IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 3 August 2024, accepted 7 September 2024, date of publication 11 September 2024, date of current version 8 October 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3458435

RESEARCH ARTICLE

Deep Learning-Based Predictive Control for Optimal Battery Management in Microgrids

SILVANA MATRONE[1], (Student Member, IEEE), ANDREA POZZI[2], (Member, IEEE), EMANUELE OGLIARI[1], (Member, IEEE), AND SONIA LEVA[1], (Senior Member, IEEE)
[1]Department of Energy, Politecnico di Milano, 20156 Milan, Italy
[2]Faculty of Mathematical, Physical and Natural Sciences, Catholic University of the Sacred Heart, 25133 Brescia, Italy

Corresponding author: Silvana Matrone (silvana.matrone@polimi.it)

This study was carried out within the MOST – Sustainable Mobility Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1033 17/06/2022, CN00000023). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.

ABSTRACT Effective microgrid management necessitates sophisticated strategies to optimally balance grid components and minimize power exchanges with the main grid. Central to this challenge is the energy storage system, typically comprised of lithium-ion batteries, which must operate within specific safety thresholds. Among the different approaches used for battery management in microgrids, model predictive control appears particularly suitable due to its ability to deal with nonlinear systems and constraints. However, the practical deployment of predictive control is often constrained by its substantial computational demands. Notably, achieving high performance typically requires a long prediction horizon, which exacerbates the computational complexity that increases superlinearly with the horizon length. To overcome these limitations, this paper exploits a neural network to approximate the predictive control law, thereby maintaining constant online time complexity regardless of the prediction horizon and facilitating real-time application. This innovative deep learning-based strategy is applied and specifically adapted for the first time to microgrid battery management, incorporating a comparative analysis of several machine learning models to identify the most efficient solution for this application. The results demonstrate that this approach can achieve performance comparable to traditional controllers while ensuring scalability and efficiency. Specifically, the proposed methodology is able to approximate the predictive control action with a mean error of 0.24A and a standard deviation of 2.11A, while reducing the required computational cost by over 200 times when considering a two-day ahead prediction horizon.

INDEX TERMS Deep neural networks, imitation learning, model predictive control, microgrids.

I. INTRODUCTION

A microgrid is a complex system comprising a network of interconnected renewable and traditional energy sources linked to consumers, managed by systems designed to optimize energy utilization and storage for efficiency. These systems are capable of operating in either grid-connected or standalone modes, tailored to specific operational goals. However, as the adoption of distributed generators increases, the unpredictability and intermittent nature of renewable energy generation pose challenges to the reliability and stability of the energy system.

The associate editor coordinating the review of this manuscript and approving it for publication was Jiefeng Hu.

Key functions of Energy Management System (EMS) include maintaining power quality and controlling costs, managing power dispatch, and determining operational strategies to ensure the system operates economically and meets demand reliably [1], [2]. This process takes into account economic, environmental, and technological aspects, such as electricity prices and weather conditions, as well as the dynamics of microgrid components, including energy storage systems and distributed energy sources [3]. The literature on microgrid applications features a wide array of optimization and management models. Lately, there has been a growing trend towards using predictive control methods for enhanced system management [4]. These controllers are designed to predict future actions and decisions, a capability that depends on the availability of forecasted data, such as power consumption and production figures. Among the predictive control schemes employed in microgrid management, Model Predictive Control (MPC) and its variations have received a lot of attention [5]. MPC is a sophisticated control strategy that relies on a mathematical model to forecast and optimize the system's performance over a future horizon [6]. In the context of power flow management, MPC exploits a mathematical model to anticipate future states of the microgrid based on actual and forecasted conditions, for the optimal control that balances objectives such as: cost reduction, efficiency improvement, and reliability enhancement. Successful applications of MPC for microgird management are proposed, for instance, by the authors in [7], [8], and [9].

While MPC has demonstrated promising outcomes in simulations, its practical implementation faces challenges due to computational demands [10]. The essence of MPC involves solving an optimization problem in real-time at each control interval, with the goal of determining the optimal actions over a specific prediction horizon. The performance of the control algorithm is heavily influenced by the length of this horizon. However, the time complexity of the MPC increases superlinearly with both the length of the prediction horizon and the size of the system state and input variables, even for linear systems [11].

Nonlinear systems present a significantly greater challenge due to their complex optimization landscapes lacking in guarantees of reaching the optimal solution or any solution at all, extending computation times considerably. This computational burden becomes a critical obstacle in real-world applications when calculations cannot be completed within the available timeframe between control actions. To address this issue, the concept of explicit MPC has been introduced, aiming to streamline the online computational effort by pre-calculating the optimal control strategies as a piece-wise function of the state-actions offline [12]. Although explicit MPC has been utilized in microgrid management, as demonstrated by the authors in [13], the approach still faces scalability issues for complex systems [14].

Besides, researchers have investigated the potential of approximating MPC control strategies through Machine Learning (ML) models [15], [16], [17]. This approach is framed as an imitation learning challenge [18], where MPC serves as the expert system, and a ML model acts as the apprentice. By training the ML model offline on collected measurements or synthetically generated data, such a technique can replicate the MPC's optimal control strategy with significantly reduced computational demands during online operation. Importantly, the time complexity of the ML model's inference phase is constant, remaining independent of both the system dimension and the length of the prediction horizon. In contrast to MPC, which requires solving complex optimization problems, the ML approach involves merely executing a predefined mathematical function, streamlining the process considerably. Deep Neural Networks (DNN), in particular, have shown exceptional performance in this imitation framework, thanks to their advanced representational abilities, leading to the development of the so-called Deep MPC (DMPC) [10]. This latter has been effectively applied in energy management tasks, including the control of DC-DC power converters [19], [20] and addressing the optimal charging issue for lithium-ion batteries [21], [22], demonstrating its versatility and efficiency in practical applications.

In this study, we propose an adaptation of the DMPC framework, specifically tailored for the efficient management of power flows within a microgrid. To our knowledge, this represents the inaugural exploration of leveraging imitation learning for microgrid energy management, aimed specifically at diminishing the computational burden inherent in traditional predictive control approaches. In our approach, the conventional predictive controller serves a dual role: it is the model of excellence that our DMPC strategy strives to emulate and the standard against which we measure our method's computational efficiency.

Furthermore, we have conducted a comprehensive comparison of several ML models to identify the most effective architecture for this application. We implement our methodology through simulations utilizing a simplified microgrid model, incorporating realistic data for both power supply and demand forecasts. This setup serves as a proof of concept, demonstrating the potential of our innovative approach in the realm of power flow management.

It is important to remark that the primary goal of using Deep Learning (DL) in this paper is to approximate and imitate the predictive controller, considered as the expert agent, in a computationally efficient manner. The performance of this predictive controller on microgrid management is assumed to be satisfactory, as discussed by [4], except for the main limitation of online computational complexity. Addressing this limitation, our approach leverages DL to reduce the computational burden associated with real-time application. It is worth noting that a completely different task, beyond the scope of the present work, would be to use DNN to discover novel management strategies, as described in [23], with a particular focus on EMS.

Summarizing, the main contributions of this paper are:

- adapting the DMPC approach to the context of microgrid management within a general framework of imitation learning;
- conducting an exhaustive model selection to identify the most suitable approximation of the MPC control law;
- testing the proposed DMPC strategy on a microgrid simulator with realistic data for power supply and load demand;

The rest of the paper is organized as follows. In Section II the main concepts regarding microgrid operation and optimal energy management are described, while in Section III the main features of MPC and DMPC are introduced. Section IV includes the description of the case study on which the simulations have been performed. Finally, the results are presented in Section V, while Section VI concludes the paper with the drawn conclusions.

II. FUNDAMENTALS OF MICROGRID OPERATION

This section introduces the foundational elements and operational principles of microgrids, structured into two pivotal subsections. The discussion begins with the introduction of the microgrid concept in II-A, elucidating its role as a localized group of electricity sources and loads that operates autonomously from the traditional grid or in conjunction with it. A straightforward mathematical model of a microgrid is provided, which will serve as the backbone for the simulation and analysis conducted throughout this paper. Following this, the definition of optimal energy management of a microgrid is outlined in II-B. This segment focuses on the requirements to ensure efficient, reliable, and sustainable operation. It delves into the balancing act of meeting energy demands, optimizing the use of renewable resources, minimizing costs, and maintaining system stability. This framework lays the groundwork for the subsequent application of advanced control strategies aimed at enhancing microgrid efficiency and sustainability.

A. MATHEMATICAL MODEL OF THE MICROGRID

A microgrid can be defined as a self-sufficient energy system that encompasses a localized group of electricity sources and loads capable of operating independently or in conjunction with the main power grid [24]. This compact network integrates various power generation sources, including renewable energy sources (such as solar panels and wind turbines), conventional power generators, and energy storage systems. It can autonomously manage energy production, consumption, and storage to ensure a reliable and continuous energy supply to its connected consumers, even during grid outages or in remote locations without grid access. Microgrids are designed to optimize the use of local energy resources, enhance energy security, and reduce environmental impact. They play a pivotal role in transitioning towards a more sustainable and resilient energy infrastructure by allowing for greater integration of renewable energy, providing energy independence, and enabling a more efficient and smarter energy management system. Their flexibility and adaptability make them an essential component in modern energy strategies, catering to the needs of diverse settings—from remote communities to urban areas and industrial sites.

