health and environmental comfort while ensuring energy efficiency in building operations. This involves improving various aspects of indoor occupant well-being, including:

- Indoor Air Quality (IAQ): DTs, through building modeling and indoor/outdoor air quality monitoring, enable optimal ventilation by switching ventilation modes and adjusting ventilation rates. This results in reduced indoor air pollutant concentrations and energy savings in ventilation operations (W. Kim et al., 2016), (Cho et al., 2015).
- Controlled Indoor Air Pollutants: Various air pollutants like CO_2, Total Volatile Organic Compounds (TVOCs), respiratory particles, and ozone are managed simultaneously (Ganesh et al., 2019), Afroz et al., 2020b), (Saini et al., 2020a) Effective deployment of indoor environmental sensors is essential for efficient indoor air pollutant management (Zeng et al., 2020)
- Occupant Behavior Modeling: Occupant behaviors play a crucial role in effective indoor air pollutant management, especially in demand-controlled ventilation strategies that rely on accurate occupancy information (W. Wang et al., 2018).

Amid the COVID-19 pandemic, there has been increased attention on methods to control airborne infections. Smart ventilation strategies, encompassing both mechanical and natural ventilation, along with occupancy adjustments, have been proposed to mitigate virus transmission risks in building operations (Xu et al., 2021).

Furthermore, in pursuit of energy efficiency, smart lighting control strategies have been developed to optimize lighting operation schedules based on environmental conditions and occupancy, ensuring satisfactory indoor illuminance levels without unnecessary energy consumption (Kandasamy et al., 2018)

In line with broader sustainability goals, building DTs are being leveraged to integrate renewable energy sources and support demand response initiatives. These DTs utilize system-sensing data to facilitate the harmonious integration of renewable energy, energy storage (e.g., thermal or battery storage), and building operations, considering both power generation and occupant requirements Biyik & Kahraman, 2019),(Kathirgamanathan et al., 2021)

In summary, DT-enabled smart building controls are instrumental in optimizing building performance across various facets, from indoor air quality and occupant comfort to energy efficiency and sustainability, thus fostering healthier and more efficient indoor environments.

In essence, a DT for a healthy indoor environment serves as a comprehensive and dynamic platform that continuously monitors, models, assesses, and controls various factors to promote both environmental sustainability and occupant well-being as shown in **Figure 15**.

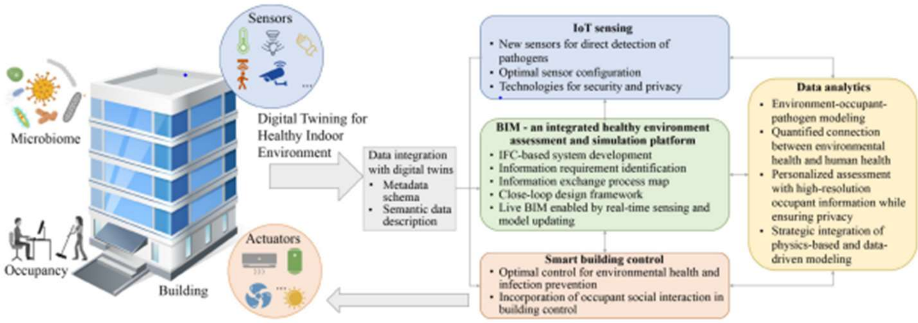


Figure 15. DT and IoT reference architecture for IAQ management

The progression of IoT sensing is anticipated to furnish unparalleled real-time data encompassing all facets of the indoor environment. To attain this, it is imperative to develop novel, economical sensors for the direct identification of microbiome populations (e.g., pathogens). Furthermore, it is necessary to devise methodologies for the optimal setup of sensors to guarantee precision in measurements while also enhancing cost-effectiveness. Modern technologies, including edge devices and privacy-preserving sensors, ought to be harnessed in IoT sensing to safeguard both security and privacy.

In the future, in the context of a healthy indoor environment, Building Information Modeling (BIM) should evolve into an integrated assessment and simulation platform. Its primary role would be to furnish actionable insights for building control systems by amalgamating information encompassing the environment, occupants, and microbiome. These insights would be derived from multi-disciplinary simulations covering domains such as energy usage, air quality, and pathogen transmission. A significant challenge in this context is ensuring data interoperability. The adoption of an Industry Foundation Classes (IFC)-based system development approach could potentially address this issue. Moving forward, it is crucial to identify information requirements and clarify the process of information exchange.

Regarding real-time sensory data, future BIM systems are expected to undergo continuous updates to ensure that simulations remain calibrated. In addition to its role in assessment and simulation for smart building control, a closed-loop design framework is essential. This framework should enable the identification of relationships between indoor environmental attributes and health indicators, ultimately guiding building design.

The prevailing paradigm of smart building control, which primarily emphasizes energy efficiency and occupant comfort, needs to expand its scope. It should acquire the capacity to ensure environmental health and mitigate infection risks. To achieve this, it is imperative to quantify the relationship between building operations (e.g., ventilation and surface disinfection) and infection risk, thereby enabling bio-informed adaptive control. Moreover, occupants play a pivotal role as active components within the indoor environment. Their behavior significantly influences the transmission of infectious diseases. Consequently, building control strategies must take into account the interactions among occupants and between occupants and the environment.

In this scenario, advanced data analytics is a fundamental component of digital twinning for a healthy indoor environment. It offers valuable insights that support the other three techniques mentioned. Novel methodologies that leverage rich, real-time sensor data, building design and operation data, and external contextual data are required. These methodologies should strategically integrate physics-based simulations and data-driven analyses for modeling the complex relationships among the environment, occupants, and pathogens. Furthermore, there is a need to establish quantifiable relationships between environmental factors and human health.

2.3.6 Open challenges

This section delves into the primary challenges associated with Digital Twins (DT) and discusses them in four distinct parts:

- **Data Analysis and Data Access:** the existing IT infrastructure often falls short in supporting data analysis within a DT environment. DTs inherently require advanced infrastructures to enable efficient data analysis. In essence, the success of DT relies on a robust and interconnected IT foundation. High-performance GPUs can be employed for this purpose. Cloud service providers like Amazon, Microsoft, and Google offer substantial resources, but using cloud services for data analysis still presents significant security challenges. Edge computing emerges as a potential solution to mitigate data transmission delays, increase bandwidth, and enable rapid and accurate data exchange and analysis.

- **Connectivity within the DT Framework:** the Internet of Things (IoT) technology facilitates connectivity within the DT framework. However, despite advancements such as the emergence of 5G technology, connectivity-related challenges persist, including software errors and power outages during real-time monitoring. The presence of missing data can substantially impact algorithm performance, necessitating methods to recover lost data and establish complete connections.

- **Security:** security and privacy pose notable challenges in the realm of advanced technologies. In the context of DT, the sheer volume of data and information exchange presents cybersecurity risks, including the potential for cyberattacks. Adherence to updates in privacy and security laws is imperative for DTs. Establishing a secure platform is paramount and results in a highly reliable digital environment. Blockchain technology, with its secure and decentralized structure, can effectively address security concerns by providing a secure framework for data and information exchange among agents.

- **Standardization:** standardization is another significant challenge associated with DTs that can potentially hinder their development. The adoption of standardized formats is essential for defining, storing, and executing DT models, fostering interoperability and integration among different DTs. In complex systems like hub energy systems, where multiple infrastructures (e.g., electricity, water, gas) coexist, a unified platform is required to enable seamless interoperability between these systems. This challenge can be addressed through technologies such as Semantic Web or by employing a DT definition language to establish common standards and ensure effective communication and integration among DTs.

Chapter 3

Methodology

Historically, managing energy grids and monitoring indoor air quality relied on manual readings, rudimentary sensors, and often, human intuition. The limitations were evident – lack of real-time monitoring, reactive instead of proactive responses, and a heavy dependence on human intervention.

Traditional approaches to energy microgrids and IAQ have had their share of inefficiencies, often reliant on manual readings, basic sensors, and human judgment. The pitfalls were evident – latency in real-time monitoring, reactive strategies, and intensive human oversight. Yet, with interconnected devices, the Internet of Things (IoT), and robust analytical platforms, the proposition of constructing an insightful digital representation of these systems has emerged.

However, with the rise of interconnected devices, the Internet of Things (IoT), and powerful analytical platforms, the prospect of creating a real-time, self-learning digital representation of these systems became plausible. This led to the exploration of integrating data platforms with machine learning to forge an intelligent Digital Twin.

As a Digital Twin, at its core, is a dynamic software representation of a physical entity or system the digital counterpart mirrors the real-world system in real-time, capturing its every nuance. By harnessing the data generated by IoT devices and sensors, it provides a platform for analysis, understanding behaviors, and predicting future trends.

In the context of energy microgrids and IAQ, envision a comprehensive digital representation that not only displays real-time data but also predicts potential failures, recommends optimization strategies, and continually learns from new data.

Data is the core component of a Digital Twin. The more data it ingests, the clearer its representation becomes. By seamlessly integrating various data platforms, we ensure that every sensor, every meter, and every IoT device feeds into the Digital Twin. This integration, often cloud-based, guarantees scalability, redundancy, and real-time data acquisition.

For instance, in monitoring a microgrid, data from solar panels, wind turbines, battery storage, and demand meters are continually funneled into the digital counterpart. Likewise, for IAQ, sensors capturing temperature, humidity, CO_2 levels, and volatile organic compounds stream data in real-time.

In this scenario, the integration of machine learning offers a transformative edge to the Digital Twin. Instead of merely showcasing real-time data, the system begins to 'learn' patterns, recognize anomalies, and make predictions.

In the microgrid context, the Digital Twin might predict a demand surge based on historical data and recommend energy storage strategies. For IAQ, it could forecast a decline in air quality and preemptively suggest ventilation strategies or adjustments to HVAC systems. This predictive capability not only ensures efficient operations but also prolongs the lifespan of physical assets by averting potential failures.

The proposed approach integrating data platforms and machine learning presents several innovations. Unlike traditional systems that viewed microgrids and IAQ as separate entities, the Digital Twin provides an integrated, holistic view, capturing the intricate interplay between energy consumption and indoor air quality.

As such, instead of reacting to failures or declines in IAQ, the system becomes proactive, offering solutions before issues manifest and as the Digital Twin ingests more data, its machine learning algorithms refine themselves, becoming more accurate in predictions and recommendations.

The Digital Twin relies heavily on quantitative modeling, a bedrock of scientific research. By mathematically representing physical systems, predictions and analyses are based on rigorous, replicable methods. A critical aspect of scientific methodology is the feedback loop – the ability to adjust hypotheses and predictions based on new data. Digital Twins inherently adopt this methodology, constantly refining their algorithms based on new data influx.

The proposed approach introduces the integration of data platforms and machine learning in constructing a Digital Twin as a paradigm shift in how we perceive and manage microgrids and IAQ. This data-driven approach not only optimizes operations but also pioneers a future where our systems are self-aware, predictive, and continually evolving.

2.4 Review of Machine Learning-based and Data-driven strategies

The present section is aimed at discussing Machine Learning algorithms usable to assess one of the main objectives of the present research, which is the energy balance of a building, considering electrical loads, power production devices, HVAC systems and people needs and habits. Nowadays, Machine Learning has been widely adopted for improving building energy efficiency and flexibility thanks to the ever-increasing availability of massive building operational data. However, it is difficult for end-users to understand and trust Machine Learning tools because of their black-box nature. To this end, the interpretability of Machine Learning models is attracting ever more attention from the scientific and industrial community because it helps users in understanding the decisions made by these models.

It is well known that building sector is one of the major contributors to global energy consumption and carbon emissions. It is accounted that about 36% of the global energy consumption and the 37% of CO_2 emissions comes from this sector (OECD/IEA, 2019; Zarco-Periñán et al., 2022). Moreover, throughout the entire life cycle of a building, the operation phase accounts for $80\% - 90\%$ of total energy consumption (Ramesh et al., 2010). Considering these data is evident how much the building sector impact on global warming, the reduction of consumption is therefore crucial for global energy-saving and to reach carbon neutrality as advised by Sustainable Development Goals (Leal Filho et al., 2022; Parra-Domínguez et al., 2022) proposed by United Nation and for the 55% CO_2 reduction target by 2030 recommended in the New Green Deal (Kazak, 2022).

In this framework, Building Automation Systems (BASs) could play an essential role in improving energy efficiency and flexibility. It is possible to implement inside a BASs various smart control strategies for building energy management, such as Heating, Ventilating, and Air Conditioning (HVAC) controls, energy storage and renewable energy monitoring systems, smart lighting and so on (Hurtado et al., 2018; H. Tang et al., 2021). Traditionally, the control approaches were developed using rule-based control strategies based on physical analyses and experience. However, this sometimes results in great challenges due to the complicated interactions among building energy systems (Y. Chen, Chen, et al., 2022). Nowadays, thanks to the wide diffusion of Internet of Things technology it is possible to get real-time data of the building environment with few efforts and expenses. Therefore, such data can be integrated inside big-data platforms to evaluate innovative control strategies using Data Science.

Machine Learning models can bring new way of data interpretation discovering and learning directly from the environment the actual needs and requirements of the building, adapting in almost real time to the changes present in the building. This new approach is called “data-driven” and is considered as the paradigm of digitalization (Capozzoli et al., 2016; Marotta et al., 2021). With such data-driven models, building can be monitored to make decisions autonomously with and without human intervention (Sengupta & Chandrashekhar, 2021). At least, the possibility of integrating real-time data with geometrical information coming from the project phase (as from Building Information Modelling) brings great value to the Digital Twin technology, as the digital construct is able to replicate and simulate the behaviour of a building in real-time. Machine Learning is therefore a fundamental technology for the future of the buildings sector and shall be considered as a building component itself.

Machine learning has effectively facilitated building energy management developed during the last years as evidenced by many studies and applications present in literature (Z. Chen et al., 2023). This includes, load and power prediction, Fault Detection and Diagnosis (FDD) and occupancy-related applications.

2.4.1 Load and Power prediction

Load prediction refers to the ability of a Machine Learning model to predict the cooling/heating/electricity demand in the future hours or days, while power prediction aims to predict the power generation of equipment such as photovoltaic (PV) panels, wind turbines or other power generation devices. The development of accurate prediction tool for power management is aimed at the optimization of the energy efficiency and flexibility inside the building (L. Zhang et al., 2021). As an example, with a demand-side model it is possible to improve the flexibility of the building energy system by balancing the supply and the demand in real-time using as much as possible renewable energy sources. Instead, with a model predictive control it is possible to operate power equipment (set-point, thermal comfort, lighting) basing on minimal cost of energy consumption logics. Therefore, the choice of the model depends on what is the model’s goal and from what it is installed in the building, if there are straight constraints on power system operations a demand-side model is advisable, instead, if the goal is to reduce the global energy usage, a model predictive control is better suited (Y. Chen, Guo, et al., 2022).

Such models require historicized data instead of a physical-based load model (energy equations, equipment information, etc). Therefore, its development is quicker, but it is necessary to have historical operative data available (Z. Chen, Chen, et al., 2022). Between the different Machine Learning algorithms Artificial Neural Networks (ANN), tree-based methods, and Deep Neural Networks (DNN) autoregressive models are the most suitable, as evidenced by many literature studies (L. Zhang et al., 2021). ANN are very suitable to cope with complex real-world problems such as for solar radiation prediction (Bahani et al., 2020) or for heating/cooling efficiency (Le et al., 2019). Moreover, ANN are quite more flexible not depending on linear approximations and being independent from binary logics as for tree-based methods. DNN instead, seems to be very promising in load demand prediction, especially if coped with Long Short-Term Memory layers (LSMT) (G. Li et al., 2022).

2.4.2 Fault Detection and Diagnosis

Fault Detection and Diagnosis is another widely discussed topic because it is essential to maintain building efficiency, especially when energy-intensive loads are installed (such as chillers and heaters). Using Machine Learning it is possible to monitor device status and life cycle using self-learning energy signature models independently from the specific device. Some examples of applications employ Support Vector Machines (SVMs), Artificial Neural Networks, Convolutional Neural Network (CNN) models, Recurrent Neural Network (RNN), and Deep Generative Systems (Ciaburro, 2022). ANN and CNN are the most suitable ones due to their flexibility and capacity to self-learning from data.

2.4.3 Occupancy-related applications

At least, occupancy-related applications refer to the problems related to the presence of people inside an environment. As an example, occupant thermal comfort is occupancy-related applications directly connected to the HVAC system and the use of accurate Machine Learning models can improve indoor comfort yet reducing energy consumptions. Using Machine Learning based models instead of Predicted Means Vote (PVM) (Yau & Chew, 2014) it was demonstrated that it is possible to reduce energy consumption more than 58% and CO_2 emissions up to 24% (Qavidel Fard et al., 2022). This is possible thanks to the occupancy level prediction and activity recognition that are two branches of Computer Vision models that uses Machine Learning and Artificial Intelligence models to determine the number of people in an environment and their activity. HVAC consumption could be reduced up to 23% using Machine Learning for occupancy predictions and up to a more 20% using activity recognition (Dai et al., 2020; Javed et al., 2018). IoT plays a fundamental role in this field obtaining real-time and historicized data of a parameter which is difficult to be evaluated.

2.4.4 Interpretability of Machine Learning models

One of the main problems occurring when a Machine Learning model is used is the capability of interpreting what data affect the process, its performances and which features influences the predictions. Sometimes this can go over the human comprehension and the efforts to get understandable data could not worth the use of Machine Learning tools (Rudin et al., 2022). An example of interpretability issue is electrical equipment fault detection whereby the use of ANN. Although, the problem is well defined but the data that affects the process input and/output is not easy to define and can vary depending on the device. This can reduce a machine learning model application due to its lack of interpretability. ANN the structure of inputs is well clear by the developer but how the data is elaborated by the model is dark. Neural networks are made of a big number of hidden layers that are trained on the dataset and which output is usually not understandable. Moreover, as the number of layers increase (Deep Neural Networks) the model becomes deeper and darker. However, it is possible to employ simpler models such as Support Vector Machine (SVM), k-NN and regressors. The tradeoff is the accuracy of the output, that is usually less if compared to most complex models. The lack of interpretability is one of the main challenges for Machine Learning in real-world applications (Burkart & Huber, 2021). This issue is commonly connected to the lack of the physical knowledge of the process and the presence of incomplete or bad structured data (Yan et al., 2018). **Figure 16** shows the correlation between the interpretability and accuracy of Machine Learning techniques for interpretable and black-box models.

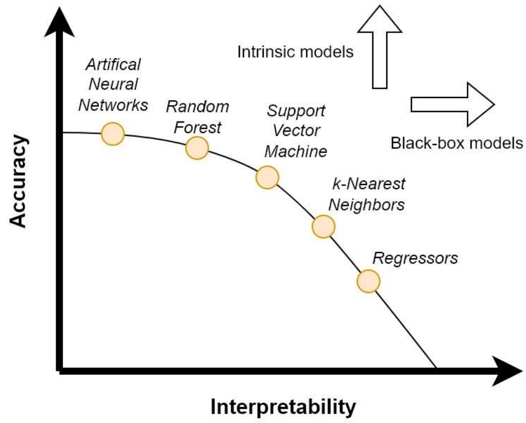


Figure 16. Correlation between the interpretability and accuracy of a Machine Learning models for interpretable and black box models

Therefore, decision-makers may find the ML untrustworthy if the models are not trained on real operational data. Moreover, the actual performance of models could be worse than on the training data. Another issue is due to the black-box nature of ML models because they produce output without any explanations. Decision-makers usually need insights into how and why models produce such predictions so that they can understand, control, and apply the models. This leads to skepticism in the building industry about the application of ML. Therefore, it is advisable to generate reasonable interpretations that explain the original ML model without oversimplifying essential details or sacrificing prediction performance. The issue is graphically explained in **Figure 17**.

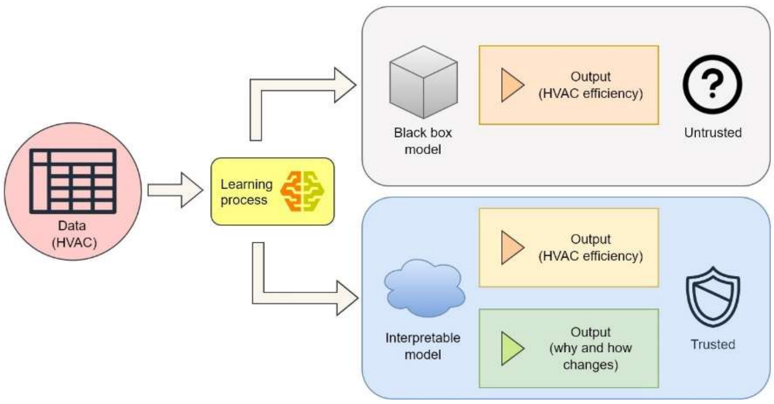


Figure 17. Interpretable Machine Learning model trust example for HVAC systems.

Interpretability can be local or global depending on the scope of model output that needs to be interpreted. Global interpretation explains an ML model based on a full view of the model structures and parameters. In contrast, local interpretation explains each prediction individually. This means that global models provide a holistic understanding of the model itself by measuring the global effects of input parameters on the predictions. This is important for decision making because it gives a macro-level understanding of the prevision process, as example, explaining what features are most significant in predicting energy consumption. Instead, local models provide a transparent understanding of the prediction for a specific input. These models focus on the effects of every single input on the result. Differently from global models, this is important for decision making to trust the output and to understand if it is current or wrong. As an example, local models can explain the effect on external air temperature on HVAC power consumption and efficiency. **Figure 18** summarize what is intended as global and local interpretability level.

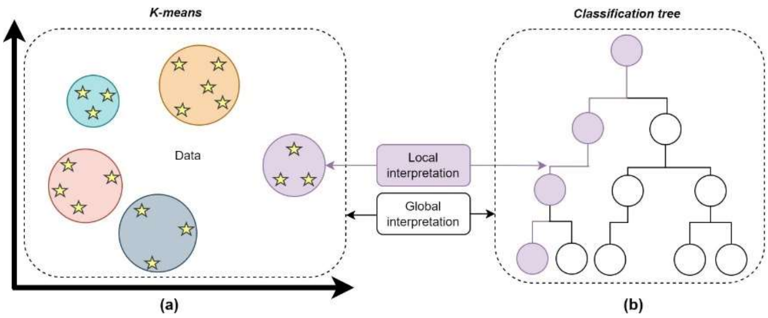


Figure 18. Local and Global interpretation using as example k-means (a) and tree-classifier (b) models.

There are two types of Machine Learning families defined as Ante-Hoc and Post-Hoc, the division is based on where the models are applied in the data elaboration process:

- Ante-hoc Machine Learning models are applied during the training process,
- Post-Hoc Machine Learning models are applied after the training process.

Ante-hoc models are also called intrinsic or transparent models. For example, a linear regression is a simple ante-hoc model for predicting a continuous outcome variable based on one or more predictor variables (El Fiorenza Caroline et al., 2019). The model is self-explanatory because it makes predictions using a linear combination of the input variables, which can be easily understood and explained (Kamath & Liu, 2021). Although linear regression is high interpretable, it is somehow too simple to address complicated problems in building energy management (Sha et al., 2021). Generalized Additive Models (GAMs) is a more suitable variant of linear regression with strong flexibility and interoperability for regression and classification roles (Zhuang et al., 2021).

Instead, Post-hoc models are applied as a black box after their training. They are used to interpret and understand the dependency and significance of an input over the output without the need to understand the internal structures. They are aimed at generating interpretation by examining the interrelationship between input features and the predictions without knowing how the process is developed (Friston & Penny, 2011; Vilone & Longo, 2021).

The graphical representation of **Figure 19** shows post-hoc and ante-hoc implementation in an example of Machine Learning working flow; On the left is visible the ante-hoc flow, where the flow skips the surrogate model fitting going directly to the evaluation step. Instead, in the right is represented the post-hoc schematic, where the flow is implemented in a black-box model that shall be evaluated somehow as a global result. It shall be noted that it is also possible to concatenate a black-box model with an interpretable model in a single logic flow where post and ante data elaboration steps are performed. However, the possibility of obtaining both the advantages of the two workflows is not guaranteed and depends on the specific surrogate model implemented.

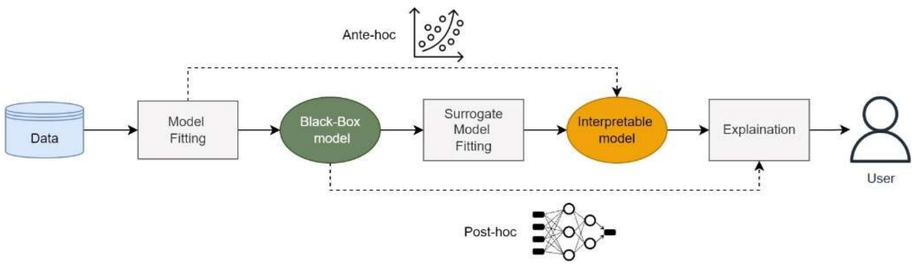


Figure 19. Ante-hoc (left) and Post-hoc (right) Machine Learning flows implementation

Going forward, Machine Learning techniques can be agnostic depending on their structure. As example, a Machine Learning technique can be intended as a single component that can be used into any Machine Learning model independently from the model as showed in **Figure 20**. Others can only be used to interpret certain types of models and are thus called model-specific techniques. The first is called model agnostic technique, the second model specific technique. Model-agnostic techniques can be applied to any ML model because they require only the input and output of the ML model without considering its inner structures. Usually, post-hoc models are also model-agnostic. An example of such technique is Local Interpretable Model-agnostic Explanations (LIME) a post-hoc model agnostic technique that can approximate any Machine Learning model locally. Instead, model-specific techniques are more specific on the architecture of the ML model, providing in-depth interpretability that may not be possible with model-agnostic methods. For example, the attention mechanism is usually employed in neural networks to improve interpretability as a model-specific technique.

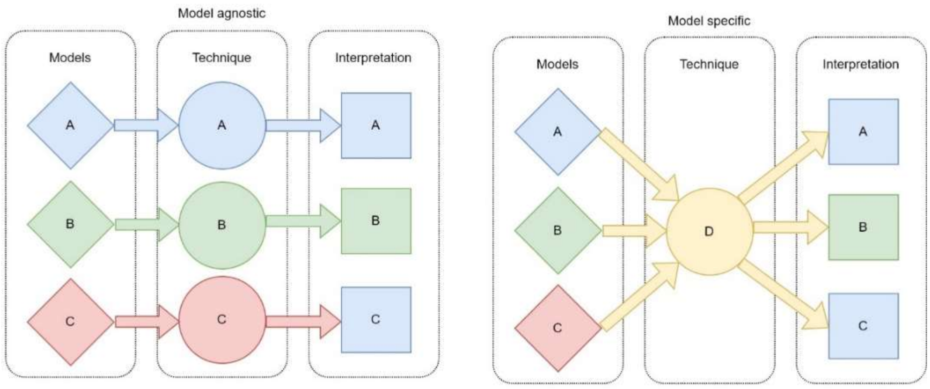


Figure 20. Model agnostic and model specific Machine Learning techniques

2.4.5 Evaluation of interpretability

Between scientific community there is no consensus about what interpretability is and how to measure it (F. L. Fan et al., 2021). However, Robnik-Šikonja & Bohanec proposed an evaluation method based on three main levels as described (**Figure 21**) (Robnik-Šikonja & Bohanec, 2018):

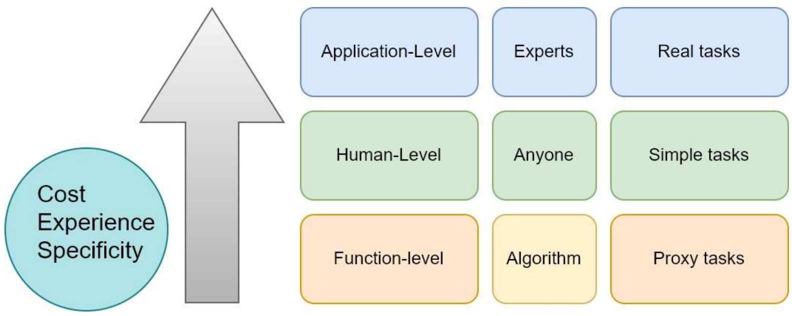


Figure 21. Interpretability levels as proposed by Robnik-Šikonja & Bohanec

1. Application-level (real task): The explanation is made at the end of the process and its quality is tested by the end user. As an example, a Machine Learning crack detection software can locate and marks them in images. At the application level, the end-user can test the crack detection software directly to evaluate the model. This requires a good experimental setup and a good understanding of how to assess quality of the output. Therefore, the quality of the evaluation relay on how good the end user is at explaining the same decision.
2. Human-level (simple task) is a simplified application-level evaluation. The difference is that these experiments are not carried out with the domain experts, but with anyone. This makes experiments cheaper (because it is not necessary to get and expert) and it is easier to find

more testers. In this case many explanations can be produced by the Machine Learning software and the uses chose the best one. However, this method somehow is limited by the user experience and its capacity to get the correct choice, therefore, it is not suitable for complex tasks where an expert is of mandatory importance for the output evaluation.

3. Function-level (proxy task) does not require a user. This method works when the model has been already evaluated by someone else in a human-level evaluation or in application-level evaluation. For example, it might be known that the end users understand decision trees. In this case, a proxy for explanation quality may be the depth of the tree. Shorter trees would get a better explain score.

2.5 Models and techniques for energy assessment and optimization

The actual panorama of ML models is quite wide, with many variants developed for specific tasks. However, they can be organized basing on the task they have to perform, the type of learning and the related input data (Sarker, 2021a). The classification is shown in **Figure 22**.

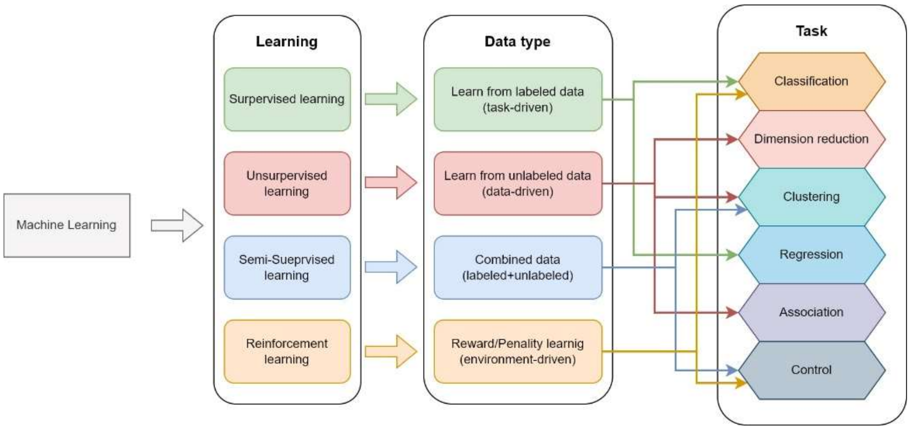


Figure 22. Machine Learning models' classification based on learning type and tasks to perform (Sarker, 2021b)

Generally, it is possible to describe data types as following:

- Structured: input data must be well-defined and structured, with information organized and described at detail level. For instance, device names, times, power, temperatures, locations, occupancy, etc. are examples of structured data.
- Unstructured: if the data has no pre-defined format or organization, it is considered as unstructured. This makes the analysis of relevant information much more difficult to perform. For instance, textual input, word processing, audio files, videos, images, etc. can be considered as unstructured data.

- Semi-structured: data are not stored in an organized structure (such as a relation-al database), but it has some organizational properties. For instance, XML, JSON documents, NoSQL databases, etc., are examples of semi-structured data.

Moreover, input data can be associated also with metadata information that can be considered either as structured, unstructured, or semi-structured. Before choosing a ML model it is important to understand the structure of input data. Moving to interpretability, the data structure is related to the knowledge of the process. Usually, it is recommendable to use structured or semi-structured data especially when using ante hoc models that are intrinsically interpretable. If data is well structured, it is possible to easily get a higher level of interpretability (level 2-3) because the process is repeatable with less issues related to the model's accuracy, and it is necessary a lesser insight on the specific process analysed (Franchini et al., 2022). However, this rise the importance of having good interpretability for the other data structure (semi-structured and unstructured). Hereafter, the most used interpretable ML models for energy management in built environment are discussed with examples based on the scientific literature on the matter.

2.5.1 Artificial Neural Networks

Artificial Neural Networks are widely diffused across the scientific and industrial community because they are quite easy to use and to adapt on different contexts. ANN are developed to mimic the neural brain structure by using mathematical models made of nodes and connections. Although, they are ante-hoc models usually difficult to interpretate due to their black-box nature. However, during the last years due to their wide diffusion and interest have been developed modified ANN with an enhanced interpretability. As an example, adding element with a physical meaning to the models. (A. Li et al., 2021) developed an innovative ANN integrating a GRAvitation model (GRA) with a Gated Recurrent Units (GRU) model to predict building consumption. The weights of the ANN have been determined using the mutual information obtained by the two models. H. Wang et al. proposed a Direct eXplainable Neural Network (DXNN) using ridge function instead of sigmoid to weight the network nodes (H. Wang et al., 2020). Using such different kernel, it was possible to obtain the direct relationship between model's input and output. They apply their model to calculate solar irradiance. Moving to occupancy detection, X. Zhang et al. combined a Takagi-Sugeno-Kang fuzzy classifier with a Deep Neural Network to generate fuzzy rules for occupancy (X. Zhang et al., 2020). Another approach to the problem has been developed by E. Kim that proposed an Interpretable Convolutional Neural Network (I-CNN) for indoor activity detection by adding temporal convolution and pooling layers into a CNN (E. Kim, 2020). Using an

Automatic Relevance Determination (ARD) network, L. Li et al. proposed a modified neural network can reveal the relationship between input features and model output (L. Li et al., 2019). As a result, they showed how the day hour was the most influential feature for hourly electricity prediction.

Another way to increase the interpretability of ANN is using the physical domain knowledge to guide networks training process. Z. Chen, Xiao, et al., used this methodology to predict the performances of air-conditioning units (Z. Chen, Xiao, et al., 2022). Z. Yu et al. concluded that a knowledge-based search strategy could significantly reduce ANN training time (Z. Yu et al., 2020). Y. Chen & Zhang proposed a modified Long Short-Term Memory (LSTM) using thermal dynamics parameters for modelling building thermal performance, the developed network can learn interpretable dynamic models from measurement data (Y. Chen & Zhang, 2021). Following this approach, Di Natale et al. proposed a physically consistent neural network by incorporating domain knowledge into an ANN for building thermal modelling (Di Natale et al., 2022). The proposed approach was proved to be physically interpretable.

2.5.2 Encoder-Decoder

Encoder models take an input sequence and creates a contextual representation (which is also called context) of it, instead, decoder models take this contextual representation as input and generates output sequence. Therefore, Enecoder-Decorer models uses textual or categorial input to generate categorical or textual output. A very famous example of Encoder-Decoder model is Google Translator that generates text decoding a textual input (Daza & Frank, 2019). Recently, the attention mechanist has been included inside Encoder-Decoder models to increase machine-machine translation performances and the human interpretability of the output (Bahdanau et al., 2015). Attention mechanism is aimed at mimicking the cognitive attention. It can be made using "soft" weights for the Machine Learning models. Moreover, the weights can change during each runtime, in contrast to “hard” weights that are pre-trained and fine-tuned. This is made by stressing a part of input feature while weakening the others. Encoder-decoder models can use time-series data therefore the attention mechanism should consider time-series data (C. Li et al., 2022). Many studies have used the attention mechanism to analyse temporal dependency in time-series data in both regression and classification tasks. A chiller fault diagnosis (FDD) model has been developed by D. Li et al. to analyse the temporal dependency of time-series data and removing redundant features, thus providing a local interpretation of the sensor data. (D. Li et al., 2019). Instead, Azam & Younis developed an encoder-decoder model for the prediction of energy consumption demonstrating the importance of historical feature inputs on model output (Azam & Younis, 2021).

2.5.3 Clustering and feature extraction

Clustering models are aimed at classifying input data into clusters that shares a common feature, instead, feature extraction is aimed at transforming raw data into numerical features that can be processed while preserving the information in the original dataset. As an example, measuring the dimension of leaves with a clustering model is possible to organize them by the plant species, using a feature extraction model is possible to analyse which feature is plant specific. Unlike ANN and encoder-decoder techniques, clustering and feature extraction can be integrated inside any black-box models without the need to modify their structure. As example, A. Grimaldo & Novak showed that it is possible to use k-Nearest Neighbors (kNN) algorithms obtaining the same performances of more sophisticated machine learning models such as Random Forest (RF) and Gradient Boosted Trees (GBT) to predict energy usage in buildings (A. Grimaldo & Novak, 2021). Due to their aim of organizing and reducing the data complexity dividing it into groups or extracting relevant information these models are highly interpretable (Vigneau, 2021). An innovative clustering technique that is aimed to simulating building thermal design data was developed by Bhatia et al. (Bhatia et al., 2019). Their approach is called axis-aligned hyper-rectangles and can cluster information dividing data into hyper-rectangle boundaries interpretable with specific rules. Using such techniques authors created rules for the range of window-to-wall ratio, to assist the design of building envelopes in different climate zones. The case study developed by Laurinec & Lucká was aimed at developing interpretable time-series for power load forecasting (Laurinec & Lucká, 2019). As a result they showed that it is possible to extract interpretable features from moving windows of time-series data improving demand forecast accuracy.

Time-series data can be analysed by Highly Comparative Time-Series Analysis (HCTSA) toolkit to generate interpretable time-series features. The works of Miller and Xiao showed that it is possible to classify the space usage using building energy consumption data by clustering models (Miller, 2019; T. Xiao et al., 2022). Kasuya obtained the load prediction of the next day using a Gaussian mixture model and a distribution-based clustering algorithm on energy usage data (Kasuya et al., 2020). Z. Chen et al. generated mode labels as input features using a novel early classification approach to enhance the interpretability and performance of building load prediction (Z. Chen et al., 2021). Data visualization is also a common tool used to improve Machine Learning interpretability for clustering and feature extraction models. A. I. Grimaldo & Novak developed a smart energy dashboard to visualize energy consumption on similar day to increase user understanding of energy consumptions (A. I. Grimaldo & Novak, 2019). The dashboard was fed by data obtained from kNN and Decision Trees (DT). They also developed a radar chart dynamic figure to compare weather parameters during similar days (A. I. Grimaldo & Novak, 2019).

2.5.4 Generalized Additive Models (GAMs)

Generalized Additive Models (GAMs) was born as a variant of Linear Regression and Logistic Regression models. Differently from them, GAMs are more generalized but maintains a good degree of interpretability, they can model any non-linear additive effects on different features (Kamath & Liu, 2021). Moreover, GAMs can incorporate irregular and volatile effects to improve flexibility in handling high-resolution data (Khamma et al., 2020). As an example, Khamma et al. showed that it is possible to predict heat and power production using ambient air temperature, solar radiation, and hour of the day as input data (Khamma et al., 2020). The outdoor air temperature had a negative linear relation with heating load prediction, while solar radiation had a negative exponential relationship. In addition, GAMs were also applied to perform sensitivity analysis of input features in thermal comfort modelling (Charalampopoulos, 2019), thermal energy storage modelling (Voss et al., 2021), to identify operational patterns of gas using HVAC systems (Pathak et al., 2018), distributed PV power prediction (Sundararajan & Ollis, 2021), and short-term energy prediction in buildings (Khamma et al., 2020). The main drawback of GAMs is their simplicity that cannot be compared to more complex models but can only approximate the real behaviour of the system analysed. Therefore, GAMs are more suitable for continuous data that don't have high variability in the dataset.

2.5.5 Local Interpretable Model-Agnostic Explanations (LIME)

In 2016, Ribeiro et al. introduced LIME as a model-agnostic method designed to provide localized interpretations for individual predictions (Ribeiro et al., 2016). This approach involves training a local surrogate model to approximate the characteristics of a black-box model in the vicinity of a specific prediction sample. LIME is particularly valuable for explaining classification problems, as it can provide both contradictory and supportive information for each input feature in relation to a given prediction. Wastensteiner et al. applied LIME to interpret machine learning-based time-series classification models in the context of building energy consumption (Wastensteiner et al., 2021). They also conducted an analysis to assess the stability and reliability of these interpretations.

Madhikermi et al., utilized LIME to increase the interpretability of context for Artificial Neural Networks (ANN) and Support Vector Machines (SVM) for Air

Handling Unit (AHU) fault diagnosis (FDD), selecting six random samples for demonstration purposes (Madhikermi et al., 2019). Srinivasan et al. conducted experiments using LIME for chiller FDD, including issues such as scaling in condenser fins, sensor errors caused by flow pulsations, and false alarms (S. Srinivasan et al., 2021). LIME's ability to provide contradicting and supporting information plays a dual role: it assists decision-makers in understanding the model's output, aiding in fault identification, and also helping for the identification of potential false alarms generated by the black-box model. While LIME is often associated with classification tasks, it is also applicable to regression problems. Carlsson et al. for instance, integrated the contradicting and supporting values into a single metric to evaluate the confidence level of individual predictions related to chiller COP (Coefficient of Performance) efficiency, distinguishing between low and high efficiency cases (Carlsson et al., 2020).

2.5.6 SHapley Additive exPlanations (SHAP)

SHAP is designed as a tool for local interpretability, however, it can be employed also for global interpretation. For instance, Carlsson et al. leveraged the average SHAP values to determine feature importance in an ANN model, uncovering key features affecting energy consumption (Carlsson et al., 2019, 2020). Similarly, Ugwuanyi employed averaged SHAP values to globally interpret CO_2 prediction (Ugwuanyi, 2021). Much like LIME, SHAP is useful for the evaluation of influential factors in fault detection scenarios. In S. Park et al. work, SHAP elucidated both local and global interpretations of Random Forest (RF) for Fault Detection and Diagnosis (FDD) for district heating systems (S. Park et al., 2020). Gao et al. used SHAP to interpret RF and Light Gradient Boosting Machine (LightGBM) models for chiller FDD (Gao et al., 2022). In a different context, Sakkas et al. employed XGBoost to detect fraudulent electricity consumption (Sakkas et al., 2021). Additionally, M. Wang et al. used a SHAP model to interpret time-series classification for building energy consumption (M. Wang et al., 2022). SHAP model can be also used for occupancy-related tasks, such as for CO_2 concentration prediction (Ugwuanyi, 2021).

