This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

A User Based HVAC System Management Through Blockchain Technology and Model Predictive Control

Giuseppe Olivieri, Gaetano Volpe, Agostino Marcello Mangini, *Senior Member, IEEE*, and Maria Pia Fanti, *Fellow, IEEE*

***Abstract*— This paper introduces an innovative approach to designing a user-based Heating, Ventilation, and Air-Conditioning (HVAC) system management connected with the District Energy Management System. By classifying the users into dynamic energy consumption classes to reward energy efficiency and penalize excessive use, users can modify their behavior to pass to a less expensive and more virtuous consumption class. To this aim, a blockchain platform determines the rewards and penalties and, by a K-means clustering algorithm, categorizes users into respective groups. Then, a Class Follower Problem is formulated and solved by a Model Predictive Control (MPC) strategy integrated with a Long Short-Term Memory network as a predictive model. If the users follow the suggestions proposed by the controller, i.e., the thermostat set-points and the time intervals in which the HVAC system must be switched off or on, the users can be located in a more virtuous consumption class. A case study conducted within an energy district in Bari (Italy) shows how the proposed architectural framework tuned thermal regulation in intelligent buildings while concurrently achieving energy optimization.**

***Note to Practitioners*—This paper addresses the challenge of efficiently managing HVAC systems in smart districts through a novel blockchain-based framework and an optimization strategy solved by an MPC approach. The objective is to incentivize users to optimize their energy consumption by introducing dynamic Consumption Classes that reward energy efficiency and penalize inefficient utilization. For practitioners, this strategy translates to a granular level of energy management that not only adapts to individual behaviors but also aligns with broader sustainability goals. Integrating the blockchain platform ensures a transparent and secure method for managing and recording energy usage. At the same time, adopting MPC with Long Short-Term Memory Networks offers accurate forecasts and adjustments to enhance system responsiveness. Although the study focuses on HVAC systems, the principles may be extended to other energy-intensive applications, providing a comprehensive tool for energy management and user engagement in smart cities. Future research could integrate renewable energy sources and explore the implications of user-driven adjustments on the overall energy distribution and efficiency.**

Manuscript received 18 January 2024; accepted 28 March 2024. This article was recommended for publication by Associate Editor F. Tedesco and Editor Q. Zhao upon evaluation of the reviewers' comments. *(Corresponding author: Maria Pia Fanti.)*

The authors are with the Department of Electrical and Information Engineering, Polytechnic University of Bari, 70126 Bari, Italy (e-mail: g.olivieri@phd.poliba.it; gaetano.volpe@poliba.it; agostinomarcello.mangini@poliba.it; mariapia.fanti@poliba.it).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TASE.2024.3397561.

Digital Object Identifier 10.1109/TASE.2024.3397561

***Index Terms*— HVAC management system, blockchain, model predictive control, long short-term memory.**

I. Introduction

THE quest for energy efficiency and consumption control is fundamental from an environmental perspective. Fifty percent of building energy consumption is used in Heating, Ventilation, and Air-Conditioning (HVAC) systems [1]. Hence, effective HVAC control techniques for energy consumption minimization and thermal comfort guarantee have attracted the attention of researchers. In order to simultaneously maximize comfort and minimize energy consumption, it is essential to manage the building network (district) in a way that optimally balances real-time energy usage.

While existing regulations have initiated some changes, traditional penalty-based systems often fail to encourage full user compliance. Emerging researchers advocate for integrating blockchain technology, like proposing a novel system that dynamically rewards or penalizes users based on their real-time energy consumption. This approach promises enhanced security and privacy, addressing key concerns in energy management [2]. The concept of bestowing complimentary energy credits to domestic end-users as a means to mitigate the demands during periods of peak loads, leveraging a direct load control mechanism envisaged by Erdinç et al. [3], has also undergone scrutiny. Furthermore, Shi et al. work [4] elucidates the efficacy of incentive-based demand response in eliciting shifts in consumer energy usage, providing a refined methodology for assessing load profiles and attenuating peak loads through economic inducements.

In the context of decentralized energy management, it is crucial to consider the market dynamics. Mnatsakanyan et al. [5] propose an innovative electricity market structure emphasizing individual pricing mechanisms for fairer demand response benefits distribution. Brahmia et al. [6] highlight the challenges of electricity price forecasting in multi-microgrid systems, advocating for advanced predictive models. Additionally, using reinforcement learning, Biemann et al. [7] explore real-time pricing optimization in data center HVAC control, underscoring the need for responsive management solutions in fluctuating market conditions.