As it can be noticed from the generic scheme depicted in Figure 1, a microgrid comprises several key components, each contributing to the system's overall functionality and stability:

- Generators: these units generate the power $P_{\text{gen}} > 0$ and can include both renewable energy sources (such as solar panels and wind turbines) and conventional generators;
- Loads: these represent the total demand within the microgrid, consuming the power $P_{\text{load}} > 0$;
- Energy Storage Systems (ESSs): an ESS can supply power to the grid or absorb power for storage, with P_{ESS} being positive in the former case and negative in the latter;
- Main Power Grid: this interface facilitates the exchange of power P_{grid} between the microgrid and the main utility grid, with the power being considered positive when imported from the main grid and negative when exported to it.

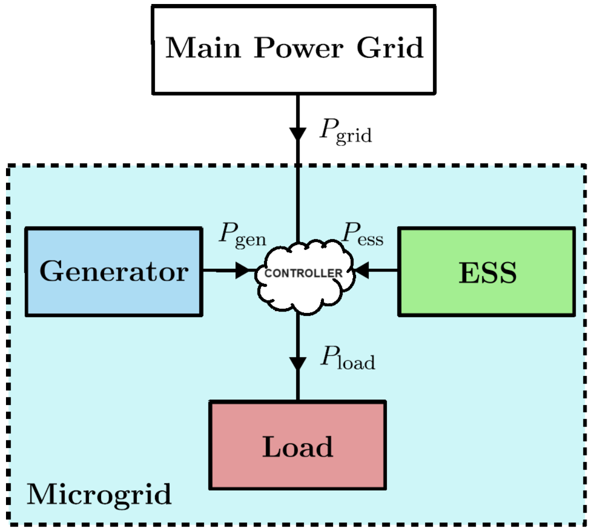


FIGURE 1. A schematic representation of a generic microgrid: the fundamental elements such as generator, energy storage system, and load are depicted together with the connection with the main power grid.

In addition to the components illustrated in Figure 1, power converters are essential elements within a microgrid. A power converter is a device that transforms electrical energy from one form to another, facilitating the conversion between alternating current (AC) and direct current (DC). This capability is critical for the integration and coexistence of both DC sources (such as batteries and solar panels) and AC sources (like electrical motors) within the same grid. Power converters ensure that these diverse power units can function synergistically, optimizing the overall efficiency and reliability of the microgrid.

The operation of a microgrid is governed by the principle of power balance, ensuring that the total generated power equals the total consumed power, including losses and storage, at every discrete time step k. The power balance equation can be expressed as:

$$P_{\text{gen},k} - P_{\text{load},k} + P_{\text{ess},k} + P_{\text{grid},k} = 0, \tag{1}$$

where the subscript k denotes the considered discrete time instant. This equation reflects the fundamental requirement that, at each time step, the sum of generated power, the power exchanged with the main grid, and the power from or to the energy storage system must exactly meet the load demand.

In this paper, the ESS is assumed to be constituted by a lithium-ion battery pack. Lithium-ion batteries are commonly utilized in ESSs due to their high energy density, long lifespan, and minimal maintenance requirements [25]. These characteristics make them an optimal choice for efficiently storing and releasing energy on demand, thereby enhancing the overall reliability and flexibility of microgrids. The basic working principle of a battery involves the conversion of chemical energy into electrical energy through electrochemical reactions. These reactions occur between the anode and cathode of the battery, facilitated by the movement of ions through an electrolyte medium while electrons flow through an external circuit, generating an electric current. Various models exist to represent the behavior and performance of batteries, including equivalent circuit models and electrochemical models. Equivalent circuit models offer a simplified and intuitive approach, making them well-suited for initial analyses and systems where high fidelity is not critical [26]. On the other hand, electrochemical models provide a more detailed and accurate representation of the internal processes of the battery. However, this accuracy comes at the cost of increased computational complexity, making these models more demanding in terms of computational resources.

For the purposes of this paper, we adopt a very simple battery model. This decision is driven by our focus on demonstrating a proof of concept for applying imitation learning techniques to microgrid control, rather than delving into the intricacies of battery dynamics.

To introduce the main equations of the simplified battery model considered in this paper, we start focusing on the variation of the State of Charge (SoC) of the battery, indicated with the variable $Z \in [0, 1]$, as a consequence of an applied electrical current I_{app}. The SoC is a critical parameter representing the current energy level of the battery relative to its maximum capacity, and its dynamics can be represented by the following equation:

$$Z_{k+1} = Z_k - \Delta t \frac{I_{\mathrm{app},k}}{3600C}, \tag{2}$$

where C is the nominal capacity of the battery in Ah and Δt is the considered sample time for the discretization of the model expressed in seconds. Note that in this paper a negative applied current is assumed to charge the battery by convention. The terminal voltage of the cell can be computed as follows:

$$V_k = V_{\mathrm{ocp}}(Z_k) - RI_{\mathrm{app},k}, \tag{3}$$

where $V_{\mathrm{ocp}}(\cdot)$ is a mathematical function relating the open circuit voltage and the state of charge and R is the internal resistance of the battery. In this simplified battery model, the open circuit potential is considered to be a linear function of the state of charge:

$$V_{\mathrm{ocp}}(Z_k) = V_0 + Z_k(V_1 - V_0), \tag{4}$$

where V_0 and V_1 represent the open circuit voltage when the battery is completely discharged and charged, respectively. Finally, the power exchanged by the energy storage system upstream of the power converter is given by:

$$\tilde{P}_{\mathrm{ess},k} = \begin{cases} \dfrac{V_k I_{\mathrm{app},k}}{\eta_{\mathrm{ess}}}, & \text{if } I_{\mathrm{app},k} < 0, \\ V_k I_{\mathrm{app},k}\, \eta_{\mathrm{ess}}, & \text{otherwise.} \end{cases} \tag{5}$$

where η_{ess} is the efficiency of the charging/discharging process, which, for the sake of simplicity, is assumed to be constant in this work.

It is crucial to note that the powers specified in (1) are calculated downstream of power conversion processes. Within the microgrid, each power converter is presumed to exhibit its own efficiency, denoted as $\eta_{\mathrm{conv},k}$, which varies over time. This efficiency is modeled in this study as an exponential function of the ratio of the input power at a specific converter $P_{\mathrm{in},k}$ to its nominal power P_{nom}:

$$\eta_{\mathrm{conv},k} = f\left(\frac{P_{\mathrm{in},k}}{P_{\mathrm{nom}}}\right). \tag{6}$$

This approach recognizes the dynamic nature of converter efficiency, which can fluctuate based on operational conditions and input power levels. For instance, the power exchanged by the energy storage system downstream of its power converter is computed as:

$$P_{\mathrm{ess},k} = \begin{cases} \dfrac{\tilde{P}_{\mathrm{ess},k}}{\eta^{\mathrm{ess}}_{\mathrm{conv},k}}, & \text{if } I_{\mathrm{app},k} < 0, \\ \tilde{P}_{\mathrm{ess},k}\, \eta^{\mathrm{ess}}_{\mathrm{conv},k}, & \text{otherwise,} \end{cases} \tag{7}$$

where $\eta^{\mathrm{ess}}_{\mathrm{conv},k}$ is the efficiency of the power converted associated with the ESS (see Section IV, for the efficiency of the other microgrid converters considered in the simulations presented in this paper).

It is also important to clarify that the study discussed in this paper assumes, for the sake of simplicity, that all the microgrid converters operate with a power factor of 1 in the simulation environment. This assumption is common in microgrid studies focused on stand-alone operation modes where the primary concern is the active power management. This simplification does not detract from the functionality of the proposed control algorithm, which is the central focus of the papers's investigation.

B. OPTIMAL ENERGY MANAGEMENT

The principal aim of optimal energy management within a microgrid is to judiciously control the power flow to minimize the long-term energy exchange with the main power grid. This strategy seeks to enhance the microgrid's autonomy, striving to operate it as independently as possible. Achieving this involves a delicate balance: the system must minimize reliance on external energy sources while simultaneously satisfying the load demand and safeguarding the energy storage system against undue stress. This entails maintaining the battery's operation within safe voltage and SoC thresholds, as well as moderating the intensity of the battery current to prolong its lifespan and ensure efficient energy use.

In scenarios, like the one considered in this paper, where the battery serves as the sole controllable component of the system (e.g. when the generators consist only of renewable energy sources), this approach is typically referred to as battery energy management [27]. Such a management strategy focuses on optimizing the battery's charging and discharging processes to align with the microgrid's energy needs, thereby reducing the need for external energy inputs. This involves strategic planning to store excess energy during periods of low demand or high renewable production and then utilizing this stored energy during peak demand times or when renewable energy generation is insufficient.

Given these premises, a crucial aspect of planning the optimal battery current over a future horizon is the ability to accurately predict future load demand [28] and generator production [29], which both exhibit high levels of unpredictability, particularly in the case in which the generators are renewable energy sources. The effectiveness of battery energy management significantly hinges on the accuracy of these forecasts since they directly influence decisions on when to store or release energy. To facilitate this strategic planning, it is assumed that estimates of future load demand ($\hat{P}_{\text{load}}$) and generator production ($\hat{P}_{\text{gen}}$) are available through sophisticated forecast predictors. These tools are essential for anticipating the variability inherent in renewable energy sources and demand patterns, thereby allowing for more informed and efficient management decisions. The methodologies and technologies employed to generate these forecasts, along with their integration into the overall energy management strategy, will be discussed in greater detail in Section IV.