W. Zhang, Wen, et al., adopted SHAP to provide local interpretations for thermal comfort evaluation calculating the individual Predicted Mean Vote (PMV) outputs in a complex numerical equations-based thermal comfort model (W. Zhang, Wen, et al., 2021). Moreover, using SHAP values, the authors proposed a potential solution to enhance thermal comfort across different weather scenarios. H. Park & Park used SHAP values to rank feature importance for predicting natural ventilation rates (H. Park & Park, 2021). The most influential features included pressure differences, outdoor temperature, and wind speed. Additionally, the authors offered individual SHAP value plots for local interpretation. Yang et al., developed a SHAP model to evaluate three thermal sensation models hot, neutral, and cold showing how air temperature and relative humidity are the most influential features across all three models (Y. Yang et al., 2022).

Moving to building energy benchmarking, SHAP helps pinpoint key features contributing to high or low energy usage intensity in individual buildings. As example in the work of Papadopoulos & Kontokosta, SHAP has been in a XGBoost-based residential building energy benchmarking (Papadopoulos & Kontokosta, 2019). SHAP values identified unit density as the strongest predictor for energy use intensity in residential buildings, followed closely by the property value and the number of floors.

SHAP is also very used in load/power prediction applications. For example, Chang et al. utilized SHAP to analyse and reveal feature importance in PV power generation models (TS-SOM and XGBoost) (Chang et al., 2020). Their results highlighted global horizontal irradiance as the most influential feature, aligning with Pearson Correlation (PC) analyses. (Movahedi & Derrible, 2021) delved into interpreting and interrelating three prediction models (electricity, water, and gas consumption) using SHAP. They discovered that the type of building (residential or commercial) and water consumption were key features influencing electricity prediction, showing a strong interrelationship between gas and water consumption due to gas usage in water heating. By examining SHAP values from an XGBoost model, Chakraborty et al. deduced that single-family homes were likely to experience a more substantial increase in building cooling energy consumption in the context of global climate change (Chakraborty et al., 2021). They also found that buildings in hot-humid zones will consume more energy for cooling due to global warming.

In a different context, Golizadeh Akhlaghi et al. used SHAP to interpret performance-related indexes (cooling capacity, Coefficient of Performance (COP), and wet/dew point efficiency) for a dew point cooler using DNN (Golizadeh Akhlaghi et al., 2021). For instance, a sample exhibited higher cooling capacity than the base value due to relatively high intake air velocity. Lastly, using SHAP model Lu et al. spotlighted load and solar generation one hour ahead, along with solar irradiance, as the top three influential features for predicting distributed PV power one hour ahead (Y. Lu et al., 2021).

2.5.7 Other techniques

There are also other techniques different from the aforementioned. For instance, permutation importance can assess feature importance by shuffling feature values and observing the impact on model predictions. A feature is deemed important if its shuffling leads to a substantial prediction error. Carlson et al.

used permutation importance to evaluate feature importance in an Artificial Neural Network (ANN) model for electricity prediction (Carlsson et al., 2019, 2020). However, like Partial Dependence Plots (PDP), permutation importance can be biased when features exhibit strong correlations. Another approach taken by C. Zhang et al. involved the use of a Dimensionless Sensitivity Index (DSI) to quantify feature importance (C. Zhang et al., 2020). Their results indicated that time-lagged cooling load features held more sway than other factors. J. Y. Kim & Cho, on the other hand, employed the Kullback–Leibler divergence method to measure feature relevance in prediction using latent states from an encoder-decoder model (J. Y. Kim & Cho, 2019).

Tree-based methods have been widely used in evaluating feature importance by calculating each feature's contribution to reducing impurity within the tree model. These methods include Random Forest (RF), gradient boosting machine (GBM), XGBoost, and Cubist (Moon et al., 2022b, 2022a). For example, Moon et al. showed how in a building load prediction scenario utilizing Cubist regression it is possible to rank the importance of features, revealing that external factors such as outdoor air temperature and holiday indexes, along with internal factors like one day-ahead and one week-ahead energy load, plays fundamental roles (Moon et al., 2022a).

Sipple proposed an unsupervised anomaly detection method for Artificial Neural Networks (ANN) to identify power meter device failures in office buildings, employing an integrated gradients approach to interpret anomalies (Sipple, 2020). M. Wang et al. developed a similar work by utilizing ANN gradients models to quantify the marginal impact of features on predictions based on a backpropagation rule for HVAC performances evaluation, the gradients were also leveraged for feature selection (M. Wang et al., 2022).

Counterfactual explanation is yet another system for generating local interpretations of individual samples. This method creates nearby samples with minimal feature changes that alter the model's output. Sakkas et al. selected features through statistical analysis and then utilized them for Diverse Counterfactual Explanation (DiCE) model to conduct counterfactual analysis to interpretate energy demand forecasting (Sakkas et al., 2021).

The ELI5 Python package, named "explain like I'm five," has been used to interpret various black-box machine learning models, including XGBoost, LightGBM, CatBoost, Keras, and Scikit-learn. Sarp et al. employed ELI5 to interpret an XGBoost model, facilitating the deployment of a machine learning-based renewable energy prediction model (Sarp et al., 2021). In their study, time and irradiance emerged as the most influential features for predicting solar power generation. In addition to global interpretations, the authors also delved into local interpretations using ELI5.

W. Zhang, Liu, et al. trained a rule-set surrogate model to replace the Random Forest (RF) model for building energy prediction (W. Zhang, Liu, et al., 2021). Another study of G. Li et al. employed a surrogate model to replace a multi-objective optimization algorithm for HVAC setpoint control (G. Li et al., 2021). Remarkably, the results demonstrated that straightforward decision-tree-like rule sets could achieve about 90% of the performance of detailed model predictive controllers while significantly reducing computational costs.

2.6 SmartLAB: Use case scenario and Proof of Concept (PoC)

The present research outlines the design and development of a control and management system for a local microgrid located within the Faculty of Architecture at Sapienza University of Rome (**Figure 23**). This initiative seeks to establish an experimental demonstration environment, called the "SmartLAB", endowed with photovoltaic energy generation capabilities and an electric vehicle charging station, leveraging advanced artificial intelligence (AI) self-learning techniques. In this scenario, the study aims to delineate an experimental methodology to develop a Digital Twin prototype, serving as a Proof of Concept (PoC), that specializes in the management and optimization of Indoor Air Quality (IAQ) and Energy Management for buildings connected to a microgrid.

Figure 23. Faculty of Architecture, Sapienza University of Rome

2.6.1 Objectives and Key Performance Indicators

The prototype aims to experiment with the use of IoT (Internet of Things) technologies and open-source tools to establish a Digital Twin architecture. This architecture is designed to manage and reduce energy consumption, enhance the utilization of spaces, and increase the knowledge of the entire real estate portfolio, ensuring air quality control in the most critical indoor environments. It involves the installation of IoT devices to measure the energy consumption of equipment and integration with existing building automation systems (BACS and BMS) to achieve energy efficiency goals. This also includes monitoring thermo-hygrometric well-being, safety, and environmental quality.

Efficient management of building technology systems translates into energy savings and CO_2 emission reduction, aligning with national and international directives. The system also aims to provide planning and decision-support tools for emergency management, including real-time data acquisition and dashboard and visual analytics for information sharing with users and building managers.

In the subsequent sections, the multifaceted approaches to Energy Management, Indoor Air Quality Control, and Space Management and Optimization representing the main targets of DT-based building management technology are described.

Energy Management

Multi-zone climate management: temperature regulation based on usage modes and independently for different areas of the building.

Automatic lighting: lights turn on and off automatically based on people's presence.

Disabling temperature control with open windows: deactivation of heating and cooling systems in individual rooms when doors and windows are opened.

Heating and cooling in economy mode in the absence of people: automation of the heating system based on the presence or absence of people.

Management and monitoring of energy consumption data: application of machine learning algorithms to highlight patterns and define strategies for energy efficiency.

Reduction of malfunctions and service disruptions: capability for remote monitoring and control of devices.

Indoor Air Quality Control

Automatic air exchange: automation of window opening based on timing and usage of critical environments such as meeting rooms and printer rooms, detecting concentrations of airborne pollutants.

Space Management and Optimization

Improving knowledge of the real estate portfolio: creating a unique and distributed data source to support decision-making processes.

Data analysis of space occupancy: defining strategies for space utilization optimization and scenario simulation in response to critical events or emergencies.

Key Performance Indicators (KPIs)

Primary Energy (kWh/m^2): Consumption in the supply chains of used energy carriers.

Energy Demand and Consumption (kWh/(m^2 month or year)): Assessment of the building's energy demand and consumption.

Energy Saving (%): Percentage reduction in energy consumption compared to the baseline.

Global Energy Performance Indicator – EpGl (kWh/m^2): Non-renewable energy consumption, like gas for heating or hot water production.

Peak Load Reduction (%): Comparison of peak demand before and after technology implementation.

Building Operational Performance KPI (%): Correlation of energy consumption, emissions, and geometric information.

Reduction in energy costs due to ICT technologies (%): Change in energy-related costs post ICT technology implementation.

EU Energy Label: Energy efficiency rating from A to G.

Increased reliability (%): Reduction in faults for greater reliability and fewer interruptions.

User involvement: Engagement in controlling energy use in the building.

Average System Interruption Duration Index: Estimate of the average duration of interruptions affecting users and maintenance costs.

Average System Interruption Frequency Index: Estimate of the average number of service interruptions experienced by a typical end user over a defined period.

Indoor Air Quality Control

Reduction of airborne pollutants (CO_2, TVOC, PM2.5, etc.): Measurement of pollutant concentration reduction in meeting rooms and printer areas.

Space Management and Optimization

Enhancing knowledge of indoor and outdoor spaces: Reducing inefficiencies in space management and utilization.

Table 6. Objectives, KPI and Enabling technologies

Domain	Objective	KPI	Indicator	Tools, Methods and Technology
1. Energy Management	1.1 Energy Breakdown - Reduction of energy consumption for lighting and air conditioning	1.1.1 Reduction of building energy need and consumption	(kWh/m^2), EpGl, EpGl,nren	Energy analysis, BIM
		1.1.2 Reduction of primary electrical energy consumption	$kWhel/m^2$, kWhel/year; kWhel/day;	Smart Building smart metering, BMS
		1.1.3 Reduction of primary thermal energy consumption	$kWhth/m^2$, kWhth/year; kWhth/day;	Smart metering
		1.1.4 Reduction of energy costs	€/m2; €/m^2 year	Data analytics, BIM
		1.1.5 Energy efficiency class attribution	A-G	Energy Performance Certificate (APE)
		1.1.6 Building Operational Performance	BOPerf = (EnEq + EnLig + EnHVAC) / Area * Time * Ems	Smart meter/plug, BMS
	1.2 Reliability and occupant satisfaction – Reduction of service disruptions and malfunction	1.2.1 Reduction of service disruptions and malfunctions	Number of tickets/month	Data analytics (facility management data)
		1.2.2 Reduction of extra maintenance costs	€/month materials, €/month personnel (extra)	Data analytics (facility management data)

	s of terminals through remote management/monitoring and automated flows	1.2.3 Average duration index of system interruptions	min./day, min./week, min./month	Data analytics, BMS (facility management)
		1.2.4 Average frequency index of system interruptions	number/day, number/week, number/month	Data analytics, BMS (facility management)
2. Indoor Air Quality Control	2.1 Reduction of airborne pollutants (CO_2, TVOC, PM2.5 etc.)	2.1.1 Monitoring of airborne pollutant concentration in critical environments	Room occupancy, CO_2 levels, Humidity levels, Average seasonal temperature	IAQ sensor, BMS, BIM, data analytics
		2.1.2 Reduction of airborne pollutant concentration in critical environments	Room occupancy, CO_2 levels, PM presence, TVOC levels, Humidity levels, Average seasonal temperature	IAQ sensor, IAQ actuator, BMS, BIM, data analytics
3. Space Management and Optimization	3.1 Improving knowledge of indoor and outdoor spaces	3.1.1 Reduction of inefficiencies related to the management and use of spaces	Number of unused spaces, Number of unrecorded environments, Number of rooms not assigned to a CC	Data analytics, presence sensors, BMS (facility management, inventory, BIM)
	3.2 Improving knowledge of indoor and outdoor space occupancy	3.1.2 Probabilistic estimation of occupant presence in environments for emergency management	Times to secure environments; Extent of damage to people and property	Data analytics, presence sensors, simulation, BIM

2.6.2 Applied Digital Twin Strategy

To understand the conceptual foundation of this research, the comprehensive definition provided by Michael Grieves is a key element:

"A Digital Twin represents a comprehensive suite of virtual constructs which encapsulate every detail of a prospective or extant physically manufactured entity. This spans from its intricate atomic structures to its overarching geometric configurations. Ideally, any intelligence gleaned from a physical inspection of such an entity should be readily retrievable from its Digital Twin counterpart". Michael Grieves, 2016, Digital Twin Institute.

The International Energy Agency (IEA) forecasts a burgeoning role for artificial intelligence in the energy sector, catalyzing a transformation in global energy systems to become more interconnected, reliable, and sustainable. Economic projections suggest that by 2040, there will be an estimated one billion "Smart Homes" and 11 billion smart devices globally. Efficiently optimizing these through AI could lead to a more than 10% reduction in energy consumption. Integrating this with photovoltaic energy production further provides end-users and managers the tools to make more informed energy management decisions.

Microgrids harvest vast amounts of data from Internet of Things (IoT) systems, particularly their sensors. This data undergoes processing and translation into actionable insights by sophisticated self-learning algorithms. Concerning renewable energy production, unpredictable weather conditions can introduce significant challenges. AI systems are well-positioned to optimize the production, transmission, and storage of energy harnessed from photovoltaic installations.

The culmination of these efforts is the realization of a small-scale smart energy grid, entirely overseen by artificial intelligence through machine learning algorithms. The AI software, at a home automation level, monitors, and interprets the unique energy requirements of individual users, proactively intervening to eliminate any unnecessary consumption.

AI's capabilities extend to monitoring multiple photovoltaic installations within a network, amalgamating data related to energy generation, maintenance needs, and generation efficiency. Within the realm of photovoltaic energy production system control and management, AI techniques can be delineated into three primary areas:

- Forecasting and modeling meteorological data
- Fundamental solar cell modeling
- Sizing of photovoltaic installations.

Consequently, through real-time AI analysis, the system can alert users of potential power plant malfunctions, anticipate energy generation, and maintain a comprehensive database for the optimal management of photovoltaic systems.

For instance, should solar energy reserves approach depletion, the system can automatically switch off a television, dim the lights, or even reduce a stereo's volume or a fan's intensity.

Such a system, when synchronized with a building-scaled photovoltaic energy production system, can provide estimates of costs and consumption. This equips a local microgrid manager or user to make more sustainable decisions.

The research endeavor is structured into three distinct implementation phases:

1. Creation of a digital information system using Building Information Modeling (BIM) methodology; design and implementation of a photovoltaic installation for electrical energy production and an electric mobility charging system;
2. Design and execution of a home automation system serving the SmartLAB to holistically manage installations and devices;
3. Development of a machine learning system that records and discerns the energy needs of the SmartLAB and vehicle charging, preemptively forecasting the actual production of the photovoltaic system. This then proactively interacts with the storage system's usage and potentially remodulates consumption.

The benchmark for assessing the research's accuracy and relevance will hinge on analyzing data from consumption monitoring software reports of the realized micro-smart grid. This is juxtaposed against daily production data from the installed photovoltaic system. The AI management system's performance will be assessed using the most contemporary criteria for building and microgrid energy efficiency and management.

The facility, depicted in **Figure 24** and **Figure 25**, consists of an electrical generation system based on photovoltaic panels mounted on a flat roof. These panels, crucial components in harnessing sunlight and converting it to electricity, operate in conjunction with an inverter, a device integral to the transformation of direct current (DC) generated by the panels into alternating current (AC) suitable for domestic or commercial use.

Figure 24. Photovoltaic system powering the SmartLAB

Figure 25. Photovoltaic system powering the SmartLAB

Connected to this energy production system are several loads, which are essentially components or sections of the facility that consume the generated power. Notable among these are:

- Lights in the corridors adjacent to the main hall
- Laboratories of the Interdepartmental Center
- Electric vehicle (EV) charging station

The establishment of the network infrastructure involves the placement of meters, as depicted in **Figure 26**. These meters are vital components for the real-time monitoring of various parameters like electricity consumption, generation, and load demand within the infrastructure.

These measuring instruments need to be interconnected through a dedicated data network using hybrid technology. This technology combines both wired and wireless connections. Specifically:

- Wired Connections: these connections often offer stability, high-speed data transmission, and reduced interference. In settings where real-time data is critical, such as power consumption metrics, a wired connection is often the preferred choice due to its consistent and reliable performance.

- Wireless Connection Components: while wired connections bring stability, wireless connections bring flexibility. They can be especially useful in areas that are hard to reach or where laying physical cables might be challenging or not cost-effective. Modern wireless communication technologies, such as Wi-Fi 6 or LoRaWAN, can ensure a relatively stable and secure connection over considerable distances.

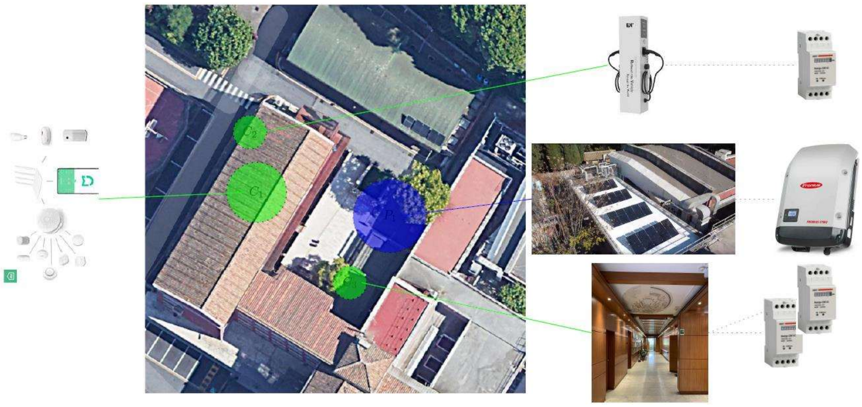


Figure 26. SmartLAB's network infrastructure

The architectural layout (**Figure 27**) of the SmartLAB laboratory is characterized by:

- **Two Distinct Macro Environments:** these zones are linked by a transitional corridor and are segregated by a translucent polycarbonate partition.
- **Functional Workspaces:** each of these environments is equipped with multiple dedicated workspaces, complete with workstations and advanced computing infrastructure. These setups are integral for facilitating both routine research and specialized experimental undertakings.

In the context of the case study's objective, these two macro environments will be approached as independent entities. The primary aim is to monitor environmental parameters encompassing temperature, relative humidity (RH), carbon dioxide (CO_2) levels, and overall energy consumption metrics.

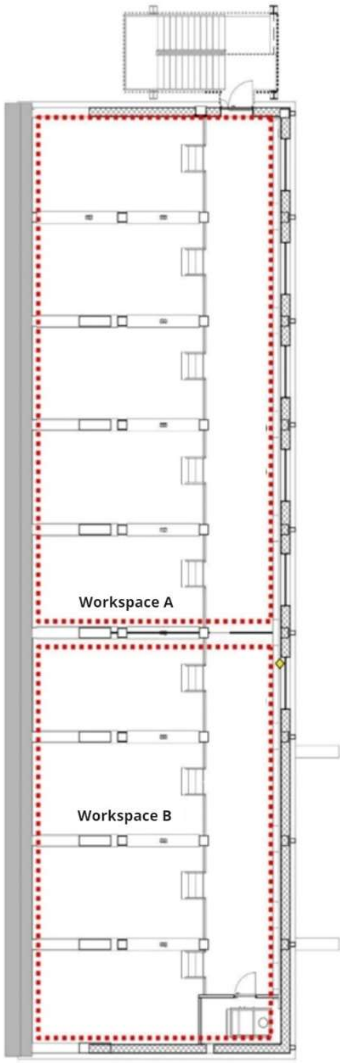


Figure 27. SmartLAB's architectural layout

Building upon this foundational understanding, the present research delineated specific requirements for the Digital Twin's PoC, summarized as:

- Data Management: real-time processing of data from IoT devices with storage in specialized databases.
- Visualization: dynamic portrayal of data through three-dimensional models in real-time.
- Data Analysis: employing Artificial Intelligence (AI) techniques for data interpretation, predictive analytics, and disseminating insights via comprehensive dashboards.
- Implementation: concrete strategies rooted in the assimilated data to enhance Indoor Air Quality (IAQ) and streamline energy management.

These requirements were then methodically aligned with capabilities inspired by "The Digital Twin Capabilities Periodic Table" (CPT) for the digital model's development. Key aspects of significant relevance were discerned based on the Clusters proposed by the research consortium (refer to **Figure 28**).

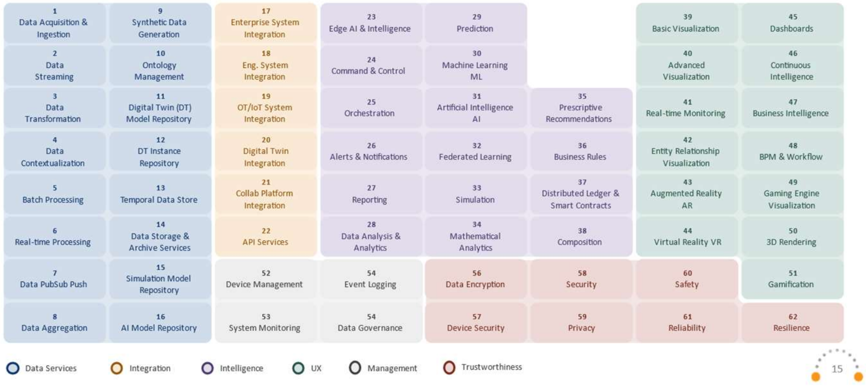


Figure 28. Digital Twin Consortium Capabilities Periodic Table

The design phase aimed to ascertain the alignment between physical assets and their digital counterparts as described below.

a. Pinpointing use-case specific objectives

Recognizing essential capabilities based on insights from the Digital Twin Consortium, 2022.

Opting for pivotal technologies and demarcating requisite information types across the asset's lifespan.

The architecture of the system for the SmartLAB Digital Twin encompasses: a BIM model for virtual representation; an IoT framework for real-time data capture; an interactive IoT platform bridging Physical and Digital Twins; a database for data storage; and a system dedicated to data visualization and analytics.

b. Tool Selection & Benchmarking

There's a research deficit regarding effective Digital Twin creation tools. Both proprietary and open-source solutions were evaluated. Given the integration gap between BIM and IoT, an IoT platform synergized with a BIM-integrative application was explored. The choice was the open-source Node-Red-based platform, favored for its user-friendliness and comprehensive toolset that facilitates remote IoT device monitoring, data storage, rule-setting for action triggers, and data analytics.

c. Proof-of-Concept Creation

The preliminary phase involved configuring the platform using simulated sensors, replicating real ones. These sensors periodically relayed predefined values, either generated through Python scripts for realistic scenarios or manually for edge cases. IoT device interaction rules were formulated, and their efficacy was validated using simulated data.

d. Synchronizing Physical and Digital Twins

Post-validation with simulated sensors, the physical IoT devices were integrated into the SmartLAB. Ensuring high-quality data from sensors was pivotal. With collaboration from manufacturer technicians, device models and optimal positioning were determined. Subsequently, connections were established between IoT devices and the data platform, with real-time data driving the rule configurations.

e. Historical Data Accumulation

Sensor data is archived in a PostgreSQL database, accessible for advanced analytics via Machine Learning. Once adequate data is amassed, it forms the foundation for training Machine Learning models.

f. Machine Learning Integration

With centralized data storage, advanced analytical models can be crafted. These models are adept at emulating and optimizing the primary parameters of the sophisticated energy grid. Utilizing location-focused machine learning techniques and rules, active devices can be discerned. AI techniques transform power consumption patterns into "energy device words", and a "Naïve Bayes classifier" discerns each energy load's category, highlighting discrepancies between digital and real-world representations.

g. Data Visualization and Analysis

Within the IoT platform, data is exhibited on user-centric dashboards. Real-time information is juxtaposed against reference benchmarks for intuitive comprehension. Time-series graphs facilitate historical data comparison across sensors.

h. Augmentation Phase

Given a Digital Twin's inherent requirement for scalability and expandability, the SmartLAB solution was conceived with adaptability in mind. Leveraging open-source software ensures seamless future tool integrations, augmenting the Digital Twin's prowess. Post-validation, this model's functionalities will extend to encompass the entire building, a testament to the platform's scalability accommodating multitudes of IoT devices.

In the contemporary landscape of digital twin technology, there's a diverse spectrum of offerings. The market currently leans towards two primary segments:

1. **Holistic Commercial Solutions:** these are all-encompassing platforms offering end-to-end solutions for digital twin creation, management, and analysis.

2. **Modular Solutions:** this approach allows the combination of multiple applications, each catering to specific functions, thereby piecing together a more bespoke digital twin solution. Within this modular framework, both commercial and open-source solutions find representation.

For the scope of the study, a hybrid methodology was carried out, emphasizing especially on leveraging the strengths of open-source platforms.

The following diagram is related to the architecture of the DT system which is divided into various components and layers that work together to create a digital counterpart of a physical asset, allowing for simulation, monitoring, and management through the use of technologies such as IoT (Internet of Things), BIM (Building Information Modeling), and databases (**Figure 29**).

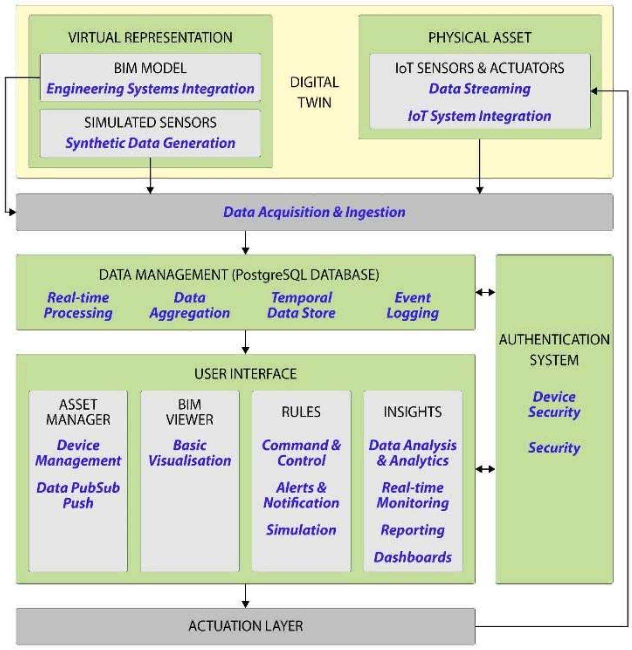


Figure 29. Proposed conceptual architecture for the CITERA SmartLAB

The main components are:

1. Virtual Representation: It includes the BIM model, the integration of engineering systems, and simulated sensors for synthetic data generation.
2. Physical Asset: Concerns IoT sensors and actuators, data streaming, and IoT system integration.
3. Data Acquisition & Ingestion: This step is responsible for the acquisition and ingestion of data from the physical level to the virtual one.

4. Data Management: Uses a PostgreSQL database for real-time processing, data aggregation, temporal data storage, and event logging.
5. User Interface: Comprises various modules such as the asset manager, BIM viewer, rules for command and control, insights for data analysis, and dashboards.
6. Authentication System: Deals with the security of devices and the system as a whole.
7. Actuation Layer: This is the output layer that allows the system to act on the physical world, likely through commands sent to actuators.

The flow of data primarily occurs from the physical asset to the virtual representation and back, enabling continuous monitoring and control of the physical asset via its digital twin. This type of architecture is typically used for the management of buildings, industrial plants, and other infrastructures to optimize operations, maintenance, and planning.

2.6.3 Data acquisition and transmission

Automated production operations can be networked together thanks to real-time data-driven technologies, producing direct communication and coordination systems and resulting in highly autonomous processes. Then the Digital Twin paradigm defines a model where computer-driven systems monitor physical processes, configuring a virtual copy of the physical world and obtaining decentralized decisions based on self-organization mechanisms (Smit, 2016)

The collaboration between 3D information models and IoT devices is highly necessary for a successful implementation of real-time DT purposes as well as for developing energy management optimizations. As such, the BIM Model containing data and information useful to the process assessments becomes able to communicate with the real system using data from sensors, developing learning capabilities able to process the received information.

The implementation of IoT in the real-world environments in a smart, ubiquitous and live-interconnected way is still partially restricted by barriers like device battery life, network capacity and the cost of maintaining both.

The core functionality of IoT devices in DT configurations for energy and IAQ management is to reliably collect and share data (such as flow rates, temperatures, pressures, physical movements, distance, mass etc.) from its designated environment to the virtual world.

The hardware element consists of a battery-powered sensor, an actuator and a communication system. The collected data is then processed on the device and consequently sent to remote servers through the communication network.

In the present methodology, the connection between the physical and virtual model is made through sensors (Raval, 2021) able to monitor and communicate electrical power data such as Power Energy Voltmeter Ammeter for lighting and HVAC systems, and Smart Plugs for electromotive equipment such as computers, televisions, washing machines and so forth (Karami et al., 2018).

AI systems then allow the DT to develop predictive capabilities, learning from the events and improving the outputs, ultimately taking and implementing autonomous decisions based on the analysis carried out without human interventions.

Then the AI system achieves a balance between energy consumption and performance parameters (Corry et al., 2015) of energy production system, adapting himself to the environment in order to achieve the predefined objectives. In other words, the system takes data from sensing devices and uses a reasoning system to generate appropriate and specific actions, modifying the behavior of the equipment in order to optimize the energy consumption.

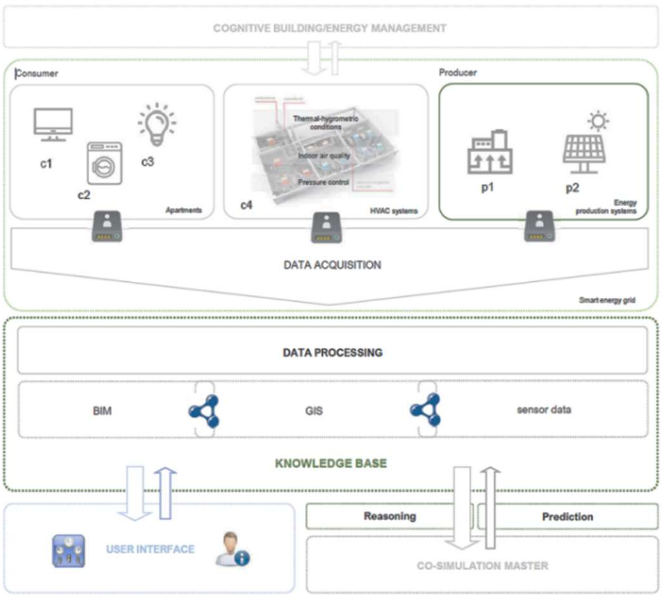


Figure 30. DT comprehensive conceptualization

The precision-driven design of modern architectural environments requires a consistent approach, especially in integrating sensor-based monitoring systems. These sensors and their peripherals are crucial for real-time feedback, enabling users to fine-tune environmental conditions for optimal performance and safety (**Figure 30**).

To achieve comprehensive monitoring of the stated environment, an array of strategically placed sensors has been deployed. These sensors are tasked with tracking various environmental parameters. Furthermore, the associated peripherals are designed not only to receive this data but also to execute specific, targeted interventions based on the received input.

2.6.4 IoT sensors network

In detail, the sensor suite comprises the following devices:

Shelly SmartPlug S: a smart plug device that can remotely control and monitor the power consumption of connected electrical devices.

Shelly Motion Sensor: a device adept at detecting motion within its field of view, crucial for understanding the occupancy and movement patterns in the space.

Shelly Relay 2.5: a versatile actuator that allows for the remote toggling of devices, making it instrumental in dynamic environmental adjustments.

S**helly Door/Window:** this sensor provides insights into the opening or closure status of doors and windows, thus playing a crucial role in safety and energy conservation.

Kolb Burmeister CO_2 Ampel: a CO_2 measurement device, ensuring that the indoor air quality ensuring safe and comfortable levels for the occupants.

A notable feature of the Shelly devices is their compatibility with the MQTT protocol (Message Queuing Telemetry Transport). This lightweight, publish/subscribe network protocol is specifically tailored for devices operating over potentially unstable networks, making it ideal for IoT applications. Furthermore, these peripherals come with an in-built web server, functioning as MQTT brokers. This design facet ensures that the devices are self-reliant, negating the need for external servers, and can directly connect to the network, rendering them instantly operational.

Complementing this setup are the XIAOMI sensors within the laboratory. These devices communicate through the proprietary MIoT-Spec protocol. They interact with dedicated gateways and use the ZIGBEE wireless standard for intra-laboratory communication.

Lastly, the Schneider Electric PowerTags have been integrated into the system. Installed within the control panel, these devices liaise through the ModBus protocol, a time-tested serial communication protocol conceptualized in 1979 for interfacing programmable logic controllers.

An intricate web of sensors, actuators, and communication protocols culminates in a responsive, adaptive, and efficient environment, reflecting the zenith of modern architectural practices described as follows (**Figure 31**).

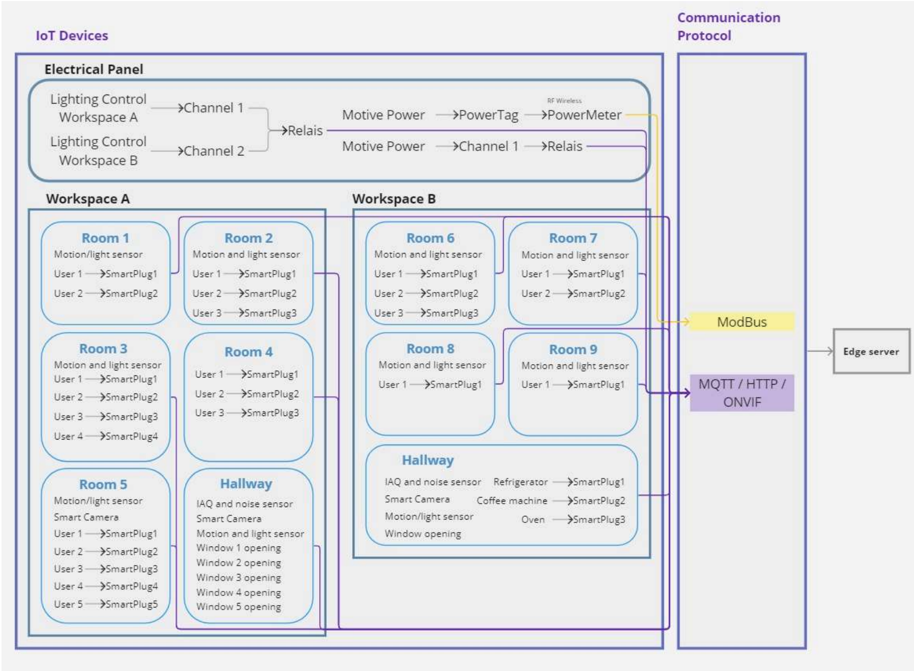


Figure 31. Indoor IoT devices and Communication protocols

Outdoor devices are shown in **Figure 32** and include sensors for photovoltaic production, consumption associated with electric vehicle charging, and the weather station.

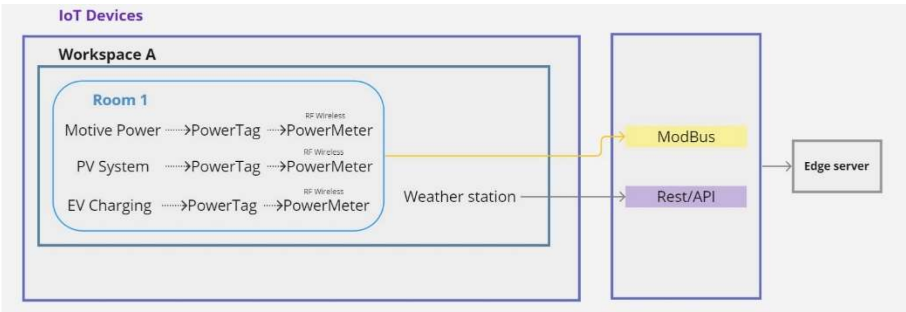


Figure 32. Outdoor IoT devices and Communication protocols

Since most sensors monitor multiple parameters simultaneously, the list of quantities monitored by the different sensors is detailed in **Figure 33** and **Figure 34**.

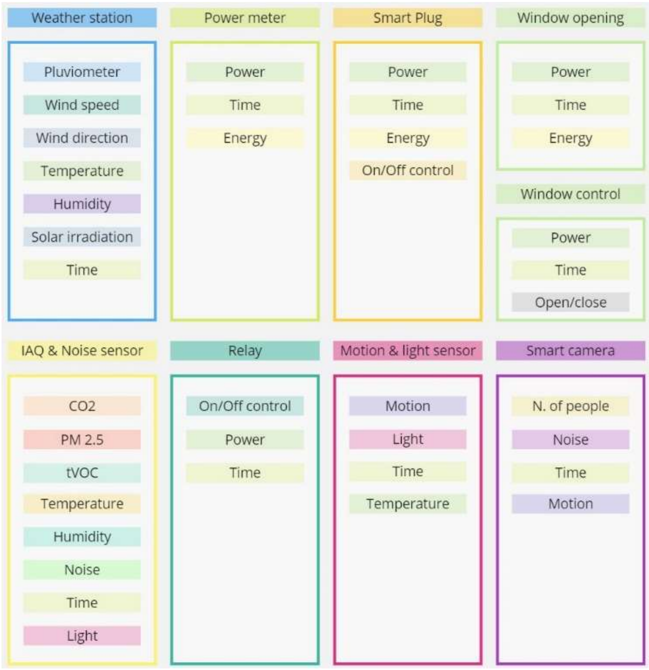


Figure 33. Quantities monitored by the different sensors

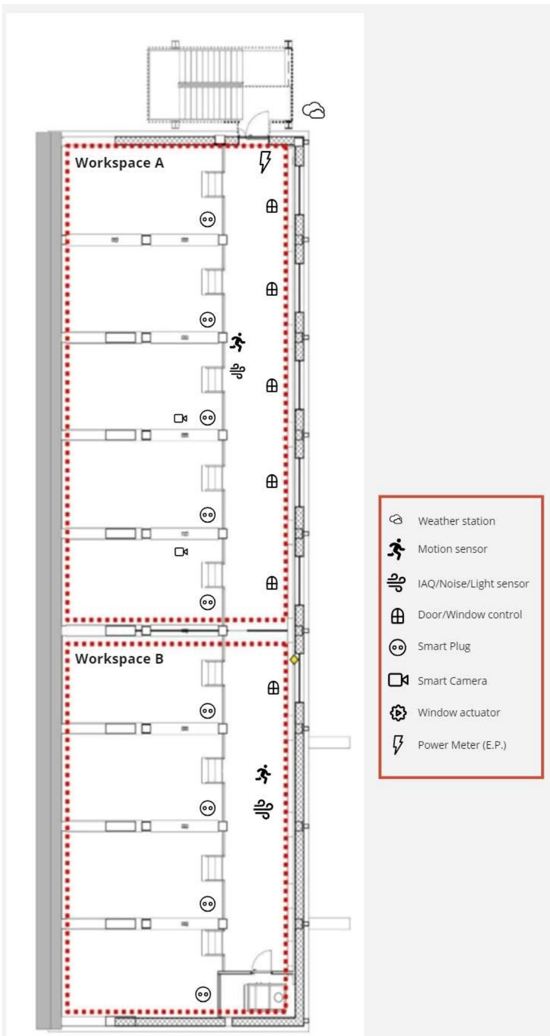


Figure 34. SmartLAB's layout and sensors' location

Electricity Consumption Sensors

The number of smart-plug electricity consumption sensors is determined by the need to monitor the consumption of individual users within the CITERA environments individually and to correlate the data with the presence data from cameras to simulate consumption profiles. The sensors can monitor the instantaneous power required by the user (W), the overall electrical consumption (kWh) by integrating power over time (min), and potentially turn on and off downstream devices to reduce consumption due to the presence of hardware in standby or service utilities (coffee machines, fixed lamps, etc.).

The presence of power-meters at the panel allows for the monitoring of overall consumption (kWh) and the required power (W), taking into account and separating data from the lighting system, air conditioning, outlets not covered by smart-plugs, and energy losses due to transportation. Furthermore, both the photovoltaic system and the electric car charging station are monitored. However, these elements are currently not active within the structure of the Valle Giulia faculty due to delays in the authorization process. The number of channels provided in the power-meters is proportional to the number of environments and the main expected loads. Specifically, the following are monitored through the different channels of the power-meters:

- CITERA's air conditioning system.
- CITERA's motive power system.
- CITERA's lighting system.
- Photovoltaic system.
- Electric charging station.

Monitoring of electrical utilities occurs via smart plugs for the analysis of instantaneous and average power (W), status (on-off), and usage time (min); it is possible to monitor the general electrical consumption of utilities, lighting system, air conditioning, electric column, and photovoltaic system from the panel via dedicated power-meters, as well as the analysis of average and instantaneous power (W), and time (min).

Presence Sensors

Occupancy monitoring is ensured by the presence of occupancy sensors and cameras capable of working with computer vision algorithms. Monitoring of all environments is carried out using presence sensors, delegating the counting of people to cameras only for environments where crowding of several people at the same time is expected.