Advanced control and prediction algorithms, such as Long Short-Term Memory (LSTM) networks, have been utilized effectively for time-series prediction in energy management, capturing temporal dependencies for accurate real-time predictions and enabling more efficient energy resource utilization [8]. In parallel, Model Predictive Control (MPC) has proven pivotal for the dynamic optimization of complex systems like multi-zone buildings, which is critical in balancing energy efficiency with user comfort [9]. Integrating energy management algorithms with blockchains for secure communication is a substantial advancement. However, there are still notable gaps, particularly in tracking financial transactions and enhancing user comprehension of their energy usage patterns.

This paper proposes an HVAC control system connected to the District Energy Management System (DEMS) and devoted to improving the HVAC consumptions' user management. In detail, the DEMS receives the electricity invoice from the energy supplier at the end of the billing cycle and other data concerning weather conditions, calendar days, and law limitations. Such data are notarized via a dedicated transaction on a blockchain platform.

Moreover, at the end of the billing period, the blockchain classifies the users into Consumption Classes, and rewards or penalties are assigned to each class. Hence, the blockchain decides all the user payments based on a virtuous classification performed by a K-means clustering algorithm proposed in our previous paper [10].

At this point, the users can decide to modify their behavior to pass to a less expensive class. To this aim, a Model Predictive Control (MPC) strategy is locally applied by the user to determine the thermostat set-points of the HVAC system and the intervals in which the system HVAC must be switched off or switched on. A lexicographic optimization allows the user to minimize energy consumption by guaranteeing the user's comfort. In addition, an LSTM-based method is presented to determine the building's thermodynamic model and HVAC energy consumption. Then, the MPC approach solves the lexicographic optimization problem by using the building thermodynamic model and the HVAC energy consumption description obtained by the LSTM pre-trained network.

The new contribution of the paper is twofold.

First, the proposed HVAC control system is connected with the DEMS pricing and classifies users in the consumption classes using a K-means clustering algorithm. The approach's novelty also lies in the use of the blockchain platform, which guarantees accountability, transparency, and data notarization.

Second, we propose a methodology that the users can implement to pass to a more virtuous class. Even if applying the MPC approaches and the LSTM-based methods are familiar in the HVAC control, we formulate and solve a novel lexicographic minimization problem. The solution can be applied by the users to suitably manage the HVAC system and reach the desired *Consumption Class* by satisfying their comfort.

The system architecture incentivizes users to shift towards more virtuous energy practices, ensuring that energy suppliers do not bear the consequences of inefficient consumption. By implementing a penalty-reward system rooted in blockchain technology, we advocate for a self-regulating user community where sustainable actions are incentivized and wasteful habits are discouraged.

The remaining structure of the paper is as follows: Section II presents the literature review about the role of blockchain technology, Machine Learning (ML), MPC and LSTM methods within Smart Grids (SG) and DEMS. Moreover, Sections III and IV describe the proposed HVAC Control System and the blockchain platform architecture, respectively. In addition, Section V formulates the district user clustering and the class follower problem. Section VI designs the control system based on the MPC strategy and the LSTM thermodynamic model. Finally, Section VII discusses the case study and Section VIII draws the conclusions.

II. Literature Review

A. Blockchain in Smart Grid and District Energy Management

With its distributed ledger architecture, blockchain technology is pivotal in enhancing the reliability and security of SG and district energy systems. Each transaction within this system is meticulously cataloged in structured blocks, linked sequentially to form a sequential or chain-like structure. Each block in the blockchain is securely linked to its predecessor by incorporating the output of the Secure Hash Algorithm 256 (SHA-256), commonly referred to as the *hash*, of the previous block, which is included in the structure of the current block [11]. The slightest variation would produce a completely different control number (**H**), also known as a hash number. The strength of storing data in this way lies in the mathematical certainty of the correctness of the data stored. The same data processed with the same hashing function will return the same number **H** in a deterministic way.

In SG, blockchain can be employed to transparently manage energy distribution, recording transactions from energy production to consumption. This ensures integrity in the trade of renewable energy certificates and facilitates real-time billing for consumers. This system could be applied to a database to certify data integrity. However, the innovation brought by blockchain lies in the fact that blocks include the **H** of the previous block within the input information [12].