Forecasting inaccuracies are a notable concern as they can lead to energy imbalances, which are particularly critical in off-grid scenarios, necessitating a spinning reserve. In this study, however, the microgrid is assumed to be connected to the main grid, allowing any imbalances caused by prediction errors to be effectively compensated. This compensation is facilitated through the power exchanged with the main grid, which is treated as a dependent variable in the balance equation (1). Such connectivity ensures that the microgrid can maintain operational stability despite the inherent unpredictability of load demand and generator outputs, especially when relying on renewable energy sources.

At a certain time instant k, the optimal energy management task described above can be formulated in terms of a constrained optimal control problem that aims to minimize the following objective function over a certain future horizon H:

$$J_{[k,\,k+H]} = q \sum_{i=k+1}^{k+H} \hat{P}^2_{\text{gird},i} + r \sum_{i=k}^{k+H-1} I^2_{\text{app},k}, \tag{8}$$

with $\hat{P}_{\text{gird},k}$ being the estimate for the power exchanged with the main grid, obtained from (1) as follows:

$$\hat{P}_{\text{gird},k} = \hat{P}_{\text{load},k} - \hat{P}_{\text{gen},k} - P_{\text{ess},k}, \tag{9}$$

subject to the following constraints for $i = k+1, \ldots, k+H$:

$$I_{\min} \leq I_{\text{app},i} \leq I_{\max}, \tag{10a}$$
$$Z_{\min} \leq Z_i \leq Z_{\min}, \tag{10b}$$
$$V_{\min} \leq V_i \leq V_{\max}, \tag{10c}$$
$$P_{\text{grid, min}} \leq P_{\text{grid, i}} \leq P_{\text{grid, min}}, \tag{10d}$$
$$P_{\text{ess, min}} \leq P_{\text{ess, i}} \leq P_{\text{ess, max}}, \tag{10e}$$

with $\phi_{\min}$ and $\phi_{\max}$ being the lower and upper bounds, respectively, for the generic variable ϕ. Note that q and r in (8) are weighting factors that allow for addressing the trade-off between effectively achieving the control objective and applying smooth control actions.

Therefore, the constrained optimal control problem to be solved at each time step k is the following:

$$I^{\star}_{\text{app},k}, \ldots, I^{\star}_{\text{app},k+H} = \underset{I_{\text{app},k},\ldots,I_{\text{app},k+H}}{\operatorname{argmin}} \; J_{[k,\,k+H]}, \tag{11}$$

subject to:

$$\text{system dynamics in (2) – (5) and (9)}, \tag{12a}$$
$$\text{constraints in (10)}, \tag{12b}$$

where $I^{\star}_{\text{app},k}, \ldots, I^{\star}_{\text{app},k+H}$ is the sequence of optimal battery currents to be applied over the next H time steps.

To streamline the notation for subsequent discussions, we define several key variables. Initially, we introduce the system state vector at time instant k, denoted as $\mathbf{s}_k$. For the purposes of our simplified model, this vector primarily consists of the battery's SoC. Next, the vector $\mathbf{a}_k$ represents the control actions applied to the system at time-step k, specifically referring to the battery current. Lastly, we consolidate the estimated future power demand and generation over the horizon H into the vector $\hat{\mathbf{d}}_k$, structured as follows:

$$\hat{\mathbf{d}}_k = [\hat{P}_{\text{load},k+1},\ \hat{P}_{\text{gen},k+1},\ \ldots,\ \hat{P}_{\text{load},k+H},\ \hat{P}_{\text{gen},k+H}]. \tag{13}$$

This approach organizes crucial data for managing the microgrid, facilitating a clear and concise analysis in the remainder of the paper.

III. METHODOLOGY

This section delves into the methodology behind the deep learning-based control strategy implemented in this study. Specifically, we revisit the essential aspects of MPC in III-A, emphasizing its defining principle of the receding horizon. Following that, subsection III-B introduces the basics of imitation learning, explaining how it can be employed to approximate the MPC control law, which serves as the expert agent in this context. Through this approach, we aim to illuminate the process of leveraging advanced computational techniques to enhance the operational efficiency and effectiveness of microgrid management.

A. MODEL PREDICTIVE CONTROL

The control challenge outlined in II-B is adeptly addressed using a model predictive control strategy. Notably effective in managing nonlinear systems constrained by specific input and state limitations, MPC leverages a receding horizon principle to navigate these complexities [30]. At every step

k, MPC formulates and solves a constrained optimization problem, as defined in equations (11)–(12), to identify the optimal sequence of control actions $\mathbf{a}_k^\star, \ldots, \mathbf{a}_{k+H}^\star$ for a given prediction horizon H. This optimization is grounded in system predictions generated by a mathematical model. Following the receding horizon paradigm, only the initial control action $\mathbf{a}_k^\star$ is executed, with the subsequent actions being recalculated at each step.

The selection of the optimal control action $\mathbf{a}_k^\star$ at each interval k is informed by the current system states $\mathbf{s}_k$ and the estimation of future disturbances $\hat{\mathbf{d}}_k$. These components guide the MPC in determining and applying the most suitable action for that moment, determining the optimal policy $\pi^\star$ that maps the current states and forecasts to the optimal control action as follows:

$$\pi^\star : (\mathbf{s}_k,\, \hat{\mathbf{d}}_k) \mapsto \mathbf{a}_k^\star. \tag{14}$$

This policy $\pi^\star$ is what we aim to approximate in our proposed imitation learning framework, positioning it as the "expert agent."

The computational demands of MPC are a critical consideration, especially when dealing with real-time applications. For linear systems with states $\mathbf{s} \in \mathbb{R}^n$ and actions $\mathbf{a} \in \mathbb{R}^m$, the complexity of solving the optimization problem at each time step using traditional dense solvers is $O(H^3(n+m)^3)$, where H is the prediction horizon. Efficient solvers that exploit the sparsity and structural properties of the problem can reduce this complexity to $O(H(n+m)^3)$ [11]. However, these improvements are primarily effective for linear systems. When dealing with nonlinear systems, the situation becomes markedly more challenging. The time complexity for standard methods increases significantly due to the non-convex nature of the optimization problems, which not only complicates the computational process but also introduces the risk of failing to converge to the global optimum.

This increase in complexity underscores the need for alternative approaches that can manage the computational load effectively while still delivering the requisite control precision. Our proposed imitation learning framework seeks to address this by approximating the optimal policy $\pi^\star$, thereby sidestepping the intensive computational demands traditionally associated with MPC in nonlinear settings.

B. AN IMITATION LEARNING APPROACH TO DEEP MPC

Imitation learning represents a method that involves a machine learning model acquiring skills through observation or examples from an expert agent, striving to emulate the expert's actions with precision.

In essence, imitation learning is considered a specialized form of supervised learning, where the goal is to develop a policy, characterized by parameters $\boldsymbol{\theta}$ and symbolized as $\pi^{\boldsymbol{\theta}}$, that accurately reflects the actions of an expert policy $\pi^\star$. The core objective is to find the optimal set of parameters $\boldsymbol{\theta}^\star$ that enables the policy $\pi^{\boldsymbol{\theta}^\star}$ to closely mimic the expert's decisions.

Viewed through the lens of supervised learning, this process entails learning a function that approximates the map from observations ($\mathbf{s}_k$ and $\hat{\mathbf{d}}_k$) to optimal actions ($\pi^\star(\mathbf{s}_k, \hat{\mathbf{d}}_k)$), based on a dataset $\mathcal{B}_{\text{tr}}$ collected from the expert policy's execution and defined as follows:

$$\mathcal{B}_{\text{tr}} = \{(\mathbf{s}_k, \hat{\mathbf{d}}_k, \pi^\star(\mathbf{s}_k, \hat{\mathbf{d}}_k))\}_{k=1}^{N}, \tag{15}$$

where N is the number of samples in the dataset. To ensure the imitation policy aligns with the expert's behavior, a loss function is minimized across $\mathcal{B}_{\text{tr}}$. Typically, the loss function is the mean square error between the actions suggested by policy $\pi_{\boldsymbol{\theta}}$ and those of the expert policy $\pi^\star$, expressed as

$$L(\boldsymbol{\theta}) = \frac{1}{N}\sum_{k=1}^{N}\left[\pi^{\boldsymbol{\theta}}(\mathbf{s}_k, \hat{\mathbf{d}}_k) - \pi^\star(\mathbf{s}_k, \hat{\mathbf{d}}_k)\right]^2. \tag{16}$$

This function motivates the imitation policy to generate actions that are congruent with the expert's, particularly in the states experienced through the execution of the expert policy. When the expert in question is a model predictive controller and the approximating machine learning model is a deep neural network, this form of imitation is termed deep MPC [31].

Regarding the online time complexity, the imitation learning approach presented in this paper fundamentally alters the computational architecture of predictive control, introducing a learned model that approximates the predictive control law and achieves a constant time complexity O(1), independent of the system's complexity and the length of the prediction horizon H.