IAQ Sensors

The assessment of air quality (IAQ) and thermal comfort is carried out through the integration of data available from a series of sensors. In particular, IAQ is guaranteed by the presence of sensors capable of analyzing the level of concentration of tVOC, PM2.5, PM10, and CO_2 in terms of ppm simultaneously. Thus, air quality sensors are expected to be one per room. Considering the air recirculation in different environments due to thermal loads, it is not necessary to plan for more, since small variations in the values within the environments are not significant for monitoring IAQ and thermal comfort (the full scale and error of these sensors are normally lower than the concentration variations naturally present in the environments). The sensors are able to provide air quality assessments also through aggregated metrics (IAQ indices) and algorithms for compliance with the relevant ISO standards will be implemented on the SW side.

Thermal Comfort Sensors

In order to evaluate the temperature, sensors capable of recording the temperature in degrees (°C) and ambient humidity in terms of relative humidity (RH) are inserted in each environment. This information is obtained from presence sensors (temperature), IAQ sensors (temperature and humidity), and sensors for the closing of openings (temperature). The data will be taken both punctually and aggregated through Machine Learning logic in order to obtain a general data capable of describing each environment also at a predictive level and performing PVM and PPD calculations as described by UNI EN ISO 7730:2006.

Lighting Sensors

The overall assessment of internal lighting is carried out through the use of Luxmeters (Lux) installed inside presence sensors, IAQ sensor, and magnetic sensors for doors/windows. As for the assessment of thermal comfort, the data will be both aggregated and averaged to create a comprehensive and punctual measure based on the information collected in the environment. Furthermore, in the future, the data thus obtained may be used for the evaluation of lighting comfort at the level of work surfaces (Standard UNI EN 12464-1) and for the dimming of the fixtures (currently not supported by the lighting system).

Actuators

Smart Plugs, the automatic opening system of the openings, and the relays represent the main actuators inserted within the case study. The actuation of the devices must be carefully evaluated, as turning off the lighting systems represents a risk for safety, turning off the electrical outlets can cause data loss from computer devices, damage the connected devices, or interrupt ongoing experiments. The expected number is sufficient to ensure the monitoring of electrical loads and remote control of the individual workstations to which they are connected. The relays are instead inserted internally to the electrical panel for monitoring the loads and control (on/off) of the lights inside CITERA and in the box for controlling the servomotor in charge of opening and closing the window. Each of the actuators is able to record general absorption in W, the system status, and the usage time. These data allow through their historicization the use in Machine Learning logics for the development of usage models and the creation of future predictions.

Finally, among the actuators, the role of maintainers and users is integrated, who through a Human-in-the-loop logic, can perform the actions recommended by the system where the installation of automatic actuators is not foreseen or not possible.

Table 7. Smart Lab's IoT equipment

Device	**Location**	**Product**	**Function**	**Com. protocol**
Power-Meter	Outdoor	Acti9 PowerTag Link Schneider	Electrical consumption for air conditioning	ModBus
	Outdoor	Acti9 PowerTag Link Schneider	Photovoltaic production	
	Outdoor	Acti9 PowerTag Link Schneider	Electric charging station	
Weather station	Outdoor	Davis Instruments Vantage Pro2	Solar radiation, rain gauge, wind speed and direction, absolute pressure, temperature, and relative humidity	REST/API Davis
Workspace A				
Power-Meter	Workspace A	Acti9 PowerTag Link Schneider	Electrical consumption and total power monitoring	ModBus
Relè	Workspace A	Shelly 2.5	Monitoring of consumption and absorbed power for turning on/off lights in environment A/B	HTTP / MQTT
	Workspace A	Shelly 1	Turning on/off	HTTP / MQTT
	Workspace A - Hallway	Shelly Motion 2		HTTP / MQTT

Motion Sensor	Workspace A – Room1		Monitoring of presence, lighting, and temperature	
	Workspace A – Room2			
	Workspace A – Room3			
	Workspace A – Room4			
	Workspace A – Room5			
IAQ	Workspace A	Awair Omni	Monitoring of temperature, relative humidity, CO_2, tVOC, PM2.5, noise, and lighting	HTTP / MQTT
Window sensor	Workspace A – Window1	Shelly Door /Window sensor 2	Opening/closing windows, lighting, and temperature	HTTP / MQTT
	Workspace A – Window2			
	Workspace A – Window3			
	Workspace A – Window4			
	Workspace A – Window5			
Smart camera	Workspace A - Hallway	Tapo C110 TpLink	People counting, noise	HTTP / ONVIF
	Workspace A – Room3			
	Workspace A – Room5			
Smart-plug	Workspace A – Room1 – Workstation 1	Shelly Plug S	Monitoring of electrical consumption and power	HTTP / MQTT
	Workspace A – Room1 – Workstation 2			
	Workspace A – Room2 – Workstation 1			
	Workspace A – Room2 – Workstation 2			
	Workspace A – Room2 – Workstation 2			

	Workspace A – Room3 – Workstation 1			
	Workspace A – Room3 – Workstation 2			
	Workspace A – Room3 – Workstation 3			
	Workspace A – Room3 – Workstation 4			
	Workspace A – Room4 – Workstation 1			
	Workspace A – Room4 – Workstation 2			
	Workspace A – Room4 – Workstation 3			
	Workspace A – Room5 – Workstation 1			
	Workspace A – Room5 – Workstation 2			
	Workspace A – Room5 – Workstation 3			
	Workspace A – Room5 – Workstation 4			
	Workspace A – Room5 – Workstation 5			
Motion sensor	Workspace B – Room6	Shelly Motion 2	Monitoring of presence, lighting, and temperature	HTTP / MQTT
	Workspace B – Room7			
	Workspace B – Room8			
	Workspace B – Room9			
	Workspace B - Hallway			

IAQ	Workspace B	Awair Omni	Monitoring temperature, relative humidity, CO_2, tVOC, PM2.5, noise, and lighting	HTTP / MQTT
Smart camera	Workspace B - Hallway	Tapo C110 TpLink	People counting, noise	HTTP / ONVIF
Window actuator	Workspace B – Window6	Shelly Door /Window sensor 2	Opening/closing windows, lighting, and temperature	HTTP / MQTT
	Workspace B - Hallway	Shelly 2.5	Window operation (on-off), consumption	HTTP / MQTT
Smart-plug	Workspace B – Room6 – Workstation 1	Shelly Plug S	Monitoring of electrical consumption and power	HTTP / MQTT
	Workspace B – Room6 – Workstation 2			
	Workspace B – Room6 – Workstation 3			
	Workspace B – Room7 – Workstation 1			
	Workspace B – Room7 – Workstation 2			
	Workspace B – Room8 – Workstation 1			
	Workspace B – Room9 – Workstation 1			
	Workspace B – Hallway 1			
	Workspace B – Hallway 2			
	Workspace B – Hallway 3			

Communication Protocols

The management of communication between devices and the EDGE server node is carried out using various IoT communication protocols. The use of HTTP allows connection and exchange with any device that has the appropriate configuration and registration on the general network, thus respecting a stable and reliable communication standard as required by Digital Twin-type infrastructures. MQTT, on the other hand, is a messaging protocol designed for Internet of Things (IoT) networks. The use of the MQTT protocol requires a specific Broker service that is performed separately through service within server nodes or concentrators. Where possible, the use of HTTP is preferred in order to reduce the number of active software services on the server. Modbus is a communication protocol used for transmitting information over serial lines between electronic devices to an industrial standard. ONVIF is a standard dedicated to managing video streams from IP cameras and is preferable to data access via FTP as it allows for the request of individual frames directly from the device and low latency. REST manages data calls over HTTP through APIs that are used for communication with local or external servers to access available data. Another preferential element is the ability to request data traffic on demand from the management platform in order to reduce network traffic, decide the update frequency for each device, and ensure system stability.

During the infrastructure design, an effort was made to standardize the type of data protocols used as much as possible, preferring hardware from the same manufacturer (where possible) and equipped with open communication libraries already integrated into the main existing Digital Twin platforms. For example, Shelly peripherals are internally equipped with a web server to allow direct network access without the need for configuration of a dedicated MQTT Broker service; such peripherals can thus send traffic related to measured data through a standard connection. The ModBus protocol is used for the connection of Schneider power-meters, which devices are wired and physically integrated into the network via Ethernet cable. The weather station operates via REST APIs through the Davis service of the manufacturer. Finally, the cameras work through ONVIF. Each peripheral is assigned a static IP during registration to the network segment; these IPs are used to connect the specific peripheral with the management platform and ensure continuity in case of system restart.

The internal management of the Schneider power-meters, which uses an internal radio protocol for wireless communication between the amperometric clamps and the data reading device (power-meter), is an exception to the previously stated. The system is configurable via the proprietary ECOStruxure software. In the case of integrating additional devices that use MQTT and are not equipped with an internal web server, Mosquitto Broker, an open-source message broker that implements the said protocol and is widely used for managing lightweight messaging in various communication scenarios, will be employed.

2.6.5 Advanced system architecture

The construction of the system's architecture was founded on a comprehensive assessment of the capabilities and tailored solutions employed for the use case. This is visualized in a general schema in **Figure 35**.

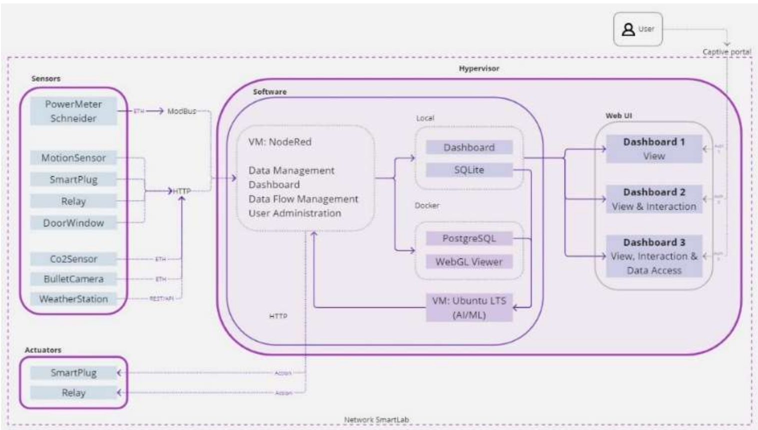


Figure 35. Key components and system architecture

Local System Configuration and Data Management:

A decision was made to deploy a localized system built upon a robust platform. This platform receives data inputs, processes it using a feed-forward mechanism, and subsequently disseminates this data. This processed data finds utility in three primary domains:

- Implementation
- Visualization
- Archival for predictive analytics

Network Configuration:

The peripheral devices are linked to an exclusive WLAN to ensure their autonomy from the primary laboratory network. The dedicated network utilizes EDGE technology for data reception and processing. This design employs data publication to segregate the peripherals from logical flows. This strategic insertion guarantees peripheral shielding, safeguarding against unauthorized or aberrant activities. Notably, the IoT devices in the SmartLAB connect to the Digital Twin platform through a specialized Wi-Fi network, with certain exceptions like Schneider Acti9 PowerTag Links and the Fronius inverter, which employ a wired connection.

At the heart of this architecture are the following components:

Edge System: interacts with the local sensor network, buffering data, and managing automated behaviors. It uses a Raspberry Pi 4 Model B with the Raspbian operating system. Various software components, both basic and application-specific, are deployed. Notably, Real T s.r.l.'s Elettra system manages sensor data.

Proxy: this intermediates the Edge and Server, ensuring uninterrupted data synchronization. For this, a standard PC with a Debian Linux distribution is employed, utilizing the Zabbix open-source software for data relay to the central system.

Server: handles computation and provides a GUI. A robust server, situated outside the SmartLAB, was chosen, running on the Debian Linux distribution. This server interfaces with the Zabbix Server, manages connections to specific IoT devices, and houses the machine learning system for energy optimization.

Energy Management System (EMS): in the context of EMS, IoT devices liaise with a local data concentrator (edge device), which further interfaces with a central data recording and processing unit. Due to the system's scale in this study, there's only one data concentrator and integration point with the sensor network, as depicted in **Figure 36**.

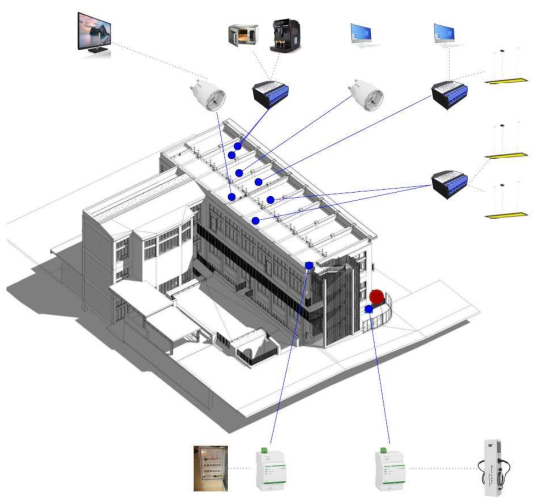


Figure 36. IoT devices and local data concentrator (edge device) for Energy Management System (EMS)

Centralized Processing and Analytics: The logical structure of the system is consolidated in Figure 29 and 30. The Central Control System on the server employs Artificial Intelligence to analyze collected data. Its objective is to craft algorithms that enhance energy efficiency without compromising comfort.

Infrastructure Design and Implementation:

According to the network configuration described above, the core elements of this architecture are three:

- *Edge:* it is responsible for interacting with the local sensor network and acting as a buffer for automatic data and behaviors;
- *Proxy:* it enables the interface between Edge and Server, ensuring connection and synchronization of data and commands (and recovery after any momentary lack of connectivity);
- *Server:* it is responsible for all computation and provides a graphical user interface (GUI).

Edge computer: a classic SoC (System on Chip) often used in research has been chosen, the Raspberry system, specifically Raspberry Pi 4 Model B. The operating system is the Raspbian (a version of the Debian Linux distribution adapted to Raspberry hardware).

A set of software, both basic and application-specific software, has been installed on this device. The basic software components are necessary for managing both wired and wireless communications with the sensor networks and the Proxy system. The communication chosen was via IP protocol and where possible via TCP/IP, as far as sensor networks are concerned, and certainly TCP/IP as far as communication with the Proxy is concerned. In this installation, being an experimental installation, a remote monitoring VPN (Virtual Private Network) was also installed: under normal use this is not expected to be a VPN installation in the Edge system.

In terms of application-specific software, the sensor management component of Real T s.r.l.'s *Elettra* system has been installed. This component has the responsibility of the retrieval of consumption information using several communication protocols used by hardware vendors (e.g., Shelly, Schneider, Fronius) and the production of standardised data in the form [timestamp, device_identifier, measurement].

The Elettra subsystem is responsible for dialogue with assorted sensors and therefore must implement the communication protocols typical of each vendor; it is therefore appropriate for it to be in close proximity to the sensors themselves to minimize transmission problems and to monitor sensor networks even in the absence of connectivity to the central system. For this reason, it has been installed on the Edge system.

Proxy: a standard Personal Computer with a Linux system Debian distribution has been used for this purpose.

The Proxy subsystem has the responsibility of normalizing the data and transmitting it, when connectivity is available, to the central system from which it can then receive commands for turn-on or turn-off devices.

A set of software has been installed on this device.

The communication software components use TCP/IP protocol through a dedicated VPN (Virtual Private Network) to guarantee the connection with the server located in a datacenter outside the SmartLAB facility.

As for the application-specific software, the Proxy component of the open-source software Zabbix has been installed: this software has the single responsibility of data transfer to the central system.

Appropriate scripts have been created to enable the communication between the Zabbix Proxy and the Elettra subsystem; specifically, they allow to query the Elettra system and to properly format the data for the Zabbix Server.

In the SmartLAB use case this solution might rightfully seem redundant; however, the architecture is designed to be scalable: the typical configuration should be one Edge system per apartment (or office) and one Proxy per building (about 20-30 apartments). In that case it is appropriate to have the Proxy responsible for proximity dialogue with the Edge systems and each Edge system responsible for dialogue with its own sensor network.

Server: a server with appropriate computational resources has been chosen in a datacentre outside the SmartLAB. The operating system Linux distribution Debian has been installed on it.

As in the proxy system a set of software, described hereafter, has been installed on the server.

The basic software components are in charge of communications with the proxy system and are mainly related to the VPN (Virtual Private Network) with the proxy located at CITERA and to the web and application server necessary for the user interface.

The Zabbix Server ensures communication with the edge device running the Zabbix Proxy.

The central IoT platform is responsible for connections to IoT devices that don't require management using Elettra, for storing data and sending commands for turn-on or turn-off devices.

The server also hosts the Machine Learning system that performs all calculations related to energy management optimization.

The logical architecture of the system is described in **Figure 37**, the software stack installed on the server is defined Central Control System.

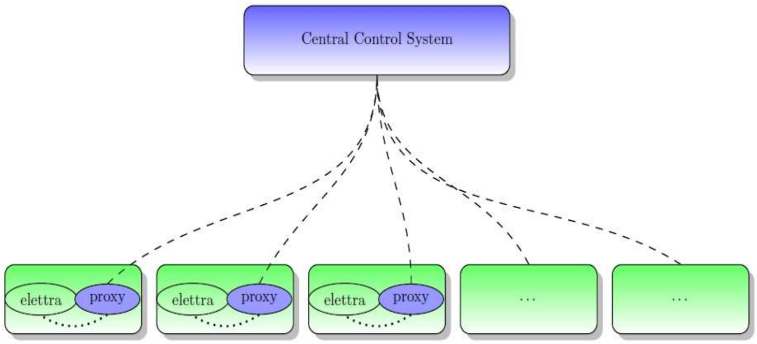


Figure 37. Logical Architecture of the System

The infrastructure was realized in two stages:

1. Infrastructure design and implementation
2. Data acquisition, model training, and validation

Considering energy data as time series recordings from sensors, each with its API, there's a need for a synchronous and unified dataset. Each workstation is equipped with sensors to monitor individual equipment, aiding in devising energy-efficient control solutions. The challenge was to centralize this data processing, necessitating a local system responsible for sensor and actuator interconnections over networks like the Internet. Two primary challenges were addressed: ensuring uniform data and having data processed at requisite locations. This led to the design of two subsystems: “elettra” and “proxy”, as shown in **Figure 38**.

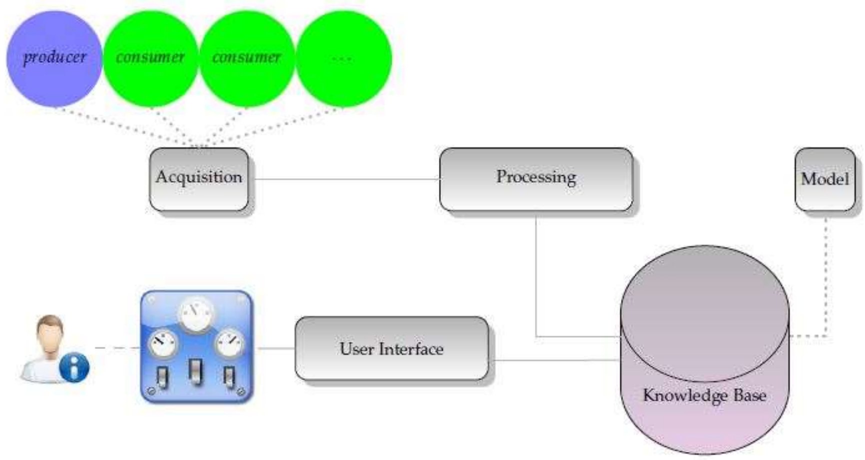


Figure 38. Design and implementation of the infrastructure

As energy data are simple time series of power consumption or production, in a given time lapse coming from real sensors, each one transmits data with its own Application Programming Interface (API); moreover, they are located close to energy loads, or near power sources.

This means that data are not all in the same place at the same time, which is a necessary condition to start the analysis which will led to the desired algorithms. The first nontrivial problem is therefore to plan and deploy a cost-effective IT infrastructure able to provide reliable data for the software to analyse.

Each single workplace has been implemented with sensors monitoring every single equipment, so that the energy consumption of every device can be considered to model the control solution for the overall energy requirement of each workplace.

All those metering sensor produce a huge amount of data requiring significative computational resources to obtain acceptable analysis performances, so the best solution for reducing installation expenses would be to control the system acquiring all the information in a data center or a service in a data center. This architecture leads to the necessity of setting a local system, each with responsibility for interconnecting sensors and actuators over a geographical network (e.g., the Internet), executing sort of local computation and buffering data in case of connection blackout, using the known “Ubiquitous and Pervasive computing” (Tomazzoli et al., 2023a) techniques to deal with the computational problems of centralized intelligence.

So, as mentioned, the first element in the infrastructure will be a sub-system, able to cope with several transmission protocols and several time frames, whose output will be the synchronized power consumption (or production) of the devices smart metered. This sub-system can accept instruction from the second element to switch on and off some of the controlled devices.

This element needs to dialogue with all sensor networks, so it has to be physically placed next to them, to minimize transmission problems and monitoring local environment even in absence of communication with the central control system. This kind of elements in the following will be called “elettra”.

The second architecture element will be another sub-system, composed of different “proxy”, and each proxy receive as input the outputs of the first sub-system. The proxies will deliver the data to the central unit, and receive back commands from the same device, taking care of bandwidth problems and unreliability of the network.

These proxies have to be physically close to the first sub-system while the central unit can be remote; the central control system, a centralized unit able to store and process data, operate building digital simulation models and delivers commands back to the proxies.

Following this logic infrastructure, a series of "cheap" small computer, or SoC (System on Chip) has to be equipped, containing both the "elettra" and "proxy" sub-systems; all those computers are connected to a high-performing server in a data center able to run the software of the central control system.

The operative concept of this infrastructure is exemplified in **Figure 39**, where are sketched only a few energy consumer devices as an example.

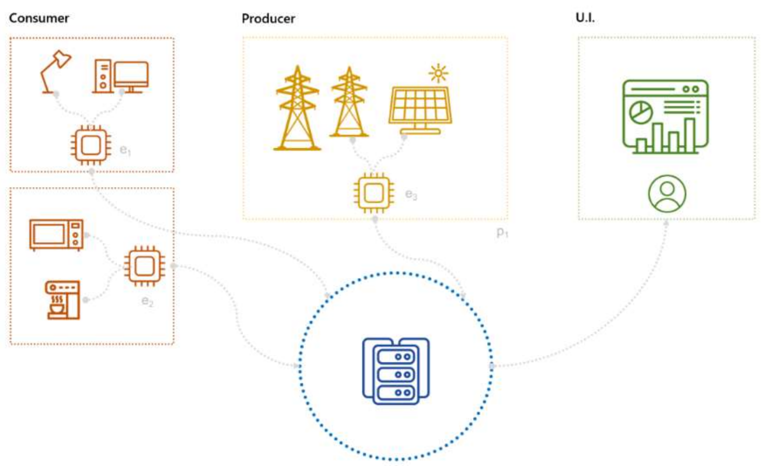


Figure 39. Operative concept

In **Figure 39** elements e1, e2, e3, e4, e5 are the cheap computing unequip containing the elettra sub-system and the proxy, while elements c1 to c10 are energy load examples, and production system is related to photovoltaic panels for electric power production.

2.6.6 Multi-service approach

Utilizing a composite methodology provides the advantage of exploring a diverse range of software solutions, each tailored with its unique attributes that address specific challenges and requirements. By analyzing software solutions through a comparative lens, one can discern the inherent strengths and limitations associated with each.

Embracing an open-source philosophy for the process not only underscores transparency and community-driven development but also distinguishes this approach markedly from proprietary commercial solutions.

Open-source platforms typically foster a more collaborative environment, encouraging shared knowledge, rapid iteration, and the ability to customize applications to suit unique needs.

In our quest to meet the outlined objectives and harness the capabilities desired, three distinct platforms were identified:

- Node-Red: an event-driven programming tool that facilitates visual coding for wiring together devices, APIs, and online services. Its flow-based development environment is specifically designed to assist with the integration of disparate systems.
- Godot: a multifaceted, open-source game engine that supports both 2D and 3D game creation. Beyond gaming, Godot's powerful scripting and visualization capabilities make it a viable tool for various interactive projects and simulations.
- KNIME: an open analytics platform that offers advanced data integration, transformation, and exploration capabilities. With its rich set of nodes for data processing and visualization, KNIME provides a flexible environment for data-driven tasks and analytics.

Although each platform excels in its domain, the true potential is realized when they operate in unison. By facilitating seamless interoperability and data exchange among these platforms, a cohesive workflow that maintains both the continuity and the integrity of the overarching process was ensured, as visualized in **Figure 40**'s comprehensive diagram. This synergy underscores the potency of an integrated multi-platform approach in addressing complex challenges as described below.

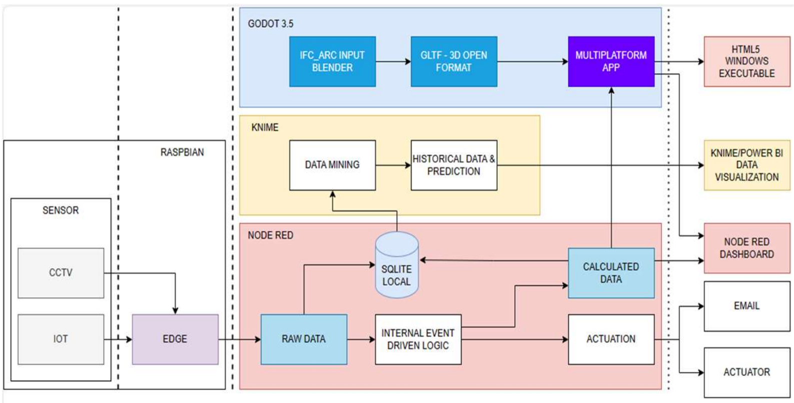


Figure 40. Interoperability and data exchange

Data Collection

- IoT: Internet of Things devices, which can range from smart appliances to industrial machines, collecting data from the environment.
- EDGE: the edge computing component where the data from sensors and IoT devices is processed. It likely preprocesses this data before it's sent to the main system for further processing.

Data Processing and Visualization

- Raspbian: This is a free operating system based on Debian, optimized for the Raspberry Pi hardware. It hosts the applications mentioned in the diagram.
- Node Red: A flow-based development tool for visual programming, mainly used for wiring together hardware devices, APIs, and online services.
- SQLite Local: A local lightweight database used to store raw data.
- Internal Event Driven Logic: Processes raw data based on specific events or conditions.
- KNIME: An open-source data analytics platform. It seems to be used here for:
- Data Mining: Extracting patterns from large data sets.
- Historical Data & Prediction: using past data to forecast future events or trends.
- KNIME/Power BI Data Visualization: Power BI is a business analytics tool by Microsoft. Together with KNIME, it's used for visual representation of the data.

Data Output and Integration

- Actuation: Represents actions or changes triggered based on the processed data.
- Node Red Dashboard: A dashboard that displays data and perhaps allows user interaction.
- Email: Notification or data sending mechanism.

Graphics and Application Development

- Godot 3.5: A game engine used for developing multiplatform games and applications.
- IFC, ARC Input Blender: Input formats for 3D modeling, likely used to create graphics or representations of the data.
- GLTF - 3D Open Format: A format for 3D scenes and models. It's a common format for three-dimensional graphics data.
- Multiplatform App: The main application developed using Godot, which can run on multiple platforms.
- HTML5 Windows Executable: A web-based application or frontend.

The architecture provides an end-to-end flow starting from raw data collection to processing, analysis, visualization, and finally to actuation. It incorporates both data science elements (like prediction and data mining) and application development components (like the multiplatform app).

2.6.6.1 Node-Red

In the expansive realm of applications tailored for managing data from peripheral devices, the Node-Red platform stands out and was consequently chosen to orchestrate a workflow adept at bi-directionally managing the IoT peripherals within the laboratory setting.

Node-Red is ingeniously built atop Node.js, an open-source, cross-platform JavaScript runtime environment. This foundational choice ensures scalability, flexibility, and a strong compatibility across diverse hardware and software ecosystems. One of the pivotal strengths of Node-Red is its emphasis on a No-code or Low-code paradigm, primarily executed via nodes. This design philosophy paves the way for users to intuitively harness the platform's capabilities, eliminating the often-daunting necessity to craft original code from scratch.

These nodes, which encapsulate specific functionalities, are predominantly authored in the ubiquitous JavaScript language. Given the vast and vibrant community behind Node-Red, there's a continual influx of novel nodes, each addressing unique challenges or streamlining existing processes. When these nodes are strategically sequenced, they culminate in automated workflows that can perpetually operate, ensuring uninterrupted data management and processing. This systematic approach has been instrumental in the laboratory, especially for the nuanced management of sensors and actuators, as depicted in **Figure 41**. The figure elucidates the strategic interplay of various nodes, each contributing to the seamless operation of the entire system.

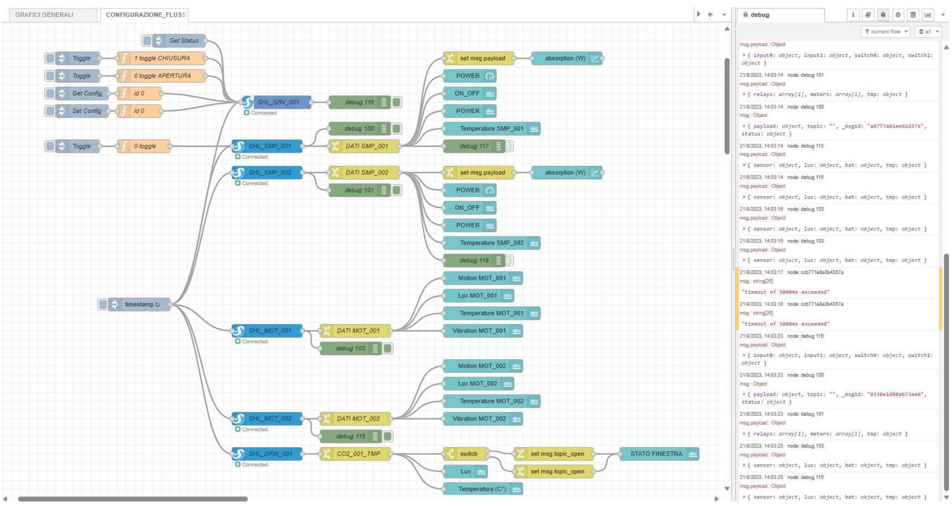


Figure 41. Sensors network implementation

Node-Red offers an array of functionalities that facilitate multifaceted interactions. Some notable features encompass:

- Data Retrieval: Capability to source data from external devices, ensuring a continuous flow of real-time information.
- Custom Processing: Leverage custom functions to process and transform incoming data to meet specific requirements.
- Active Output Production: Generate outputs that can dynamically modify the states of interlinked machinery and peripherals, ensuring real-time response and adaptability.
- Data Visualization: Feature-rich integrated dashboards that allow intuitive visualization and monitoring of the system's operations.
- Image Recognition with Computer Vision: Integration with robust frameworks like TensorFlow and implementation of models like YOLO facilitates precise image recognition tasks.

The hands-on experimentation conducted in the laboratory involved interfacing peripherals from diverse manufacturers such as SHELLY and KOLB BURMEISTER via the MQTT protocol. This flexible approach also postulated the potential for connecting peripherals that employ proprietary protocols or ones divergent from MQTT.

A noteworthy application revolves around environmental sensors that monitor CO_2 concentration. These sensors are adept not only at gauging the primary CO_2 levels but also provide supplementary data like ambient temperature, relative humidity, and illuminance (Lux) measurements.

In a practical scenario implemented within the laboratory, data from CO_2 sensors undergoes a comparative analysis with real-time occupancy data. If the ambient CO_2 concentration surpasses a predefined threshold (measured in parts per million or ppm), an automated process is triggered. This initiates the actuation of a servo motor to open a window, facilitating air circulation and renewal. The system is programmed to conclude this ventilation phase after a 15-minute duration. Post this, a magnetic sensor verifies the window's closure status, ensuring environmental safety and energy conservation.

Furthermore, the lab's setup incorporated Smart Plugs and relays that could be remotely actuated by the platform to perform specific tasks associated with them. Unique nodes, tailored for each peripheral, were designed to enable seamless bidirectional communication through the platform, adapting to the requisite protocols. This intricate yet efficient arrangement is visually represented in **Figure 42**, illustrating the synergy among various components.

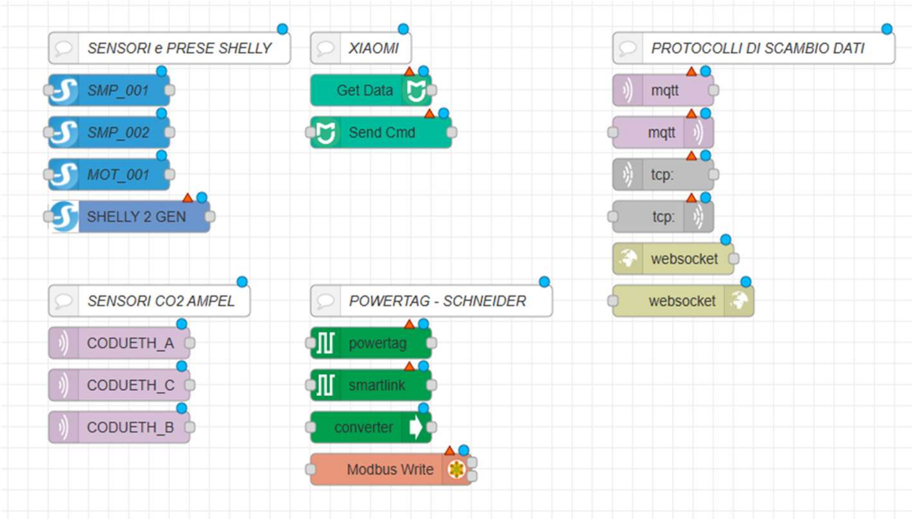


Figure 42. Sensor's network and communication protocols

Node-Red, with its versatile and modular architecture, establishes a systematic flow of data through nodes. Here's a more scientific elucidation of the platform's operation:

Initialization through the INJECT Node: The workflow commences via the INJECT node, which periodically triggers the process based on predefined intervals. The signal thus initiated is designated as "Payload", a packet of data that traverses through subsequent nodes in the workflow.

Payload Transformation: as the Payload progresses, its constitution can be altered based on the interactions and operations specified. **Figure 43** exemplifies a Payload associated with interconnected sensors. Possible transformations include data enrichment—where data is augmented, revised, or aggregated to align with specific requirements.

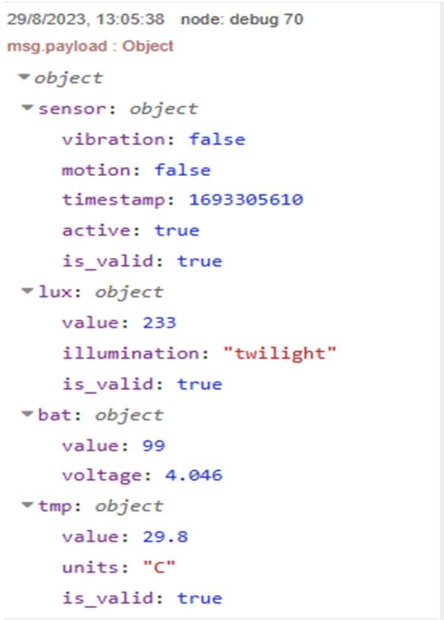


Figure 43. Payload associated with interconnected sensors

Visualization through Custom Dashboards: the processed data is funneled through nodes to intuitive dashboards. These dashboards, exemplified in **Figure 44**, are tailored within the platform using pre-compiled themes, providing stakeholders with a consolidated view of the system's status and insights.

Figure 44. Data dashboard

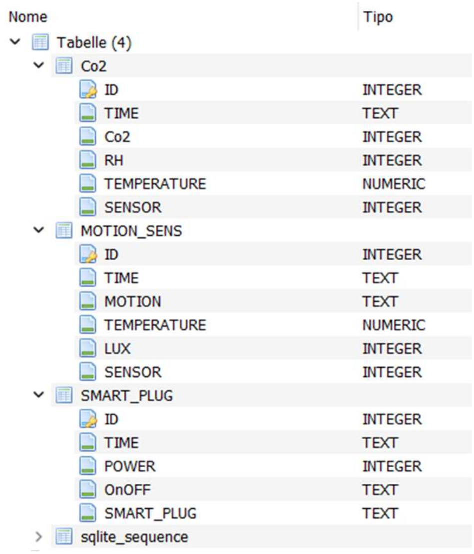


Figure 45. Export to RDBMS & SQLite Integration

Export to RDBMS & SQLite Integration: data is not merely visualized but is also persisted in a Relational Database Management System (RDBMS) configured as per the SQLite standard. SQLite, an embedded database library written in C, is renowned for its lightweight characteristics. It champions a serverless architecture, allowing the database to reside on the executing machine, facilitating in-place updates. **Figure 45** outlines the structured database with its triad of tables, each tailored to accommodate data from distinct sensor types.

Versatile Data Export Options: beyond relational databases, Node-Red offers flexibility in data export formats. Data can be exported in structured formats like JSON, CSV, XML, and YAML, or semi-structured formats like HTML, ensuring compatibility with diverse downstream applications and systems. **Figure 46** illustrate the gamut of data export modalities available on the platform.

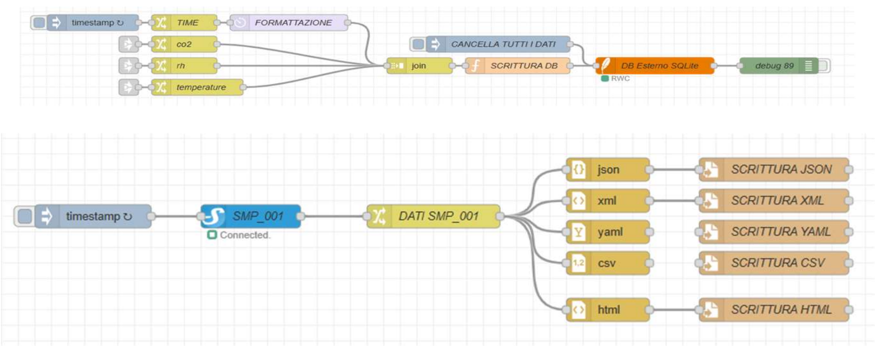


Figure 46. Data export modalities

Integration with Game Engine: to ensure a more immersive experience, real-time IoT data was rendered accessible in textual formats (.txt) for visualization within a three-dimensional model inside a Game Engine. However, this solution is slated for an upgrade. The future roadmap envisions leveraging dedicated WebSockets for broadcasting data, streamlining their integration and consumption in various applications.

2.6.6.2 Godot Game Engine

The intricate world of game development hinges on powerful game engines, and in this context, the chosen game engine emerges as a robust yet lightweight solution, offering a plethora of features. Here's an enriched elucidation of the described integration process, enhanced with scientific and technical detail.

The game engine under discussion is a potent, cross-platform engine designed for the creation of both 2D and 3D games. A hallmark of this engine is its impressively light footprint. Released under the MIT license, the engine is remarkable for its platform-agnostic export capabilities, with native support spanning platforms such as HTML5, Android, Windows, Mac, and more.

Transition from IFC to glTF: A critical aspect of the model integration was the conversion from the Industry Foundation Classes (IFC) format, an open data model in the architecture and construction sector, to a format amenable to the game engine. Out of the multiple formats like FBX, OBJ, and 3ds, the Graphics Library Transmission Format (glTF) was singled out. Developed by the Khronos Group, glTF stands out as an open standard, purpose-built for the seamless transmission of 3D assets, especially in web environments.

Blender & Blender BIM for Conversion: The conversion endeavor was facilitated by Blender, a renowned open-source 3D computer graphics software toolset. To specifically cater to IFC files, the Blender BIM plugin was employed, an open-source extension enhancing IFC capabilities within Blender. Post-conversion, the resultant file was funneled into the project's resource directory, triggering an automatic import into the game engine.

Model Visualization & GUI Development: The model's visualization was anchored on the laboratory floor, leaning towards a more abstract rendition of the integrated geometries. Post-import, essential project elements were discerned and classified into distinct categories, accessible and filterable via a Graphic User Interface (GUI). This delineation encompassed machinery, CO_2 metrics, and occupancy data (Refer **Figure 47** for a snapshot of the operational GUI).

Data Visualization via Particle Systems: To translate data into visual form within the model, a novel approach was employed. The various data outputs were tethered to particle systems, which excel at dynamically rendering geometric content based on data input. In this ecosystem, real-time data from the Node Red platform steers the behavior of these particle systems. Custom scripts, crafted in the GDScript language, process this data to manifest a dynamic visualization reflecting the laboratory's real-world parameters.

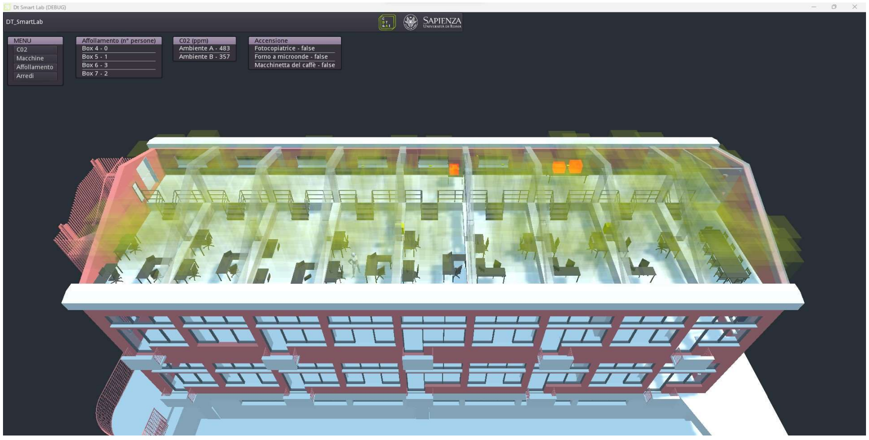


Figure 47. Operational GUI

The developed work was then exported in two different formats, the first in an executable format for a Windows environment as a stand-alone application and the second in HTML5 format for integration into web pages.