In the context of SG security and privacy, various studies have highlighted the challenges and opportunities arising from emerging technologies, such as ML and blockchain [13], [14]. Mirzaee et al. [13] explored security and privacy challenges in SGs, including vulnerabilities and potential attacks in evolving power networks, emphasizing the need for additional research into security and privacy mechanisms. Furthermore, the authors discussed the growing use of ML algorithms in SG components for attack detection and threat analysis, highlighting the susceptibility of ML systems to adversarial attacks.

As previously mentioned, blockchain, a distributed technology that, thanks to its structure, enhances system redundancy and resilience to failures and cyberattacks, has emerged as a promising application within the SG (Smart Grid) paradigm [14]. In the pursuit of building Smart Cities, a smart district model has been designed, leveraging new technologies and efficient energy management systems integrated into an Internet of Things (IoT) and blockchain platform [15]. Christidis and Devetsikiotis [16] delved into the integration of blockchains and smart contracts with IoT, illustrating how these technologies could foster a marketplace of services between devices and automate multi-step processes in a cryptographically verifiable manner. In a similar vein, Benedict et al. [17] introduced an IoT blockchain solution, i.e., an IoT-enabled blockchain for air quality monitoring systems in smart cities. Implementing *chaincodes* for air quality monitoring systems, the proposed architecture addressed prevailing security and performance challenges associated with IoT cloud solutions.

Moreover, blockchain has been applied in the SG for cybersecurity [18]. Kosba et al. [19] presented a decentralized smart contract system that ensures transactional privacy in decentralized cryptocurrencies, enabling programmers to write private smart contracts without implementing cryptography directly, as the compiler automatically generates an efficient cryptographic protocol. Li et al. [20] provided a quantitative and qualitative review of blockchain research from 2015 to 2021, identifying six research hotspots and five research frontiers to offer a comprehensive view of recent trends in the field. Malla et al. [21] conducted a state-of-the-art review on the status, challenges, and future directions of blockchain technology in power systems, discussing interfaces and possibilities that can ensure trust, security, and transparency, facilitating a decentralized power system and power market. The Hyperledger Fabric, a modular and extensible open-source blockchain system, is a promising solution for supply chain management, allowing customization for specific use cases and trust models without relying on a native cryptocurrency [22]. Although it is important to acknowledge the inherent limitations associated with its nature as a closed and non-public platform, the Hyperledger platform enables a secure and efficient way to track and manage transactions and assets throughout the entire supply chain.

B. Optimization Models for Energy Management

In recent years, ML algorithms have been applied to energy consumption prediction in smart buildings, examining the performance of Support Vector Regression, Artificial Neural Networks, and Random Forest algorithms. Wu and Chu identified in their study *Random Forest* as the best-performing algorithm and investigated the impact of sampling strategy on prediction accuracy. They discovered that increasing sampling density in high variance data enhanced prediction results, which can be employed to optimize ML algorithms for building energy consumption prediction, ultimately contributing to energy conservation, environmental protection, and smart city development [23].

Another study by Roccotelli et al. addressed the energy management issue in cooperative microgrids within a smart energy district [24]. It proposed an optimization model that aims to maximize the use of energy purchased at the day-ahead market, minimizes the need for expensive real-time energy, and optimizes the integration of renewable energy sources, energy storage systems, and electric vehicle batteries. In order to tackle the uncertainties of key parameters, the proposed optimization model was solved using two approaches: one deterministic and one stochastic.

Muralidar et al. [25] emphasized the need for integrating blockchain and ML technologies in energy management systems to enhance efficiency, reduce costs, and support the implementation of renewable technologies in smart buildings. In their review, the energy management systems are at the center of monitoring and controlling energy needs in industrial buildings, underlining the necessity for such systems to address energy use efficiency improvement, energy cost reduction, and renewable energy technology implementation to cater to local energy loads in structures with distributed resources. Rajith et al. [26] pioneered the development of a real-time optimized HVAC control system using IoT, which was built upon an IoT framework that collected thermal parameters from sensors and user feedback information for real-time processing in a distributed cloud environment. Incorporating optimization techniques, demand response, and predictive models in their system led to a 20%–40% reduction in energy consumption while maintaining user thermal comfort.