IV. CASE STUDY

This section presents a detailed overview of the methodology employed to assess the efficacy of the deep MPC strategy within a structured case study. Emphasis is placed on the setup and procedural aspects, reserving the discussion of results for Section V. The initial focus, detailed in Subsection IV-A, revolves around the microgrid environment. Here, we delineate the configuration of key components, including generators, loads, ESS, and power converters, with their parameters systematically tabulated for clarity. Additionally, we elaborate on the selection of weights for the optimal energy management problem and define the bounds for critical variables, shedding light on the strategic considerations integral to effective energy management.

Subsequently, Subsection IV-B outlines the preparatory steps necessary for evaluating the computational advantages of employing imitation learning techniques. This includes a comprehensive description of the process for generating a synthetic dataset through the execution of the expert policy, which serves as a foundation for training the imitative neural network agent. The architecture of this neural network, pivotal to the imitation learning approach, is also discussed, providing a technical snapshot of the system designed to replicate the expert policy's decision-making process.

TABLE 1. Microgrid elements characteristics.

Microgrid Parameter	Value
PV 1 Nominal Power	24 kW
PV 2 Nominal Power	26 kW
Electrical Load peak	52 kW
Battery Nominal Capacity (C)	95 Ah
Battery Internal Resistance (R)	0.1 Ω
Battery Efficiency (η_{ess})	0.97
PV Inverter maximum power output	28 kW
Battery Inverter maximum power output	56 kW
Open Circuit Potential at 0% of SoC (V_0)	634 V
Open Circuit Potential at 100% of SoC (V_1)	822 V

TABLE 2. Forecast models performances.

Forecaster	Mean Absolute Error
PV Production	2.26 kW
Load Demand	0.388 kW

A. MICROGRID TESTBED CONFIGURATION

This study utilizes a streamlined microgrid model for simulation purposes, comprising a selection of essential components. Illustrated in Figure 2, the microgrid model includes a connection to the main electrical grid, enabling it to both absorb and inject power to maintain energy equilibrium. While it possesses the theoretical capability to operate both in islanded or grid-connected modes, our analysis presumes constant interaction with the national grid. Switching between power injection and absorption modes occurs seamlessly, without associated transitory effects.

Energy storage is facilitated by a lithium-ion battery pack, exhibiting a nominal capacity of 94 Ah and an internal resistance of 0.1 Ω, with a charging/discharging efficiency of 97%. The open-circuit voltage of the battery linearly correlates with the SoC, ranging from 634 V when completely discharged to 822 V when fully charged. The battery is connected to the grid via a converter with a nominal maximum power output of 56 kW.

Power generation within the microgrid is exclusively provided by two photovoltaic arrays, with nominal powers of 24 kW and 26 kW, respectively. These PV arrays are connected to the grid through two power inverters, each with a maximum output of 28 kW. Note that each of such power converters exhibits an efficiency $\eta^{\text{PV}}_{\text{conv},k}$ which is dynamically computed through (6). Finally, the microgrid supports a non-dispatchable load typical of an industrial setting, peaking at 52 kW.

Table 1 compiles the pertinent specifications of these microgrid components for reference.

The dataset utilized to construct the temporal profiles of photovoltaic power production is sourced from the SolarTech Lab at Politecnico di Milano [32]. Similarly, the simulation of the industrial electric load time series leverages modified data from an actual Italian paper factory, with the peak load refined to 52kW to ensure compatibility with the microgrid's component capacities. As outlined in prior discussions, this study operates under the premise that perfect foresight into future PV production and load demand is unavailable to the controller (i.e. the decision-making agent). Instead, forecasts are generated using a recurrent neural network model as suggested by [33] and further enhanced with an attention layer [34], which takes in input a window historical measurements with a length of 672 steps. This enhancement enables the model to adeptly capture the intricate historical patterns characteristic of the time series data under consideration in this research. The accuracy and uncertainty associated with this forecast model are detailed in Table 2 for both PV generation and load predictions. Despite achieving high accuracy, a degree of uncertainty persists, underscoring the realistic nature of the scenario portrayed in our study. This approach ensures the simulated environment closely mirrors real-world conditions, providing a robust foundation for evaluating the proposed control strategies.

TABLE 3. Lower and upper bounds for microgrid's variables.

Variable	Lower Bound	Upper Bound
Applied Battery Current ($I_{\text{app},k}$)	-36 A	36 A
Battery Voltage (V)	634 V	822 V
Battery SoC (Z)	0.25	0.85
Battery Power (P_{ess})	-30 kW	30 kW
Grid Power (P_{grid})	-40 kW	40 kW

Regarding the configuration of the optimal energy management problem, it is essential to highlight the significance of the trade-off within the cost function (8), which is managed through the selection of weights q and r. These weights are crucial for balancing the objectives of minimizing grid energy exchange and ensuring smooth control action. For the purposes of this study, the weights are set to $q = 10$ and $r = 10^{-2}$, respectively. Additionally, to ensure the microgrid operates within safe and reliable parameters, the lower and upper bounds for the key variables are meticulously defined. These operational constraints are summarized in Table 3, providing a comprehensive overview of the limits established to safeguard the microgrid's functionality.

To fully encapsulate the dynamics of the microgrid in our simulations, it is crucial to specify the time-step used for model discretization. In this context, we have selected a time-step, Δt, of 900 seconds, equivalent to 15 minutes. This duration is considered a practical and reasonable sampling interval for the operational dynamics of a power network.

B. IMITATION LEARNING FRAMEWORK

This subsection introduces the setup utilized to develop a DMPC system, designed to approximate the decision-making process of an expert policy, specifically, a standard MPC. Since this paper focuses on assessing the advantages in terms of computational requirements of DMPC techniques with respect to the standard predictive controllers, several MPC algorithms with different horizon lengths need to be approximated. In particular, we define $\pi^{\star}_{H}$ the expert policy associated with the MPC with horizon H, and we consider $H \in \{4, 12, 24, 48, 96, 192, 384\}$. For each horizon H, our approach involves generating a dataset by executing the corresponding expert MPC agent across

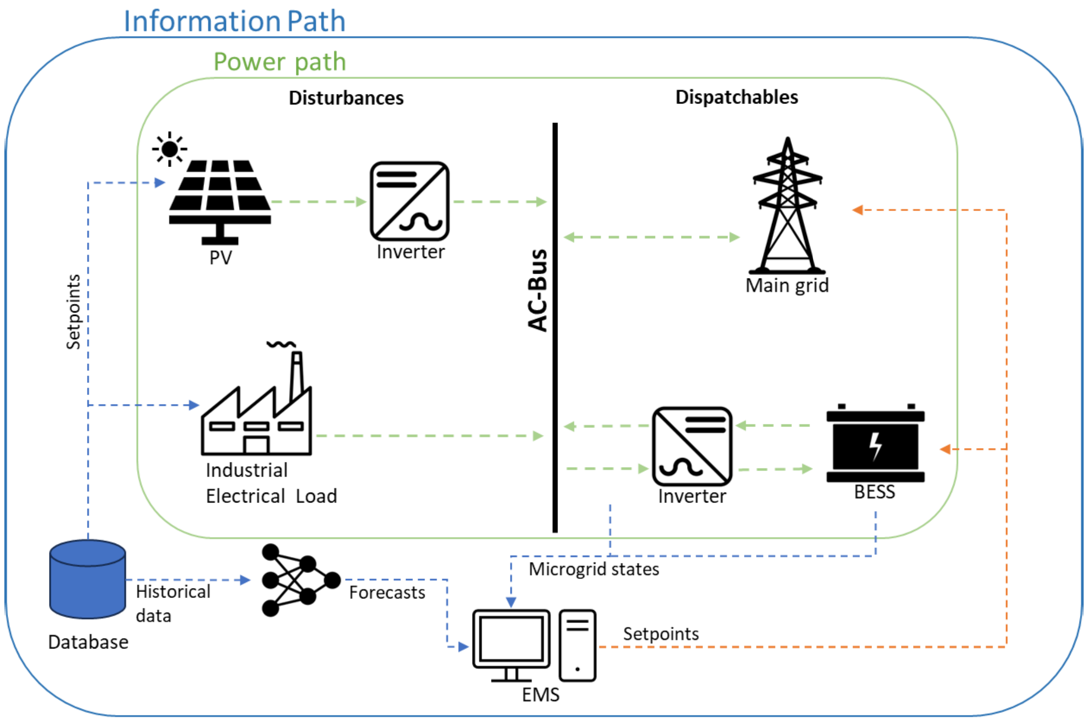


FIGURE 2. Graphical representation of the microgrid used to perform simulations.

various operational scenarios, encapsulated as $\mathcal{B}_{\mathrm{tr}}^{H}$. This dataset forms the foundation for training the corresponding DMPC algorithm represented by the neural network $\pi_{H}^{\boldsymbol{\theta}}$. The intricacies of dataset generation and the neural network architecture, including hyperparameter selection for the training algorithm, are further discussed in the following subsubsections.

1) DATASET GENERATION

For each value of H, the dataset used for training the corresponding DMPC system is synthetically generated by recording the optimal actions taken by the expert agent (the standard MPC with horizon H) and the corresponding observations that inform these decisions during its interaction with the microgrid model.

To adequately prepare the DMPC for bringing the microgrid to operational equilibrium from various initial conditions, we created a substantial collection of short sequences, with $n_e = 1000$ episodes designed to represent diverse starting scenarios. Each episode spans $n_s = 288$ steps, equivalent to 72 hours, ensuring that a comprehensive range of microgrid states are examined. For every episode, specific segments from two extensive datasets—one detailing PV power production and the other electric load forecasts—are selected to simulate the PV power and load demand over time. These datasets encompass a total of 22080 records, roughly covering nearly 8 months of measurements.