The export in HTML5 format was then inserted into the general Node Red dashboard in order to create a single webpage for the reception of data and their consultation.

This development process resulted in dual export formats, reflecting a broad vision to make the work both standalone and web interactive.

Exportation Flexibility: in today's dynamic software landscape, adaptability is crucial. Reflecting this, the project was tailored to be versatile in its exportation capabilities. The game engine's robustness played a pivotal role in facilitating this flexibility. The first export format is a standalone executable tailored for the Windows operating system. This format is beneficial for users who prefer a dedicated application, ensuring seamless performance without reliance on internet browsers or connectivity. It offers an encapsulated environment, streamlining the user experience by eliminating potential browser-based constraints or compatibility issues.

HTML5 Format for Web Integration: in stark contrast to the standalone format, the second export, in HTML5, emphasizes accessibility and universal reach. HTML5, a modern web standard, is recognized for its multimedia capabilities and cross-platform nature. It's especially suited for rendering 2D and 3D graphics within web browsers without necessitating additional plugins. The choice of HTML5 ensures that users, irrespective of their device or operating system, can access and interact with the model seamlessly via a web interface.

Unified Dashboard Integration: to elevate the user experience, the HTML5 export was seamlessly integrated into the overarching Node Red dashboard. This strategic integration meant that users could access a unified web portal, amalgamating real-time data visualization with the 3D model. This singular webpage serves as a nexus for data reception and consultation, offering users a holistic view of the data interplay within the 3D environment (**Figure 48**).

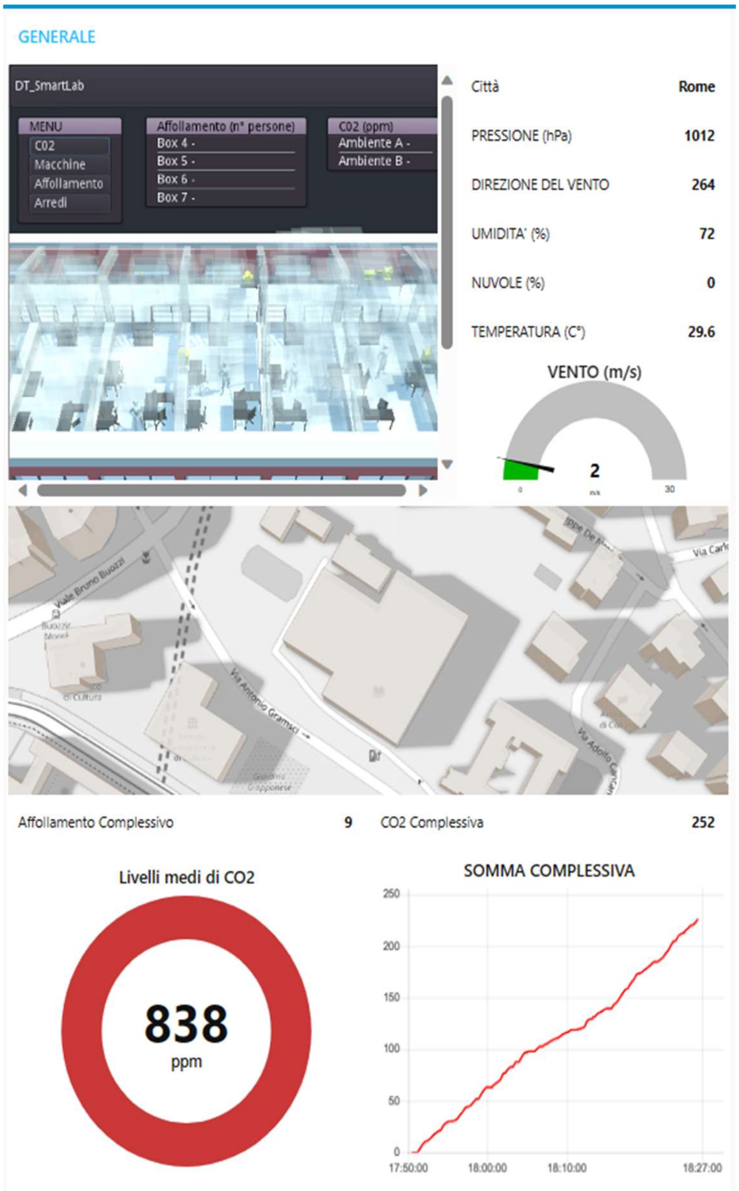


Figure 48. Web view of the dashboard integrated with the model updated in real-time

2.6.6.3 Knime

KNIME, standing for Konstanz Information Miner, is a hallmark in the realm of open-source data analytics, reporting, and integration platform. Emerging from the academic corridors of the University of Konstanz, KNIME offers an interactive platform that allows users to visually create data workflows (often referred to as "nodes"), allowing for an intuitive process of data analytics and modeling.

With its evolution into version 5, KNIME has continuously enhanced its capabilities, refined its tools, and expanded its node library. This evolution signifies its persistent commitment to address the escalating complexities of the data world.

At the heart of KNIME lies its intricate node management system. This offers the advantage of segmenting various data processes, be it extraction, transformation, or modeling. Moreover, its prowess to handle voluminous data sets paves the way for comprehensive data analytics, catering to diverse objectives.

The synergy between Node-Red and KNIME is evident when data exported from the former, specifically in the SQLite database format (.db), is ingested into KNIME seamlessly. The platform possesses dedicated nodes that facilitate the connection to specific format databases, as illustrated in **Figure 49**.

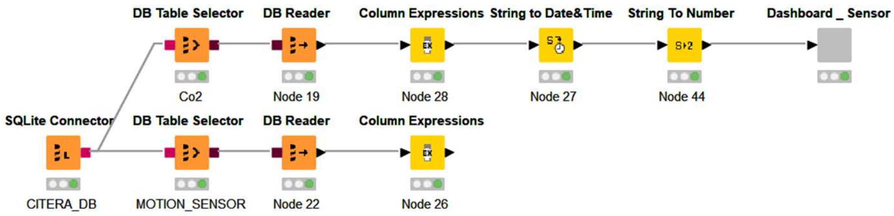


Figure 49. KNIME's nodes connection

Once the data is ingested, the platform transitions into its cleaning and preprocessing phase. Leveraging its diverse nodes, specific tables can be cherry-picked from the database, ensuring the data aligns with the requirements of the subsequent analytic processes.

Post-processing, visualization capabilities come into play. Data is orchestrated into dedicated dashboards, offering a visual narrative of the analytics. This not only enhances data interpretability but also aids in strategic decision-making.

The true potential of data lies not just in understanding the present, but in predicting the future. With this vision, once the data is archived annually, the platform plans to delve into advanced ML (Machine Learning) and deep learning algorithms.

The objectives achieved by the experimentation have led to the creation within the Node-Red platform of a series of logical flows capable of communicating bidirectionally with the peripherals; the data exchange thus defined is archived in a SQLite database for processing in Knime and display in the Game Engine.

The connection was achieved for the IoT components through the use of dedicated nodes, it was possible to control their operation and design an on and off scheme linked to logics internal and external to the system.

The real-time data is transmitted locally to the application developed by the game engine for visualization, so the data relating to CO_2, switching on/off, absorption and crowding of the environments can be viewed.

The data connected within the Game Engine in this phase only dealt with some portions of the available environments, the visualization led to the monitoring of an office area and a corridor area for presence checks, while the CO_2 was monitored extensively for both environments.

The platform is currently connected and ready for any experiments, such as data reading or an increase in physical peripherals.

The case study offers the possibility of further implementations for both the hardware and software sides.

Among the objectives to be achieved, we will try to connect the PowerTags through the ModBus protocol, for a correct reading of the laboratory's energy consumption, and to increase the number of peripherals and register them on an ad hoc network service, segregated from the external network, in in order to make monitoring more widespread with respect to the environment and current solutions.

On the software side, the missing nodes will be integrated to develop and consolidate the overall functionality, improve the connectivity of the model by replacing the current exchange based on textual input with an ad hoc websocket service and increase its functionality by allowing the visualization not only of real-time data but also of historical data directly in the model.

The historicization of at least one month of progressive operation is to be completed in order to begin the implementation of the ML algorithms for the development of forecasts relating to the use of spaces and energy within the environments.

Once suitable times are reached, they can then be used widely for possible monitoring and forecasting on an annual basis.

For comparison two similar works, developed by X. Zhang et al. (X. Zhang et al., 2023) and Muthiah-Nakarajan et al. (Muthiah-Nakarajan et al., 2021) discuss about the utilization of Machine Learning tools to reduce energy consumption in building also considering the presence of an Electric Veicle charging station. However, the two papers do not consider the integration of Machine Learning approach in a Digital Twin system. Moreover, the present study discusses about the integration of many physical data comings from different sensors (air quality, temperatures, weather, pressure, humidity, occupancy) and from the building itself (materials, HVAC, geometries, electric powered devices). Therefore, the presented approach can be considered innovative and aimed at developing a holistic model for building energy optimization purposes.

2.7 Results

An Energy Management System (EMS) is a suite of integrated tools and strategies designed to monitor, control, and optimize the performance of the generation and load profiles in real-time. Deployed initially in 2022, the EMS at the SmartLAB has been a pivotal instrument in understanding the energy consumption patterns and making improvements based on quantitative experimental results, which will be detailed in subsequent sections.

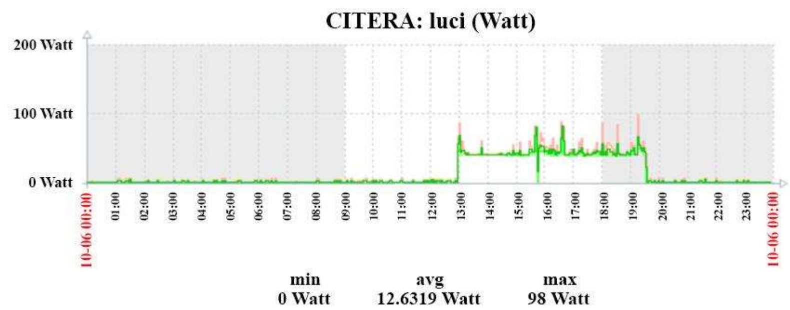
While the EMS focuses primarily on optimizing energy utilization, it's essential to underscore the holistic approach to sustainable operations. Parallelly, the SmartLAB has an operational Air Quality Management System (AQMS) aligned with pre-established management protocols. It monitors air contaminants and ensures optimal indoor air quality, crucial for both human health and efficient operation of laboratory equipment. The AQMS is in the process of collating historical data, which, upon reaching a significant volume, will serve as training datasets for a Machine Learning (ML) algorithm. This approach leverages ML's potential in predicting and improving indoor air quality based on patterns and trends.

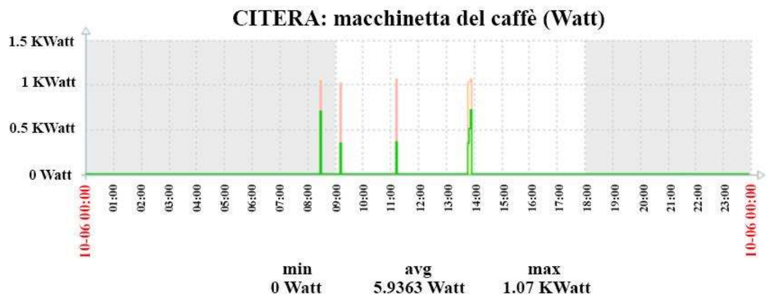
Over a span exceeding one-year, continuous monitoring has been conducted in the SmartLAB, capitalizing on sophisticated sensors. These sensors, integrated into a network, relay their readings to a Digital Twin platform. Upon comprehensive analysis of this data, the lab's energy consumption was delineated. A benchmark was set using a certified software which predicted an annual energy intake of 15.612 kWh.

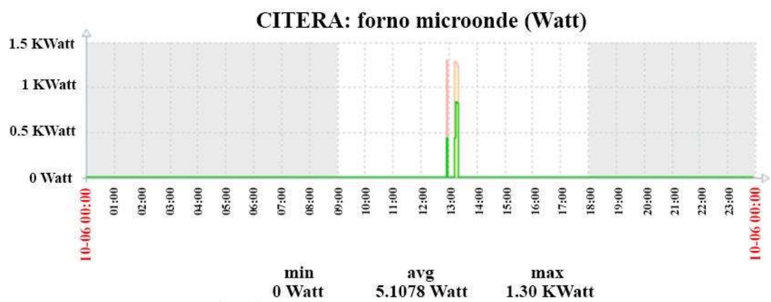
Diving deeper into the granularities of the data, we discerned the consumption metrics associated with individual socket points and the integrated lighting system. It's pivotal to note that each socket point was tagged to specific apparatuses, ranging from computational devices to routine appliances such as coffee machines.

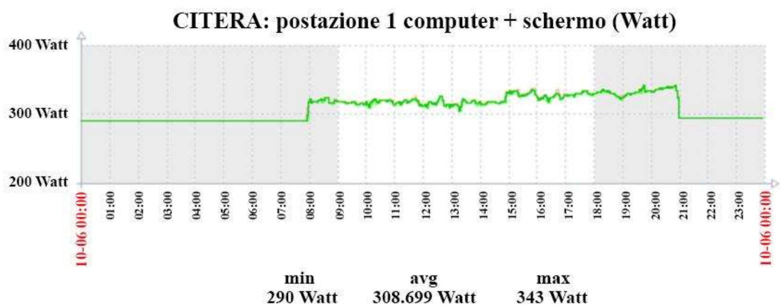
Given the architectural uniformity of the two workspaces examined, characterized by an identical configuration, equipment count, and operational hours, it provided a unique opportunity. This uniformity allowed for an extrapolation of the data, thereby estimating the energy consumption for the entire laboratory infrastructure.

Furthermore, a methodical observation of the lab's lighting system was executed, shedding light on its consumption metrics that were automatically detected through machine learning load recognition algorithms, as demonstrated in the following paragraphs. Subsequent graphical representations will illustrate device-wise consumption across a typical day (**Figure 50-53**) and week .









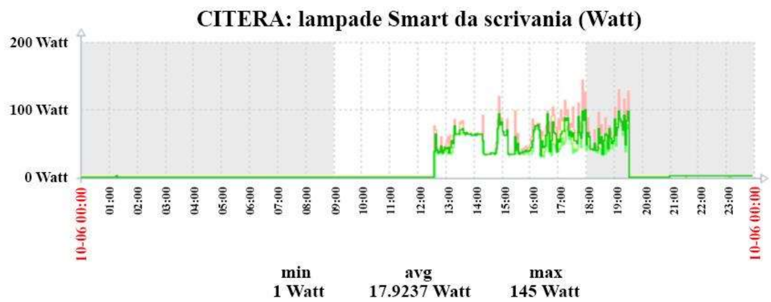
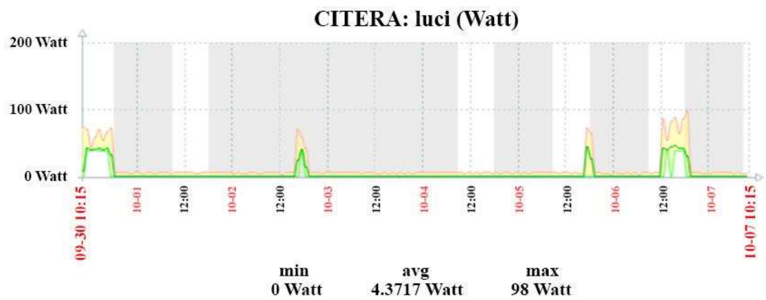
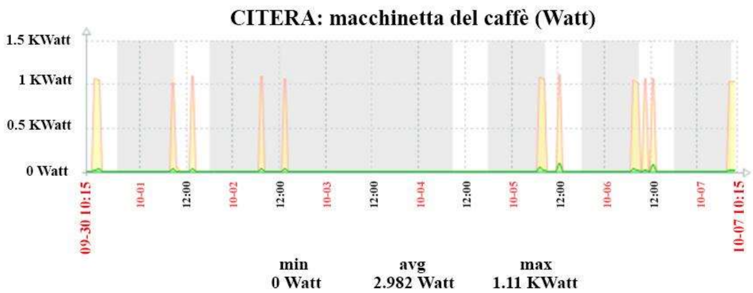
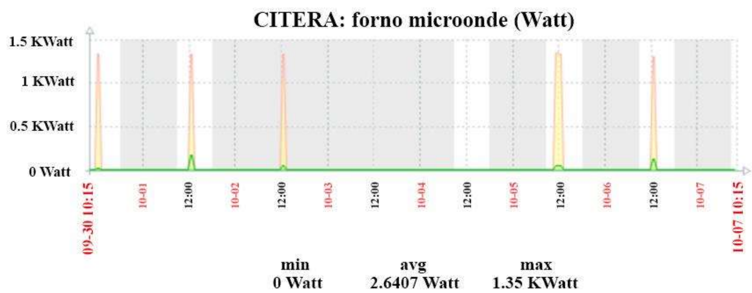
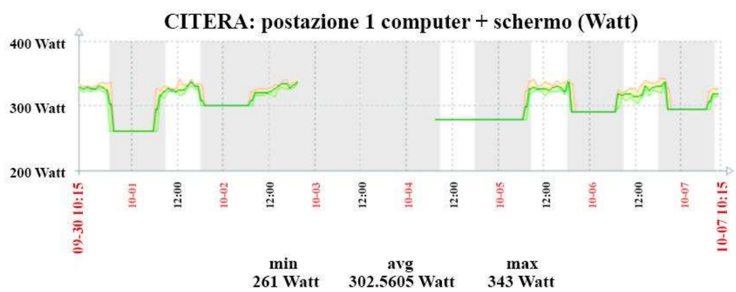


Figure 50. Energy consumption monitored over a typical day









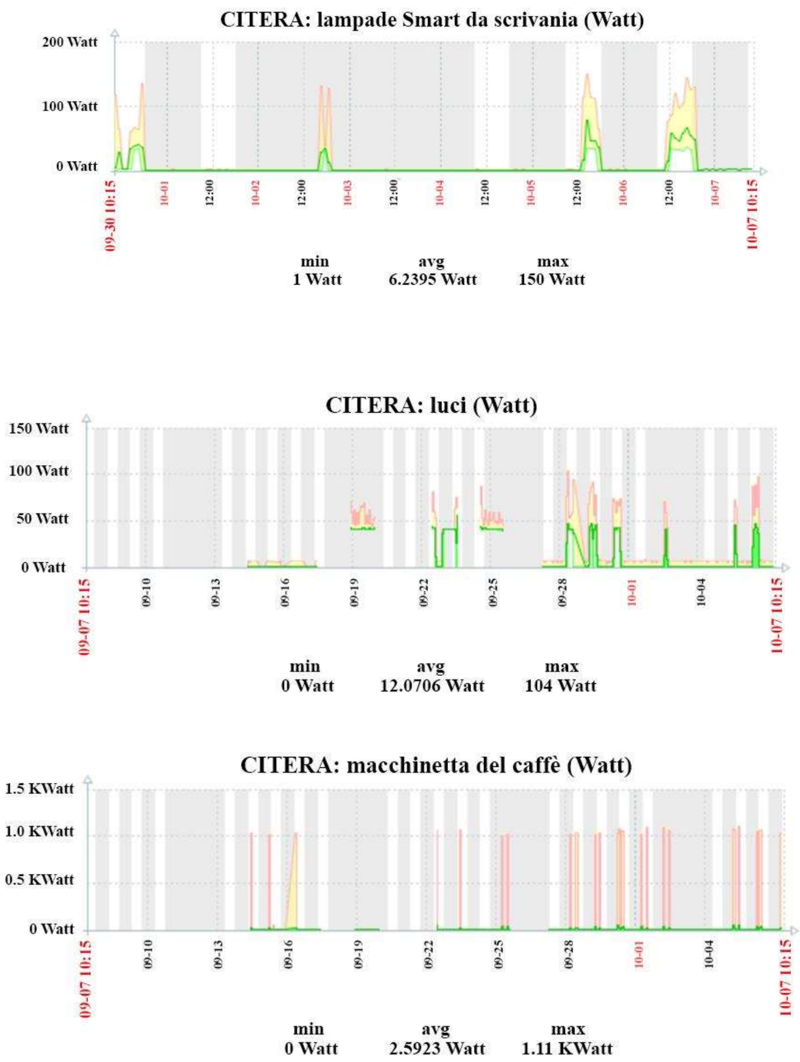
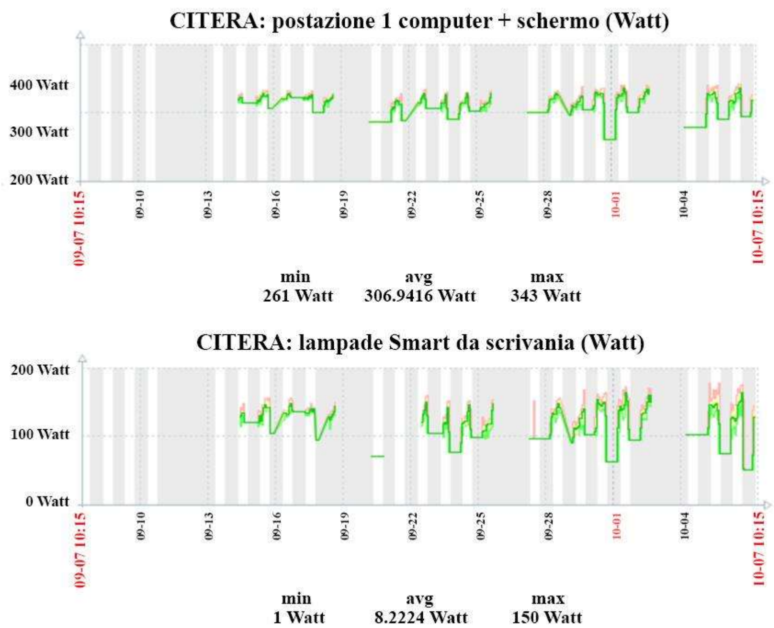
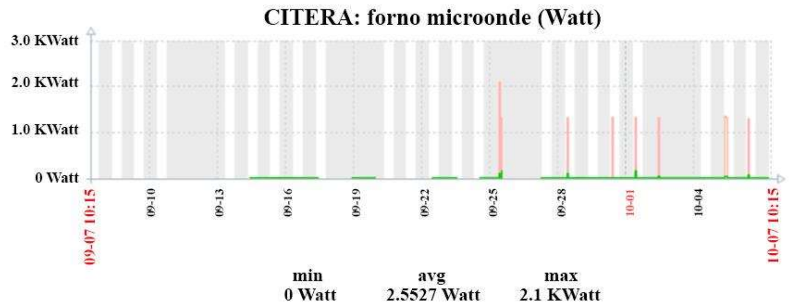


Figure 51. Energy consumption monitored over a typical week





From the development of the specific consumption obtained from the initial data monitored in the two workspaces, an average daily consumption of 35Wh was achieved during the operating hours of the Smart Lab.

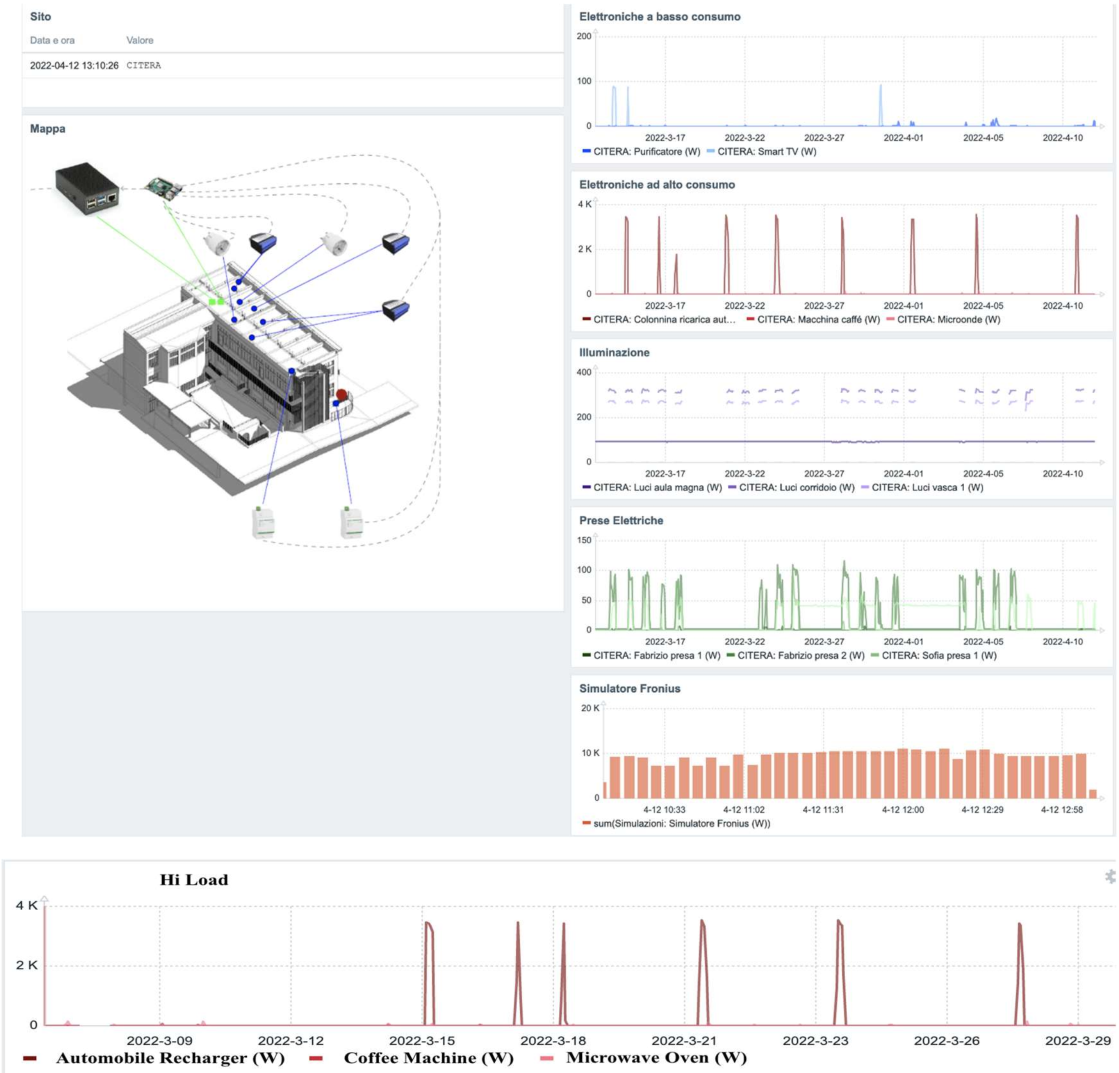


Figure 52. SmartLAB overall consumption on a month (April 2023)

2.7.1 Intelligent Energy Management

Machine learning (ML), a subfield of Artificial Intelligence (AI), is fundamentally rooted in the concept of enabling systems to acquire knowledge not from explicit code but from vast swathes of data. This paradigm shift is transformative: systems evolve based on patterns and regularities within datasets (Scannapieco & Tomazzoli, 2018).

Drawing parallels with human cognition, just as young children refine their understanding of the world through experiential learning, ML models "learn" and "refine" through training data. The underlying algorithms, often referred to as "machine-learning algorithms," iterate and improve their performance as they process more data (Bock, 2007).

The efficacy of an ML model is contingent upon both the quality and quantity of the training data. A richer dataset invariably lends itself to more robust model performance. Once the training phase culminates, these models can make predictions or classifications given new input data.

It's pivotal to note that the computational intensity during the training phase far eclipses that during the model's execution or inference stage. Training, especially with deep neural networks, demands resources that are often orders of magnitude greater, necessitating specialized hardware like Graphics Processing Units (GPUs) or Tensor Processing Units (TPUs) (Marques et al., 2019), (Marques & Pitarma, 2016).

In the context of SmartLAB, the proposal to deploy a ML architecture stems from the imperative to derive insights from the vast troves of data it generates. This data, when centralized, takes the form of structured tuples, typically represented as {location, date-time, object, value}. This structured approach aids in segregating data by individual locations or units.

With this granularity, it becomes feasible to discern the operational status of individual objects - be they energy consumers like appliances or producers like solar panels. Such temporal and spatial granularity allows stakeholders to gauge, for each monitored interval, which devices are in operation at specific locales.

A powerful method to extract such insights is through rule-based machine learning techniques, notably Association Rule Learning. Such methods unravel correlations, helping delineate operational patterns and, subsequently, optimize energy consumption and production dynamics.

2.7.2 Association Rule Learning

The pursuit of an optimal energy management strategy underscores the need for an intelligent system that harmonizes electrical loads with anticipated selfenergy-production. When bolstered with insights from Earth observation mechanisms, such as the Copernicus system, the potential of this synergy leaps forward. Preliminary estimations posit an efficiency augmentation of over 10%, translating to a stark reduction in the building complex's reliance on external distribution networks to a mere 15% of its overall energy consumption.

The intricacies of energy efficiency within building structures are manifold, largely due to the cumulative effect of myriad small loads. True energy conservation, without compromising on functionality, demands a sophisticated approach. This entails meticulous monitoring and adaptive control of energy loads, leveraging real-time data on their state, power level, and dynamic interactions within the consumption-production continuum.

Given the heterogeneity and sheer volume of devices in play, invoking principles from "Ubiquitous and Pervasive Computing" emerges as a cogent strategy to manage the computational demands of this multifaceted challenge.

While the energy dynamics of each building are uniquely tethered to its structural design and user preferences, one cannot overgeneralize energy-saving strategies. Yet, there exists a realm of "best practices" tailored for energetically congruent environments.

To actualize an autonomous energy management framework, a pivotal step involves the derivation of behavioral heuristics for diverse energy consumers. By closely observing and learning from varied energy behaviors, the system can dynamically evolve these "best practices", harnessing the prowess of machine learning techniques whose efficacy is well-established in academic literature.

Such an ambitious venture mandates addressing a trifold set of sub-challenges (Tomazzoli et al., 2023b).

1. Categorization Strategy: Formulating methodologies to cluster installations based on energy efficiency commonalities.
2. Rule Source Determination: Identifying the pivotal installation within each cluster that serves as the learning paradigm.
3. Rule Derivation: Engaging in profound analytics to extract actionable rules from the consumption patterns of the designated installation.

Complicating this mission is the inherent volatility of household energy loads, which flow with technological evolution and shifting human behaviors. Hence, a system's digital mapping must be perpetually rejuvenated, ensuring fidelity to the ground reality. This introduces two additional complexities (Tomazzoli et al., 2023b)

a. Dynamic Load Representation: Crafting an adaptive schema that lucidly captures the current energy loads within a system.

b. Device Recognition: Incorporating mechanisms that intuitively discern and catalog devices integral to an installation.

2.7.3 Clustering energy environments

The landscape of machine learning is dotted with a plethora of techniques aimed at deciphering patterns within data. One of the cornerstones of this analytical arsenal is the conceptualization of data points within a multidimensional vector space. This paradigm paints a holistic picture of the dataset, mapping each unique entity (in our case, an energy environment or plant) as a distinct point in this expansive space.

Imagine an energy environment as a complex amalgam of features: power usage, type of energy sources, consumption patterns, efficiency metrics, and more. Each of these features can be conceived as a dimension in our vector space. Consequently, an energy plant with a specific set of these features can be pinpointed as a unique vector in this space. The positioning of this vector is dictated by the values of its features, effectively transforming abstract characteristics into quantifiable coordinates.

The use of representing energy plants in this geometric format is the inherent potential to gauge similarity. Plants that share analogous energy features will inevitably reside closer within the vector space, whereas those with divergent characteristics will be spaced further apart. In mathematical parlance, this spatial closeness is often quantified using distance metrics, such as the Euclidean distance.

Clustering, in essence, is grouping entities based on their similarity. In our multidimensional vector space, a cluster is envisaged as a congregation of points (energy environments) that reside within close proximity to one another. The criterion for this proximity is often dictated by a predetermined threshold. If the distance between two points falls below this threshold, they are deemed similar and thus belong to the same cluster.

K-means and hierarchical clustering

Understanding the structure of data, especially in complex datasets like energy environments, is critical. Clustering algorithms facilitate this by grouping data points based on inherent similarities. Among the vast array of clustering techniques, K-means and hierarchical clustering are notably prominent. Here's a comprehensive exploration of both:

K-means clustering aims to partition a dataset into K distinct, non-overlapping subsets (or clusters). It does this by minimizing the variance within each cluster and maximizing the variance between clusters.

- Initialization: K cluster centroids are randomly chosen.
- Assignment: Each data point is assigned to the closest centroid, and it becomes a member of that cluster.
- Update: The centroid of each cluster is recalculated as the mean of all points in that cluster.
- Repeat: Steps 2 and 3 are iterated until the centroids no longer change significantly, indicating that the algorithm has converged.

Unlike K-means, which partitions the dataset outright, hierarchical clustering creates a tree of clusters. This tree, often visualized as a dendrogram, can be dissected at various levels to yield different clustering structures.

- Initialization: Each data point is treated as a single cluster, meaning there are N clusters at the start (where N is the number of data points).
- Agglomeration: In each of the subsequent stages, the two clusters that are closest to each other are merged into a single cluster.
- Completion: This merging process is repeated until there is only one single cluster containing all data points.

Both K-means and hierarchical clustering are heavily reliant on distance metrics to gauge the similarity between data points. Common metrics include:

- Euclidean Distance: Geometric distance in the multidimensional space.
- Manhattan Distance: Sum of absolute differences between coordinates.
- Cosine Similarity: Measures cosine of the angle between two vectors.

Without loss of generality, we can say that every energy load belongs to a specific type such as, for instance: light bulb, microwave induction plate, TV set, refrigerator. We define T = {t1 . . . tn} as the set of all possible types of devices belonging to any energy environment of a system S.

The feature representation of a plant pi (denoted by ~pi) is a n-dimensional vector whose j-th component ~pij , with $1 \leq j \leq n$, is the number of devices of type tj in pi so that it can be considered a point in a vector space, as represented in **Figure 53**.

Consider, for instance, a system S with three consumers p1, p2, and p3 where:

- plant p1 consists of two computers, ten light bulbs and one microwave induction plate;
- plant p2 consists of one refrigerator, two TV sets, one washing machine, and one induction plate;
- plant p3 consists of five computers, one microwave induction plate, one refrigerator and one light bulb.

The set of all device types in S is thus T = {light bulb, Computer, Refrigerator, Microwave induction plate, TV set, Washing machine, Induction plate} and the feature representation of the aforementioned plants is

$$\rightarrow p1 = [10\ 2\ 0\ 1\ 0\ 0\ 0]$$

$$\rightarrow p2 = [0\ 0\ 1\ 0\ 2\ 1\ 1]$$

$$\rightarrow p3 = [1\ 5\ 1\ 1\ 0\ 0\ 0]$$

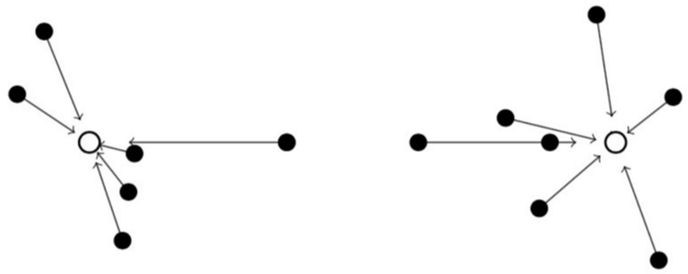


Figure 53. Vector plant representation and clustering

As a straightforward consequence, plants may be univocally represented in the n-dimensional vector space induced by |T| and a clustering algorithm such as k-means can be applied to isolate groups of similar plants.

2.7.4 Selecting exemplary performers in Energy Clusters

The process of pinpointing high-performing energy plants within their respective clusters is foundational to the establishment of benchmarks and best practices. Let's delve into the rationale and methods behind this approach:

- **Benchmarking:** Identifying the top-performing plant within a cluster provides a standard against which others in the group can be compared. This fosters a competitive environment and encourages plants to strive for efficiency.
- **Learning from the Best**: By scrutinizing the operational strategies and technologies of the top performers, valuable insights can be gleaned that can be disseminated to other members of the cluster, promoting energy-saving practices.
- **Motivation for Continuous Improvement**: Recognizing and celebrating top-performers can incentivize other plants to constantly seek improvement in their operations.

The methodology is described below.

1. **Gather Data:** Obtain energy consumption data for all plants in a cluster for a specific duration. This data can be sourced from utility bills, meter readings, or energy management systems.
2. **Standardize the Time Frame:** Ensure that the comparison is made over a consistent period, such as monthly, quarterly, or annually, to guarantee accuracy.
3. **Compare and Rank:** Rank plants based on their energy consumption. The plant with the lowest consumption, which implies the highest efficiency, emerges as the leader.
4. **Adjust for External Factors:** Ensure that the comparison is fair by considering external factors that might affect energy usage. These can include variations in production volume, weather patterns affecting heating or cooling needs, or periods of maintenance downtime.
5. **Validation:** To further validate the results, one could also examine the energy cost. Since energy prices can fluctuate based on demand, time-of-use, and other market dynamics, the energy bill provides a tangible metric that reflects both consumption and cost-effectiveness.
6. **Feedback Loop:** Communicate the results to all plants within the cluster. Encourage knowledge sharing and collaboration to lift the performance of all members.

While the "best" plants serve as role models, it's essential to remember that every plant operates within its unique constraints and opportunities. The goal isn't necessarily for every plant to emulate the leader but to understand the efficient practices and apply them in ways that make sense for their specific context.

2.7.5 Extraction of Behavioral Rules from Energy consumption data

Understanding the energy behavior of the best-performing environments is fundamental to extracting best practices and formulating strategies for efficient energy consumption. When dealing with extensive datasets that capture configurations and time series of energy usage data, rule-based methods like association rule learning are beneficial. However, a detailed methodology is required to extract meaningful insights:

Data Representation:

1. Energy Consumption Dataset (D)

Represents the energy consumption of a system. Each data point captures:

- Plant: the specific energy environment or setup.
- Device: the specific gear or equipment consuming energy.
- Power Value: energy level or power consumption of the device.

- Timestamp: the time when the measurement was taken.

2. Identification of relevant moments

In any energy environment, not all moments are equally significant. Some moments reflect a change in configuration – a device being turned on/off, or a significant surge or drop in energy usage. These moments, called "relevant moments," are pivotal for understanding energy behavior.

Method**:**

- From the dataset **D**, isolate records for each installation in the form: {device, power value, timestamp}.
- Identify changes in power values. The corresponding timestamps become the candidate relevant moments.
- Mark devices as active (power value > threshold) or inactive (power value <= threshold) based on the recorded power value.

3. Forming Configurations at relevant moments

From the isolated records, derive a set of binary configurations corresponding to each relevant moment.

Dataset D_0: an itemset of configurations at relevant moments.

- Each element d_0_i is labeled with a relevant moment.
- Contains binary variables representing the activity status of each device at that moment.

Variable Set V_0: Contains binary variables for each device in system **S**, indicating active (1) or inactive (0) status.

4. Association Rule Mining

With the data in the form of D_0, association rule mining, such as the Apriori algorithm, can be applied to identify patterns and relationships between different devices' operational statuses.

Example of a rule:

- $computer \rightarrow \sim TV\ set$
  - Meaning: If the computer is on, the TV set is likely off.

5. Rule Refinement and Learning

Extracted rules can be further refined by:

- Combining rules that overlap or are closely related.
- Identifying exceptions where general rules may not apply.

- Assigning priorities to different rules based on frequency, importance, or other metrics.

2.7.6 Automatic load recognition using Text Mining and Machine Learning

Ensuring the real-time accuracy of the composition of an installation is pivotal for the effective implementation of the method. Maintaining an updated database of connected devices is not only labor-intensive but is also susceptible to accumulating errors over time. Such inconsistencies pose a significant risk to the validity of the entire best-practices-based method. Thus, there's an imperative need to devise a solution that accurately mirrors the actual load in an energy system.

In a theoretical framework, the power consumption of a device can be delineated as a continuous power-over-time function. In practical scenarios, however, measurements are typically acquired at discrete intervals. Thus, a pragmatic approximation of the power consumption curve can be obtained using linear interpolation between these discrete data points.

When translating this sequence of energy measurements into a symbolic representation (as depicted in **Figure 54**), patterns such as "OOOOVOOOOOOOPOOVOOOOOPO" can be observed. To further refine this representation, consecutive similar symbols can be condensed, resulting in sequences like "O VO O PO VO O PO", thereby transmuting the power curve of a device into a textual footprint, termed "energy words".

Drawing parallels from the established domain of text mining provides valuable insights. In the realm of text analytics, the bag-of-words model (Harris, 1954) suggests that textual entities sharing a greater number of common "words" or tokens are intrinsically more similar than those with fewer overlaps. Extending this analogy to our context, the energy word sequence of a device can be conceptualized as its unique "bag of energy words."

Consequently, the process of automated energy load classification can be succinctly encapsulated as: given a repository of predefined (labeled) devices and a newly introduced, unclassified device, the classification of this new device is ascertained by juxtaposing its energy word sequence with the energy word bags of the known devices. Through this comparative analysis, the unknown device is assigned to the most congruent category or class.

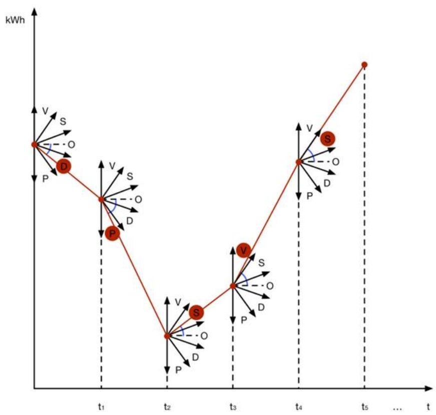


Figure 54. Alphabetic mapping applied to power measure

In the realm of energy consumption optimization, the employment of best practices becomes paramount. However, defining a universally applicable set of best practices is challenging, given the heterogeneity of installations which can vary based on their intended use, be it residential (home), commercial (office), or hybrid (mixed-use), and the unique layouts they encompass.