C. Model Predictive Control for Building Climate Management

The utility of MPC in sustainable building management is increasingly recognized, especially when integrated with data-driven methodologies. Chen et al. [27] introduce a Data-Driven Robust MPC framework, which tackles the prevalent issue of weather forecast uncertainty. This work aligns with the thrust of our proposal to enhance climate control strategies in buildings. By incorporating ML techniques for constructing uncertainty sets, paper [27] establishes a foundation upon which our research builds, particularly in developing a tailored predictive model that accounts for the unique climatic and architectural characteristics of our focus buildings.

On the topic of learning-based approaches, the work of Eini and Abdelwahed [28] is noteworthy for its integration of Artificial Neural Networks (ANNs) with the MPC framework, resulting in notable energy savings and improved occupant comfort. Their method offers a proof of concept that resonates with our proposal's objective to optimize energy management while maintaining thermal comfort. Our research seeks to bridge the gap identified in their study by extending the learning-based control scheme to a wider range of building types and climatic conditions, ensuring broader applicability and scalability of the proposed solutions.

Recent research has explored the optimization of building energy management and indoor thermal comfort. Homod et al. [29] presented a hybrid model highlighting the importance of non-temperature factors in HVAC system references. The potential of MPC in HVAC systems has been underscored by multiple studies, emphasizing its utility in ensuring energy efficiency and thermal comfort. However, from a pricing perspective, the limitations on the operational time-frames of HVAC systems imposed by some nations are not optimal solutions, as they risk leading to exceeding energy peaks [30].

While these studies lay the groundwork for innovative climate control through MPC, our research intends to expand on these methodologies. We aim to explore the intersection of advanced control algorithms and emerging technologies, such as the IoT and edge computing, to enable more responsive and adaptive building management systems. By leveraging the strengths of the existing models and identifying areas for improvement, our proposal aspires to contribute to the evolution of smart building energy systems that are both efficient and responsive to the occupants' needs.

D. Long Short-Term Memory Networks in HVAC Systems

LSTM networks are special kinds of Recurrent Neural Networks (RNNs) that are gaining attention in the field of HVAC systems for their ability to model and predict time series data with long-term dependencies. LSTMs are particularly well-suited for HVAC load prediction because of their capability to remember information for long periods, which is essential for capturing the dynamics of energy consumption in buildings.

Friansa et al. [31] compared LSTM and bi-directional LSTM models for the prediction of HVAC electricity load based on daily datasets. Their findings showed that Bi-LSTM models yielded higher accuracy, with a Mean Absolute Percentage Error of 15.35%, suggesting that LSTMs could significantly improve the prediction accuracy for HVAC electricity load management.

In a similar way, Wang et al. [32] proposed a Distributed Fusion LSTM model to forecast temperature and relative humidity in smart buildings. Their model, which utilizes distributed data-fusion technology, outperformed other forecasting methods, including Support Vector Regression and classical LSTM, highlighting the efficacy of LSTMs in predicting key environmental variables that affect HVAC performance.

Alden et al. [33] explored the use of LSTM networks to separate HVAC energy use from total residential load, which can be pivotal for enhancing energy management systems in smart homes. They developed LSTM encoder-decoder models using future weather data, which proved to be effective in providing accurate day-ahead HVAC energy forecasts, thus facilitating more efficient energy management.

III. HVAC Control System

In this section, we propose an HVAC control system at the DEMS level to improve the user management of HVAC consumption. The system architecture is described in Fig. 1, which points out the main system components.

In detail, the DEMS receives the electricity invoice from the energy supplier at the end of the billing cycle and other data concerning weather conditions, calendar days, and law limitations. Such data are notarized via a dedicated transaction on the blockchain platform. On the other side, the IoT devices transmit HVAC states to the blockchain, such as the temperature at which the thermostat has been set.

The start-up and closing times of the billing cycle are managed in a distributed manner through the use of a Smart Contract. The blockchain also notarises the data collected into the Smart Contract with a time-stamp through the same code of the blockchain platform, which includes the transaction containing the information within a time-stamped block.

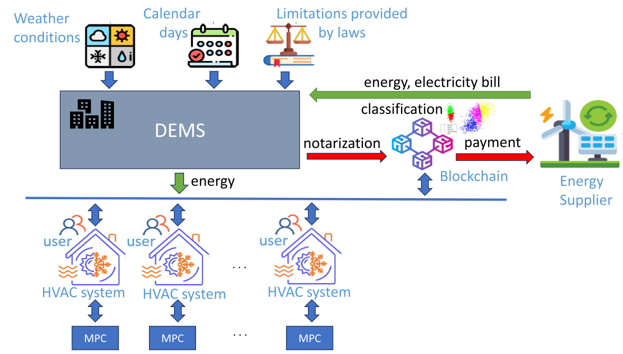


Fig. 1. System architecture.