At each step j within episode i, the data tuple $(\mathbf{s}_k, \hat{\mathbf{d}}_k, \pi_H^{\star}(\mathbf{s}_k, \hat{\mathbf{d}}_k))$ is stored and insert in the dataset $\mathcal{B}_{\mathrm{tr}}^{H}$, with $k = i n_e + j$. Note that $\hat{\mathbf{d}}_k$ represents the forecasted vectors of PV power and load demand over the horizon H, derived using the forecaster models that rely on historical data for prediction.

2) TRAINING PHASE, MODEL SELECTION, AND HYPERPARAMETERS

To validate the suitability of employing a deep learning model for approximating the predictive control law, we compared the performance of several models using the dataset $\mathcal{B}_{\mathrm{tr}}^{H}$ outlined in the previous section, with $H = 48$. To improve the learning potential of all models under consideration, we implemented a preprocessing pipeline encompassing both feature scaling and standardization of the dataset. The models employed to conduct the comparison are:

- Linear Regression;
- Support Vector Regression;
- Random Forest;
- Artificial Neural Network (ANN) with different configurations for the number of layers and neurons in each layer.

Given that the microgrid states are always measurable, and the predictions of PV power and load demand are reliably available through the properly trained forecaster, there is no need to consider recurrent neural networks. Such networks are typically used for partially observable environments, where sequences of past measurements are used in view of the system state. For each configuration of the ANN, a rectified linear unit is used as the activation function for all hidden layers, while the output layer utilizes a hyperbolic tangent to automatically constrain the network's output within the continuous range of feasible actions.

TABLE 4. Mean Square Error (MSE) over the validation set of the machine learning models considered as possible candidates for the MPC.

Model	MSE
Linear Regression	4.47 kW
Support Vector Regression	17.3 kW
Random Forest	10.3 kW
ANN $(1 \times 50, 1 \times 10)$	1.40 kW
ANN $(1 \times 50, 1 \times 20)$	1.36 kW
ANN $(1 \times 100, 1 \times 20)$	1.09 kW
ANN $(2 \times 100, 2 \times 20)$	1.19 kW
ANN $(1 \times 100, 1 \times 50, 1 \times 20)$	1.20 kW
ANN $(2 \times 200, 2 \times 100, 1 \times 20)$	**0.95 kW**
ANN $(2 \times 200, 2 \times 100, 1 \times 50)$	1.06 kW
ANN $(2 \times 300, 2 \times 200, 1 \times 20)$	1.08 kW
ANN $(2 \times 300, 2 \times 200, 1 \times 50)$	1.01 kW
ANN $(3 \times 200, 3 \times 100, 1 \times 20)$	1.11 kW
ANN $(2 \times 200, 2 \times 100, 1 \times 250, 1 \times 20)$	0.96 kW
ANN $(2 \times 300, 2 \times 200, 1 \times 100, 1 \times 50)$	1.10 kW

The results of the model comparison are summarized in Table 4, where only those configurations that showed significant performance differences are included. The training of the ANN-based models was conducted using the stochastic gradient descent algorithm over the dataset for 50 epochs, with an early stopping mechanism triggered by validation loss and a patience parameter set to 3. Specifically, the Adam optimizer was employed with a mean squared error as the loss function and a learning rate set to 0.0005. As it can be noticed in Table 4, deeper models consistently outperformed shallower ones, indicating that an effective ANN configuration for microgrid management should ideally consist of at least 3-4 layers. Non-deep models, in contrast, were inadequate for the task, exhibiting significantly higher validation losses.

Consequently, the model demonstrating superior performance and thus chosen for our simulations features a neural network architecture with five fully connected hidden layers. The configuration consists of two layers with 200 neurons each, followed by two layers with 100 neurons each, and a final layer with 20 neurons. This model, designated as $\pi_{48}^{\theta^\star}$, has been used to compute the results of this study presented in Subsection V-B.

V. RESULTS

In this section, the main results obtained in the current paper are presented. First of all, a description of the MPC performance in addressing the optimal energy management problem is provided in subsection V-A. For this specific scenario, a horizon $H = 48$ is considered. The same setting is used in subsection V-B, where the approximation capabilities of the DMPC are shown. Finally, in subsection V-C the comparison between MPC and DMPC for different prediction horizons is discussed in terms of computational complexity.

A. OPTIMAL ENERGY MANAGEMENT THROUGH MPC

In this study, we seek to validate the efficacy of the MPC strategy for optimizing power flow within the microgrid under consideration, particularly through strategic management of the battery's role in minimizing energy exchange with the main grid. To achieve this, we conduct two distinct simulations over a period of 288 steps, equivalent to 72 hours. The first simulation operates under the assumption that the battery current remains fixed at zero, effectively simulating a scenario devoid of any active control system. The second simulation, in contrast, employs an MPC with a planning horizon of $H = 48$ to dynamically manage the battery current, thereby optimizing the microgrid's power flow according to the policy $\pi_{48}^{\star}$.

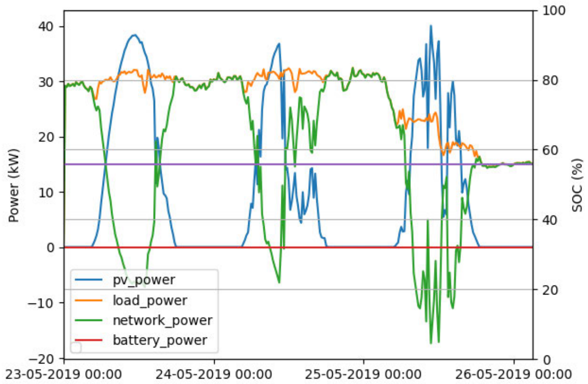
The outcomes of these simulations are illustrated in Figure 3, where the left panel represents the scenario without controller intervention, and the right panel showcases the scenario with MPC-driven control. In the absence of control, the battery plays a passive role, merely responding to the microgrid's natural dynamics without strategic energy management. In contrast, the MPC scenario demonstrates the active role of the battery in balancing energy within the microgrid, which involves strategic charging and discharging to reduce power exchanges with the main grid. Both scenarios start at midnight on May 23, 2019, with an initial SoC of 0.5. In the no-control scenario, the battery's SoC fluctuates minimally, as the system lacks a coordinated strategy for managing energy. Conversely, in the MPC-controlled scenario, the battery actively charges during peak PV production hours in the daytime and discharges during the night to meet the load demand, effectively reducing reliance on external power sources.

The choice of a springtime period for the simulation captures high PV production, covering the load during peak solar hours and highlighting the effectiveness of the MPC in managing seasonal variations in solar generation. While an industrial load profile remains consistent throughout the year, PV production varies significantly with weather and seasonal changes. The MPC adapts to these variations by prioritizing battery management, aiming to maintain grid stability and minimize power exchange with the main grid. Although winter conditions with low PV output or summer conditions with excess PV generation present the most challenging scenarios, the MPC demonstrates its capability to optimize microgrid performance across a range of operational conditions.

B. APPROXIMATION CAPABILITIES OF DMPC

In this analysis, the DMPC algorithm, symbolized by the neural network $\pi_{48}^{\theta^\star}$, aims to closely replicate the decision-making process of the expert policy $\pi_{48}^{\star}$ as implemented by the conventional MPC, as outlined in subsection V-A. We examine the system's behavior over an identical period of 3 days starting from 23^{th} May 2019 at midnight, to facilitate a direct comparison between the two control strategies.

The performance of both methodologies are depicted in Figures 4, 5, 6, and 7. In these figures, the DMPC's operational profiles are denoted by a blue line, whereas the standard MPC's actions are represented by a red dotted line.



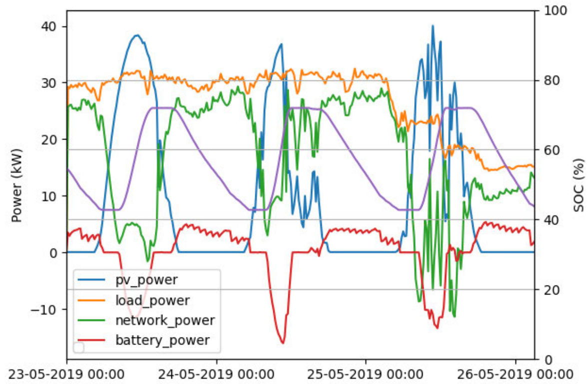


FIGURE 3. Comparison of the microgrid operation with no controller intervention (left) and with MPC control (right) during a springtime simulation, illustrating the distinct impact of predictive control on energy management and battery usage.

Overall, the trajectories produced by both control strategies demonstrate a high degree of similarity, showcasing the effectiveness of the DMPC in replicating the behavior of the standard MPC. Specifically, Figure 4 compares the SoC profiles under both control strategies. The two profiles nearly coincide, reflecting the DMPC's ability to emulate the MPC's decision-making in maintaining optimal battery usage throughout the operational period. Figure 5 shows the battery power profiles, where the similarity between the two methods further confirms DMPC's effectiveness in managing energy flows within the microgrid. Figure 6 presents the power exchanged with the main grid, with both control strategies exhibiting comparable interaction levels, while Figure 7 highlights the current applied by each controller. It is evident across all depicted metrics that the operational constraints detailed in subsection II-B are consistently adhered to by both methodologies.