To address this challenge, leveraging machine learning, particularly clustering algorithms, can be instrumental. Clustering allows us to automatically categorize installations based on predefined similarity metrics, facilitating the management of large-scale energy systems. In creating a feature vector for each unit, every energy consumer and producer is meticulously cataloged and classified by type, ensuring a comprehensive representation of all energy elements within the structure (Tomazzoli et al., 2023b).

For this undertaking, the K-means algorithm, an unsupervised machine learning technique, was employed. The algorithm's efficacy lies in its ability to discern and cluster units that exhibit analogous energy consumption patterns.

Subsequent to the clustering phase, an observational period is designated. Post this phase, for each identified cluster, a representative location is chosen based on its energy performance metrics. This representative model, or a centroid, is then utilized to derive behavioral energy consumption rules, which can then be applied across all units within that specific cluster.

Given real-time data acquisition capabilities, at any given instant, the configuration of an arbitrary apartment, say "Ai", can be juxtaposed against the centroid model "As". A potential application of this could be in the formation of actionable rules. For instance, consider the rule: "At time instance 'tk', compare the state of device type 'dj' of apartment 'Ai' (dAij) with its counterpart in the centroid model 'As' (dAsj). If states are congruent, no action is necessitated; otherwise, toggle the state of 'dAij' to mirror that of 'dAsj'."

Such rules can be meticulously formulated into machine-executable formats, termed Association Rules. A quintessential association rule in an energy grid context might appear as: TheSolarPanel IsOn $\rightarrow$ TheWashingMachineIsOn. To derive such rules, the Apriori Algorithm, renowned for mining frequent item sets for boolean association rules, was employed. This algorithm operates by recognizing recurrent individual items within the dataset and successively aggregates them until a predefined frequency threshold is met.

Once formulated, these models can be integrated into the control modules of each unit, providing either advisory recommendations or mandating specific energy consumption patterns.

However, an intrinsic limitation of such automated systems is their susceptibility to inconsistencies between the digital representation and the actual built environment. For instance, during the lifecycle of an establishment, the relocation or replacement of smart plugs can inadvertently introduce discrepancies in the energy model.

To counteract such discrepancies, AI-driven techniques have been devised. One such technique involves converting the power consumption curve of a device into a series of symbolic representations, termed "energy words". Thereafter, the Naïve Bayes classifier, a probabilistic supervised learning method, discerns the nature of each energy load. This assists the system in identifying disparities between the digital representation and the actual connected load.

Notably, the dictionary cataloging these "energy words" has expanded to over 60,000 entries. To mitigate potential dimensionality issues, any word appearing fewer than three times within the energy footprint was excluded.

The predictive accuracy of the model, constructed using the Naïve Bayes classifier, underwent rigorous validation. It was subjected to both a 66% training and 33% testing data split, and a ten-fold cross-validation method. For this purpose, the open-source machine learning software "Weka" (specifically the class weka.classifier.bayes.NaiveBayes) was employed.

Chapter 4

Discussions

Although energy consumption in industry has been studied in deep, when dealing with residential compounds or SOHO (Small Office Home Office) buildings we cannot directly borrow solutions from research experiences. In fact, the overall consumption is the sum of small contributions by a considerable amount and variety of devices, while in industrial environments there are generally few big powers draining that can be controlled one by one.

Therefore, the problem of energy savings in buildings is strictly connected to the need of measuring and controlling energy loads in an efficient way, which can evolve complex scenarios. Several sensors and actuators must be involved, and their data shall be interconnected so that an ad-hoc algorithm derives the correct energy saving policy (e.g., a motion sensor shares data with electrical relays able to switch on/off the correct devices).

A possible general solution is the adoption of best practices, which are hard to define due to the nature of all installations being quite different according to final uses (home, office, mixed use) and layouts; if grouped by location and similarities parameters (Westermann et al., 2020) Artificial Intelligence becomes able to automatize processes attributing each location to the most appropriate group, or cluster.

This aim can be achieved using machine techniques known as "Unsupervised learning": these techniques refer to machine learning algorithms used to draw inferences from datasets consisting of input data without labelled responses

(Uhlemann et al., 2017). Unsupervised learning conducts an iterative process, analysing data without human intervention.

The most common unsupervised learning method are cluster analysis and neural networks (recently led to deep learning). K-means is one of the simplest unsupervised learning algorithms that solve the well-known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed in a "a priori" logic. The main idea is to define k centroids, one for each cluster. These centroids should be placed specifically and in a cunning way since the result is strongly influenced by their position.

The first step is to define the most efficient and performing context for each group or cluster by considering the energy bill over a few months confirmed by the energy data collected over a given interval of time.

The desired "local sample" is the one with higher energy performances regarding to all other component of its cluster.

Moreover, every automated system can easily fail if the digital representation of built environment doesn't match reality. Assuming that, inevitably during the lifetime of a building, some smart plug will be connected to different devices, this load variation can affect the reliability and accuracy of the digital model (Saini et al., 2020b).

In order to keep the digital model continuously up-to-date, Artificial Intelligence techniques, using analytical processes similar to those of text analyses, transform a power absorption curve of a single device in a sequence of characters; than a supervised learning method named "Naïve Bayes classifier" automatically identify the type of each energy load, so that the system can detect a mismatch between the digital representation and what is actually connected to the network.

2.8 AI for Energy Management

Basing on the aforementioned studies, a Machine Learning model architecture has been proposed in order to analyse the data coming from the Interdepartmental Research Centre for Territory, Construction, Restoration and Environment (CITERA) SmartLAB located in Rome at the faculty of Architecture "Valle Giulia". The full architecture of models proposed is reported in the schematic of **Figure 55**.

As can be seen the data architecture consists of a physical layer that contains the main physical data coming from sensors and the building, a Machine Learning layer that contains the data to be extracted from the physical layer, and a physical action layer that is user oriented and is aimed at defining the main output for an Intelligent Digital Twin (IDT).

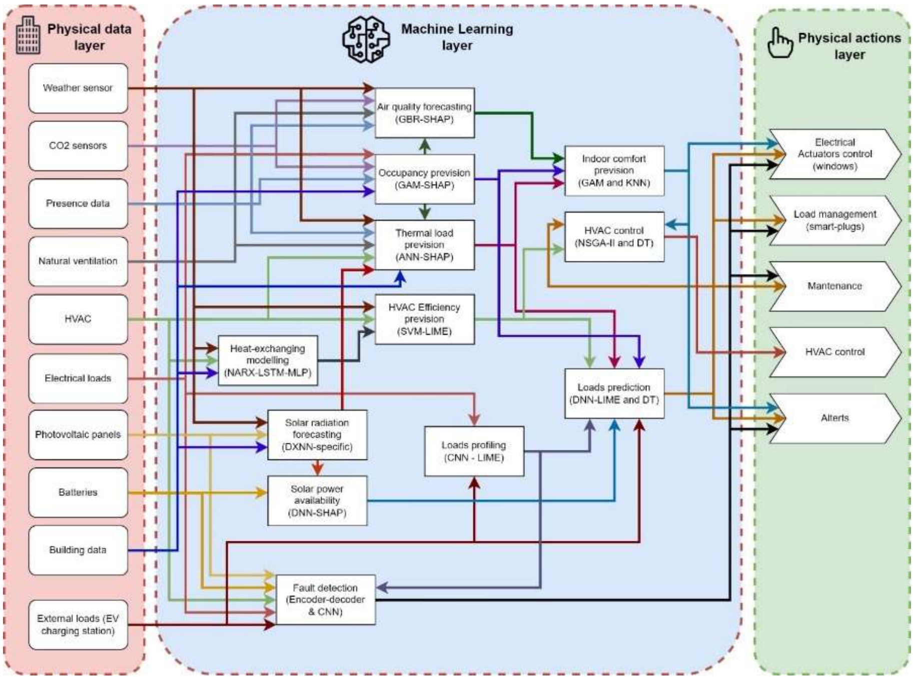


Figure 55. IDT's layers architecture

In the chart all data coming as input/output from the previous node/model shall be used as input in the next note. Moreover, Machine Learning layer shows also the most suitable and promising models to apply for the different tasks, the choices are based on the literature studies and are listed below:

Air quality forecasting: Air quality can be evaluated by Gradient Boosting Regression (GBR) with SHapley Additive exPlanations (SHAP) (Ugwuanyi, 2021). The model can be also used as classifier (for bad, average, and good air quality forecasting) and as regressor to give an estimation of the previsioned value. The model can be used both for indoor and outdoor air quality and it can be used as an input for comfort estimation and loads prediction (combined with the power needed by the ventilation system). For natural ventilation Deep Q-Network (DQN) model can be used to compare the contribution of opening windows on air quality in case of expected high outdoor values of PM (An et al., 2021). End users can also directly visualize data to understand the air quality expected in the next hours. Considering its importance, a good interpretability is fundamental to guide the choices of the decision-maker upon any eventual intervention on the building as for installing new air purifiers. Input data can be also integrated with and from the local weather services.

Heat-Exchanging modelling: The purpose of modelling heat-exchangers should be underestimated by the end user; however, this data is fundamental to determining the efficiency of the HVAC system and to monitor the status of the exchangers. If the efficiency lowers too much there could be some issues related to fouling or leakages, therefore a maintenance intervention shall be scheduled. Knowing efficiencies is also useful for future intervention where the destination of spaces changes (as for the number of people inside the environment) or in case of renovations that can alter the geometry of the building. This model can be neglected if no one of this element is deemed important by the decision maker. The best model to calculate Heat-Exchanging efficiencies is Nonlinear AutoRegressive Exogenous (NARX) Long Short-Term Memory (LSTM) MultiLayer Perpetron (MLP) as showed in the work of Z. Chen, Xiao, et al. (Z. Chen, Xiao, et al., 2022) and C. Yu et al. (C. Yu et al., 2020).

HVAC efficiency estimation: Support Vector Machine (SVD) with Local Interpretable Model-Agnostic Explanations (LIME) can be used to estimate HVAC efficiency basing on physical data as demonstrated in the work of C. Fan et al. (C. Fan et al., 2019). Using the data coming from the model it is possible to control the air conditioning using on a best efficiency approach and thus reducing the total energy usage of the device. Moreover, the output can be also used as input for the load prediction model to calculate the expected electrical consumption of the building.

Solar radiation forecasting: The solar radiation can be obtained from local weather service data but also calculated in the specific position (that is subject to shading or partial coverage due to the placement of the photovoltaic panels). Deep Neural Networks with Local Interpretable Model-Agnostic Explanations (LIME) can be used for the task (H. Wang et al., 2020). The data interpretability is useful to understand if there are shading elements that interacts with the surfaces and to program intervention, as example to increase the energy production obtained from the panels or to reduce or increase the solar heat contribution for the building. The tool can be also used to model the effect of increasing the panel surface on the building by increasing the number of panels. Moreover, solar radiation in an input for the photovoltaic energy production model and for the thermal load model.

Thermal load prevision: Calculating thermal loads contributes to the determination of the building power consumption. Moreover, using Machine Learning tools is theoretically possible to assess the effect of outdoor condition to the total heat loss of the building Also any eventual intervention aimed at increasing the building insulation could be evaluated. The best previsioning Machine Learning model for thermal loads prevision is probably Artificial Neural Network (ANN) improved with a Local Interpretable Model-Agnostic Explanation (LIME) for the interpretation of results (Di Natale et al., 2022). In case of high natural ventilation, as for Near Zero Energy Buildings (NZEB), it is possible to divide the contribution of natural ventilation on the thermal load using an ANN model with SHapley Additive exPlanations (SHAP) to increase model interpretability (H. Park & Park, 2021). Knowing the most influential factor of thermal loads permits to evidence the criticalities that needs to be addressed to increase the global energy performance of the building. Thermal load prevision uses as input the data coming from weather, occupancy, natural ventilation, HVAC, and the building geometries and materials. Instead, the model output is used as input for the electrical power provisioning model and for the air quality model.

Occupancy prevision: Predict the number of people present in an environment in a specific time is a challenge, especially if there are no information about the number of people obtained by presence sensors. However, despite the difficulties, monitoring the number of people is a fundamental data that influences many parameters such as the air quality, thermal comfort, and power consumption. Therefore, Machine Learning could be very useful in addressing such issue as reported in the review of Dai et al. (Dai et al., 2020). A suitable Machine Learning tool for the task is Generalized Additive Models (GAMs) enhanced by a SHapley Additive exPlanations (SHAP) model for interpretability. As input the model it is possible to use building data, CO_2 concentration (or direct air quality measurements), power loads and presence sensors data. Then, occupancy data can be used as input for air quality forecasting, thermal load prevision, and indoor comport calculation. It must consider that, occupancy data is also a sensible information for security and safety, its knowledge must be controlled and not open to the public. Moreover, decision makers could use the previsions coming from the occupancy model also to dimension building services (as for calculating the number of people that need to eat at lunchroom, determining the number of parking needed, etc.) and to globally increase the building ergonomic.

HVAC Control: Controlling HVAC permits to reduce and adapt the energy consumption of the building by changing the power used by the air conditioning system. An effective predictive control model can be developed using by using a Non-Dominated Sorting Genetic Algorithm II as suggested by M. G. Yu & Pavlak (M. G. Yu & Pavlak, 2022), moreover, if an alert system is implemented it is possible to use a Decision Tree algorithm to deliver textual useful information to users and stakeholder (Mollo Neto et al., 2020; Tai et al., 2020) such as if the HVAC system delivers enough cooling, if it is recommended to move from the building due to the excessive load and other such information. The Control model takes as input the HVAC efficiency model, the indoor comfort prevision model, and the load prediction model to calculate the optimal response for controlling the HVAC. As output the model can relate to an alert and maintenance system, with the load management (deactivating unused electrical equipment by smart-plugs controls in case of high loads), with the electrical actuator control (opening windows when the HVAC is off and the air quality and thermal comfort is good), and with the HVAC system itself.

Indoor comfort prevision: This model is aimed at evaluating the thermal comfort of the indoor area in the building. Generalized Additive Models (GAMs) are very suitable for comfort prevision tasks (Charalampopoulos, 2019). To develop alerts and recommendations to users and stakeholders both K-Nearest Neighbors and Decision Tree models can be used (Mollo Neto et al., 2020; Tai et al., 2020; Xie et al., 2022; Yusuf Akbar et al., 2022). It must be considered that thermal comfort is a matter of legislation as expressed in EN ISO 7730 (Cheng et al., 2012; Fanger, 1986; ISO, 2005). Therefore, some empirical rules are recommended to aid the comfort model to maintain comfort values inside an acceptable range. As model input shall be used the data coming from the air quality forecasting model and the occupancy model and the thermal load model. As output the model can deliver alerts, act on building electrical actuators (as for opening windows), and on the HVAC control model.

Solar power availability: To model the power output coming from solar panels Lu et al. developed a Deep Neural Network integrated enhanced by a SHapley Additive exPlanations (SHAP) to increase model for interpretability (Y. Lu et al., 2021). The model is useful to provide prevision on the expected power developed by solar panels and managing the power usage inside the building looking for energy neutrality (when consumption is equal to load). In this case the building efficiency is the highest and consumption are null, therefore it is the best operative case. However, the balance between load and production is usually different from zero and energy shall be stored in batteries or taken/delivered from the electrical network. Therefore, the model takes as input the solar radiation forecasting and the battery capacity to optimize the energy management and reduce consumptions. As output the model contributes to the input the loads prediction model.

Load profiling: Profiling loads is fundamental for the understanding of what is connected to the electrical network and to monitor if there are some faults in the devices. For this tasks Wastensteiner et al. proposed a Machine Learning Model that employs a Convolutional Neural Network (CNN) and a Local Interpretable Model-Agnostic Explanations (LIME) to increase the CNN interpretability (Wastensteiner et al., 2021). The use of a load profiling model can be also applied to monitor fraudulent energy usages identifying malevolent users as evidenced by (M. Wang et al., 2022). Knowing what is connected to the electrical system is usually a requirement made by decision makers for building management purposes. The model takes as input the data coming electrical loads and external loads (the EV charging station in the case of CITERA Smart-Lab). As output, energy profiles are an input of the load prediction model.

Loads prediction: Loads prediction is the main goal of any energy monitoring machine learning network. There are many literature studies that are focused on this issue; therefore, the solution is not unique. In the present work a Deep Neural Network model with a Local Interpretable Model-Agnostic Explanations (LIME) is proposed. To develop alerts and recommendations to users and stakeholders a Decision Tree models can be used (Mollo Neto et al., 2020; Tai et al., 2020; Xie et al., 2022; Yusuf Akbar et al., 2022). Interpreting load prediction is considered a fundamental requirement for stakeholders because they need to know how to operate the power systems, which load is prevalent and to develop energy reports for the assessment of the energy class of the building. The model takes as input the HVAC efficiency model, the loads profiling model, the solar power availability, the thermal load model, and the occupancy prevision model. As output the model can deliver alerts, act on actuators, recommending maintenance activities and reducing the HVAC power usage.

Fault detection: The last model is related to FDD, a good maintenance keeps the building efficient avoiding issue related to shot-circuit, End-of-Life (EoL) of devices, efficiency reduction, etc. There are many models that are specific to different equipment, as for air handling units (C. Fan et al., 2021). The proposed approach is to use two different models, one to detect numerically the efficiency reduction and one to interpretate errors and metadata data coming from sensors and devices. For numerical data a Convolutional Neural Network (CNN) model can be used tanks to its capability in interpretating data coming from very different devices (G. Li et al., 2021). For the metadata it is recommended to use an encoding-decoding model (C. Li et al., 2022). As input the model takes data coming from the physical layer devices and as output the model deliver alerts and can shut down faulty devices using the actuators.

4.2 Towards Smart Districts and Energy Communities

The integration of digital technologies is key to initiating optimization processes with data-driven methodologies that lead to real-time monitoring of all the functional parameters that characterize the energy and environmental behavior of buildings and infrastructures.

Indeed, only through continuous and real-time in-depth knowledge of physical, operational, usage, and performance parameters, combined with knowledge of boundary conditions, the system is capable of autonomously reacting to maintain optimal operating parameters both under standard conditions and in the event of unforeseen events.

Such an approach becomes increasingly relevant for managing assets of considerable size and number, becoming essential for the governance of energy districts, where the significant amount of data requires management techniques based on Artificial Intelligence and machine learning.

The main objective of Digital Twin-based and data-driven approaches on an urban scale concern improving the energy efficiency of buildings and infrastructures, as well as implementing strategies for defining energy districts through a system architecture able to interface with users through a cloud-based platform.

In the perspective of decentralizing electric energy production, the energy district can take on the role of the central core of a more extensive interconnected and scalable urban network, efficiently managed through Artificial Intelligence.

In this scenario, the virtuous management of energy production requires a strategy for migrating consumption by the end-user. It is essential to align the network's maximum consumption with the peak renewable energy production time. This can be done on individual infrastructure, as well as through the creation of a smart grid capable of feeding the energy surplus into the network. Considering the heterogeneity of the energy district in terms of usage destinations and prosumer habits, non-renewable energy consumption can tend towards zero.

The goal is to connect users of neighboring buildings in a single energy district, creating a cloud-based data platform that allows the integrated management of consumption and electric energy production, creating a virtuous, autonomous energy ecosystem capable of efficiently and effectively managing resources, distributing surplus energy among the prosumers, improving network efficiency, and minimizing waste.

As such, the smart district's Digital Twin-based data platform will allow:

- Archiving and profiling the habits of prosumers in the energy district;
- Managing the amounts of energy produced from different renewable sources to be distributed to different prosumers based on their habits;
- Sharing, dissemination, disclosure, and awareness-raising on energy-saving issues.

To ensure the platform's usability, a scalable user interface is needed to implement the system architecture with different services and functionalities based on different needs of different users.

Specifically, the cloud-based platform will need to interface with two types of users:

- Technical staff and Infrastructure Management;

- Users of the Smart Energy District and Energy Community (tertiary, residential, commercial, etc.), local administrations, stakeholders, etc.

The interpretation of the vast amount of data collected and updated in real time, and the subsequent profiling of users' energy consumption, through the use of specific machine learning (ML) algorithms validated following a monitoring period, allow the identification of preferential energy efficiency interventions, achieving the primary essential goal in the broader vision of a virtuous energy management of the built environment and the definition of a Smart Energy District and consequently an Energy Community.

The amount of structured and unstructured data, directly proportional to the breadth and heterogeneity of the system's application area, results in the definition of Artificial Intelligence algorithms aimed at automatically implementing Intelligent Load Management operations, such as Load Shed and Load Shift (**Figure 56**), as well as reporting faults to the assets involved in the system through the analysis of anomalies in electrical loads, the prediction of consumption, and the estimation of future electrical costs.

Specifically, the platform allows:

- Real-time visualization of the infrastructure's electrical consumption through dynamic dashboards;
- Analysis of loads and automatic recognition of consumption profiles to implement Intelligent Load Management actions:
    - Efficiency: Improvement of the energy efficiency of systems and components.
    - Load Shed: Reduction of load by distributing energy demand across multiple energy sources during peak usage periods.
    - Load Shift: Management of energy supply and demand in such a way that the peak consumption is shifted to periods of low energy demand.

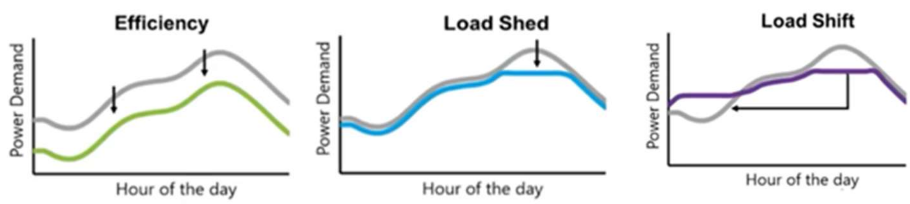


Figure 56. Load shift - shed

- Analysis of energy consumption and the production of electricity from renewable sources through integrated reports with possible intervention solutions;

- Reporting faults to the assets involved in the system through the analysis of anomalies of electrical loads connected to the specific asset;
- Creating a searchable database accessible at any time with the history of energy consumption connected to the infrastructure;
- Archiving and profiling of user habits;
- Forecasting consumption and estimating the future electrical costs of the building.

To connect and incentivize the highest number of users with heterogeneous energy profiles to join the creation of a near-zero impact Smart Energy District and the Energy Community, a cloud-based system scaled based on the specific information and usage required by the user is necessary (**Figure 57**).

This system can be translated into various services for users of the energy district that can monitor their consumption in real time and establish a relationship of sharing best practices, as well as take advantage of services such as data visualization, metaverse user experience, mixed reality, customer assistance, etc.

Technical Manager & Stakeholders

The data platform provides technical managers and stakeholders (owners and managers of real estate assets, local administrations, etc.) with:

- Basic and technical data on assets;
- Analytics on energy consumption and related costs;
- Alerts and notifications for optimized interaction between the system and the user.

People & Community

Users of the Smart Energy District and the Energy Community can benefit from the services offered by the data platform through:

- Web Apps;
- Mobile Apps;
- Immersive Reality & the Metaverse

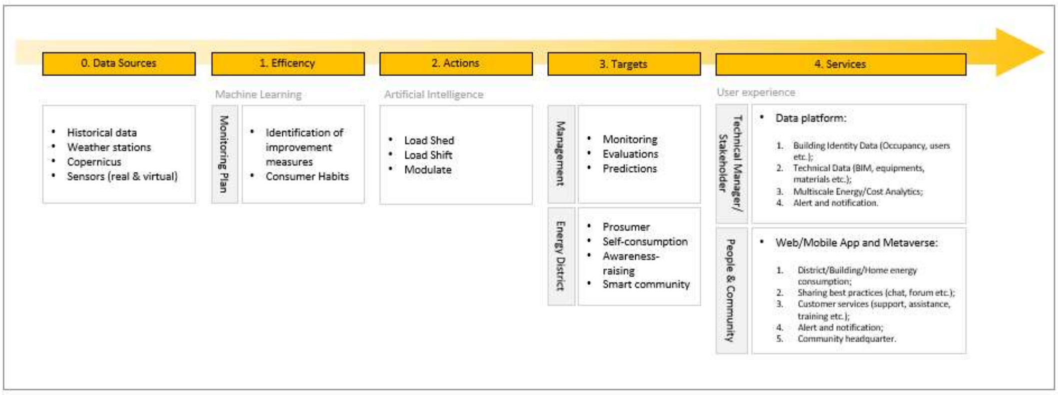


Figure 57. Multi-service data platform

Below are some of the objectives to be achieved towards the definition of a Digital Twin-based Data Platform for Energy Management in Smart Districts and Energy Communities in terms of system architecture.

The system is based on a Multi-Service Data Platform with a cloud architecture, integrating a data lake component for data ingestion and storage and a repository for unstructured data, as well as a data warehouse for managing certified data, and a set of specifically built data marts to meet the publishing needs through the Data Visualization and reporting component.

Specifically, distinct data marts are planned where the calculation rules for the metrics needed to process the KPIs for monitoring and control areas will be implemented.

The data ingestion component, using specific connectors, orchestrates data flows in streaming (near real-time) and batch modes, both structured and unstructured (data from sensors, documents, videos, images). The multi-service data platform will also have centralized governance that facilitates and simplifies the creation, protection, and management of data; the integrated data cataloging component allows secure publishing and management, defining the appropriate access levels and implementing data protection policies.

The system comes with monitoring functions for the correct operation of application components and data transformation jobs and provides all the necessary auditing functions to verify access logs and configuration changes.

Supporting the data processing functions, the platform will integrate an Artificial Intelligence Engine for energy management. Moreover, all necessary integrations towards the Data Presentation, BIM, Analytics, Command & Control, Mixed Reality, other external market systems and for exposing Open Data externally will be managed (**Figure 58**).

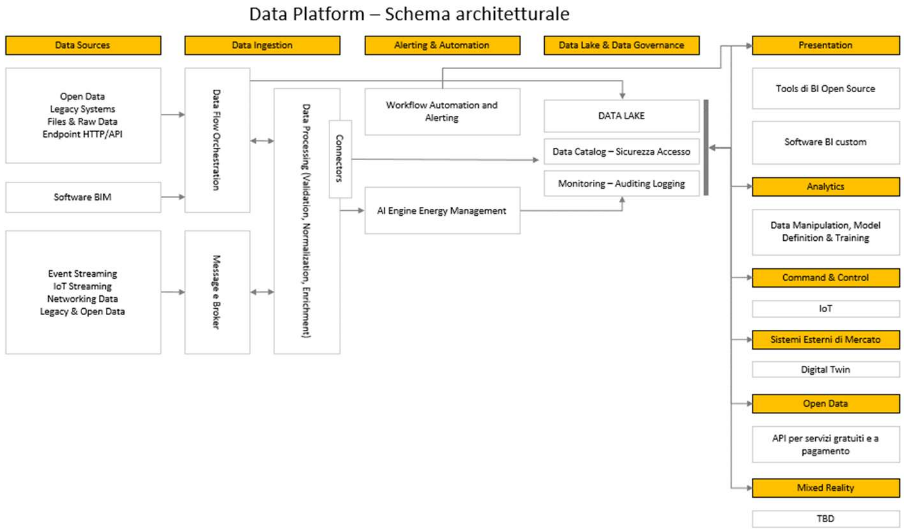


Figure 58. Data platform system architecture

Data publication will occur through a data presentation component where monitoring and control dashboards can be created. The same component allows for the creation of reports that can be distributed to recipient lists either on a scheduled basis or upon the occurrence of specific events. Furthermore, the Analytics component will allow the user to perform data discovery and processing activities autonomously on the published data, either through the integration with office automation tools or by customizing the dashboards in a self-service manner. The Data Visualization component allows the integration of published dashboards into third-party applications.

Within the multi-service data platform, a workflow management component allowing metadata configuration and information supporting the necessary monitoring dashboards, management of business rules for metric and monitoring KPI processing, as well as the definition of integration operation methods, the definition of alerts, and activation thresholds will be implemented.

To support the publication and distribution of data, the platform also integrates the generation of alert events towards third-party applications using a service bus architecture sending push notifications when alert events occur, and querying all acquired data available in the data lake.

4.3 Energy storage optimization towards a Vehicle-to-Grid (V2G) scenario

The contemporary urban environment is becoming more shaped by the pursuit of eco-friendliness and energy conservation. Through the incorporation of Renewable Energy Sources (RES) into smart grid management, urban areas can tap into a mix of energy resources to guarantee even nocturnal energy demands are met without relying on storage systems.

Nevertheless, the volatile characteristics of essential renewable technologies, such as solar and wind, in conjunction with a rise in electrification, threaten the balance and reliability of the power grid. Various solutions to these issues include energy storage, utilizing steady power generators like nuclear or geothermal energy, long-range power transmission, excessive dependence on renewables causing surplus during low-consumption times, and transitioning electricity into gas forms. As expenses reduce, the prominence of battery storage as a method for energy conservation has risen, improving grid efficiency in multiple scenarios.

Vehicle-to-Grid (V2G) represents the two-way energy exchange between an electric vehicle's (EV) battery and the grid. To achieve the kind of stability offered by fixed batteries, genuine bidirectional V2G is essential, enabling energy to be both supplied to and drawn from EV batteries in a controlled manner.

In this regard, the study by P.H. Kydd suggests a future where electric vehicles will play a pivotal role in energy storage and supply (Kydd, 2023). The storage capacity and power output of electric vehicles will see a significant rise over time, with a substantial potential impact on the electrical grid and the way energy is produced, distributed, and consumed.

Recent studies examined the potential of EVs in meeting grid storage needs investigating the utilization of EVs connected to the grid and batteries from retired EVs. Currently, most EV owners charge their cars at home overnight. When connected, these batteries could be repurposed for grid storage. By striking agreements with energy providers, EV owners can allow controlled charging and discharging of their batteries for grid support, earning financial incentives in return.

When EV batteries deteriorate to 70-80% of their original capacity, they're typically unfit for transportation but can still serve in grid storage. The combined capacity from vehicle-to-grid and expired EV batteries is predicted to reach 32 to 62 terawatt-hours by 2050. In contrast, short-term grid storage needs might only be 3.4 to 19.2 TWh by that year, implying a potential surplus of battery storage supply.

The study from Kydd reveals that only 12-43% of all EVs need to engage in vehicle-to-grid operations to meet global short-term grid-storage needs. This drops to under 10% if half the expired EV batteries are reused for grid storage. The adoption rate of these practices and the role of government incentives are crucial for maximizing EV contributions to grid storage. Effective participation requires user-friendly tools and strong regulations to recover and reintegrate batteries after their primary vehicle life (Kydd, 2023).

The chart below displays the projections for energy storage in electric vehicles (EV) in the United States from 2020 to 2050 (**Figure 59**).

Electric Vehicle Energy Storage 2020- 2050

Total fleet	253,000,000	Light duty vehicles in US		Total US Gen.	4116	Million MWh/202:
Growth rate	1%					1%
Year	**% EV**	**Number of EV Millions(1)**	**Battery cap. kWh, each(2)**	**EV storage cap. kWh, Millions**	**EV Power cap. MW at disc. rate**	**US Capacity MW(3)**
2020	1	3	55	138	68,750	456,308
2022	1.4	4	60	214	107,111	465,480
2025	3	8	80	631	315,303	479,585
2030	10	28	100	2,762	1,380,778	504,048
2035	25	73	120	8,707	4,353,634	529,760
2040	50	153	150	22,879	11,439,282	556,783
2050	75	253	150	37,908	18,954,125	615,035

1 Estimated from Bloomberg News EV sales forecast.
2 Estimated based on current capacities and trends.
3 US EIA Annual Energy Outlook, 2022, total generation / 8766 hours growing at 1%.

Figure 59. Projections for energy storage in electric vehicles (EV) in the United States from 2020 to 2050 (Kydd, 2023)

- Total Fleet and Growth Rate: There are 253 million light vehicles in the United States in 2020, with an anticipated annual growth rate of 1%.
- EV Percentage of Total: The proportion of electric vehicles out of the total vehicle count is expected to see a significant rise over time, going from 1% in 2020 to 75% by 2050.
- Number of EVs: The total number of electric vehicles is projected to increase from 3 million in 2020 to 253 million by 2050.
- Battery Capacity: The average battery capacity of electric vehicles is anticipated to grow over time. It starts at 55 kWh per vehicle in 2020 and is projected to reach 150 kWh per vehicle by 2050.
- Total Storage Capacity: The overall energy storage capacity of electric vehicles will see a substantial increase, moving from 55 million kWh in 2020 to 37,908 million kWh by 2050.
- EV Power Capacity: This refers to the maximum rate at which energy can be drawn from electric vehicle batteries. It's expected to rise from 68,750 MW in 2020 to 18,954,125 MW by 2050.
- US Capacity: The total capacity of the United States will see a modest growth, from 456,308 MW in 2020 to 615,035 MW by 2050.

The chart below shows the projected growth of V2G (Vehicle-to-Grid) technology from 2020 to 2040 (**Figure 60**).

V2G Growth 2020-2040

Year	Ford	Chevy	VW	Nissan	Others	Total sales	Total EV	MWh
	CAGR Sales	20%		Vehicles, thousands				
2020	Sales							
2022	200					200	200	20,000
2025	346	150	100	100	100	796	1,693	169,340
2030	860	373	249	249	200	1,931	8,510	850,959
2035	2,140	929	619	619	498	4,805	25,348	2,534,837
2040	2,140	2,311	1,541	1,541	1,238	8,771	59,287	5,928,663

Figure 60. Projected growth of V2G (Vehicle-to-Grid) technology from 2020 to 2040 (Kydd, 2023)

Sales by Brand:

- Ford: Ford sales are projected to rise from 200 thousand in 2020 to 2,140 thousand by 2040.
- Chevy: Sales for Chevy are expected to increase from 150 thousand in 2025 to 3,211 thousand by 2040.
- VW: VW's sales are anticipated to grow from 100 thousand in 2022 to 1,541 thousand in 2040.
- Nissan: Nissan's sales are forecasted to align with VW's, increasing from 100 thousand in 2022 to 1,541 thousand by 2040.
- Others: Sales from other brands are projected to rise from 100 thousand in 2022 to 1,238 thousand in 2040.
- Total Sales: The combined sales from all brands indicate substantial growth during the period in focus. Total sales will increase from 200 thousand in 2020 to 8,771 thousand by 2040.
- Total EV: The total number of electric vehicles (EV) in thousands will see a significant rise, from 200 in 2020 to 59,287 in 2040.
- MWh: The overall capacity in MWh will experience a significant hike, moving from 20,000 MWh in 2020 to 5,928,663 MWh in 2040.

In summary, studies depict a strong projected growth in V2G technology from 2020 to 2040, with a notable increase in V2G vehicle sales and the total MWh capacity. This implies a growing adoption and relevance of V2G technology in the transportation and energy sectors.

By 2040, the energy storage capacity of V2G-compatible vehicles is estimated to account for 15% of the entire electric vehicle storage. This points to a growth equivalent to twenty times the daily solar energy supply and a similar amount to the daily usage of the EV fleet. This hints at a substantial potential for bidirectional V2G capabilities in the near future.

Recent studies are examining the potential of electric vehicles (EVs) in meeting grid storage needs, investigating the utilization of EVs connected to the grid and batteries from retired EVs. Currently, most EV owners charge their cars at home overnight. When connected, these batteries could be repurposed for grid storage.

By striking agreements with energy providers, EV owners can allow controlled charging and discharging of their batteries for grid support, earning financial incentives in return.

When EV batteries deteriorate to 70-80% of their original capacity, they're typically unfit for transportation but can still serve in grid storage. The researchers predict that the combined capacity from vehicle-to-grid and expired EV batteries could amass 32 to 62 terawatt-hours by 2050. In contrast, short-term grid storage needs might only be 3.4 to 19.2 TWh by that year, implying a potential surplus of battery storage supply.

The study reveals that only 12-43% of all EVs need to engage in vehicle-to-grid operations to meet global short-term grid-storage needs. This drops to under 10% if half the expired EV batteries are reused for grid storage. The adoption rate of these practices and the role of government incentives are crucial for maximizing EV contributions to grid storage. Effective participation requires user-friendly tools and strong regulations to recover and reintegrate batteries after their primary vehicle life.

The vehicle-to-grid capacity refers to the portion of EV battery stock capacity suitable for vehicle-to-grid applications, accounting for the capacity allocated for EV driving, the capacity of plug-in hybrid electric vehicles (PHEVs) not participating in vehicle-to-grid due to limited capacity, and capacity decline resulting from battery degradation. EV batteries can serve a dual purpose, remaining within the vehicle for vehicle-to-grid utilization or being repurposed after the vehicle's end-of-life phase, where they are removed and utilized independently for stationary energy storage. "Smart" vehicle-to-grid charging strategies can facilitate dynamic EV charging and load-shifting grid services. EVs are capable of storing electricity and redistributing it to the grid during peak demand periods.

As mentioned above, vehicle battery EoL is conventionally defined as the point at which the remaining battery capacity falls within the range of 70-80% of its original capacity.

One crucial limitation is that these calculations presume that the rated capacity per vehicle will remain constant in the future and anticipate that a limited number of large battery electric vehicles (BEVs) could offer substantial actual vehicle-to-grid capacity. These capacities may undergo alterations in the future due to policy incentives, advancements in vehicle design, shifts in consumer preferences, enhancements in charging infrastructure, and various other influential factors. Moreover, the transportation ecosystem might experience profound and fundamental transformations, including a significant and rapid shift away from individual car usage toward mass transit, the adoption of shared electric vehicles, the implementation of autonomous driving technologies, and the success of battery swapping systems, all of which could reshape the available capacity landscape by 2050.

Using the Sustainable Development (SD) scenario, which envisions an electric vehicle fleet in line with the climate objectives of the Paris Agreement, the following graphs have been created by (C. Xu et al., 2023)to illustrate how the overall electric battery storage capacity is roughly evenly distributed between electric vehicles (EVs) and Energy Storage Systems (EoS) as in **Figure 61**.

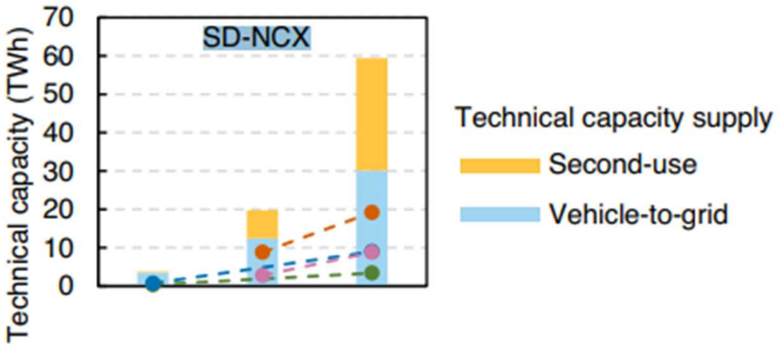


Figure 61. Technical capacity (TWh) distributed between electric vehicles (EVs) and Energy Storage Systems (EoS) (C. Xu et al., 2023).

The second graph (**Figure 62**), on the other hand, illustrates how, depending on the required power for storage in response to surpluses and the need to supply loads during a production deficit, Energy Storage Systems (EoS) can cover 100% of the demand up to a power requirement of 12 TWh (C. Xu et al., 2023). As the demand increases beyond 12 TWh and reaches 32 TWh, a portion of Electric Vehicles (EVs) becomes necessary for meeting the demand.

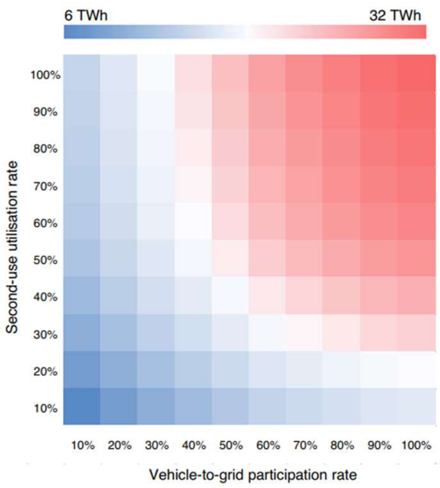


Figure 62. Second-use utilization rate vs. Vehicle-to-grid participation rate (C. Xu et al., 2023)

In the context of the issue within Italy, and by focusing on batteries that are actually installed in the current vehicle fleet, excluding Energy Storage Systems

(EoS) for which current legislation may serve as a basis for future discussions, the following analysis can be made.

Assuming the constant size of the circulating vehicle fleet in Italy until 2050, which is equal to about 39 million vehicles, and multiplying this number by the average annual mileage per vehicle in Italy, which is 6,800 kilometers per year, results in a total annual distance driven by our cars equal to [265.200.000.000 kilometers].

Given the current efficiency of electric cars, which consume about 20 kWh to travel 100 kilometers, there is an energy requirement for electric mobility equal to 53,04 TWh.

According to the most recent document from the GSE (Gestore dei Servizi Energetici) on energy scenarios related to renewable sources and [insert relevant data], it is inferred that by 2050, the electricity produced will be equal to 321,5TWh (**Figure 63**).