At the end of the billing period, the blockchain classifies the users in n Consumption Classes and a reward or penalty is assigned to each class. Hence, the blockchain decides all the user payments based on a virtuous classification performed by a K-means clustering algorithm proposed in [10]. In particular, the users are classified into consumption classes by a K-means clustering algorithm, and each class is characterized by a multiplier coefficient that increases (penalty) or decreases (reward) the user energy costs. Such coefficients are notarized on the blockchain, which provides payment at the end of a billing cycle.

Obviously, a user belonging to a virtuous class contributes less to the payment of an electricity invoice than a user belonging to a less virtuous class due to rewards and penalties. For this reason, a user can be encouraged to move from a less virtuous class to a more virtuous one. The strategy to enable a user to change their behavior to pass a less expensive class is implemented by an MPC approach.

In particular, the MPC uses a predictive model to forecast future performance and determines a control action to meet predefined constraints and objectives over a designated horizon.

The significance of MPC in smart DEMS lies in its predictive power and adaptability. It allows the system to preemptively adjust HVAC settings in response to user behavior and external factors, ensuring that the energy consumption is simultaneously efficient and economical. The proposed MPC scheme provides the thermostat set-points of the HVAC system in the different zones of the building to allow the user to reach a less expensive consumption class.

IV. Blockchain Architecture

This section describes the blockchain platform and the design of the related Smart Contracts.

A. Blockchain Platform

It is worth recalling that reading information on the blockchain has no cost; on the contrary, writing can be expensive. This leads to the first critical issue to be addressed, which is the choice regarding the blockchain platform to be

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

implemented in order to prevent the user from incurring costs that exceed what is necessary.

On the client side, the system is required to record a substantial amount of data on the blockchain, as the consumption data needs to be notarized over time. To address this challenge, several solutions exist: one local approach involves processing the data locally, regularly generating hashes of the data, and then notarizing only these hashes on the blockchain. In this way, the control over the data would not be direct but would still guarantee high reliability. This is because the optimization tool would acquire the data via the Application Programming Interface (API) directly from the server, then submit them to the same hashing algorithm (deterministic), and eventually compare the hash calculated with the notarized one in the blockchain. If data inconsistency occurs, the system assumes manipulation and applies the maximum possible coefficient. This hybrid solution involves a high-risk factor, given the poor resilience of the system to errors. A single variation of data in the period under consideration would heavily penalize the user, and it cannot be assumed that all of these errors are attributable to system manipulations.

Conversely, systems that employ on-chain data storage demonstrate enhanced resilience, as exemplified by the subsequent proposed solutions. One alternative is a blockchain that is compatible with the Ethereum Virtual Machine (EVM), which benefits from reduced costs when compared to notarization fees associated with the main Ethereum Blockchain.

Alternatively, when full on-chain data writing and optimization execution are required, a more innovative approach is preferable. This is exemplified by an *Avalanche subnet*. To date, the Avalanche Blockchain offers a solution with lower access costs and more immediate usability, largely due to its governance policy managed by a Proof of Stake (PoS) consensus algorithm. In PoS blockchains, part of the nodes that contribute to archiving the blockchain's history can also write to it, composing the new blocks. The requirement is to lock (stake) some native cryptocurrencies so that any malicious actors can be penalized by eroding the staked capital (with different methodologies depending on the parameters of the consensus algorithm implemented).

The role of these specific nodes on a PoS blockchain is called "*validator*". The Avalanche Blockchain's validators can write contextually to the main Blockchain and also to different blockchains called *subnets*, and each subnet can be programmed individually [34]. An evolved use case envisages that a particularly advanced blockchain application such as the one being studied here can be implemented on a proprietary *subnet*. This solution cuts transaction costs, bringing them to a minimum. In order to have the subnet working smoothly, the developed blockchain application must provide incentives for the validators to attract them to validate the subnet.