Similar approximation capability for the DMPC can be obtained for initial conditions in the intervals considered during the training phase. We performed a statistical analysis of the difference between MPC and DMPC action profiles in 300 simulations with randomly extracted initial states. The results of the analysis can be observed in Figure 8. Additionally, Table 5 summarizes the comparative results between MPC and DMPC in terms of number of iterations required for convergence, approximation error achieved by the DMPC, and time required for the offline training of the neural network. As it can be seen from the table, the average number of iterations of the MPC is 5.6 with a standard deviation of 1.2, while computing the number of iterations for the DMPC is not applicable since the DMPC execution consists only of the evaluation of the pre-trained neural network. The average approximation error for the applied current is 0.24A, with a standard deviation of 2.11A, while the time required offline for the neural network training is 1.8s with a standard deviation of 0.2s. Note that the approximation error and the offline required time are only applicable to the DMPC approach, as the MPC is considered as the expert agent to be approximated.

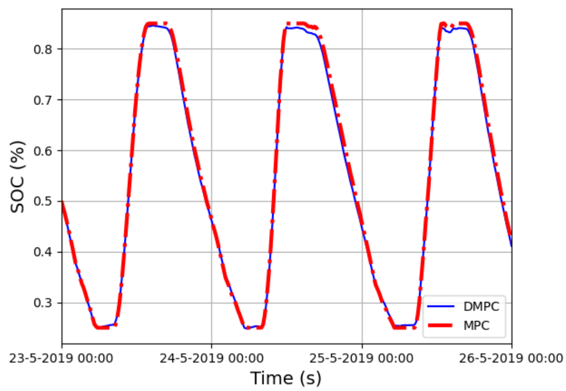


FIGURE 4. SOC profiles of the two considered methodologies.

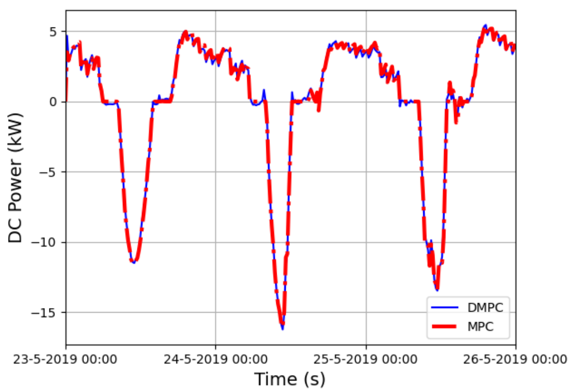


FIGURE 5. Battery power profiles of the two considered methodologies.

C. COMPUTATIONAL TIME COMPARISON

The primary objective of the DMPC approach is to significantly reduce the computational burden during online control, which is a major challenge in the deployment of traditional MPC while still maintaining high performance. In this paragraph, we compare the two models (DMPC and MPC) in terms of computational cost with increasing length of the control horizon. As previously mentioned, we consider the horizons $H \in \{4, 12, 24, 48, 96, 192, 384\}$, and for

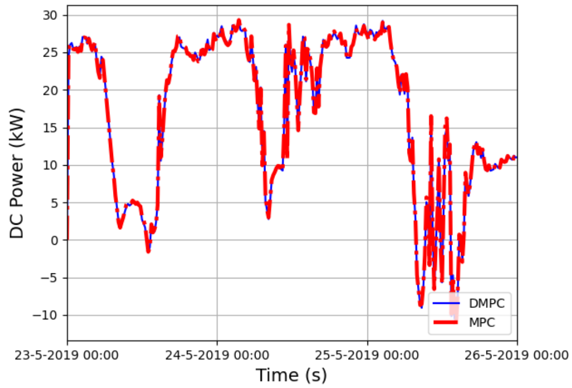


FIGURE 6. Network power profiles of the two considered methodologies.

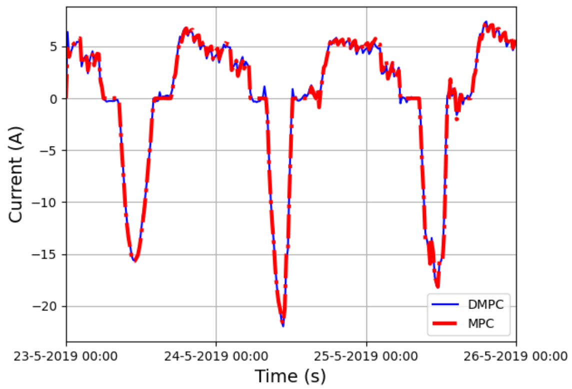


FIGURE 7. Applied current profiles of the two considered methodologies.

each value of the horizon, we collect the computational time required to retrieve the optimal action using both DMPC and MPC over a distribution of 10 simulations each of them consisting of 288 steps. In this way, it is possible to obtain statistical information related to the computational burden, e.g. its mean and standard deviation. The results are graphically represented in Figure 9, where the solid lines represent the average computational cost (green for the MPC, and orange for DMPC), while the numerical value for the average computational costs of the two discussed methodologies for different prediction horizons is reported in while Table 6.

From the graph, it is evident that MPC's computational time increases superlinearly with the prediction horizon due to the escalating complexity of the optimization problem solved at each step. In contrast, DMPC's computational time remains nearly constant, with a mean of 53.4ms and a standard deviation of 7.8ms, reflecting the algorithm's advantage of having a fixed complexity, $O(1)$, that is independent of both the prediction horizon and the system's dimensions. This is achieved as the DMPC algorithm primarily involves executing a pre-trained neural network model during online operations.

It is interesting to notice that for a very small horizon (around 4 steps) the computational cost of MPC may be lower than the one of the proposed approach since the evaluation

TABLE 5. Comparative parameters for MPC and DMPC for the case of $H = 48$.

	MPC	DMPC
Number of Iterations	5.6 ± 1.2	N/A
Approximation Error	N/A	$0.24\text{A} \pm 2.1\text{A}$
Offline Training Time	N/A	$1.8\text{s} \pm 0.2\text{s}$

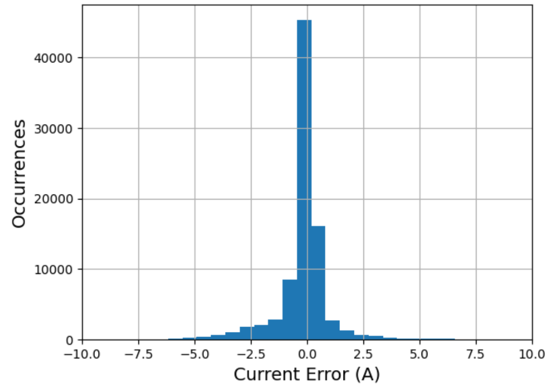


FIGURE 8. Histogram of the errors observed in 300 simulations by the DMPC approach in approximating the expert MPC controller.

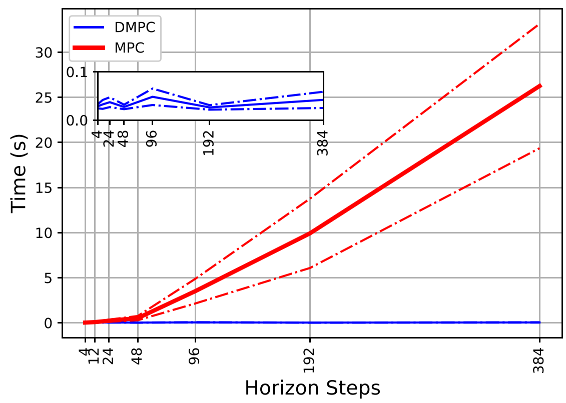


FIGURE 9. This figure presents a comparative analysis of the computational time required online by the MPC (depicted in red) and DMPC (shown in blue) as the prediction horizon expands. Data is derived from an average of 10 simulations, each with varying initial States of Charge (SoCs) and spanning 288 steps. A detailed zoom on the DMPC's computational time is also included for enhanced visibility.

TABLE 6. Average computational times for MPC and DMPC for different values of the prediction horizon H.

H	4	12	24	48	96	192	384
MPC	0.03s	0.08s	0.24s	0.52s	3.51s	9.93s	26.3s
DMPC	0.05s	0.05s	0.06s	0.05s	0.07s	0.04s	0.06s

of the neural network exhibits a certain fixed computational cost. However, a short prediction horizon is rarely used in a practical scenario since the planning of the microgrid management needs to be on the long run, especially regarding the energy that has to be stored in advance i.e. the energy we needed during a probable cloudy day has to be stored in the battery the day before. Therefore, the proposed methodology appears suitable for the common cases in which the prediction horizon is higher or equal to 48 steps.

It is important to highlight that while the use of the neural network-based algorithm reduces the online computational

TABLE 7. MPC-based control strategies comparison in similar applications.