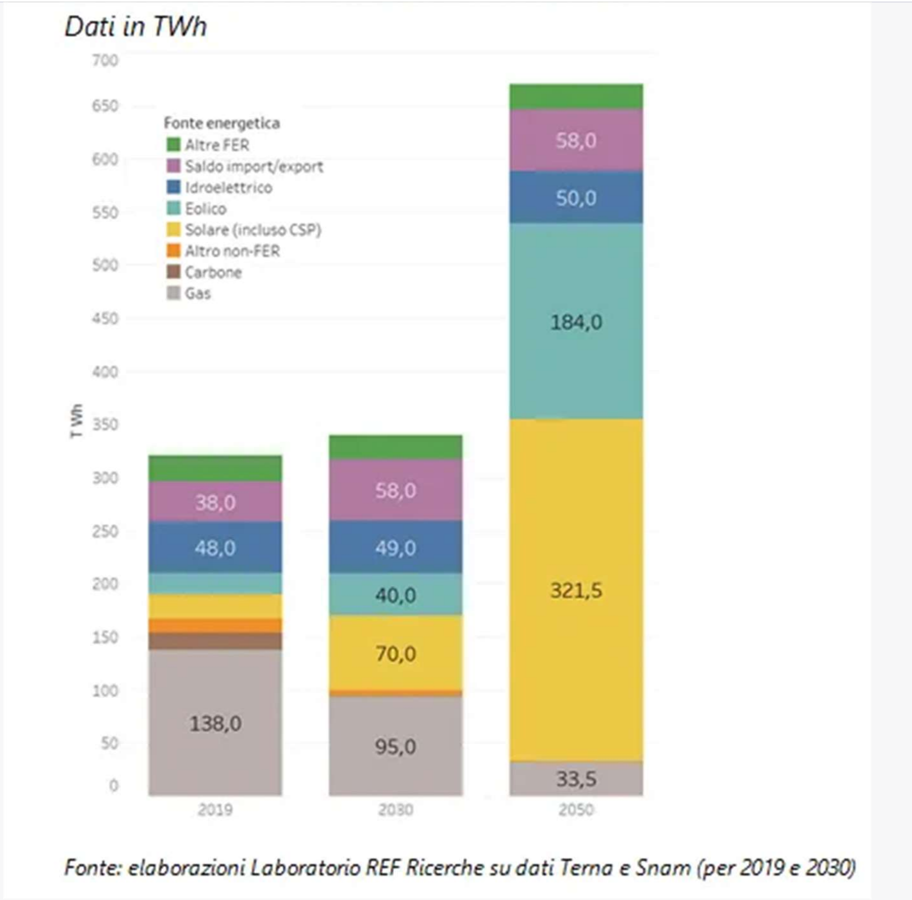


Figure 63. The evolution of the electric mix by 2050

Assuming the current surplus production percentages by self-consumption data of year 2022 (as shown in **Figure 64**) of about 13% of the overall FV production, suitably reduced by 20% due to the progressive improvement in the distribution network's efficiency, it can be hypothesized that by 2050, approximately 10% of the photovoltaic energy produced will be surplus to the load and will require storage.

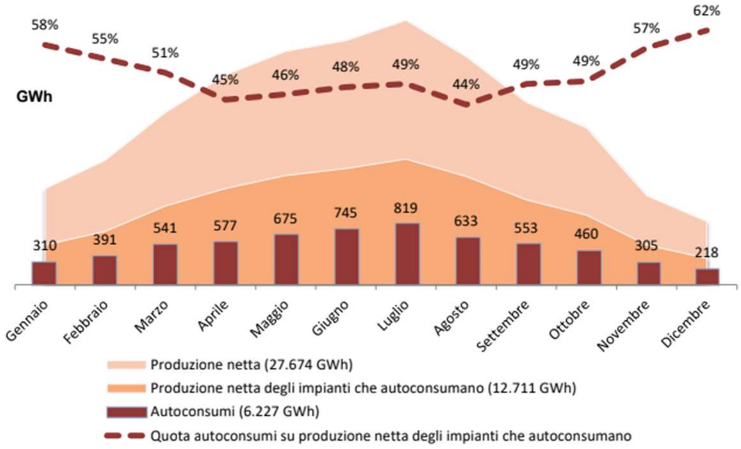


Figure 64. Surplus production percentages by self-consumption data of year 2022

This surplus is equivalent to 32,5 TWh, which is less than the energy requirements of electric vehicles in circulation. Consequently, the present framework is oriented to manage energy production surplus through electric vehicles instead of traditional storage systems.

4.4 Final remarks

The nexus between the built environment and energy utilization presents significant ramifications for global energy consumption, emissions, and the ongoing transition to more sustainable systems. Recent analyses reveal that buildings are responsible for consuming approximately 40% of global energy, contributing to over 70% of electricity consumption and nearly a third of all carbon emissions.

As the energy landscape evolves, so does the interaction between the built environment and energy sources. The emergence of Distributed Energy Resources (DERs) such as microgrids, nanogrids, and behind-the-meter energy storage introduces complexity but also opportunity. Other innovations, like grid-interactive efficient buildings and Electric Vehicle (EV) integration mechanisms (EV-to-grid, EV-to-home), are expanding the scope of demand response, ushering in a new era characterized by enhanced demand flexibility.

With advancements on the Internet of Things (IoT) over the past decade, once rudimentary equipment is now equipped with sophisticated sensors and interconnected devices. This shift from minimal data to vast data repositories enables advanced anomaly detection, predictive maintenance, and other intelligent functionalities, leveraging machine learning and other advanced techniques.

The inception of Digital Twins has revolutionized the modeling of context and environment in which new capabilities operate. Transitioning from smart offices to smart buildings and further to comprehensive smart spaces, showcasing its versatility across various scenarios including grid interactions and intelligent ecosystems, enhancing thermoigrometrical and IAQ real time parameters control.

Modern grids are increasingly integrating renewable sources such as wind and solar, creating challenges associated with demand fluctuations and weather-dependent production. With new additions like local solar and electric vehicles, predicting energy dynamics becomes complex.

As the built environment plays a crucial role in this energy landscape, impacting both consumption patterns and infrastructure resilience. This study aims to contribute to the research on the topic of introducing digital twin ecosystems in the management of the built environment and energy smart grids, providing an operational framework supported by open-source tools and methods for the creation of intelligent digital twins capable of optimizing the energy management of buildings.

As such it emerges from the study how the built environment intersects with multiple industry verticals. Whether it's a healthcare facility, a retail distribution center, or a manufacturing plant, each has distinct energy needs and interfaces uniquely with the grid. The rapid digitization and IoT implementation, evident in innovations like smart poles integrated with EV charging, further blurs these boundaries.

Incorporating digital twins, AI, and other advanced technologies enhances the ability to unlock significant capacities in the built environment. Embracing active efficiency — a blend of traditional energy-saving measures and digital transformation — ensures multi-dimensional benefits ranging from energy savings to improved grid resilience and building health.

As the built environment continues to integrate more advanced technologies and aligns more closely with evolving energy systems, it will increasingly function as a dynamic component of broader energy networks. This transformation promises significant potential savings, both in terms of energy and costs, emphasizing the importance of ongoing research, collaboration, and policy support in this domain.

Intelligent load management plays a pivotal role in optimizing energy efficiency. This revolves around several foundational principles:

Load Shedding & Shifts: simply cutting down the energy consumption during peak times, or redistributing loads at various times, can lead to significant energy savings using at the same time electric car fleet as a bidirectional exchange between an electric vehicle's (EV) battery and the grid. This capability becomes particularly potent when integrated with advanced models and AI-driven capabilities. These technologies not only allow for better energy management but can also adjust based on varying external inputs, creating a dynamic system that reacts in real-time.

Real Estate & Incentive Alignment: a challenge in energy management is the misalignment of incentives between property owners and tenants. Property owners may not feel the urge to invest in efficient systems as the cost savings are often passed to the tenants. Conversely, tenants might not be motivated to push for energy-efficient solutions for properties they don't own. Integrating AI-driven load management can create a win-win scenario. By optimizing energy usage, everyone benefits from the cost savings. Moreover, these savings can be integrated into a Distributed Energy Resource (DER) aggregation, adding another layer of benefits and incentives for various stakeholders.

Advanced Control Systems & Reinforcement Learning: traditional control systems for managing energy loads have reached their capacity. While Fault Detection and Diagnostics (FDD) systems have made substantial progress, they still operate on static rules. This is where reinforcement learning, a subset of AI, offers transformative potential. By utilizing a technique that constantly learns and adapts, energy systems can optimize beyond what traditional algorithms permit. This real-time adaptability means that energy systems can operate more efficiently, leading to substantial savings. An example is Microsoft's incorporation of Project Bonsai, which used reinforcement learning to increase energy savings from 15% to over 30%.

Digital Twins & Federated Systems: the concept of a Digital Twin refers to a digital representation of a physical system. This can range from components of a building to an entire city grid. When these digital representations are interconnected, they create an intelligent ecosystem. This ecosystem can analyze vast amounts of data, from weather forecasts to energy prices, to optimize energy usage in real-time. The challenge lies in creating harmonized ontologies that allow for seamless data exchange between various digital twins. Open-source initiatives are working towards making this federation more streamlined.

Equity in Energy Management: beyond efficiency, there's a need to ensure energy systems are equitable. As more sophisticated energy management solutions emerge, it's vital to ensure they benefit all sections of society. Furthermore, focusing on energy efficiency can particularly benefit low-income households, who proportionally spend more of their income on energy. Integrating AI and other advanced technologies in energy management can mitigate some of these disparities.

In conclusion, while alternative energy sources like wind and solar are essential, optimizing our existing systems through Intelligent Digital Twins based on AI and advanced load management can lead to a more efficient and equitable energy landscape. The integration of these technologies is still a work in progress, but the potential benefits are immense.

References

Abramovici, M., Göbel, J. C., & Savarino, P. (2017). Reconfiguration of smart products during their use phase based on virtual product twins. *CIRP Annals*, *66*(1), 165–168. https://doi.org/10.1016/j.cirp.2017.04.042

Adams, R. I., Miletto, M., Lindow, S. E., Taylor, J. W., & Bruns, T. D. (2014). Airborne Bacterial Communities in Residences: Similarities and Differences with Fungi. *PLoS ONE*, *9*(3), e91283. https://doi.org/10.1371/journal.pone.0091283

Afroz, Z., Higgins, G., Shafiullah, G. M., & Urmee, T. (2020a). Evaluation of real-life demand-controlled ventilation from the perception of indoor air quality with probable implications. *Energy and Buildings*, *219*, 110018. https://doi.org/10.1016/j.enbuild.2020.110018

Afroz, Z., Higgins, G., Shafiullah, G. M., & Urmee, T. (2020b). Evaluation of real-life demand-controlled ventilation from the perception of indoor air quality with probable implications. *Energy and Buildings*, *219*, 110018. https://doi.org/10.1016/j.enbuild.2020.110018

Aheleroff, S., Xu, X., Zhong, R. Y., & Lu, Y. (2021). Digital Twin as a Service (DTaaS) in Industry 4.0: An Architecture Reference Model. *Advanced Engineering Informatics*, *47*, 101225. https://doi.org/10.1016/j.aei.2020.101225

Ahmadian, S., Malki, H., & Han, Z. (2018). Cyber Attacks on Smart Energy Grids Using Generative Adverserial Networks. *2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 942–946. https://doi.org/10.1109/GlobalSIP.2018.8646424

Ala-Laurinaho, R., Autiosalo, J., Nikander, A., Mattila, J., & Tammi, K. (2020). Data Link for the Creation of Digital Twins. *IEEE Access*, *8*, 228675–228684. https://doi.org/10.1109/ACCESS.2020.3045856

Alam, K. M., & El Saddik, A. (2017). C2PS: A Digital Twin Architecture Reference Model for the Cloud-Based Cyber-Physical Systems. *IEEE Access*, *5*, 2050–2062. https://doi.org/10.1109/ACCESS.2017.2657006

Alavi, H., Forcada, N., Bortolini, R., & Edwards, D. J. (2021). Enhancing occupants' comfort through BIM-based probabilistic approach. *Automation in Construction*, *123*, 103528. https://doi.org/10.1016/j.autcon.2020.103528

Alcaraz, C., & Zeadally, S. (2015). Critical infrastructure protection: Requirements and challenges for the 21st century. *International Journal of Critical Infrastructure Protection*, *8*, 53–66. https://doi.org/10.1016/j.ijcip.2014.12.002

Allied Market Research. (2021). *Allied Market Research. AEC Market Insights, Segment Analysis and Industry Forecast-2027.* https://www.alliedmarketresearch.com

Al-Sarawi, S., Anbar, M., Alieyan, K., & Alzubaidi, M. (2017a). Internet of Things (IoT) communication protocols: Review. *2017 8th International Conference on Information Technology (ICIT)*, 685–690. https://doi.org/10.1109/ICITECH.2017.8079928

Al-Sarawi, S., Anbar, M., Alieyan, K., & Alzubaidi, M. (2017b). Internet of Things (IoT) communication protocols: Review. *2017 8th International Conference on Information Technology (ICIT)*, 685–690. https://doi.org/10.1109/ICITECH.2017.8079928

An, Y., Xia, T., You, R., Lai, D., Liu, J., & Chen, C. (2021). A reinforcement learning approach for control of window behavior to reduce indoor PM2.5 concentrations in naturally ventilated buildings. *Building and Environment*, *200*. https://doi.org/10.1016/j.buildenv.2021.107978

Andryushkevich, S. K., Kovalyov, S. P., & Nefedov, E. (2019). Composition and Application of Power System Digital Twins Based on Ontological Modeling. *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*, 1536–1542. https://doi.org/10.1109/INDIN41052.2019.8972267

ANSI/ASHRAE Standard 62.1-2022. (2022). *Ventilation and Acceptable Indoor Air Quality,*.

Arup. (2019). *Digital Twin*.

Autiosalo, J., Vepsalainen, J., Viitala, R., & Tammi, K. (2020). A Feature-Based Framework for Structuring Industrial Digital Twins. *IEEE Access*, *8*, 1193–1208. https://doi.org/10.1109/ACCESS.2019.2950507

Azam, M. F., & Younis, M. S. (2021). Multi-Horizon Electricity Load and Price Forecasting Using an Interpretable Multi-Head Self-Attention and EEMD-Based Framework. *IEEE Access*, *9*. https://doi.org/10.1109/ACCESS.2021.3086039

Baboli, P. T., Babazadeh, D., & Kumara Bowatte, D. R. (2020). Measurement-based Modeling of Smart Grid Dynamics: A Digital Twin Approach. *2020 10th Smart Grid Conference (SGC)*, 1–6. https://doi.org/10.1109/SGC52076.2020.9335750

Bahani, K., Ali-Ou-Salah, H., Moujabbir, M., Oukarfi, B., & Ramdani, M. (2020). A Novel Interpretable Model for Solar Radiation Prediction based on Adaptive Fuzzy Clustering and Linguistic Hedges. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3419604.3419807

Bahdanau, D., Cho, K. H., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*.

Baheti, R. and G. H. (2011). *Cyber-Physical Systems*.

Bai, L., Huang, L., Wang, Z., Ying, Q., Zheng, J., Shi, X., & Hu, J. (2020). Long-term field Evaluation of Low-cost Particulate Matter Sensors in Nanjing. *Aerosol and Air Quality Research*, *20*(2), 242–253. https://doi.org/10.4209/aaqr.2018.11.0424

Balocco, C., Petrone, G., Cammarata, G., Vitali, P., Albertini, R., & Pasquarella, C. (2014). Indoor Air Quality in a Real Operating Theatre under Effective Use Conditions. *Journal of Biomedical Science and Engineering*, *07*(11), 866–883. https://doi.org/10.4236/jbise.2014.711086

Banerjee, A., Dalal, R., Mittal, S., & Joshi, K. P. (2017). Generating Digital Twin Models using Knowledge Graphs for Industrial Production Lines. *Proceedings of* *the 2017 ACM on Web Science Conference*, 425–430. https://doi.org/10.1145/3091478.3162383

Barricelli, B. R., Casiraghi, E., & Fogli, D. (2019). A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access*, *7*, 167653–167671. https://doi.org/10.1109/ACCESS.2019.2953499

Bartolini, A., Corti, F., Reatti, A., Ciani, L., Grasso, F., & Kazimierczuk, M. K. (2020). Analysis and Design of Stand-Alone Photovoltaic System for precision agriculture network of sensors. *2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, 1–5. https://doi.org/10.1109/EEEIC/ICPSEurope49358.2020.9160554

Batty, M. (2018). Digital twins. *Environment and Planning B: Urban Analytics and City Science*, *45*(5), 817–820. https://doi.org/10.1177/2399808318796416

Bazmohammadi, N., Madary, A., Vasquez, J. C., Mohammadi, H. B., Khan, B., Wu, Y., & Guerrero, J. M. (2022a). Microgrid Digital Twins: Concepts, Applications, and Future Trends. *IEEE Access*, *10*, 2284–2302. https://doi.org/10.1109/ACCESS.2021.3138990

Bazmohammadi, N., Madary, A., Vasquez, J. C., Mohammadi, H. B., Khan, B., Wu, Y., & Guerrero, J. M. (2022b). Microgrid Digital Twins: Concepts, Applications, and Future Trends. *IEEE Access*, *10*, 2284–2302. https://doi.org/10.1109/ACCESS.2021.3138990

Bhatia, A., Garg, V., Haves, P., & Pudi, V. (2019). Explainable Clustering Using Hyper-Rectangles for Building Energy Simulation Data. *IOP Conference Series: Earth and Environmental Science*, *238*(1). https://doi.org/10.1088/1755-1315/238/1/012068

Biyik, E., & Kahraman, A. (2019). A predictive control strategy for optimal management of peak load, thermal comfort, energy storage and renewables in multi-zone buildings. *Journal of Building Engineering*, *25*, 100826. https://doi.org/10.1016/j.jobe.2019.100826

Bock, H.-H. (2007). *Clustering Methods: A History of k-Means Algorithms* (pp. 161–172). https://doi.org/10.1007/978-3-540-73560-1_15

Bolton, A. , Butler, L. , Dabson, I. , Enzer, M. , Evans, M. , Fenemore, T. , Harradence, F. , Keaney, E. , Kemp, A. , Luck, A. , Pawsey, N. , Saville, S. , Schooling, J. , Sharp, M. , Smith, T. , Tennison, J. , Whyte, J. , Wilson, A. , & Makri, C. (2018). *Gemini Principles*. https://doi.org/https://doi.org/10.17863/CAM.32260

Borowski, P. (2021). Digitization, Digital Twins, Blockchain, and Industry 4.0 as Elements of Management Process in Enterprises in the Energy Sector. *Energies*, *14*(7), 1885. https://doi.org/10.3390/en14071885

Bortolini, R., Rodrigues, R., Alavi, H., Vecchia, L. F. D., & Forcada, N. (2022). Digital Twins' Applications for Building Energy Efficiency: A Review. *Energies*, *15*(19), 7002. https://doi.org/10.3390/en15197002

Boschert, S., & Rosen, R. (2016a). Digital Twin—The Simulation Aspect. In *Mechatronic Futures* (pp. 59–74). Springer International Publishing. https://doi.org/10.1007/978-3-319-32156-1_5

Boschert, S., & Rosen, R. (2016b). Digital Twin—The Simulation Aspect. In *Mechatronic Futures* (pp. 59–74). Springer International Publishing. https://doi.org/10.1007/978-3-319-32156-1_5

Boyes, H., Hallaq, B., Cunningham, J., & Watson, T. (2018). The industrial internet of things (IIoT): An analysis framework. *Computers in Industry*, *101*, 1–12. https://doi.org/10.1016/j.compind.2018.04.015

Brilakis, I., Pan, Y., Borrmann, A., & Mayer, H.-G. (2019). *Built Environment Digital Twinning* (International Workshop on Built Environment Digital Twinning Presented by TUM Institute for Advanced Study and Siemens AG). https://doi.org/https://doi.org/10.17863/CAM.65445

Bruno, S., De Fino, M., & Fatiguso, F. (2018). Historic Building Information Modelling: performance assessment for diagnosis-aided information modelling and management. *Automation in Construction*, *86*, 256–276. https://doi.org/10.1016/j.autcon.2017.11.009

Bruynseels, K., Santoni de Sio, F., & van den Hoven, J. (2018). Digital Twins in Health Care: Ethical Implications of an Emerging Engineering Paradigm. *Frontiers in Genetics*, *9*. https://doi.org/10.3389/fgene.2018.00031

Buckman, A. H., Mayfield, M., & B.M. Beck, S. (2014). What is a Smart Building? *Smart and Sustainable Built Environment*, *3*(2), 92–109. https://doi.org/10.1108/SASBE-01-2014-0003

Buildings Performance Institute Europe. (2016). *Driving Transformational Change in the Construction Value Chain*.

Buildings Performance Institute Europe (BPIE). (2019). *HOW TO INTEGRATE INDOOR ENVIRONMENTAL QUALITY WITHIN NATIONAL LONG-TERM RENOVATION STRATEGIES*.

Burkart, N., & Huber, M. F. (2021). A survey on the explainability of supervised machine learning. In *Journal of Artificial Intelligence Research* (Vol. 70). https://doi.org/10.1613/JAIR.1.12228

C40 Cities. (2023). *C40 Cities*. https://www.c40.org/what-we-do/scaling-up-climate-action/energy-and-buildings/clean-construction-forum/

Cagnano, A., De Tuglie, E., & Mancarella, P. (2020a). Microgrids: Overview and guidelines for practical implementations and operation. *Applied Energy*, *258*, 114039. https://doi.org/10.1016/j.apenergy.2019.114039

Cagnano, A., De Tuglie, E., & Mancarella, P. (2020b). Microgrids: Overview and guidelines for practical implementations and operation. *Applied Energy*, *258*, 114039. https://doi.org/10.1016/j.apenergy.2019.114039

Cai, J., Chen, J., Hu, Y., Li, S., & He, Q. (2023). Digital twin for healthy indoor environment: A vision for the post-pandemic era. *Frontiers of Engineering Management*. https://doi.org/10.1007/s42524-022-0244-y

Cai, Z., Zheng, X., & Wang, J. (2021a). Efficient Data Trading for Stable and Privacy Preserving Histograms in Internet of Things. *2021 IEEE International Performance, Computing, and Communications Conference (IPCCC)*, 1–10. https://doi.org/10.1109/IPCCC51483.2021.9679420

Cai, Z., Zheng, X., & Wang, J. (2021b). Efficient Data Trading for Stable and Privacy Preserving Histograms in Internet of Things. *2021 IEEE International*

Performance, Computing, and Communications Conference (IPCCC), 1–10. https://doi.org/10.1109/IPCCC51483.2021.9679420

Canedo, A. (2016). Industrial IoT lifecycle via digital twins. *Proceedings of the Eleventh IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis*, 1–1. https://doi.org/10.1145/2968456.2974007

Capozzoli, A., Cerquitelli, T., & Piscitelli, M. S. (2016). Enhancing energy efficiency in buildings through innovative data analytics technologies. In *Pervasive Computing: Next Generation Platforms for Intelligent Data Collection*. https://doi.org/10.1016/B978-0-12-803663-1.00011-5

Caprari, G., Castelli, G., Montuori, M., Camardelli, M., & Malvezzi, R. (2022). Digital Twin for Urban Planning in the Green Deal Era: A State of the Art and Future Perspectives. *Sustainability*, *14*(10), 6263. https://doi.org/10.3390/su14106263

Cariou, S., Guillot, J.-M., Pepin, L., Kaluzny, P., & Faure, L.-P. (2005). A global indicator as a tool to follow airborne molecular contamination in a controlled environment. *Analytical and Bioanalytical Chemistry*, *381*(4), 850–853. https://doi.org/10.1007/s00216-004-2953-y

Carlsson, L. S., Samuelsson, P. B., & Jönsson, P. G. (2019). Using interpretable machine learning to predict the electrical energy consumption of an electric arc furnace. *Stahl Und Eisen*, *139*(9).

Carlsson, L. S., Samuelsson, P. B., & Jönsson, P. G. (2020). Interpretable Machine Learning—Tools to Interpret the Predictions of a Machine Learning Model Predicting the Electrical Energy Consumption of an Electric Arc Furnace. *Steel Research International*, *91*(11). https://doi.org/10.1002/srin.202000053

Carvalho, G., Cabral, B., Pereira, V., & Bernardino, J. (2020). Computation offloading in Edge Computing environments using Artificial Intelligence techniques. *Engineering Applications of Artificial Intelligence*, *95*, 103840. https://doi.org/10.1016/j.engappai.2020.103840

Castell, N., Dauge, F. R., Schneider, P., Vogt, M., Lerner, U., Fishbain, B., Broday, D., & Bartonova, A. (2017). Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? *Environment International*, *99*, 293–302. https://doi.org/10.1016/j.envint.2016.12.007

Centre for Digital Built Britain. (2019). *Monthly Paper: On the Governance of City Digital Twins*. https://www.cdbb.cam.ac.uk/news/monthly-paper-governance-city-digital-twins

Chakraborty, D., Alam, A., Chaudhuri, S., Başağaoğlu, H., Sulbaran, T., & Langar, S. (2021). Scenario-based prediction of climate change impacts on building cooling energy consumption with explainable artificial intelligence. *Applied Energy*, *291*. https://doi.org/10.1016/j.apenergy.2021.116807

Chang, X., Li, W., Ma, J., Yang, T., & Zomaya, A. Y. (2020). Interpretable Machine Learning in Sustainable Edge Computing: A Case Study of Short-Term Photovoltaic Power Output Prediction. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, *2020-May*. https://doi.org/10.1109/ICASSP40776.2020.9054088

Charalampopoulos, I. (2019). A comparative sensitivity analysis of human thermal comfort indices with generalized additive models. *Theoretical and Applied Climatology*, *137*(1–2). https://doi.org/10.1007/s00704-019-02900-1

Chen, B., Xiong, R., Li, H., Sun, Q., & Yang, J. (2019). Pathways for sustainable energy transition. *Journal of Cleaner Production*, *228*, 1564–1571. https://doi.org/10.1016/j.jclepro.2019.04.372

Chen, Y. (2017). Integrated and Intelligent Manufacturing: Perspectives and Enablers. *Engineering*, *3*(5), 588–595. https://doi.org/10.1016/J.ENG.2017.04.009

Chen, Y., Chen, Z., Yuan, X., Su, L., & Li, K. (2022). Optimal Control Strategies for Demand Response in Buildings under Penetration of Renewable Energy. *Buildings*, *12*(3). https://doi.org/10.3390/buildings12030371

Chen, Y., Guo, M., Chen, Z., Chen, Z., & Ji, Y. (2022). Physical energy and data-driven models in building energy prediction: A review. In *Energy Reports* (Vol. 8). https://doi.org/10.1016/j.egyr.2022.01.162

Chen, Y., Shen, H., Smith, K. R., Guan, D., Chen, Y., Shen, G., Liu, J., Cheng, H., Zeng, E. Y., & Tao, S. (2018). Estimating household air pollution exposures and health impacts from space heating in rural China. *Environment International*, *119*, 117–124. https://doi.org/10.1016/j.envint.2018.04.054

Chen, Y., & Zhang, D. (2021). Theory-guided deep-learning for electrical load forecasting (TgDLF) via ensemble long short-term memory. *Advances in Applied Energy*, *1*. https://doi.org/10.1016/j.adapen.2020.100004

Chen, Z., Chen, Y., He, R., Liu, J., Gao, M., & Zhang, L. (2022). Multi-objective residential load scheduling approach for demand response in smart grid. *Sustainable Cities and Society*, *76*. https://doi.org/10.1016/j.scs.2021.103530

Chen, Z., Chen, Y., Xiao, T., Wang, H., & Hou, P. (2021). A novel short-term load forecasting framework based on time-series clustering and early classification algorithm. *Energy and Buildings*, *251*. https://doi.org/10.1016/j.enbuild.2021.111375

Chen, Z., Xiao, F., Guo, F., & Yan, J. (2023). Interpretable machine learning for building energy management: A state-of-the-art review. In *Advances in Applied Energy* (Vol. 9). https://doi.org/10.1016/j.adapen.2023.100123

Chen, Z., Xiao, F., Shi, J., & Li, A. (2022). Dynamic model development for vehicle air conditioners based on physics-guided deep learning. *International Journal of Refrigeration*, *134*. https://doi.org/10.1016/j.ijrefrig.2021.11.021

Cheng, Y., Niu, J., & Gao, N. (2012). Thermal comfort models: A review and numerical investigation. *Building and Environment*, *47*(1). https://doi.org/10.1016/j.buildenv.2011.05.011

Cheong, K. H., Teo, Y. H., Koh, J. M., Acharya, U. R., & Man Yu, S. C. (2020). A simulation-aided approach in improving thermal-visual comfort and power efficiency in buildings. *Journal of Building Engineering*, *27*, 100936. https://doi.org/10.1016/j.jobe.2019.100936

Chiesa, G., Cesari, S., Garcia, M., Issa, M., & Li, S. (2019). Multisensor IoT Platform for Optimising IAQ Levels in Buildings through a Smart Ventilation System. *Sustainability*, *11*(20), 5777. https://doi.org/10.3390/su11205777

Cho, W., Song, D., Hwang, S., & Yun, S. (2015). Energy-efficient ventilation with air-cleaning mode and demand control in a multi-residential building. *Energy and Buildings*, *90*, 6–14. https://doi.org/10.1016/j.enbuild.2015.01.002

Choe, Y., Shin, J., Park, J., Kim, E., Oh, N., Min, K., Kim, D., Sung, K., Cho, M., & Yang, W. (2022). Inadequacy of air purifier for indoor air quality improvement in classrooms without external ventilation. *Building and Environment*, *207*, 108450. https://doi.org/10.1016/j.buildenv.2021.108450

Chui, M., Löffler, M., & Roberts, R. (2010). *The Internet of Things*. https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-internet-of-things#/

Ciaburro, G. (2022). Machine fault detection methods based on machine learning algorithms: A review. In *Mathematical Biosciences and Engineering* (Vol. 19, Issue 11). https://doi.org/10.3934/mbe.2022534

Collier, Z. A., DiMase, D., Walters, S., Tehranipoor, M. M., Lambert, J. H., & Linkov, I. (2014). Cybersecurity Standards: Managing Risk and Creating Resilience. *Computer*, *47*(9), 70–76. https://doi.org/10.1109/MC.2013.448

Colombo, A. W., Karnouskos, S., Shi, Y., Yin, S., & Kaynak, O. (2016). Industrial Cyber–Physical Systems [Scanning the Issue]. *Proceedings of the IEEE*, *104*(5), 899–903. https://doi.org/10.1109/JPROC.2016.2548318

Conde, J., Munoz-Arcentales, A., Alonso, A., Lopez-Pernas, S., & Salvachua, J. (2022). Modeling Digital Twin Data and Architecture: A Building Guide With FIWARE as Enabling Technology. *IEEE Internet Computing*, *26*(3), 7–14. https://doi.org/10.1109/MIC.2021.3056923

Corry, E., Pauwels, P., Hu, S., Keane, M., & O'Donnell, J. (2015). A performance assessment ontology for the environmental and energy management of buildings. *Automation in Construction*, *57*, 249–259. https://doi.org/10.1016/j.autcon.2015.05.002

Cronrath, C., Aderiani, A. R., & Lennartson, B. (2019). Enhancing Digital Twins through Reinforcement Learning. *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*, 293–298. https://doi.org/10.1109/COASE.2019.8842888

Dai, X., Liu, J., & Zhang, X. (2020). A review of studies applying machine learning models to predict occupancy and window-opening behaviours in smart buildings. In *Energy and Buildings* (Vol. 223). https://doi.org/10.1016/j.enbuild.2020.110159

Damjanovic-Behrendt, V., & Behrendt, W. (2019). An open-source approach to the design and implementation of Digital Twins for Smart Manufacturing. *International Journal of Computer Integrated Manufacturing*, *32*(4–5), 366–384. https://doi.org/10.1080/0951192X.2019.1599436

Danilczyk, W., Sun, Y., & He, H. (2019). ANGEL: An Intelligent Digital Twin Framework for Microgrid Security. *2019 North American Power Symposium (NAPS)*, 1–6. https://doi.org/10.1109/NAPS46351.2019.9000371

Danilczyk, W., Sun, Y. L., & He, H. (2021a). Smart Grid Anomaly Detection using a Deep Learning Digital Twin. *2020 52nd North American Power Symposium (NAPS)*, 1–6. https://doi.org/10.1109/NAPS50074.2021.9449682

Danilczyk, W., Sun, Y. L., & He, H. (2021b). Smart Grid Anomaly Detection using a Deep Learning Digital Twin. *2020 52nd North American Power Symposium (NAPS)*, 1–6. https://doi.org/10.1109/NAPS50074.2021.9449682

Danilczyk, W., Sun, Y. L., & He, H. (2021c). Smart Grid Anomaly Detection using a Deep Learning Digital Twin. *2020 52nd North American Power Symposium (NAPS)*, 1–6. https://doi.org/10.1109/NAPS50074.2021.9449682

David, J., Lobov, A., & Lanz, M. (2018). Learning Experiences Involving Digital Twins. *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*, 3681–3686. https://doi.org/10.1109/IECON.2018.8591460

Davila Delgado, J. M., & Hofmeyer, H. (2013). Automated generation of structural solutions based on spatial designs. *Automation in Construction*, *35*, 528–541. https://doi.org/10.1016/j.autcon.2013.06.008

Daza, A., & Frank, A. (2019). Translate and label! An encoder-decoder approach for cross-lingual semantic role labeling. *EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference*. https://doi.org/10.18653/v1/d19-1056

Dedinec, A., Filiposka, S., Dedinec, A., & Kocarev, L. (2016). Deep belief network based electricity load forecasting: An analysis of Macedonian case. *Energy*, *115*, 1688–1700. https://doi.org/10.1016/j.energy.2016.07.090

Deloitte. (2020). *The 2030 decarbonization challenge: The path to the future of energy*. https://www2.deloitte.com/xe/en/pages/energy-and-resources/articles/the-2030-decarbonization-challenge.html

Demanega, I., Mujan, I., Singer, B. C., Anđelković, A. S., Babich, F., & Licina, D. (2021a). Performance assessment of low-cost environmental monitors and single sensors under variable indoor air quality and thermal conditions. *Building and Environment*, *187*, 107415. https://doi.org/10.1016/j.buildenv.2020.107415

Demanega, I., Mujan, I., Singer, B. C., Anđelković, A. S., Babich, F., & Licina, D. (2021b). Performance assessment of low-cost environmental monitors and single sensors under variable indoor air quality and thermal conditions. *Building and Environment*, *187*, 107415. https://doi.org/10.1016/j.buildenv.2020.107415

Deng, M., Menassa, C. C., & Kamat, V. R. (2021). From BIM to digital twins: a systematic review of the evolution of intelligent building representations in the AEC-FM industry. *Journal of Information Technology in Construction*, *26*, 58–83. https://doi.org/10.36680/j.itcon.2021.005

Deren, L., Wenbo, Y., & Zhenfeng, S. (2021). Smart city based on digital twins. *Computational Urban Science*, *1*(1), 4. https://doi.org/10.1007/s43762-021-00005-y

Di Natale, L., Svetozarevic, B., Heer, P., & Jones, C. N. (2022). Physically Consistent Neural Networks for building thermal modeling: Theory and analysis. *Applied Energy*, *325*. https://doi.org/10.1016/j.apenergy.2022.119806

Di Silvestre, M. L., Favuzza, S., Riva Sanseverino, E., & Zizzo, G. (2018). How Decarbonization, Digitalization and Decentralization are changing key power infrastructures. *Renewable and Sustainable Energy Reviews*, *93*, 483–498. https://doi.org/10.1016/j.rser.2018.05.068

Digital Twin Consortium. (2020). *The Definition of a Digital Twin*. https://www.digitaltwinconsortium.org/hot-topics/the-definition-of-a-digital-twin/#:~:text=On%20December%203%2C%202020%2C%20to,a%20specified%20frequency%20and%20fidelity.

Dileep, G. (2020). A survey on smart grid technologies and applications. *Renewable Energy*, *146*, 2589–2625. https://doi.org/10.1016/j.renene.2019.08.092

Din, G. M. U., & Marnerides, A. K. (2017). Short term power load forecasting using Deep Neural Networks. *2017 International Conference on Computing, Networking and Communications (ICNC)*, 594–598. https://doi.org/10.1109/ICCNC.2017.7876196

Eastman, C., Teicholz, P., Sacks, R., & Liston, K. (2011). *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engi- neers and Contractors.* Hoboken, .

Ebrahimi, A. (2019). Challenges of developing a digital twin model of renewable energy generators. *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*, 1059–1066. https://doi.org/10.1109/ISIE.2019.8781529

Ekberg, L. E. (1994). Outdoor Air Contaminants and Indoor Air Quality under Transient Conditions. *Indoor Air*, *4*(3), 189–196. https://doi.org/10.1111/j.1600-0668.1994.t01-1-00007.x

El Fiorenza Caroline, J., Parmar, P., Tiwari, S., Dixit, A., & Gupta, A. (2019). Accuracy prediction using analysis methods and f-measures. *Journal of Physics: Conference Series*, *1362*(1). https://doi.org/10.1088/1742-6596/1362/1/012040

El Saddik, A. (2018). Digital Twins: The Convergence of Multimedia Technologies. *IEEE MultiMedia*, *25*(2), 87–92. https://doi.org/10.1109/MMUL.2018.023121167

EN 16798-1:2018. (2018). *EN 16798-1:2018, Energy performance of buildings - Part 1: Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics - Module M1-6.*

Estebsari, A., & Rajabi, R. (2020). Single Residential Load Forecasting Using Deep Learning and Image Encoding Techniques. *Electronics*, *9*(1), 68. https://doi.org/10.3390/electronics9010068

European Commission. Competence Centre on Foresigh. (2022). *Global population is growing.* https://knowledge4policy.ec.europa.eu/foresight/global-population-growing_en

European Commission. Directorate General for Research and Innovation. (2020). *Enabling Technologies for Industry 5.0: Results of a Workshop with Europe's Technology Leaders.*

European Commission. Directorate General for Research and Innovation. (2021). *Industry 5.0, a Transformative Vision for Europe: Governing Systemic Transformations towards a Sustainable Industry.*

European Commission, J. R. C. (2019). *Digital transformation in transport, construction, energy, government and public administration.*

European Commission. Joint Research Centre. Institute for Energy and Transport. (2015). *Energy Renovation :The Trump Card for the New Start of Europe.*

Fan, C., Liu, X., Xue, P., & Wang, J. (2021). Statistical characterization of semi-supervised neural networks for fault detection and diagnosis of air handling units. *Energy and Buildings*, *234*. https://doi.org/10.1016/j.enbuild.2021.110733

Fan, C., Xiao, F., Yan, C., Liu, C., Li, Z., & Wang, J. (2019). A novel methodology to explain and evaluate data-driven building energy performance models based on interpretable machine learning. *Applied Energy*, *235*. https://doi.org/10.1016/j.apenergy.2018.11.081

Fan, F. L., Xiong, J., Li, M., & Wang, G. (2021). On Interpretability of Artificial Neural Networks: A Survey. *IEEE Transactions on Radiation and Plasma Medical Sciences*, *5*(6). https://doi.org/10.1109/TRPMS.2021.3066428

Fanger, P. O. (1986). Thermal environment - Human requirements. *The Environmentalist*, *6*(4). https://doi.org/10.1007/BF02238059

Fanger, P. O. (1988). Introduction of the olf and the decipol units to quantify air pollution perceived by humans indoors and outdoors. *Energy and Buildings*, *12*(1), 1–6. https://doi.org/10.1016/0378-7788(88)90051-5

Farhana, K., Kadirgama, K., Mohammed, H. A., Ramasamy, D., Samykano, M., & Saidur, R. (2021). Analysis of efficiency enhancement of flat plate solar collector using crystal nano-cellulose (CNC) nanofluids. *Sustainable Energy Technologies and Assessments*, *45*, 101049. https://doi.org/10.1016/j.seta.2021.101049

Federico Cinquepalmi. (2019). *Towards (R)evolving Cities. Urban Fragilities and Prospects in the 21st century*. Dida press.

Franchini, G., Ruggiero, V., Porta, F., & Zanni, L. (2022). Neural architecture search via standard machine learning methodologies. *Mathematics in Engineering*, *5*(1), 1–21. https://doi.org/10.3934/mine.2023012

Francisco, A., Mohammadi, N., & Taylor, J. E. (2020). Smart City Digital Twin–Enabled Energy Management: Toward Real-Time Urban Building Energy Benchmarking. *Journal of Management in Engineering*, *36*(2). https://doi.org/10.1061/(ASCE)ME.1943-5479.0000741

Friston, K., & Penny, W. (2011). Post hoc Bayesian model selection. *NeuroImage*, *56*(4). https://doi.org/10.1016/j.neuroimage.2011.03.062

Fuller, A., Fan, Z., Day, C., & Barlow, C. (2020). Digital Twin: Enabling Technologies, Challenges and Open Research. *IEEE Access*, *8*, 108952–108971. https://doi.org/10.1109/ACCESS.2020.2998358

Gabor, T., Belzner, L., Kiermeier, M., Beck, M. T., & Neitz, A. (2016). A Simulation-Based Architecture for Smart Cyber-Physical Systems. *2016 IEEE International Conference on Autonomic Computing (ICAC)*, 374–379. https://doi.org/10.1109/ICAC.2016.29

Gales, B., Kander, A., Malanima, P., & Rubio, M. (2007). North versus South: Energy transition and energy intensity in Europe over 200 years. *European Review of Economic History*, *11*(2), 219–253. https://doi.org/10.1017/S1361491607001967

Gan, V. J. L., Luo, H., Tan, Y., Deng, M., & Kwok, H. L. (2021). BIM and Data-Driven Predictive Analysis of Optimum Thermal Comfort for Indoor Environment. *Sensors*, *21*(13), 4401. https://doi.org/10.3390/s21134401

Ganesh, H. S., Fritz, H. E., Edgar, T. F., Novoselac, A., & Baldea, M. (2019). A model-based dynamic optimization strategy for control of indoor air pollutants. *Energy and Buildings*, *195*, 168–179. https://doi.org/10.1016/j.enbuild.2019.04.022

Gao, Y., Han, H., Lu, H., Jiang, S. X., Zhang, Y., & Luo, M. W. (2022). Knowledge mining for chiller faults based on explanation of data-driven diagnosis. *Applied Thermal Engineering*, *205*. https://doi.org/10.1016/j.applthermaleng.2021.118032

Garetti, M., Rosa, P., & Terzi, S. (2012). Life Cycle Simulation for the design of Product–Service Systems. *Computers in Industry*, *63*(4), 361–369. https://doi.org/10.1016/j.compind.2012.02.007

Gauthier, J., & Wooldridge, B. (2012). Influences on Sustainable Innovation Adoption: Evidence from Leadership in Energy and Environmental Design. *Business Strategy and the Environment*, *21*(2), 98–110. https://doi.org/10.1002/bse.716

Gharaei, A., Lu, J., Stoll, O., Zheng, X., West, S., & Kiritsis, D. (2020). *Systems Engineering Approach to Identify Requirements for Digital Twins Development* (pp. 82–90). https://doi.org/10.1007/978-3-030-57993-7_10

Ghenai, C., Husein, L. A., Al Nahlawi, M., Hamid, A. K., & Bettayeb, M. (2022). Recent trends of digital twin technologies in the energy sector: A comprehensive review. *Sustainable Energy Technologies and Assessments*, *54*, 102837. https://doi.org/10.1016/j.seta.2022.102837

GitHub. (n.d.). *Azure. Digital Twin Definition Language.* Retrieved 15 October 2023, from https://github.com/Azure/opendigitaltwins-dtdl

Glaessgen, E., & Stargel, D. (2012, April 23). The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles. *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference
20th AIAA/ASME/AHS Adaptive Structures Conference
14th AIAA*. https://doi.org/10.2514/6.2012-1818

Glatt, M., Sinnwell, C., Yi, L., Donohoe, S., Ravani, B., & Aurich, J. C. (2021). Modeling and implementation of a digital twin of material flows based on physics simulation. *Journal of Manufacturing Systems*, *58*, 231–245. https://doi.org/10.1016/j.jmsy.2020.04.015

Godish, T., & Spengler, J. D. (1996). Relationships Between Ventilation and Indoor Air Quality: A Review. *Indoor Air*, *6*(2), 135–145. https://doi.org/10.1111/j.1600-0668.1996.00010.x

Goia, B., Cioara, T., & Anghel, I. (2022). Virtual Power Plant Optimization in Smart Grids: A Narrative Review. *Future Internet*, *14*(5), 128. https://doi.org/10.3390/fi14050128

Golizadeh Akhlaghi, Y., Aslansefat, K., Zhao, X., Sadati, S., Badiei, A., Xiao, X., Shittu, S., Fan, Y., & Ma, X. (2021). Hourly performance forecast of a dew point cooler using explainable Artificial Intelligence and evolutionary optimisations by 2050. *Applied Energy*, *281*. https://doi.org/10.1016/j.apenergy.2020.116062

Gonzalez, R., Ahmed, S., & Alamaniotis, M. (2023). Implementing Very-Short-Term Forecasting of Residential Load Demand Using a Deep Neural Network Architecture. *Energies*, *16*(9), 3636. https://doi.org/10.3390/en16093636

Gonzalo, F. del A., Griffin, M., Laskosky, J., Yost, P., & González-Lezcano, R. A. (2022). Assessment of Indoor Air Quality in Residential Buildings of New England through Actual Data. *Sustainability*, *14*(2), 739. https://doi.org/10.3390/su14020739

Grieves, M. (2006). *Product lifecycle management: driving the next generation of lean thinking*. McGraw-Hill.

Grieves M. (2011). *Virtually Perfect: Driving Innovative and Lean Products through Product Lifecycle Management*. Space Coast Press.

Grieves, M. (2016). Origins of the Digital Twin Concept. *Unpublished*.

Grieves, M. (2022). Intelligent digital twins and the development and management of complex systems. *Digital Twin*, *2*, 8. https://doi.org/10.12688/digitaltwin.17574.1

Grieves, M., & Vickers, J. (2017). Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In *Transdisciplinary Perspectives on Complex Systems* (pp. 85–113). Springer International Publishing. https://doi.org/10.1007/978-3-319-38756-7_4

Grieves, M. W. (2005). Product lifecycle management: the new paradigm for enterprises. *International Journal of Product Development*, *2*(1/2), 71. https://doi.org/10.1504/IJPD.2005.006669

Grieves, M. W. (2019). Virtually Intelligent Product Systems: Digital and Physical Twins. In *Complex Systems Engineering: Theory and Practice* (pp. 175–200). American Institute of Aeronautics and Astronautics, Inc. https://doi.org/10.2514/5.9781624105654.0175.0200

Grimaldo, A. I., & Novak, J. (2019). User-Centered Visual Analytics Approach for Interactive and Explainable Energy Demand Analysis in Prosumer Scenarios. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11754 LNCS*. https://doi.org/10.1007/978-3-030-34995-0_64

Grimaldo, A., & Novak, J. (2021). Explainable Needn't Be (Much) Less Accurate: Evaluating an Explainable AI Dashboard for Energy Forecasting. *IFIP Advances in Information and Communication Technology*, *628*. https://doi.org/10.1007/978-3-030-79157-5_28

Guan, J., Li, Y., Liu, J., Duan, X., Shen, D., Jia, D., & Ku, C. (2021). Experimental and numerical research on the performance characteristics of OPLVCR engine based on the NSGA II algorithm using digital twins. *Energy Conversion and Management*, *236*, 114052. https://doi.org/10.1016/j.enconman.2021.114052

Guan, Z., Sun, N., Xu, Y., & Yang, T. (2015). A comprehensive survey of false data injection in smart grid. *International Journal of Wireless and Mobile Computing*, *8*(1), 27. https://doi.org/10.1504/IJWMC.2015.066756

Gugliermetti, L., & Astiaso Garcia, D. (2018). A cheap and third-age-friendly home device for monitoring indoor air quality. *International Journal of Environmental Science and Technology*, *15*(1), 185–198. https://doi.org/10.1007/s13762-017-1382-3

Guo, J., Zhao, N., Sun, L., & Zhang, S. (2019). Modular based flexible digital twin for factory design. *Journal of Ambient Intelligence and Humanized Computing*, *10*(3), 1189–1200. https://doi.org/10.1007/s12652-018-0953-6

Haag, S., & Anderl, R. (2018). Digital twin – Proof of concept. *Manufacturing Letters*, *15*, 64–66. https://doi.org/10.1016/j.mfglet.2018.02.006

Harris, Z. S. (1954). Distributional Structure. WORD, *10*(2–3), 146–162. https://doi.org/10.1080/00437956.1954.11659520

Hatcher, W. G., Qian, C., Gao, W., Liang, F., Hua, K., & Yu, W. (2021). Towards Efficient and Intelligent Internet of Things Search Engine. *IEEE Access*, *9*, 15778–15795. https://doi.org/10.1109/ACCESS.2021.3052759

Hatcher, W. G., & Yu, W. (2018). A Survey of Deep Learning: Platforms, Applications and Emerging Research Trends. *IEEE Access*, *6*, 24411–24432. https://doi.org/10.1109/ACCESS.2018.2830661

He, Y., Deng, J., & Li, H. (2017). Short-Term Power Load Forecasting with Deep Belief Network and Copula Models. *2017 9th International Conference on Intelligent*

Human-Machine Systems and Cybernetics (IHMSC), 191–194. https://doi.org/10.1109/IHMSC.2017.50

Hlady, J., Glanzer, M., & Fugate, L. (2018, September 24). Automated Creation of the Pipeline Digital Twin During Construction: Improvement to Construction Quality and Pipeline Integrity. *Volume 2: Pipeline Safety Management Systems; Project Management, Design, Construction, and Environmental Issues; Strain Based Design; Risk and Reliability; Northern Offshore and Production Pipelines*. https://doi.org/10.1115/IPC2018-78146

Hochhalter J.D., Leser W.P., & Newman J.A. (2014). *Coupling Damage-Sensing Particles to the Digital Twin Concept.* . https://ntrs.nasa.gov/api/citations/20140006408/downloads/20140006408.pdf

Hossein Motlagh, N., Taleb, T., & Arouk, O. (2016). Low-Altitude Unmanned Aerial Vehicles-Based Internet of Things Services: Comprehensive Survey and Future Perspectives. *IEEE Internet of Things Journal*, *3*(6), 899–922. https://doi.org/10.1109/JIOT.2016.2612119

Huang, W., Zhang, Y., & Zeng, W. (2022). Development and application of digital twin technology for integrated regional energy systems in smart cities. *Sustainable Computing: Informatics and Systems*, *36*, 100781. https://doi.org/10.1016/j.suscom.2022.100781

Humayed, A., Lin, J., Li, F., & Luo, B. (2017). Cyber-Physical Systems Security—A Survey. *IEEE Internet of Things Journal*, *4*(6), 1802–1831. https://doi.org/10.1109/JIOT.2017.2703172

Hurtado, L. A., Mocanu, E., Nguyen, P. H., Gibescu, M., & Kamphuis, R. I. G. (2018). Enabling Cooperative Behavior for Building Demand Response Based on Extended Joint Action Learning. *IEEE Transactions on Industrial Informatics*, *14*(1). https://doi.org/10.1109/TII.2017.2753408

Imbulana Arachchi, J., & Managi, S. (2021). Preferences for energy sustainability: Different effects of gender on knowledge and importance. *Renewable and Sustainable Energy Reviews*, *141*, 110767. https://doi.org/10.1016/j.rser.2021.110767

International Energy Agency. (2022). *World Energy Outlook*.

Islam, M. T., Jain, S., Chen, Y., Chowdhury, B. D. B., & Son, Y.-J. (2022). An Agent-Based Simulation Model to Evaluate Contacts, Layout, and Policies in Entrance, Exit, and Seating in Indoor Activities Under a Pandemic Situation. *IEEE Transactions on Automation Science and Engineering*, *19*(2), 603–619. https://doi.org/10.1109/TASE.2021.3118008

ISO. (2005). ISO 7730: Ergonomics of the thermal environment Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria. *Management*, *3*.

Jaensch, F., Csiszar, A., Scheifele, C., & Verl, A. (2018). Digital Twins of Manufacturing Systems as a Base for Machine Learning. *2018 25th International Conference on Mechatronics and Machine Vision in Practice (M2VIP)*, 1–6. https://doi.org/10.1109/M2VIP.2018.8600844

Jain, P., Poon, J., Singh, J. P., Spanos, C., Sanders, S. R., & Panda, S. K. (2020). A Digital Twin Approach for Fault Diagnosis in Distributed Photovoltaic Systems. *IEEE Transactions on Power Electronics*, *35*(1), 940–956. https://doi.org/10.1109/TPEL.2019.2911594

Jaloudi, S. (2019a). Communication Protocols of an Industrial Internet of Things Environment: A Comparative Study. *Future Internet*, *11*(3), 66. https://doi.org/10.3390/fi11030066

Jaloudi, S. (2019b). Communication Protocols of an Industrial Internet of Things Environment: A Comparative Study. *Future Internet*, *11*(3), 66. https://doi.org/10.3390/fi11030066

Javed, A., Larijani, H., & Wixted, A. (2018). Improving Energy Consumption of a Commercial Building with IoT and Machine Learning. *IT Professional*, *20*(5). https://doi.org/10.1109/MITP.2018.053891335

Jokl, M. V. (2000). Evaluation of indoor air quality using the decibel concept based on carbon dioxide and TVOC. *Building and Environment*, *35*(8), 677–697. https://doi.org/10.1016/S0360-1323(99)00042-6

Joseph, A., Cvetkovic, M., & Palensky, P. (2018). Prediction of Short-Term Voltage Instability Using a Digital Faster than Real-Time Replica. *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*, 3582–3587. https://doi.org/10.1109/IECON.2018.8592818

Joshi, S. (2008). The sick building syndrome. *Indian Journal of Occupational and Environmental Medicine*, *12*(2), 61. https://doi.org/10.4103/0019-5278.43262

Justo Alonso, M., Dols, W. S., & Mathisen, H. M. (2022). Using Co-simulation between EnergyPlus and CONTAM to evaluate recirculation-based, demand-controlled ventilation strategies in an office building. *Building and Environment*, *211*, 108737. https://doi.org/10.1016/j.buildenv.2021.108737

Justo Alonso, M., Liu, P., Marman, S. F., Jørgensen, R. B., & Mathisen, H. M. (2023). Holistic methodology to reduce energy use and improve indoor air quality for demand-controlled ventilation. *Energy and Buildings*, *279*, 112692. https://doi.org/10.1016/j.enbuild.2022.112692

Justo Alonso, M., Wolf, S., Jørgensen, R. B., Madsen, H., & Mathisen, H. M. (2022). A methodology for the selection of pollutants for ensuring good indoor air quality using the de-trended cross-correlation function. *Building and Environment*, *209*, 108668. https://doi.org/10.1016/j.buildenv.2021.108668

Kallio, J., Tervonen, J., Räsänen, P., Mäkynen, R., Koivusaari, J., & Peltola, J. (2021). Forecasting office indoor CO2 concentration using machine learning with a one-year dataset. *Building and Environment*, *187*, 107409. https://doi.org/10.1016/j.buildenv.2020.107409

Kamath, U., & Liu, J. (2021). Explainable Artificial Intelligence: An Introduction to Interpretable Machine Learning. In *Explainable Artificial Intelligence: An Introduction to Interpretable Machine Learning*. https://doi.org/10.1007/978-3-030-83356-5

Kandasamy, N. K., Karunagaran, G., Spanos, C., Tseng, K. J., & Soong, B.-H. (2018). Smart lighting system using ANN-IMC for personalized lighting control and daylight harvesting. *Building and Environment*, *139*, 170–180. https://doi.org/10.1016/j.buildenv.2018.05.005

Karami, M., McMorrow, G. V., & Wang, L. (2018). Continuous monitoring of indoor environmental quality using an Arduino-based data acquisition system. *Journal of Building Engineering*, *19*, 412–419. https://doi.org/10.1016/j.jobe.2018.05.014

Kasuya, T., Takeshi, T., & Esaki, H. (2020). Building Activity Profiling: Explainable and Predictive Modeling for Building Automation. *2020 International Conference on Artificial Intelligence in Information and Communication, ICAIIC 2020*. https://doi.org/10.1109/ICAIIC48513.2020.9065268

Kathirgamanathan, A., De Rosa, M., Mangina, E., & Finn, D. P. (2021). Data-driven predictive control for unlocking building energy flexibility: A review. *Renewable and Sustainable Energy Reviews*, *135*, 110120. https://doi.org/10.1016/j.rser.2020.110120

Kazak, T. (2022). European Green Deal. *Yearbook of the Law Department*, *9*(10). https://doi.org/10.33919/yldnbu.20.9.12

Keoh, S. L., Kumar, S. S., & Tschofenig, H. (2014). Securing the Internet of Things: A Standardization Perspective. *IEEE Internet of Things Journal*, *1*(3), 265–275. https://doi.org/10.1109/JIOT.2014.2323395

Khaled N., P. B. S. A. (2020). *Digital Twin Development and Deployment on the Cloud. Developing Cloud-Friendly Dynamic Models Using Simulink®/SimscapeTM and Amazon AWS*. Elsevier.

Khamma, T. R., Zhang, Y., Guerrier, S., & Boubekri, M. (2020). Generalized additive models: An efficient method for short-term energy prediction in office buildings. *Energy*, *213*. https://doi.org/10.1016/j.energy.2020.118834

Khan, A., Sepasgozar, S., Liu, T., & Yu, R. (2021). Integration of BIM and Immersive Technologies for AEC: A Scientometric-SWOT Analysis and Critical Content Review. *Buildings*, *11*(3), 126. https://doi.org/10.3390/buildings11030126

Khodayar, M., Wang, J., & Manthouri, M. (2019). Interval Deep Generative Neural Network for Wind Speed Forecasting. *IEEE Transactions on Smart Grid*, *10*(4), 3974–3989. https://doi.org/10.1109/TSG.2018.2847223

Kim, E. (2020). Interpretable and Accurate Convolutional Neural Networks for Human Activity Recognition. *IEEE Transactions on Industrial Informatics*, *16*(11). https://doi.org/10.1109/TII.2020.2972628

Kim, J. Y., & Cho, S. B. (2019). Electric energy consumption prediction by deep learning with state explainable autoencoder. *Energies*, *12*(4). https://doi.org/10.3390/en12040739

Kim, W., Jeon, S. W., & Kim, Y. (2016). Model-based multi-objective optimal control of a VRF (variable refrigerant flow) combined system with DOAS (dedicated outdoor air system) using genetic algorithm under heating conditions. *Energy*, *107*, 196–204. https://doi.org/10.1016/j.energy.2016.03.139

Kome, M. L., Cuppens, F., Cuppens-Boulahia, N., & Frey, V. (2018). CoAP Enhancement for a Better IoT Centric Protocol: CoAP 2.0. *2018 Fifth International Conference on Internet of Things: Systems, Management and Security*, 139–146. https://doi.org/10.1109/IoTSMS.2018.8554494

Koufi, L., Younsi, Z., Cherif, Y., Naji, H., & El Ganaoui, M. (2017). A numerical study of indoor air quality in a ventilated room using different strategies of ventilation. *Mechanics & Industry*, *18*(2), 221. https://doi.org/10.1051/meca/2016043

Kritzinger, W., Karner, M., Traar, G., Henjes, J., & Sihn, W. (2018). Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, *51*(11), 1016–1022. https://doi.org/10.1016/j.ifacol.2018.08.474

Kroposki, B., Johnson, B., Zhang, Y., Gevorgian, V., Denholm, P., Hodge, B.-M., & Hannegan, B. (2017). Achieving a 100% Renewable Grid: Operating Electric Power Systems with Extremely High Levels of Variable Renewable Energy. *IEEE Power and Energy Magazine*, *15*(2), 61–73. https://doi.org/10.1109/MPE.2016.2637122

Kueppers, M., Paredes Pineda, S. N., Metzger, M., Huber, M., Paulus, S., Heger, H. J., & Niessen, S. (2021). Decarbonization pathways of worldwide energy systems – Definition and modeling of archetypes. *Applied Energy*, *285*, 116438. https://doi.org/10.1016/j.apenergy.2021.116438

Kwok, H. H. L., Cheng, J. C. P., Li, A. T. Y., Tong, J. C. K., & Lau, A. K. H. (2020). Multi-zone indoor CFD under limited information: An approach coupling solar analysis and BIM for improved accuracy. *Journal of Cleaner Production*, *244*, 118912. https://doi.org/10.1016/j.jclepro.2019.118912

Kwok, H. H. L., Cheng, J. C. P., Li, A. T. Y., Tong, J. C. K., & Lau, A. K. H. (2022). Impact of shaft design to thermal comfort and indoor air quality of floors using BIM technology. *Journal of Building Engineering*, *51*, 104326. https://doi.org/10.1016/j.jobe.2022.104326

Kydd, P. H. (2023). Electric Vehicles as Electric Energy Storage for a Zero Emission Grid. *EarthArXiv*.

Lamagna, M., Groppi, D., Nezhad, M. M., & Piras, G. (2021). A comprehensive review on digital twins for smart energy management system. *International Journal of Energy Production and Management*, *6*(4), 323–334. https://doi.org/10.2495/EQ-V6-N4-323-334

Laurinec, P., & Lucká, M. (2019). Interpretable multiple data streams clustering with clipped streams representation for the improvement of electricity consumption forecasting. *Data Mining and Knowledge Discovery*, *33*(2). https://doi.org/10.1007/s10618-018-0598-2

Le, L. T., Nguyen, H., Dou, J., & Zhou, J. (2019). A comparative study of PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN in estimating the heating load of buildings' energy efficiency for smart city planning. *Applied Sciences (Switzerland)*, *9*(13). https://doi.org/10.3390/app9132630

Leal Filho, W., Vidal, D. G., Chen, C., Petrova, M., Dinis, M. A. P., Yang, P., Rogers, S., Álvarez-Castañón, L., Djekic, I., Sharifi, A., & Neiva, S. (2022). An assessment of requirements in investments, new technologies, and infrastructures to achieve the SDGs. *Environmental Sciences Europe*, *34*(1). https://doi.org/10.1186/s12302-022-00629-9

Lee, E.-K., Gerla, M., Pau, G., Lee, U., & Lim, J.-H. (2016). Internet of Vehicles: From intelligent grid to autonomous cars and vehicular fogs. *International Journal of Distributed Sensor Networks*, *12*(9), 155014771666550. https://doi.org/10.1177/1550147716665500

Lee, J., Lapira, E., Bagheri, B., & Kao, H. (2013). Recent advances and trends in predictive manufacturing systems in big data environment. *Manufacturing Letters*, *1*(1), 38–41. https://doi.org/10.1016/j.mfglet.2013.09.005

Lehtola, V. V., Koeva, M., Elberink, S. O., Raposo, P., Virtanen, J.-P., Vahdatikhaki, F., & Borsci, S. (2022). Digital twin of a city: Review of technology serving city needs. *International Journal of Applied Earth Observation and Geoinformation*, *114*, 102915. https://doi.org/10.1016/j.jag.2022.102915

Li, A., Xiao, F., Zhang, C., & Fan, C. (2021). Attention-based interpretable neural network for building cooling load prediction. *Applied Energy*, *299*. https://doi.org/10.1016/j.apenergy.2021.117238

Li, C., Dong, Z., Ding, L., Petersen, H., Qiu, Z., Chen, G., & Prasad, D. (2022). Interpretable Memristive LSTM Network Design for Probabilistic Residential Load Forecasting. *IEEE Transactions on Circuits and Systems I: Regular Papers*, *69*(6). https://doi.org/10.1109/TCSI.2022.3155443

Li, D., Li, D., Li, C., Li, L., & Gao, L. (2019). A novel data-temporal attention network based strategy for fault diagnosis of chiller sensors. *Energy and Buildings*, *198*. https://doi.org/10.1016/j.enbuild.2019.06.034

Li, G., Li, F., Xu, C., & Fang, X. (2022). A spatial-temporal layer-wise relevance propagation method for improving interpretability and prediction accuracy of LSTM building energy prediction. *Energy and Buildings*, *271*. https://doi.org/10.1016/j.enbuild.2022.112317

Li, G., Yao, Q., Fan, C., Zhou, C., Wu, G., Zhou, Z., & Fang, X. (2021). An explainable one-dimensional convolutional neural networks based fault diagnosis method for building heating, ventilation and air conditioning systems. *Building and Environment*, *203*. https://doi.org/10.1016/j.buildenv.2021.108057

Li, J., Mattewal, S. K., Patel, S., & Biswas, P. (2020). Evaluation of Nine Low-cost-sensor-based Particulate Matter Monitors. *Aerosol and Air Quality Research*, *20*(2), 254–270. https://doi.org/10.4209/aaqr.2018.12.0485

Li, L., Yan, J., Yang, X., & Jin, Y. (2019). Learning interpretable deep state space model for probabilistic time series forecasting. *IJCAI International Joint Conference on Artificial Intelligence*, *2019-August*. https://doi.org/10.24963/ijcai.2019/402

Li, S., Xu, Y., Cai, J., Hu, D., & He, Q. (2021). Integrated environment-occupant-pathogen information modeling to assess and communicate room-level outbreak risks of infectious diseases. *Building and Environment*, *187*, 107394. https://doi.org/10.1016/j.buildenv.2020.107394

Li, X., Lu, W., Xue, F., Wu, L., Zhao, R., Lou, J., & Xu, J. (2022). Blockchain-Enabled IoT-BIM Platform for Supply Chain Management in Modular Construction. *Journal of Construction Engineering and Management*, *148*(2). https://doi.org/10.1061/(ASCE)CO.1943-7862.0002229

Li, Y., O'Neill, Z., Zhang, L., Chen, J., Im, P., & DeGraw, J. (2021). Grey-box modeling and application for building energy simulations - A critical review. *Renewable and Sustainable Energy Reviews*, *146*, 111174. https://doi.org/10.1016/j.rser.2021.111174

Li, Y., & Shen, X. (2022). A Novel Wind Speed-Sensing Methodology for Wind Turbines Based on Digital Twin Technology. *IEEE Transactions on Instrumentation and Measurement*, *71*, 1–13. https://doi.org/10.1109/TIM.2021.3139698

Liang, F., Hatcher, W. G., Liao, W., Gao, W., & Yu, W. (2019). Machine Learning for Security and the Internet of Things: The Good, the Bad, and the Ugly. *IEEE Access*, *7*, 158126–158147. https://doi.org/10.1109/ACCESS.2019.2948912

Liu, S., Bao, J., Lu, Y., Li, J., Lu, S., & Sun, X. (2021). Digital twin modeling method based on biomimicry for machining aerospace components. *Journal of Manufacturing Systems*, *58*, 180–195. https://doi.org/10.1016/j.jmsy.2020.04.014

Lu, Q., Parlikad, A. K., Woodall, P., Don Ranasinghe, G., Xie, X., Liang, Z., Konstantinou, E., Heaton, J., & Schooling, J. (2020). Developing a Digital Twin at Building and City Levels: Case Study of West Cambridge Campus. *Journal of Management in Engineering*, *36*(3). https://doi.org/10.1061/(ASCE)ME.1943-5479.0000763

Lu, Y., Murzakhanov, I., & Chatzivasileiadis, S. (2021). Neural network interpretability for forecasting of aggregated renewable generation. *2021 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids, SmartGridComm 2021*. https://doi.org/10.1109/SmartGridComm51999.2021.9631993

Lu, Y., & Xu, X. (2019). Cloud-based manufacturing equipment and big data analytics to enable on-demand manufacturing services. *Robotics and Computer-Integrated Manufacturing*, *57*, 92–102. https://doi.org/10.1016/j.rcim.2018.11.006

Ma, N., Aviv, D., Guo, H., & Braham, W. W. (2021). Measuring the right factors: A review of variables and models for thermal comfort and indoor air quality. *Renewable and Sustainable Energy Reviews*, *135*, 110436. https://doi.org/10.1016/j.rser.2020.110436

Mabkhot, M., Al-Ahmari, A., Salah, B., & Alkhalefah, H. (2018). Requirements of the Smart Factory System: A Survey and Perspective. *Machines*, *6*(2), 23. https://doi.org/10.3390/machines6020023

Maddalena, R., Mendell, M. J., Eliseeva, K., Chan, W. R., Sullivan, D. P., Russell, M., Satish, U., & Fisk, W. J. (2015). Effects of ventilation rate per person and per floor area on perceived air quality, sick building syndrome symptoms, and decision-making. *Indoor Air*, *25*(4), 362–370. https://doi.org/10.1111/ina.12149

Madhikermi, M., Malhi, A. K., & Främling, K. (2019). Explainable artificial intelligence based heat recycler fault detection in air handling unit. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11763 LNAI*. https://doi.org/10.1007/978-3-030-30391-4_7

Madni, A., Madni, C., & Lucero, S. (2019). Leveraging Digital Twin Technology in Model-Based Systems Engineering. *Systems*, *7*(1), 7. https://doi.org/10.3390/systems7010007

Mahmud, Md. S., Wang, H., Esfar-E-Alam, A. M., & Fang, H. (2017). A Wireless Health Monitoring System Using Mobile Phone Accessories. *IEEE Internet of Things Journal*, *4*(6), 2009–2018. https://doi.org/10.1109/JIOT.2016.2645125

Majd, E., McCormack, M., Davis, M., Curriero, F., Berman, J., Connolly, F., Leaf, P., Rule, A., Green, T., Clemons-Erby, D., Gummerson, C., & Koehler, K. (2019). Indoor air quality in inner-city schools and its associations with building characteristics and environmental factors. *Environmental Research*, *170*, 83–91. https://doi.org/10.1016/j.envres.2018.12.012

Marco Casini. (2022). *Construction 4.0*. Elsevier.

Marino, D. L., Amarasinghe, K., & Manic, M. (2016). Building energy load forecasting using Deep Neural Networks. *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*, 7046–7051. https://doi.org/10.1109/IECON.2016.7793413

Marinov, M. B., Djermanova, N., Ganev, B., Nikolov, G., & Janchevska, E. (2018). Performance Evaluation of Low-cost Carbon Dioxide Sensors. *2018 IEEE XXVII*

International Scientific Conference Electronics - ET, 1–4. https://doi.org/10.1109/ET.2018.8549621

Marotta, I., Guarino, F., Cellura, M., & Longo, S. (2021). Investigation of design strategies and quantification of energy flexibility in buildings: A case-study in southern Italy. *Journal of Building Engineering*, *41*. https://doi.org/10.1016/j.jobe.2021.102392

Marques, G., Pires, I. M., Miranda, N., & Pitarma, R. (2019). Air Quality Monitoring Using Assistive Robots for Ambient Assisted Living and Enhanced Living Environments through Internet of Things. *Electronics*, *8*(12), 1375. https://doi.org/10.3390/electronics8121375

Marques, G., & Pitarma, R. (2016). An Indoor Monitoring System for Ambient Assisted Living Based on Internet of Things Architecture. *International Journal of Environmental Research and Public Health*, *13*(11), 1152. https://doi.org/10.3390/ijerph13111152

Maschler, B., Braun, D., Jazdi, N., & Weyrich, M. (2021). Transfer learning as an enabler of the intelligent digital twin. *Procedia CIRP*, *100*, 127–132. https://doi.org/10.1016/j.procir.2021.05.020

McKinsey Global Institute. (2022). *Digital twins: The foundation of the enterprise metaverse*. https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-internet-of-things#/

Menegon, J., & Isatto, E. L. (2023). Digital twins as enablers of structure inspection and maintenance. *Gestão & Produção*, *30*. https://doi.org/10.1590/1806-9649-2022v30e4922

Michael Grieves. (2014). *Digital Twin: Manufacturing Excellence through Virtual Factory Replication*. https://www.researchgate.net/publication/275211047_Digital_Twin_Manufacturing_Excellence_through_Virtual_Factory_Replication

Militello-Hourigan, R. E., & Miller, S. L. (2018). The impacts of cooking and an assessment of indoor air quality in Colorado passive and tightly constructed homes. *Building and Environment*, *144*, 573–582. https://doi.org/10.1016/j.buildenv.2018.08.044

Miller, C. (2019). What's in the box?! Towards explainable machine learning applied to non-residential building smart meter classification. *Energy and Buildings*, *199*. https://doi.org/10.1016/j.enbuild.2019.07.019

Ministry of Economic Development. (2019). *Integrated national energy and climate plan*. European Commission. https://commission.europa.eu/energy-climate-change-environment/implementation-eu-countries/energy-and-climate-governance-and-reporting/national-energy-and-climate-plans_en

Mohammadi, N., & Taylor, J. E. (2017). Smart city digital twins. *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1–5. https://doi.org/10.1109/SSCI.2017.8285439

Mollo Neto, M., Silva Montanhani, M. E., Gobbo de Freitas Bueno, L., Harada, É. dos S., & Florentino Pereira, D. (2020). Data Mining Generating Decision Trees to Alert System Against Death and Losses in Egg Production. *International Journal for Innovation Education and Research*, *8*(8). https://doi.org/10.31686/ijier.vol8.iss8.2584

Montiel-Santiago, F. J., Hermoso-Orzáez, M. J., & Terrados-Cepeda, J. (2020). Sustainability and Energy Efficiency: BIM 6D. Study of the BIM Methodology Applied to Hospital Buildings. Value of Interior Lighting and Daylight in Energy Simulation. *Sustainability*, *12*(14), 5731. https://doi.org/10.3390/su12145731

Moon, J., Park, S., Rho, S., & Hwang, E. (2022a). Interpretable Short-Term Electrical Load Forecasting Scheme Using Cubist. *Computational Intelligence and Neuroscience*, *2022*. https://doi.org/10.1155/2022/6892995

Moon, J., Park, S., Rho, S., & Hwang, E. (2022b). Robust building energy consumption forecasting using an online learning approach with R ranger. *Journal of Building Engineering*, *47*. https://doi.org/10.1016/j.jobe.2021.103851

Morawska, L., Allen, J., Bahnfleth, W., Bluyssen, P. M., Boerstra, A., Buonanno, G., Cao, J., Dancer, S. J., Floto, A., Franchimon, F., Greenhalgh, T., Haworth, C., Hogeling, J., Isaxon, C., Jimenez, J. L., Kurnitski, J., Li, Y., Loomans, M., Marks, G., … Yao, M. (2021). A paradigm shift to combat indoor respiratory infection. *Science*, *372*(6543), 689–691. https://doi.org/10.1126/science.abg2025

Motamedi, H., Shirzadi, M., Tominaga, Y., & Mirzaei, P. A. (2022). CFD modeling of airborne pathogen transmission of COVID-19 in confined spaces under different ventilation strategies. *Sustainable Cities and Society*, *76*, 103397. https://doi.org/10.1016/j.scs.2021.103397

Movahedi, A., & Derrible, S. (2021). Interrelationships between electricity, gas, and water consumption in large-scale buildings. *Journal of Industrial Ecology*, *25*(4). https://doi.org/10.1111/jiec.13097

Murtagh, N., Scott, L., & Fan, J. (2020). Sustainable and resilient construction: Current status and future challenges. *Journal of Cleaner Production*, *268*, 122264. https://doi.org/10.1016/j.jclepro.2020.122264

Muthiah-Nakarajan, V., Cherukuri, S. H. C., Saravanan, B., & Palanisamy, K. (2021). Residential energy management strategy considering the usage of storage facilities and electric vehicles. *Sustainable Energy Technologies and Assessments*, *45*. https://doi.org/10.1016/j.seta.2021.101167

Mysen, M., Berntsen, S., Nafstad, P., & Schild, P. G. (2005). Occupancy density and benefits of demand-controlled ventilation in Norwegian primary schools. *Energy and Buildings*, *37*(12), 1234–1240. https://doi.org/10.1016/j.enbuild.2005.01.003

National Infrastructure Commission (UK). United Kingdom. (2017). *Data for the public good*. https://policycommons.net/artifacts/2475918/data-for-the-public-good-national-infrastructure-commission/3497975/

Nazaroff, W. W. (2013). Exploring the consequences of climate change for indoor air quality. *Environmental Research Letters*, *8*(1), 015022. https://doi.org/10.1088/1748-9326/8/1/015022

Negri, E., Fumagalli, L., & Macchi, M. (2017). A Review of the Roles of Digital Twin in CPS-based Production Systems. *Procedia Manufacturing*, *11*, 939–948. https://doi.org/10.1016/j.promfg.2017.07.198

Nirlo, E. L., Crain, N., Corsi, R. L., & Siegel, J. A. (2015). Field evaluation of five volatile organic compound measurement techniques: Implications for green building decision making. *Science and Technology for the Built Environment*, *21*(1), 67–79. https://doi.org/10.1080/10789669.2014.969172

OECD/IEA. (2019). Global Energy & CO2 Status Report 2019 – Analysis - IEA. In *Iea*.

Olesen B. W. (2012). Revision of EN 15251: Indoor Environmental Criteria. *REHVA Journal*.

Oliver, D. , A. D. H.-S. A. (2018, January). *"Living with a Digital Twin: Operational management and engagement using IoT and Mixed Realities at UCL's Here East Campus on the Queen Elizabeth Olympic Park." (2018).*

Othmane, L. Ben, Jacobson, D., & Weippl, E. (2019). Special Issue on Security and Privacy in Smart Cyber-physical Systems. *Computers & Security*, *87*, 101611. https://doi.org/10.1016/j.cose.2019.101611

Palensky, P., Cvetkovic, M., Gusain, D., & Joseph, A. (2022). Digital twins and their use in future power systems. *Digital Twin*, *1*, 4. https://doi.org/10.12688/digitaltwin.17435.2

Papadopoulos, S., & Kontokosta, C. E. (2019). Grading buildings on energy performance using city benchmarking data. *Applied Energy*, *233–234*. https://doi.org/10.1016/j.apenergy.2018.10.053

Park, H., & Park, D. Y. (2021). Comparative analysis on predictability of natural ventilation rate based on machine learning algorithms. *Building and Environment*, *195*. https://doi.org/10.1016/j.buildenv.2021.107744

Park, H.-A., Byeon, G., Son, W., Jo, H.-C., Kim, J., & Kim, S. (2020). Digital Twin for Operation of Microgrid: Optimal Scheduling in Virtual Space of Digital Twin. *Energies*, *13*(20), 5504. https://doi.org/10.3390/en13205504

Park, K. T., Nam, Y. W., Lee, H. S., Im, S. J., Noh, S. Do, Son, J. Y., & Kim, H. (2019). Design and implementation of a digital twin application for a connected micro smart factory. *International Journal of Computer Integrated Manufacturing*, *32*(6), 596–614. https://doi.org/10.1080/0951192X.2019.1599439

Park, S., Moon, J., & Hwang, E. (2020). Explainable Anomaly Detection for District Heating Based on Shapley Additive Explanations. *IEEE International Conference on Data Mining Workshops, ICDMW*, *2020-November*. https://doi.org/10.1109/ICDMW51313.2020.00111

Parra-Domínguez, J., Gil-Egido, A., & Rodríguez-González, S. (2022). SDGs as One of the Drivers of Smart City Development: The Indicator Selection Process. *Smart Cities*, *5*(3), 1025–1038. https://doi.org/10.3390/smartcities5030051

Pathak, N., Ba, A., Ploennigs, J., & Roy, N. (2018). Forecasting gas usage for big buildings using generalized additive models and deep learning. *Proceedings - 2018 IEEE International Conference on Smart Computing, SMARTCOMP 2018*. https://doi.org/10.1109/SMARTCOMP.2018.00092

Peng, S., Chen, Q., & Liu, E. (2020). The role of computational fluid dynamics tools on investigation of pathogen transmission: Prevention and control. *Science of The Total Environment*, *746*, 142090. https://doi.org/10.1016/j.scitotenv.2020.142090

Piasecki, M., & Kostyrko, K. (2020). Development of Weighting Scheme for Indoor Air Quality Model Using a Multi-Attribute Decision Making Method. *Energies*, *13*(12), 3120. https://doi.org/10.3390/en13123120

Ponnusamy, V. K., Kasinathan, P., Madurai Elavarasan, R., Ramanathan, V., Anandan, R. K., Subramaniam, U., Ghosh, A., & Hossain, E. (2021). A

Comprehensive Review on Sustainable Aspects of Big Data Analytics for the Smart Grid. *Sustainability*, *13*(23), 13322. https://doi.org/10.3390/su132313322

Qavidel Fard, Z., Zomorodian, Z. S., & Korsavi, S. S. (2022). Application of machine learning in thermal comfort studies: A review of methods, performance and challenges. In *Energy and Buildings* (Vol. 256). https://doi.org/10.1016/j.enbuild.2021.111771

Qi, Q., & Tao, F. (2018a). Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. *IEEE Access*, *6*, 3585–3593. https://doi.org/10.1109/ACCESS.2018.2793265

Qi, Q., & Tao, F. (2018b). Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. *IEEE Access*, *6*, 3585–3593. https://doi.org/10.1109/ACCESS.2018.2793265

Qian, C., Liu, X., Ripley, C., Qian, M., Liang, F., & Yu, W. (2022). Digital Twin—Cyber Replica of Physical Things: Architecture, Applications and Future Research Directions. *Future Internet*, *14*(2), 64. https://doi.org/10.3390/fi14020064

Ramesh, T., Prakash, R., & Shukla, K. K. (2010). Life cycle energy analysis of buildings: An overview. In *Energy and Buildings* (Vol. 42, Issue 10). https://doi.org/10.1016/j.enbuild.2010.05.007

Rani, H. A., Farouk, A. M., Anandh, K. S., Almutairi, S., & Rahman, R. A. (2022). Impact of COVID-19 on Construction Projects: The Case of India. *Buildings*, *12*(6), 762. https://doi.org/10.3390/buildings12060762

Rathore, M. M., Shah, S. A., Shukla, D., Bentafat, E., & Bakiras, S. (2021). The Role of AI, Machine Learning, and Big Data in Digital Twinning: A Systematic Literature Review, Challenges, and Opportunities. *IEEE Access*, *9*, 32030–32052. https://doi.org/10.1109/ACCESS.2021.3060863

Raval, M. ; A. A. ; D. J. ; G. H. (2021). Smart energy optimization for massive IoT using artificial intelligence. *Internet Things*, *13*(100354).