Reference (year)	Control Strategy	Application	Numerical Results	Main Findings
[35] Alarcón et al. (2022)	Economic MPC (EMPC)	Microgrid connected to main grid	EMPC vs MPC: - Run Time (s): 120; 150 - Iterations: 1000; 1500 - Accuracy (%): 95; 93	EMPC demonstrated superior economic performance and system efficiency compared to other methods.
[36] Umaid et al. (2023)	Consensus-based MPC	DC Microgrid	MPC vs Sliding Mode Control (SMC): - Rise Time (s): 0.03; 0.04 - Settling Time (s): 10e-06; 2x10e-04 - Overshoot (%): 4.2; 9.9	Consensus-based MPC improved system stability, power sharing, and voltage regulation compared to traditional methods.
[37] Pereira et al. (2021)	Nonlinear MPC (NMPC)	Fuel Cell Hybrid Electric Vehicle	NMPC: - Prediction Horizon (Steps): 5, 10, 15 - Execution Time (ms): 20, 100, 250	Nonlinear MPC improved fuel economy and emissions compared to other methods, especially under dynamic driving conditions.
[38] Real et al. (2024)	Deep Reinforcement Learning (DRL)	Photovoltaic-Battery System	DLR vs Mixed Integer Linear Programming (MILP) Execution time in seconds (s): - Weekly testing sets: 1.3; 45.6 - Continuous testing set: 3.6; 312.0 - Testing house: 45.7; 4468.5	Deep Reinforcement Learning optimized photovoltaic-battery system operation considering load forecasting.
[39] Sen and Kumar (2023)	Distributed Adaptive-MPC	Isolated Microgrid	Distributed Adaptive-MPC Simulation Horizon (h): 24, 48, 72 Computational Time (min): 7.43, 9.86, 14.03	Distributed Adaptive-MPC improved system performance and robustness in isolated microgrids with uncertain renewable generation.

time, it incurs a fixed initial computational effort to generate the training dataset and train the neural network itself. However, this offline computational cost is considered secondary compared to the high cost of online computations. In this context, the deep learning-based method is effective as it transfers the majority of the computational burden offline, thus facilitating real-time applications. This trade-off ensures that the proposed DMPC approach is practical for scenarios requiring long prediction horizons, where real-time performance is critical.

Besides, to strengthen the obtained results, the comparative Table 7 based on recent bibliography employing MPC energy management within microgrid systems is included.

MPC has emerged as one of the main techniques in this domain. For instance, Alarcón et al. [35] employed Economic MPC (EMPC) to enhance the economic performance and efficiency of microgrids connected to the main electrical grid. In the context of DC microgrids, Ali et al. [36] proposed a consensus-based MPC approach to improve system stability, power sharing, and voltage regulation. Beyond microgrids, Pereira et al. [37] demonstrated the potential of nonlinear MPC (NMPC) in optimizing the energy management of fuel cell hybrid electric vehicles.

The integration of renewable energy sources, particularly photovoltaic systems, has stimulated the advancement of innovative control methods. Real et al. [38] showcased the effectiveness of deep Reinforcement Learning (DLR) in optimizing photovoltaic-battery system operation, incorporating load forecasting for enhanced performance. Furthermore, Sen and Kumar [39] introduced a distributed adaptive-MPC approach to address the challenges of energy management in isolated microgrids, emphasizing system robustness and performance under uncertain conditions.

D. IMPLEMENTATION DETAILS

The simulations detailed in this paper were executed on a Windows 11 Pro computer equipped with 64 GB of RAM and an i7-13700F processor, within a Python environment (version 3.11). The equations of the model predictive control were integrated and solved using CasADi [40], a symbolic framework well-suited for automatic differentiation and optimization tasks. The neural network central to the DMPC algorithm was implemented and trained using TensorFlow 2.0 [41].

Importantly, while the specific computational times in seconds are influenced by the hardware configuration used for the simulations, the principal advantage of the DMPC approach—its constant computational complexity with respect to variations in the prediction horizon and in the system's dimensions—does not depend on the specific implementation setting. This ensures that DMPC performance remains robust against increases in system complexity or extension of the prediction horizon, independently on the hardware used. This hardware-agnostic characteristic highlights our approach's scalability and adaptability, making it particularly effective for diverse real-world applications where computational resources vary.

VI. CONCLUSION

In this paper, the concept of deep model predictive control is applied for the first time to the optimal management of a microgrid. The strength of such methodology lies in its ability to approximate the predictive control, thanks to the learning capabilities of neural networks, thus reducing the real-time computational cost. While traditional predictive control faces a computational complexity that grows superlinearly with both the length of the prediction horizon and the system's dimensions, DMPC achieves a constant computational complexity irrespective of these factors. This fundamental change from a superlinear to a constant computational curve dramatically reduces the time required to compute the optimal action, streamlining the process significantly in real-time applications. A proper dataset generation is required to enhance the generalization capabilities of the deep-learning algorithm. Set aside the offline computational cost of this operation, the time required by DMPC to perform online has proven to be a small fraction of the time spent by the standard MPC. This latter in fact relies on the solution of an optimization problem at each step, which may be particularly intensive especially for longer control horizons, as it is typical in real-world scenarios. As shown by the presented results, the DMPC approach is capable of replicating the states and the input profiles of the standard methodology, except for an approximation error which is negligible from an application point of view. Note that the objective was not to outperform MPC in managing the microgrid, as MPC is already demonstrated to work well in this context, but rather to mimic its behavior through imitation learning to eliminate one of its main limitations: the online computational burden.

Looking ahead, this study serves as a preliminary investigation, setting the stage for subsequent applications to real microgrid infrastructures. Future expansions of this research could include the integration of other storage technologies, such as hydrogen, which could provide additional flexibility and resilience. Additionally, exploring the fault ride-through capabilities of the energy management system in the context of fault conditions could offer valuable insights into improving system robustness and reliability. These directions will help refine and extend the applicability of DMPC, potentially revolutionizing microgrid management under varying and challenging conditions.

REFERENCES

[1] A. H. Fathima and K. Palanisamy, "Optimization in microgrids with hybrid energy systems—A review," *Renew. Sustain. Energy Rev.*, vol. 45, pp. 431–446, May 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032115000696

[2] M. A. Jirdehi, V. S. Tabar, S. Ghassemzadeh, and S. Tohidi, "Different aspects of microgrid management: A comprehensive review," *J. Energy Storage*, vol. 30, Aug. 2020, Art. no. 101457. https://www.sciencedirect.com/science/article/pii/S2352152X19305109

[3] A. Vasilakis, I. Zafeiratou, D. T. Lagos, and N. D. Hatziargyriou, "The evolution of research in microgrids control," *IEEE Open Access J. Power Energy*, vol. 7, pp. 331–343, 2020.

[4] A. Villalón, M. Rivera, Y. Salgueiro, J. Muñoz, T. Dragičević, and F. Blaabjerg, "Predictive control for microgrid applications: A review study," *Energies*, vol. 13, no. 10, p. 2454, May 2020.

[5] J. Hu, Y. Shan, J. M. Guerrero, A. Ioinovici, K. W. Chan, and J. Rodriguez, "Model predictive control of microgrids—An overview," *Renew. Sustain. Energy Rev.*, vol. 136, Feb. 2021, Art. no. 110422.

[6] E. F. Camacho, C. Bordons, E. F. Camacho, and C. Bordons, *Model Predictive Controllers*. Cham, Switzerland: Springer, 2007.

[7] M. Petrollese, L. Valverde, D. Cocco, G. Cau, and J. Guerra, "Real-time integration of optimal generation scheduling with MPC for the energy management of a renewable hydrogen-based microgrid," *Appl. Energy*, vol. 166, pp. 96–106, Mar. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261916000337

[8] S. Batiyah, N. Zohrabi, S. Abdelwahed, and R. Sharma, "An MPC-based power management of a PV/battery system in an islanded DC microgrid," in *Proc. IEEE Transp. Electrific. Conf. Expo. (ITEC)*, Jun. 2018, pp. 231–236. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8450155

[9] Y. Zhang, F. Meng, R. Wang, W. Zhu, and X.-J. Zeng, "A stochastic MPC based approach to integrated energy management in microgrids," *Sustain. Cities Soc.*, vol. 41, pp. 349–362, Aug. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210670718303378

[10] B. Karg and S. Lucia, "Efficient representation and approximation of model predictive control laws via deep learning," *IEEE Trans. Cybern.*, vol. 50, no. 9, pp. 3866–3878, Sep. 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9123670/

[11] Y. Wang and S. Boyd, "Fast model predictive control using online optimization," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 2, pp. 267–278, Mar. 2010.

[12] T. Parisini and R. Zoppoli, "A receding-horizon regulator for nonlinear systems and a neural approximation," *Automatica*, vol. 31, no. 10, pp. 1443–1451, Oct. 1995. [Online]. Available: https://www.sciencedirect.com/science/article/pii/000510989500044W

[13] A. Cabrera-Tobar, N. Blasuttigh, A. M. Pavan, and G. Spagnuolo, "Demand response of an electric vehicle charging station using a robust-explicit model predictive control considering uncertainties to minimize carbon intensity," *Sustain. Energy, Grids Netw.*, vol. 38, Jun. 2024, Art. no. 101381. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352467724001103

[14] A. Bemporad, A. Oliveri, T. Poggi, and M. Storace, "Ultra-fast stabilizing model predictive control via canonical piecewise affine approximations," *IEEE Trans. Autom. Control*, vol. 56, no. 12, pp. 2883–2897, Dec. 2011. [Online]. Available: https://ieeexplore.ieee.org/document/5746505

[15] L. H. Cseko, M. Kvasnica, and B. Lantos, "Explicit MPC-based RBF neural network controller design with discrete-time actual Kalman filter for semiactive suspension," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 5, pp. 1736–1753, Sep. 2015. [Online]. Available: https://ieeexplore.ieee.org/document/7024892

[16] S. Chen, K. Saulnier, N. Atanasov, D. D. Lee, V. Kumar, G. J. Pappas, and M. Morari, "Approximating explicit model predictive control using constrained neural networks," in *Proc. Annu. Amer. Control Conf. (ACC)*, Jun. 2018, pp. 1520–1527.