Reischauer, G. (2018). Industry 4.0 as policy-driven discourse to institutionalize innovation systems in manufacturing. *Technological Forecasting and Social Change*, *132*, 26–33. https://doi.org/10.1016/j.techfore.2018.02.012

Reka, S. S., & Dragicevic, T. (2018). Future effectual role of energy delivery: A comprehensive review of Internet of Things and smart grid. *Renewable and Sustainable Energy Reviews*, *91*, 90–108. https://doi.org/10.1016/j.rser.2018.03.089

Ren, B., Li, Q., Jia, Y., Zhou, Q., Wang, C., & Zou, X. (2023). A machine learning method for locating subsynchronous oscillation source of VSCs in wind farm induced by open-loop modal resonance based on measurement. *Frontiers in Energy Research*, *10*. https://doi.org/10.3389/fenrg.2022.1047624

Renewables REN21. Global status report. (2018). *Renewables REN21. Global status report.*

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). 'why should i trust you?' explaining the predictions of any classifier. *NAACL-HLT 2016 - 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Demonstrations Session*. https://doi.org/10.18653/v1/n16-3020

Ribéron, J. , D. M. L. M. , & K. S. (n.d.). Impact of airing behaviour on air stuffiness in schools and daycares centres: Development of a specific tool for ventilation management. *2 Th International Conference on Indoor Air Quality and Climate, Indoor Air.*

Rice, L. (2020). Healthy BIM: the feasibility of integrating architecture health indicators using a building information model (BIM) computer system. *Archnet-IJAR: International Journal of Architectural Research*, *15*(1), 252–265. https://doi.org/10.1108/ARCH-07-2020-0133

Roberto Baldoni, L. M. (2016). *Italian Cyber Security Report - Un Framework Nazionale per la Cybersecurity.*

Robnik-Šikonja, M., & Bohanec, M. (2018). *Perturbation-Based Explanations of Prediction Models*. https://doi.org/10.1007/978-3-319-90403-0_9

Rojas, G. , P. R. , Feist. W. (2016). Ventilation concepts for energy efficient housing in Central European climate—a simulation study comparing IAQ, mold risk and ventilation losses. *Proceedings of the 14th International Conference on Indoor Air Quality and Climate.*

Roman, R., Zhou, J., & Lopez, J. (2013). On the features and challenges of security and privacy in distributed internet of things. *Computer Networks*, *57*(10), 2266–2279. https://doi.org/10.1016/j.comnet.2012.12.018

Rosen, R., von Wichert, G., Lo, G., & Bettenhausen, K. D. (2015). About The Importance of Autonomy and Digital Twins for the Future of Manufacturing. *IFAC-PapersOnLine*, *48*(3), 567–572. https://doi.org/10.1016/j.ifacol.2015.06.141

Rudin, C., Chen, C., Chen, Z., Huang, H., Semenova, L., & Zhong, C. (2022). Interpretable machine learning: Fundamental principles and 10 grand challenges. *Statistics Surveys*, *16*. https://doi.org/10.1214/21-SS133

Ruusu, R., Cao, S., Manrique Delgado, B., & Hasan, A. (2019). Direct quantification of multiple-source energy flexibility in a residential building using a new model predictive high-level controller. *Energy Conversion and Management*, *180*, 1109–1128. https://doi.org/10.1016/j.enconman.2018.11.026

Saad, A., Faddel, S., Youssef, T., & Mohammed, O. A. (2020a). On the Implementation of IoT-Based Digital Twin for Networked Microgrids Resiliency Against Cyber Attacks. *IEEE Transactions on Smart Grid*, *11*(6), 5138–5150. https://doi.org/10.1109/TSG.2020.3000958

Saad, A., Faddel, S., Youssef, T., & Mohammed, O. A. (2020b). On the Implementation of IoT-Based Digital Twin for Networked Microgrids Resiliency Against Cyber Attacks. *IEEE Transactions on Smart Grid*, *11*(6), 5138–5150. https://doi.org/10.1109/TSG.2020.3000958

Saeed, Z., Mancini, F., Glusac, T., & Izadpanahi, P. (2022a). Future City, Digital Twinning and the Urban Realm: A Systematic Literature Review. *Buildings*, *12*(5), 685. https://doi.org/10.3390/buildings12050685

Saeed, Z., Mancini, F., Glusac, T., & Izadpanahi, P. (2022b). Future City, Digital Twinning and the Urban Realm: A Systematic Literature Review. *Buildings*, *12*(5), 685. https://doi.org/10.3390/buildings12050685

Saini, J., Dutta, M., & Marques, G. (2020a). A comprehensive review on indoor air quality monitoring systems for enhanced public health. *Sustainable Environment Research*, *30*(1), 6. https://doi.org/10.1186/s42834-020-0047-y

Saini, J., Dutta, M., & Marques, G. (2020b). Indoor Air Quality Monitoring Systems Based on Internet of Things: A Systematic Review. *International Journal of Environmental Research and Public Health*, *17*(14), 4942. https://doi.org/10.3390/ijerph17144942

Sakkas, N., Yfanti, S., Daskalakis, C., Barbu, E., & Domnich, M. (2021). Interpretable forecasting of energy demand in the residential sector. *Energies*, *14*(20). https://doi.org/10.3390/en14206568

Sapronov, A. , Belavin, V. , Arzymatov, K. , Karpov, M. , Nevolin, A. , & Ustyuzhanin, A. (2018). Tuning hybrid distributed storage system digital twins by reinforcement learning. *Advances in Systems Science and Applications* , *18(4)*, 1–12.

Sarbu, I., & Sebarchievici, C. (2013). Aspects of indoor environmental quality assessment in buildings. *Energy and Buildings*, *60*, 410–419. https://doi.org/10.1016/j.enbuild.2013.02.005

Sarker, I. H. (2021a). Machine Learning: Algorithms, Real-World Applications and Research Directions. In *SN Computer Science* (Vol. 2, Issue 3). https://doi.org/10.1007/s42979-021-00592-x

Sarker, I. H. (2021b). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, *2*(3), 160. https://doi.org/10.1007/s42979-021-00592-x

Sarp, S., Knzlu, M., Cali, U., Elma, O., & Guler, O. (2021). An interpretable solar photovoltaic power generation forecasting approach using an explainable artificial intelligence tool. *2021 IEEE Power and Energy Society Innovative Smart Grid Technologies Conference, ISGT 2021*. https://doi.org/10.1109/ISGT49243.2021.9372263

Scannapieco, S., & Tomazzoli, C. (2018). *Ubiquitous and Pervasive Computing for Real-Time Energy Management and Saving* (pp. 3–15). https://doi.org/10.1007/978-3-319-61542-4_1

Schleich, B., Anwer, N., Mathieu, L., & Wartzack, S. (2017). Shaping the digital twin for design and production engineering. *CIRP Annals*, *66*(1), 141–144. https://doi.org/10.1016/j.cirp.2017.04.040

Schluse, M., Atorf, L., & Rossmann, J. (2017). Experimentable digital twins for model-based systems engineering and simulation-based development. *2017 Annual IEEE International Systems Conference (SysCon)*, 1–8. https://doi.org/10.1109/SYSCON.2017.7934796

Schroeder, G. N., Steinmetz, C., Pereira, C. E., & Espindola, D. B. (2016). Digital Twin Data Modeling with AutomationML and a Communication Methodology for Data Exchange. *IFAC-PapersOnLine*, *49*(30), 12–17. https://doi.org/10.1016/j.ifacol.2016.11.115

Scutaru, A. M., & Witterseh, T. (2020). Risk mitigation for indoor air quality using the example of construction products – efforts towards a harmonization of the health-related evaluation in the EU. *International Journal of Hygiene and Environmental Health*, *229*, 113588. https://doi.org/10.1016/j.ijheh.2020.113588

Semeraro, C., Lezoche, M., Panetto, H., & Dassisti, M. (2021). Digital twin paradigm: A systematic literature review. *Computers in Industry*, *130*, 103469. https://doi.org/10.1016/j.compind.2021.103469

Sengupta, P. P., & Chandrashekhar, Y. S. (2021). Building Trust in AI. *JACC: Cardiovascular Imaging*, *14*(2). https://doi.org/10.1016/j.jcmg.2021.01.002

Sepasgozar, S. M. E. (2020). Digital Twin and Web-Based Virtual Gaming Technologies for Online Education: A Case of Construction Management and Engineering. *Applied Sciences*, *10*(13), 4678. https://doi.org/10.3390/app10134678

Sha, H., Xu, P., Lin, M., Peng, C., & Dou, Q. (2021). Development of a multi-granularity energy forecasting toolkit for demand response baseline calculation. *Applied Energy*, *289*. https://doi.org/10.1016/j.apenergy.2021.116652

Shafto M., Conroy M., Doyle R., Glaessgen E, Kemp C, LeMoigne J, & Wang L. (2010). *Modeling, Simulation, Information Technology & Processing Roadmap*.

Shahat, E., Hyun, C. T., & Yeom, C. (2021). City Digital Twin Potentials: A Review and Research Agenda. *Sustainability*, *13*(6), 3386. https://doi.org/10.3390/su13063386

Shao, G., Jain, S., Laroque, C., Lee, L. H., Lendermann, P., & Rose, O. (2019a). Digital Twin for Smart Manufacturing: The Simulation Aspect. *2019 Winter Simulation Conference (WSC)*, 2085–2098. https://doi.org/10.1109/WSC40007.2019.9004659

Shao, G., Jain, S., Laroque, C., Lee, L. H., Lendermann, P., & Rose, O. (2019b). Digital Twin for Smart Manufacturing: The Simulation Aspect. *2019 Winter Simulation Conference (WSC)*, 2085–2098. https://doi.org/10.1109/WSC40007.2019.9004659

Sharma, A., Kosasih, E., Zhang, J., Brintrup, A., & Calinescu, A. (2020). *Digital Twins: State of the Art Theory and Practice, Challenges, and Open Research Questions*.

Shi, M. H. , & Tao, A. R. (2000). Investigation on the comfort evaluation index CPD of indoor environments. *Int. J. Architec.*

Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge Computing: Vision and Challenges. *IEEE Internet of Things Journal*, *3*(5), 637–646. https://doi.org/10.1109/JIOT.2016.2579198

Shubenkova, K., Valiev, A., Shepelev, V., Tsiulin, S., & Reinau, K. H. (2018). Possibility of Digital Twins Technology for Improving Efficiency of the Branded Service System. *2018 Global Smart Industry Conference (GloSIC)*, 1–7. https://doi.org/10.1109/GloSIC.2018.8570075

Silva, B. N., Khan, M., & Han, K. (2020). Integration of Big Data analytics embedded smart city architecture with RESTful web of things for efficient service provision and energy management. *Future Generation Computer Systems*, *107*, 975–987. https://doi.org/10.1016/j.future.2017.06.024

Silva, D., Carvalho, L. I., Soares, J., & Sofia, R. C. (2021a). A Performance Analysis of Internet of Things Networking Protocols: Evaluating MQTT, CoAP, OPC UA. *Applied Sciences*, *11*(11), 4879. https://doi.org/10.3390/app11114879

Silva, D., Carvalho, L. I., Soares, J., & Sofia, R. C. (2021b). A Performance Analysis of Internet of Things Networking Protocols: Evaluating MQTT, CoAP, OPC UA. *Applied Sciences*, *11*(11), 4879. https://doi.org/10.3390/app11114879

Singer, B. C., & Delp, W. W. (2018). Response of consumer and research grade indoor air quality monitors to residential sources of fine particles. *Indoor Air*, *28*(4), 624–639. https://doi.org/10.1111/ina.12463

Sipple, J. (2020). Interpretable, multidimensional, multimodal anomaly detection with negative sampling for detection of device failure. *37th International Conference on Machine Learning, ICML 2020*, *PartF168147-12*.

Sivalingam, K., Sepulveda, M., Spring, M., & Davies, P. (2018). A Review and Methodology Development for Remaining Useful Life Prediction of Offshore Fixed and Floating Wind turbine Power Converter with Digital Twin Technology Perspective. *2018 2nd International Conference on Green Energy and Applications (ICGEA)*, 197–204. https://doi.org/10.1109/ICGEA.2018.8356292

Smit, J. ; K. S. ; M. C. ; C. M. (2016). *Industry 4.0*. https://www.europarl.europa.eu/RegData/etudes/STUD/2016/570007/IPOL_STU(2016)570007_EN.pdf

Society's Indoor Carbon Dioxide Position Document Committee. (2022). *ASHRAE Position Document on Indoor Carbon Dioxide*.

Song, X., Cai, H., Kircheis, J., Jiang, T., Schlegel, S., & Westermann, D. (2020). Application of Digital Twin Assistant-System in State Estimation for Inverter Dominated Grid. *2020 55th International Universities Power Engineering Conference (UPEC)*, 1–6. https://doi.org/10.1109/UPEC49904.2020.9209876

Song, X., Jiang, T., Schlegel, S., & Westermann, D. (2020). Parameter tuning for dynamic digital twins in inverter-dominated distribution grid. *IET Renewable Power Generation*, *14*(5), 811–821. https://doi.org/10.1049/iet-rpg.2019.0163

Sporr, A., Zucker, G., & Hofmann, R. (2019). Automated HVAC Control Creation Based on Building Information Modeling (BIM): Ventilation System. *IEEE Access*, *7*, 74747–74758. https://doi.org/10.1109/ACCESS.2019.2919262

Srinivasan, R. S., Manohar, B., & Issa, R. R. A. (2020). Urban Building Energy CPS (UBE-CPS): Real-Time Demand Response Using Digital Twin. In *Cyber-Physical Systems in the Built Environment* (pp. 309–322). Springer International Publishing. https://doi.org/10.1007/978-3-030-41560-0_17

Srinivasan, S., Arjunan, P., Jin, B., Sangiovanni-Vincentelli, A. L., Sultan, Z., & Poolla, K. (2021). Explainable AI for Chiller Fault-Detection Systems: Gaining Human Trust. *Computer*, *54*(10). https://doi.org/10.1109/MC.2021.3071551

Stankovic, J. A. (2014). Research Directions for the Internet of Things. *IEEE Internet of Things Journal*, *1*(1), 3–9. https://doi.org/10.1109/JIOT.2014.2312291

Stark, R., Fresemann, C., & Lindow, K. (2019). Development and operation of Digital Twins for technical systems and services. *CIRP Annals*, *68*(1), 129–132. https://doi.org/10.1016/j.cirp.2019.04.024

Steindl, G., & Kastner, W. (2021). Semantic Microservice Framework for Digital Twins. *Applied Sciences*, *11*(12), 5633. https://doi.org/10.3390/app11125633

Stephen Lopez, N., Kay Galeos, S., Raphael Calderon, B., Roy Dominguez, D., Joseph Uy, B., & Iyengar, R. (2021). Computational Fluid Dynamics Simulation of Indoor Air Quality and Thermal Stratification of an Underfloor Air Distribution System (UFAD) with Various Vent Layouts. *Fluid Dynamics & Materials Processing*, *17*(2), 333–347. https://doi.org/10.32604/fdmp.2021.011213

Stouffer, K. (2023). *Guide to Operational Technology (OT) Security*. https://doi.org/10.6028/NIST.SP.800-82r3

Šujanová, P., Rychtáriková, M., Sotto Mayor, T., & Hyder, A. (2019). A Healthy, Energy-Efficient and Comfortable Indoor Environment, a Review. *Energies*, *12*(8), 1414. https://doi.org/10.3390/en12081414

Sundararajan, A., & Ollis, B. (2021). Regression and generalized additive model to enhance the performance of photovoltaic power ensemble predictors. *IEEE Access*, *9*. https://doi.org/10.1109/ACCESS.2021.3103126

Tai, C. T., Sue, K. L., & Hu, Y. H. (2020). Machine learning in high-alert medication treatment: A study on the cardiovascular drug. *Applied Sciences (Switzerland)*, *10*(17). https://doi.org/10.3390/app10175798

Talkhestani, B. A., Jazdi, N., Schloegl, W., & Weyrich, M. (2018). Consistency check to synchronize the Digital Twin of manufacturing automation based on anchor points. *Procedia CIRP*, *72*, 159–164. https://doi.org/10.1016/j.procir.2018.03.166

Tan, J., Sha, X., Dai, B., & Lu, T. (2020). Wireless Technology and Protocol for IIoT and Digital Twins. *2020 ITU Kaleidoscope: Industry-Driven Digital Transformation (ITU K)*, 1–8. https://doi.org/10.23919/ITUK50268.2020.9303189

Tang, H., Wang, S., & Li, H. (2021). Flexibility categorization, sources, capabilities and technologies for energy-flexible and grid-responsive buildings: State-of-the-art and future perspective. In *Energy* (Vol. 219). https://doi.org/10.1016/j.energy.2020.119598

Tang, S., Shelden, D. R., Eastman, C. M., Pishdad-Bozorgi, P., & Gao, X. (2019a). A review of building information modeling (BIM) and the internet of things (IoT) devices integration: Present status and future trends. *Automation in Construction*, *101*, 127–139. https://doi.org/10.1016/j.autcon.2019.01.020

Tang, S., Shelden, D. R., Eastman, C. M., Pishdad-Bozorgi, P., & Gao, X. (2019b). A review of building information modeling (BIM) and the internet of things (IoT) devices integration: Present status and future trends. *Automation in Construction*, *101*, 127–139. https://doi.org/10.1016/j.autcon.2019.01.020

Tao F, Z. M. C. J. et al. (2017). Digital twin workshop: a new paradigm for future workshop. . *Comput Integr Manuf Syst 23(1):1–9*.

Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., & Sui, F. (2018). Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology*, *94*(9–12), 3563–3576. https://doi.org/10.1007/s00170-017-0233-1

Tao, F., Sui, F., Liu, A., Qi, Q., Zhang, M., Song, B., Guo, Z., Lu, S. C.-Y., & Nee, A. Y. C. (2019). Digital twin-driven product design framework. *International Journal of Production Research*, *57*(12), 3935–3953. https://doi.org/10.1080/00207543.2018.1443229

Tao, F., Zhang, M., Liu, Y., & Nee, A. Y. C. (2018). Digital twin driven prognostics and health management for complex equipment. *CIRP Annals*, *67*(1), 169–172. https://doi.org/10.1016/j.cirp.2018.04.055

Terkaj, W., Gaboardi, P., Trevisan, C., Tolio, T., & Urgo, M. (2019). A digital factory platform for the design of roll shop plants. *CIRP Journal of Manufacturing Science and Technology*, *26*, 88–93. https://doi.org/10.1016/j.cirpj.2019.04.007

Tomazzoli, C., Scannapieco, S., & Cristani, M. (2023a). Internet of Things and artificial intelligence enable energy efficiency. *Journal of Ambient Intelligence and*

Humanized Computing, *14*(5), 4933–4954. https://doi.org/10.1007/s12652-020-02151-3

Tomazzoli, C., Scannapieco, S., & Cristani, M. (2023b). Internet of Things and artificial intelligence enable energy efficiency. *Journal of Ambient Intelligence and Humanized Computing*, *14*(5), 4933–4954. https://doi.org/10.1007/s12652-020-02151-3

Tomko, M., & Winter, S. (2019). Beyond digital twins – A commentary. *Environment and Planning B: Urban Analytics and City Science*, *46*(2), 395–399. https://doi.org/10.1177/2399808318816992

Tong, C., Li, J., Lang, C., Kong, F., Niu, J., & Rodrigues, J. J. P. C. (2018). An efficient deep model for day-ahead electricity load forecasting with stacked denoising auto-encoders. *Journal of Parallel and Distributed Computing*, *117*, 267–273. https://doi.org/10.1016/j.jpdc.2017.06.007

Trauer, J., Schweigert-Recksiek, S., Engel, C., Spreitzer, K., & Zimmermann, M. (2020). WHAT IS A DIGITAL TWIN? – DEFINITIONS AND INSIGHTS FROM AN INDUSTRIAL CASE STUDY IN TECHNICAL PRODUCT DEVELOPMENT. *Proceedings of the Design Society: DESIGN Conference*, *1*, 757–766. https://doi.org/10.1017/dsd.2020.15

Tsoutsanis, E., & Meskin, N. (2019). Dynamic performance simulation and control of gas turbines used for hybrid gas/wind energy applications. *Applied Thermal Engineering*, *147*, 122–142. https://doi.org/10.1016/j.applthermaleng.2018.09.031

Tuegel, E. J., Ingraffea, A. R., Eason, T. G., & Spottswood, S. M. (2011a). Reengineering Aircraft Structural Life Prediction Using a Digital Twin. *International Journal of Aerospace Engineering*, *2011*, 1–14. https://doi.org/10.1155/2011/154798

Tuegel, E. J., Ingraffea, A. R., Eason, T. G., & Spottswood, S. M. (2011b). Reengineering Aircraft Structural Life Prediction Using a Digital Twin. *International Journal of Aerospace Engineering*, *2011*, 1–14. https://doi.org/10.1155/2011/154798

Tzanis, N., Andriopoulos, N., Magklaras, A., Mylonas, E., Birbas, M., & Birbas, A. (2020a). A Hybrid Cyber Physical Digital Twin Approach for Smart Grid Fault Prediction. *2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS)*, 393–397. https://doi.org/10.1109/ICPS48405.2020.9274723

Tzanis, N., Andriopoulos, N., Magklaras, A., Mylonas, E., Birbas, M., & Birbas, A. (2020b). A Hybrid Cyber Physical Digital Twin Approach for Smart Grid Fault Prediction. *2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS)*, 393–397. https://doi.org/10.1109/ICPS48405.2020.9274723

Tzanis, N., Andriopoulos, N., Magklaras, A., Mylonas, E., Birbas, M., & Birbas, A. (2020c). A Hybrid Cyber Physical Digital Twin Approach for Smart Grid Fault Prediction. *2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS)*, 393–397. https://doi.org/10.1109/ICPS48405.2020.9274723

Tzanis, N., Andriopoulos, N., Magklaras, A., Mylonas, E., Birbas, M., & Birbas, A. (2020d). A Hybrid Cyber Physical Digital Twin Approach for Smart Grid Fault Prediction. *2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS)*, 393–397. https://doi.org/10.1109/ICPS48405.2020.9274723

Tzanis, N., Andriopoulos, N., Magklaras, A., Mylonas, E., Birbas, M., & Birbas, A. (2020e). A Hybrid Cyber Physical Digital Twin Approach for Smart Grid Fault

Prediction. *2020 IEEE Conference on Industrial Cyberphysical Systems (ICPS)*, 393–397. https://doi.org/10.1109/ICPS48405.2020.9274723

Ugwuanyi, C. (2021). . *Using interpretable machine learning for indoor CO 2 level prediction and occupancy estimation* [Doctoral Thesis]. University of Strathclyde.

Uhlemann, T. H.-J., Lehmann, C., & Steinhilper, R. (2017). The Digital Twin: Realizing the Cyber-Physical Production System for Industry 4.0. *Procedia CIRP*, *61*, 335–340. https://doi.org/10.1016/j.procir.2016.11.152

UK Department for Environment Food & Rural Affairs. (2010). *UK Air Quality Limits*.

United Nations. Department of Economic and Social Affairs. (2018). *68% of the world population projected to live in urban areas by 2050, says UN*. https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html

Utkucu, D., & Sözer, H. (2020). Interoperability and data exchange within BIM platform to evaluate building energy performance and indoor comfort. *Automation in Construction*, *116*, 103225. https://doi.org/10.1016/j.autcon.2020.103225

Vachalek, J., Bartalsky, L., Rovny, O., Sismisova, D., Morhac, M., & Loksik, M. (2017). The digital twin of an industrial production line within the industry 4.0 concept. *2017 21st International Conference on Process Control (PC)*, 258–262. https://doi.org/10.1109/PC.2017.7976223

Valinejadshoubi, M., Moselhi, O., Bagchi, A., & Salem, A. (2021). Development of an IoT and BIM-based automated alert system for thermal comfort monitoring in buildings. *Sustainable Cities and Society*, *66*, 102602. https://doi.org/10.1016/j.scs.2020.102602

Valk, H. V. , H. H. , M. F. , A. M. , H. J. , & O. B. (2020). A. T. of D. Twins. A. C. on I. Systems. (2020). A Taxonomy of Digital Twins. *Americas Conference on Information Systems*.

Vardoulakis, S., Dimitroulopoulou, C., Thornes, J., Lai, K.-M., Taylor, J., Myers, I., Heaviside, C., Mavrogianni, A., Shrubsole, C., Chalabi, Z., Davies, M., & Wilkinson, P. (2015). Impact of climate change on the domestic indoor environment and associated health risks in the UK. *Environment International*, *85*, 299–313. https://doi.org/10.1016/j.envint.2015.09.010

Vigneau, E. (2021). Clustering of Variables for Enhanced Interpretability of Predictive Models. *Informatica (Slovenia)*, *45*(4). https://doi.org/10.31449/inf.v45i4.3283

Vilone, G., & Longo, L. (2021). A Quantitative Evaluation of Global, Rule-Based Explanations of Post-Hoc, Model Agnostic Methods. *Frontiers in Artificial Intelligence*, *4*. https://doi.org/10.3389/frai.2021.717899

Vivi, Q. L., Parlikad, A. K., Woodall, P., Ranasinghe, G. D., & Heaton, J. (2019). Developing a dynamic digital twin at a building level: Using Cambridge campus as case study. *ICE Publishing*.

Volk, R., Stengel, J., & Schultmann, F. (2014). Building Information Modeling (BIM) for existing buildings — Literature review and future needs. *Automation in Construction*, *38*, 109–127. https://doi.org/10.1016/j.autcon.2013.10.023

Voss, M., Heinekamp, J. F., Krutzsch, S., Sick, F., Albayrak, S., & Strunz, K. (2021). Generalized Additive Modeling of Building Inertia Thermal Energy Storage for

Integration into Smart Grid Control. *IEEE Access*, *9*. https://doi.org/10.1109/ACCESS.2021.3078802

Vrabič, R., Erkoyuncu, J. A., Butala, P., & Roy, R. (2018). Digital twins: Understanding the added value of integrated models for through-life engineering services. *Procedia Manufacturing*, *16*, 139–146. https://doi.org/10.1016/j.promfg.2018.10.167

Wang, H., Cai, R., Zhou, B., Aziz, S., Qin, B., Voropai, N., Gan, L., & Barakhtenko, E. (2020). Solar irradiance forecasting based on direct explainable neural network. *Energy Conversion and Management*, *226*. https://doi.org/10.1016/j.enconman.2020.113487

Wang, J., Ye, L., Gao, R. X., Li, C., & Zhang, L. (2019). Digital Twin for rotating machinery fault diagnosis in smart manufacturing. *International Journal of Production Research*, *57*(12), 3920–3934. https://doi.org/10.1080/00207543.2018.1552032

Wang, M., Wang, Z., Geng, Y., & Lin, B. (2022). Interpreting the neural network model for HVAC system energy data mining. *Building and Environment*, *209*. https://doi.org/10.1016/j.buildenv.2021.108449

Wang, M., Zhang, G., Zhang, C., Zhang, J., & Li, C. (2013). An IoT-based appliance control system for smart homes. *2013 Fourth International Conference on Intelligent Control and Information Processing (ICICIP)*, 744–747. https://doi.org/10.1109/ICICIP.2013.6568171

Wang, P., & Luo, M. (2021). A digital twin-based big data virtual and real fusion learning reference framework supported by industrial internet towards smart manufacturing. *Journal of Manufacturing Systems*, *58*, 16–32. https://doi.org/10.1016/j.jmsy.2020.11.012

Wang, W., Shan, X., Hussain, S. A., Wang, C., & Ji, Y. (2020). Comparison of Multi-Control Strategies for the Control of Indoor Air Temperature and CO2 with OpenModelica Modeling. *Energies*, *13*(17), 4425. https://doi.org/10.3390/en13174425

Wang, W., Wang, J., Chen, J., Huang, G., & Guo, X. (2018). Multi-zone outdoor air coordination through Wi-Fi probe-based occupancy sensing. *Energy and Buildings*, *159*, 495–507. https://doi.org/10.1016/j.enbuild.2017.11.041

Wang, Y., Wang, S., Yang, B., Zhu, L., & Liu, F. (2020). Big data driven Hierarchical Digital Twin Predictive Remanufacturing paradigm: Architecture, control mechanism, application scenario and benefits. *Journal of Cleaner Production*, *248*, 119299. https://doi.org/10.1016/j.jclepro.2019.119299

Wang, Z., Delp, W. W., & Singer, B. C. (2020). Performance of low-cost indoor air quality monitors for PM2.5 and PM10 from residential sources. *Building and Environment*, *171*, 106654. https://doi.org/10.1016/j.buildenv.2020.106654

Wärmefjord, K., Söderberg, R., Lindkvist, L., Lindau, B., & Carlson, J. S. (2017, November 3). Inspection Data to Support a Digital Twin for Geometry Assurance. *Volume 2: Advanced Manufacturing*. https://doi.org/10.1115/IMECE2017-70398

Wastensteiner, J., Michael Weiss, T., Haag, F., & Hopf, K. (2021). Explainable AI for tailored electricity consumption feedback–An experimental evaluation of visualizations experimental evaluation of visualizations. *ECIS 2021 Research Papers*, *Ecis*.

Wen, L., Zhou, K., & Yang, S. (2020). Load demand forecasting of residential buildings using a deep learning model. *Electric Power Systems Research*, *179*, 106073. https://doi.org/10.1016/j.epsr.2019.106073

Westermann, P., Deb, C., Schlueter, A., & Evins, R. (2020). Unsupervised learning of energy signatures to identify the heating system and building type using smart meter data. *Applied Energy*, *264*, 114715. https://doi.org/10.1016/j.apenergy.2020.114715

White, G., Zink, A., Codecá, L., & Clarke, S. (2021). A digital twin smart city for citizen feedback. *Cities*, *110*, 103064. https://doi.org/10.1016/j.cities.2020.103064

Wong, K., & Fan, Q. (2013). Building information modelling (BIM) for sustainable building design. *Facilities*, *31*(3/4), 138–157. https://doi.org/10.1108/02632771311299412

World Health Organization. Regional Office for Europe. (2010). *WHO guidelines for indoor air quality: selected pollutants*.

Wu, D., Shi, H., Wang, H., Wang, R., & Fang, H. (2019a). A Feature-Based Learning System for Internet of Things Applications. *IEEE Internet of Things Journal*, *6*(2), 1928–1937. https://doi.org/10.1109/JIOT.2018.2884485

Wu, D., Shi, H., Wang, H., Wang, R., & Fang, H. (2019b). A Feature-Based Learning System for Internet of Things Applications. *IEEE Internet of Things Journal*, *6*(2), 1928–1937. https://doi.org/10.1109/JIOT.2018.2884485

Wu, W., Chen, K., Qiao, Y., & Lu, Z. (2016). Probabilistic short-term wind power forecasting based on deep neural networks. *2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, 1–8. https://doi.org/10.1109/PMAPS.2016.7764155

Xiao, T., Xu, P., Ding, R., & Chen, Z. (2022). An interpretable method for identifying mislabeled commercial building based on temporal feature extraction and ensemble classifier. *Sustainable Cities and Society*, *78*. https://doi.org/10.1016/j.scs.2021.103635

Xiao, Z., & Xiao, Y. (2013). Security and Privacy in Cloud Computing. *IEEE Communications Surveys & Tutorials*, *15*(2), 843–859.

Xie, Y., Ishida, Y., Hu, J., & Mochida, A. (2022). A backpropagation neural network improved by a genetic algorithm for predicting the mean radiant temperature around buildings within the long-term period of the near future. *Building Simulation*, *15*(3). https://doi.org/10.1007/s12273-021-0823-6

Xu, C., Behrens, P., Gasper, P., Smith, K., Hu, M., Tukker, A., & Steubing, B. (2023). Electric vehicle batteries alone could satisfy short-term grid storage demand by as early as 2030. *Nature Communications*, *14*(1), 119. https://doi.org/10.1038/s41467-022-35393-0

Xu, L. Da, He, W., & Li, S. (2014). Internet of Things in Industries: A Survey. *IEEE Transactions on Industrial Informatics*, *10*(4), 2233–2243. https://doi.org/10.1109/TII.2014.2300753

Xu, G., Yu, W., Griffith, D., Golmie, N., & Moulema, P. (2016). Towards Integrating Distributed Energy Resources and Storage Devices in Smart Grid. *IEEE Internet of Things Journal*, 1–1. https://doi.org/10.1109/JIOT.2016.2640563

Xu, H., Yu, W., Liu, X., Griffith, D., & Golmie, N. (2020). On Data Integrity Attacks against Industrial Internet of Things. *2020 IEEE Intl Conf on Dependable,* *Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*, 21–28. https://doi.org/10.1109/DASC-PICom-CBDCom-CyberSciTech49142.2020.00020

XU, Y., CAI, J., LI, S., HE, Q., & ZHU, S. (2021). Airborne infection risks of SARS-CoV-2 in U.S. schools and impacts of different intervention strategies. *Sustainable Cities and Society*, *74*, 103188. https://doi.org/10.1016/j.scs.2021.103188

Xu, Y., Sun, Y., Liu, X., & Zheng, Y. (2019). A Digital-Twin-Assisted Fault Diagnosis Using Deep Transfer Learning. *IEEE Access*, *7*, 19990–19999. https://doi.org/10.1109/ACCESS.2018.2890566

Yan, K., Zhong, C., Ji, Z., & Huang, J. (2018). Semi-supervised learning for early detection and diagnosis of various air handling unit faults. *Energy and Buildings*, *181*. https://doi.org/10.1016/j.enbuild.2018.10.016

Yang, H., Cheng, L., & Chuah, M. C. (2019). Deep-Learning-Based Network Intrusion Detection for SCADA Systems. *2019 IEEE Conference on Communications and Network Security (CNS)*, 1–7. https://doi.org/10.1109/CNS.2019.8802785

Yang, K., Zhang, B., Zhang, J., & Zhu, J. (2021). Design of Remote Control Inverter Based on MQTT Communication Protocol. *2021 IEEE International Conference on Mechatronics and Automation (ICMA)*, 1374–1378. https://doi.org/10.1109/ICMA52036.2021.9512665

Yang, Y., Yuan, Y., Han, Z., & Liu, G. (2022). Interpretability analysis for thermal sensation machine learning models: An exploration based on the SHAP approach. *Indoor Air*, *32*(2). https://doi.org/10.1111/ina.12984

Yau, Y. H., & Chew, B. T. (2014). A review on predicted mean vote and adaptive thermal comfort models. In *Building Services Engineering Research and Technology* (Vol. 35, Issue 1). https://doi.org/10.1177/0143624412465200

Yu, C., Ren, G., & Dong, Y. (2020). Supervised-actor-critic reinforcement learning for intelligent mechanical ventilation and sedative dosing in intensive care units. *BMC Medical Informatics and Decision Making*, *20*. https://doi.org/10.1186/s12911-020-1120-5

Yu, M. G., & Pavlak, G. S. (2022). Extracting interpretable building control rules from multi-objective model predictive control data sets. *Energy*, *240*. https://doi.org/10.1016/j.energy.2021.122691

Yu, Z., Yang, X., Gao, F., Huang, J., Tu, R., & Cui, J. (2020). A Knowledge-based reinforcement learning control approach using deep Q network for cooling tower in HVAC systems. *Proceedings - 2020 Chinese Automation Congress, CAC 2020*. https://doi.org/10.1109/CAC51589.2020.9327385

Yusuf Akbar, I., Faisal, M., & Pagalay, U. (2022). Analisis Perbandingan Metode K-Nn Dan Decision Tree Dalam Klasifikasi Kenyamanan Thermal Bangunan. *Jurnal Syntax Fusion*, *2*(06). https://doi.org/10.54543/fusion.v2i06.197

Zarco-Periñán, P. J., Zarco-Soto, F. J., Zarco-Soto, I. M., Martínez-Ramos, J. L., & Sánchez-Durán, R. (2022). CO2 Emissions in Buildings: A Synopsis of Current Studies. In *Energies* (Vol. 15, Issue 18). https://doi.org/10.3390/en15186635

Zeng, L., Gao, J., Lv, L., Zhang, R., Tong, L., Zhang, X., Huang, Z., & Zhang, Z. (2020). Markov-chain-based probabilistic approach to optimize sensor network against deliberately released pollutants in buildings with ventilation systems. *Building and Environment*, *168*, 106534. https://doi.org/10.1016/j.buildenv.2019.106534

Zhang, C., Li, J., Zhao, Y., Li, T., Chen, Q., & Zhang, X. (2020). A hybrid deep learning-based method for short-term building energy load prediction combined with an interpretation process. *Energy and Buildings*, *225*. https://doi.org/10.1016/j.enbuild.2020.110301

Zhang, J., Seet, B.-C., & Lie, T. (2015). Building Information Modelling for Smart Built Environments. *Buildings*, *5*(1), 100–115. https://doi.org/10.3390/buildings5010100

Zhang, K., Blum, D., Cheng, H., Paliaga, G., Wetter, M., & Granderson, J. (2022). Estimating ASHRAE Guideline 36 energy savings for multi-zone variable air volume systems using Spawn of EnergyPlus. *Journal of Building Performance Simulation*, *15*(2), 215–236. https://doi.org/10.1080/19401493.2021.2021286

Zhang, L., Wen, J., Li, Y., Chen, J., Ye, Y., Fu, Y., & Livingood, W. (2021). A review of machine learning in building load prediction. *Applied Energy*, *285*. https://doi.org/10.1016/j.apenergy.2021.116452

Zhang, T., Li, X., Zhao, Q., & Rao, Y. (2019). Control of a novel synthetical index for the local indoor air quality by the artificial neural network and genetic algorithm. *Sustainable Cities and Society*, *51*, 101714. https://doi.org/10.1016/j.scs.2019.101714

Zhang, W., Liu, F., Wen, Y., & Nee, B. (2021). Toward explainable and interpretable building energy modelling: An explainable artificial intelligence approach. *BuildSys 2021 - Proceedings of the 2021 ACM International Conference on Systems for Energy-Efficient Built Environments*. https://doi.org/10.1145/3486611.3491127

Zhang, W., Wen, Y., Tseng, K. J., & Jin, G. (2021). Demystifying Thermal Comfort in Smart Buildings: An Interpretable Machine Learning Approach. *IEEE Internet of Things Journal*, *8*(10). https://doi.org/10.1109/JIOT.2020.3042783

Zhang, X., Chung, F. L., & Wang, S. (2020). An interpretable fuzzy DBN-Based classifier for indoor user movement prediction in ambient assisted living applications. *IEEE Transactions on Industrial Informatics*, *16*(1). https://doi.org/10.1109/TII.2019.2912625

Zhang, X., Kong, X., Yan, R., Liu, Y., Xia, P., Sun, X., Zeng, R., & Li, H. (2023). Data-driven cooling, heating and electrical load prediction for building integrated with electric vehicles considering occupant travel behavior. *Energy*, *264*. https://doi.org/10.1016/j.energy.2022.126274

Zhang, Y., Huang, T., & Bompard, E. F. (2018). Big data analytics in smart grids: a review. *Energy Informatics*, *1*(1), 8. https://doi.org/10.1186/s42162-018-0007-5

Zhang, Z., Lu, J., Xia, L., Wang, S., Zhang, H., & Zhao, R. (2020). Digital twin system design for dual-manipulator cooperation unit. *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 1431–1434. https://doi.org/10.1109/ITNEC48623.2020.9084652

Zheng, P., & Sivabalan, A. S. (2020a). A generic tri-model-based approach for product-level digital twin development in a smart manufacturing environment. *Robotics and Computer-Integrated Manufacturing*, *64*, 101958. https://doi.org/10.1016/j.rcim.2020.101958

Zheng, P., & Sivabalan, A. S. (2020b). A generic tri-model-based approach for product-level digital twin development in a smart manufacturing environment. *Robotics and Computer-Integrated Manufacturing*, *64*, 101958. https://doi.org/10.1016/j.rcim.2020.101958

Zhou, X., Xu, L., Zhang, J., Niu, B., Luo, M., Zhou, G., & Zhang, X. (2020). Data-driven thermal comfort model via support vector machine algorithms: Insights from ASHRAE RP-884 database. *Energy and Buildings*, *211*, 109795. https://doi.org/10.1016/j.enbuild.2020.109795

Zhou, Y., Nikolaev, A., Bian, L., Lin, L., & Li, L. (2021). Investigating transmission dynamics of influenza in a public indoor venue: An agent-based modeling approach. *Computers & Industrial Engineering*, *157*, 107327. https://doi.org/10.1016/j.cie.2021.107327

Zhuang, H., Wang, X., Bendersky, M., Grushetsky, A., Wu, Y., Mitrichev, P., Sterling, E., Bell, N., Ravina, W., & Qian, H. (2021). Interpretable Ranking with Generalized Additive Models. *WSDM 2021 - Proceedings of the 14th ACM International Conference on Web Search and Data Mining*. https://doi.org/10.1145/3437963.3441796

List of Figures and Tables

Annexes

Annex I

Agostinelli, S., Cumo, F., Guidi G., Tomazzoli, C. (2020) **The Potential of Digital Twin Model integrated with Artificial Intelligence Systems**, 2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe), Madrid, Spain, 2020, pp. 1-6.

DOI: https://doi:10.1109/EEEIC/ICPSEurope49358.2020.9160810

Annex II

Agostinelli, S., Cumo, F., Guidi, G., Tomazzoli, C. (2021) **Cyber-Physical Systems Improving Building Energy Management: Digital Twin and Artificial Intelligence.** Energies 2021, 14, 2338.

DOI: https://doi.org/10.3390/en14082338

Annex III

Agostinelli, S., Neshat, M., Majidi Nezhad, M., Piras, G., Astiaso Garcia, D. (2022) **Integrating Renewable Energy Sources in Italian Port Areas towards Renewable Energy Communities.** Sustainability 2022, 14, 13720.

DOI: https://doi.org/10.3390/su142113720

Annex IV

Agostinelli, S., Cumo, F., Nezhad, M.M., Orsini, G., Piras, G. (2022) **Renewable Energy System Controlled by Open-Source Tools and Digital Twin Model: Zero Energy Port Area in Italy.** Energies 2022, 15, 1817.

DOI: https://doi.org/10.3390/en15051817

Annex V

Agostinelli, S., Cumo, F., Marzo, R., Muzi, F. (2023) **Digital Construction Strategy for Project Management Optimization in a Building Renovation Site: Machine Learning and Big Data Analysis.** In: Gomes Correia, A., Azenha, M., Cruz, P.J.S., Novais, P., Pereira, P. (eds) Trends on Construction in the Digital Era. ISIC 2022. Lecture Notes in Civil Engineering, vol 306. Springer, Cham.
DOI: https://doi.org/10.1007/978-3-031-20241-4_2

Annex VI

Agostinelli, S., Nastasi, B. (2023). **Immersive Facility Management—A Methodological Approach Based on BIM and Mixed Reality for Training and Maintenance Operations.** In: Arbizzani, E., et al. Technological Imagination in the Green and Digital Transition. CONF.ITECH 2022. The Urban Book Series. Springer, Cham.
DOI: https://doi.org/10.1007/978-3-031-29515-7_13

Annex VII

Agostinelli, S. (2023). COGNIBUILD: **Cognitive Digital Twin Framework for Advanced Building Management and Predictive Maintenance.** In: Arbizzani, E., et al. Technological Imagination in the Green and Digital Transition. CONF.ITECH 2022. The Urban Book Series. Springer, Cham.
DOI: https://doi.org/10.1007/978-3-031-29515-7_8