[17] M. Hertneck, J. Köhler, S. Trimpe, and F. Allgöwer, "Learning an approximate model predictive controller with guarantees," *IEEE Control Syst. Lett.*, vol. 2, no. 3, pp. 543–548, Jul. 2018.

[18] B. Zheng, S. Verma, J. Zhou, I. W. Tsang, and F. Chen, "Imitation learning: Progress, taxonomies and challenges," *IEEE Trans. Neural Netw. Learn. Syst.*, pp. 1–16, 2022.

[19] H. Sorouri, A. Oshnoei, M. Novak, F. Blaabjerg, and A. Anvari-Moghaddam, "Learning-based model predictive control of DC–DC buck converters in DC microgrids: A multi-agent deep reinforcement learning approach," *Energies*, vol. 15, no. 15, p. 5399, Jul. 2022. [Online]. Available: https://www.mdpi.com/1996-1073/15/15/5399

[20] M. Andalibi, M. Hajihosseini, S. Teymoori, M. Kargar, and M. Gheisarnejad, "A time-varying deep reinforcement model predictive control for DC power converter systems," in *Proc. IEEE 12th Int. Symp. Power Electron. Distrib. Gener. Syst. (PEDG)*, Jun. 2021, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9494214

[21] A. Pozzi, S. Moura, and D. Toti, "A deep learning-based predictive controller for the optimal charging of a lithium-ion cell with non-measurable states," *Comput. Chem. Eng.*, vol. 173, May 2023, Art. no. 108222. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0098135423000923

[22] A. Pozzi and D. Toti, "Imitation learning for agnostic battery charging: A DAGGER-based approach," *IEEE Access*, vol. 11, pp. 115190–115203, 2023.

[23] M. S. H. Lipu, S. Ansari, M. S. Miah, S. T. Meraj, K. Hasan, A. S. M. Shihavuddin, M. A. Hannan, K. M. Muttaqi, and A. Hussain, "Deep learning enabled state of charge, state of health and remaining useful life estimation for smart battery management system: Methods, implementations, issues and prospects," *J. Energy Storage*, vol. 55, Nov. 2022, Art. no. 105752.

[24] R. H. Lasseter and P. Paigi, "Microgrid: A conceptual solution," in *Proc. IEEE 35th Annu. Power Electron. Spec. Conf.*, Sep. 2004, pp. 4285–4290.

[25] M. Faisal, M. A. Hannan, P. J. Ker, A. Hussain, M. B. Mansor, and F. Blaabjerg, "Review of energy storage system technologies in microgrid applications: Issues and challenges," *IEEE Access*, vol. 6, pp. 35143–35164, 2018.

[26] A. Pozzi and D. Toti, "Lexicographic model predictive control strategy in ageing-aware optimal charging procedure for lithium-ion batteries," *Comput. Chem. Eng.*, vol. 163, Jul. 2022, Art. no. 107847.

[27] B. Mbuwir, F. Ruelens, F. Spiessens, and G. Deconinck, "Battery energy management in a microgrid using batch reinforcement learning," *Energies*, vol. 10, no. 11, p. 1846, Nov. 2017. [Online]. Available: https://www.mdpi.com/1996-1073/10/11/1846

[28] A. Incremona, G. De Nicolao, F. Fusco, B. J. Eck, and S. Tirupathi, "Aggregation of nonlinearly enhanced experts with application to electricity load forecasting," *Appl. Soft Comput.*, vol. 112, Nov. 2021, Art. no. 107857.

[29] V. Prema, M. S. Bhaskar, D. Almakhles, N. Gowtham, and K. U. Rao, "Critical review of data, models and performance metrics for wind and solar power forecast," *IEEE Access*, vol. 10, pp. 667–688, 2022.

[30] A. Bemporad and M. Morari, "Robust model predictive control: A survey," in *Robustness in Identification and Control*. Cham, Switzerland: Springer, 1999, pp. 207–226.

[31] B. M. Åkesson and H. T. Toivonen, "A neural network model predictive controller," *J. Process Control*, vol. 16, no. 9, pp. 937–946, Oct. 2006. [Online]. Available: https://api.semanticscholar.org/CorpusID:16016011

[32] A. Nespoli, A. Niccolai, E. Ogliari, G. Perego, E. Collino, and D. Ronzio, "Machine learning techniques for solar irradiation nowcasting: Cloud type classification forecast through satellite data and imagery," *Appl. Energy*, vol. 305, Jan. 2022, Art. no. 117834. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261921011600

[33] M. Wood, E. Ogliari, A. Nespoli, T. Simpkins, and S. Leva, "Day ahead electric load forecast: A comprehensive LSTM-EMD methodology and several diverse case studies," *Forecasting*, vol. 5, no. 1, pp. 297–314, Mar. 2023. [Online]. Available: https://www.mdpi.com/2571-9394/5/1/16

[34] S. Matrone, E. Ogliari, A. Nespoli, and S. Leva, "Electric vehicle supply equipment day-ahead power forecast based on deep learning and the attention mechanism," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 8, pp. 9563–9571, Aug. 2024.

[35] M. A. Alarcón, R. G. Alarcón, A. H. González, and A. Ferramosca, "Economic model predictive control for energy management of a microgrid connected to the main electrical grid," *J. Process Control*, vol. 117, pp. 40–51, Sep. 2022.

[36] S. U. Ali, A. Waqar, M. Aamir, S. M. Qaisar, and J. Iqbal, "Model predictive control of consensus-based energy management system for DC microgrid," *PLoS ONE*, vol. 18, no. 1, Jan. 2023, Art. no. e0278110.

[37] D. F. Pereira, F. d. C. Lopes, and E. H. Watanabe, "Nonlinear model predictive control for the energy management of fuel cell hybrid electric vehicles in real time," *IEEE Trans. Ind. Electron.*, vol. 68, no. 4, pp. 3213–3223, Apr. 2021.

[38] A. C. Real, G. P. Luz, J. M. C. Sousa, M. C. Brito, and S. M. Vieira, "Optimization of a photovoltaic-battery system using deep reinforcement learning and load forecasting," *Energy AI*, vol. 16, May 2024, Art. no. 100347.

[39] S. Sen and V. Kumar, "Distributed adaptive-MPC type optimal PMS for PV-battery based isolated microgrid," *IEEE Syst. J.*, vol. 17, no. 1, pp. 546–557, Mar. 2023.

[40] J. A. E. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "CasADi: A software framework for nonlinear optimization and optimal control," *Math. Program. Comput.*, vol. 11, no. 1, pp. 1–36, Mar. 2019, doi: 10.1007/s12532-018-0139-4.

[41] M. Abadi et al., "TensorFlow: A system for large-scale machine learning," in *Proc. 12th USENIX Conf. Operating Syst. Design Implement.*, Nov. 2016, pp. 1–28.

SILVANA MATRONE (Student Member, IEEE) received the B.Sc. degree in mechanical engineering from the University of Naples Federico II, in 2019, and the M.Sc. degree in energy engineering from the Politecnico di Milano, in 2022, where she is currently pursuing the Ph.D. degree with the Electrical Division, Energy Department. Her research interest includes machine learning techniques with applications in energy dispatching systems.

ANDREA POZZI (Member, IEEE) received the bachelor's degree in industrial engineering, the master's degree in electrical engineering, and the Ph.D. degree in electronics, informatics, and electrical engineering from the University of Pavia, in 2015, 2017, and 2021, respectively. He has held visiting scholar positions at TU Braunschweig, in 2016, and UC Berkeley, in 2019. After working as a Postdoctoral Researcher with the University of Pavia, he joined the Faculty of Mathematical, Physical, and Natural Sciences, Catholic University of the Sacred Heart, Brescia, Italy, as an Assistant Professor of machine learning, in January 2022. His research interests include reinforcement learning, imitation learning, machine learning, approximate dynamic programming, and advanced control theory.

EMANUELE OGLIARI (Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from the Politecnico di Milano, Italy, in 2016. He has been working on photovoltaic power plant design and their optimization, since 2010, and RES expected power by means of computational intelligence techniques, since 2012. He has collaborated with international universities and research centers (The University of Danang-University of Science and Technology, the Ben-Gurion University of the Negev, and Global Energy Interconnection Research Institute Europe), in 2013, he won the joint laboratories between Italy and Israel on Solar Energy ILSE. He is currently an Associate Professor with the Department of Energy, Politecnico di Milano, where he teaches with the Electrical Engineering and Photovoltaic-Based Systems Laboratory.

SONIA LEVA (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the Politecnico di Milano, Italy, in 2001. She is currently a Full Professor of electrical engineering with the Department of Energy, Politecnico di Milano. She is also the Director of the Solar Tech Laboratory and the Laboratory of Microgrids, Politecnico di Milano. She is a member of the IEEE Working Group "Distributed Resources: Modeling and Analysis" and a Senior Member of the IEEE Power and Energy Society.

Open Access funding provided by 'Politecnico di Milano' within the CRUI CARE Agreement