Furthermore, while the blockchain itself is not anonymous, it offers a high degree of pseudonymity by allowing transactions and interactions through addresses that are not directly linked to the users' identities. This ensures verifiability and integrity without compromising privacy. Moreover, the advanced customizability of Avalanche's subnets permits the system to be tailored to adhere to specific legal and regulatory requirements. Adjusting the parameters of a subnet can ensure compliance with data protection laws, such as the GDPR in the European Union, thereby managing permissions and safeguarding sensitive data. This flexible configuration underscores our commitment to ethical and legal responsibility, balancing technological innovation with due diligence in a dynamic regulatory landscape.

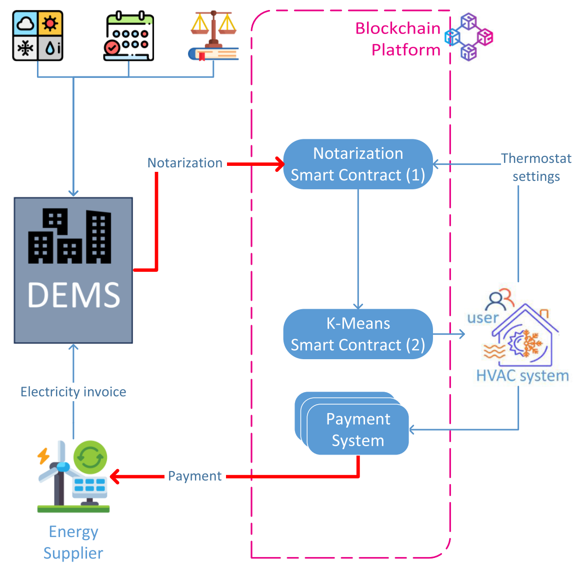


Fig. 2. The blockchain platform connected with the DEMS of Fig. 1.

It should also be remembered that *Avalanche* blockchain supports EVM and adopting the EVM allows for the seamless migration and verification of smart contracts' logic and state, thus reinforcing the advisability of developing on an EVM-based platform for enhanced flexibility and interoperability [35]. In the proposed scheme, we deploy a diptych Smart Contracts system to record the connected user's data and to manage the payment of the invoices.

B. Smart Contracts Design

Smart contracts, a revolutionary feature introduced with the advent of the Ethereum Blockchain, automate the execution of agreements, effectively preventing the need for traditional intermediaries. These contracts are considered "smart" because they can self-execute and self-enforce contract terms embedded in the blockchain's immutable ledger. Leveraging Ethereum's decentralized architecture, smart contracts encode obligations and conditions in code, thus creating a trustless environment where transactions, operations, or agreements are automatically executed once certain conditions are met [36].

In the presented system architecture, the DEMS sends and receives information and data through the blockchain platform that is connected with the building HVAC systems and the energy supplier (see Fig. 2). Utilizing smart contracts, the DEMS initiates the process by securely sending data to the blockchain, a step we refer to as "notarization". This process ensures that all energy and electricity bill-related data are immutable and verifiable, fostering trust and transparency in the energy trading market.

The notarization process started by the DEMS is performed by Smart Contract 1 that is designed to handle the complexity and variety of data generated by the DEMS. The contract acts as a notary, storing the relevant information transmitted by each dedicated transaction, thus ensuring data integrity and non-repudiation. This step lays the groundwork for accurate classification and subsequent billing.

The pseudo-code for the Smart Contract 1, named *Data Notarization*, is described in the following.

Code 1 Smart Contract 1: Data Notarization

```
contract Data Notarization {
    struct UserData {
        uint256 indoorTemp;
        uint256 outdoorTemp;
        uint256 seasonalTemp;
        uint256 lawTemp;
        uint256 activeIntv;
        uint256 totalIntv;
    }

    mapping(address => UserData)
      public userData;

    // Store user's data into the contract
    function store(
      address _user,
      UserData memory _data
    ) public {
        userData[_user] = _data;
    }
}
```

A structure called *UserData* is defined to encapsulate all the key metrics needed for each user. These metrics include both indoor and outdoor temperatures, seasonal averages, and the number of intervals during which the HVAC system is operational. The contract incorporates a function, denominated `store`, which enables the secure archival of data corresponding to each user on the blockchain. This step furnishes a reliable and immutable ledger that can be subsequently accessed for analytical endeavors.

Upon the notarization of energy data, Smart Contract 2, named *Kmeans Clustering*, determines the payment classes, the incentives and the penalties, also in collaboration with the energy supplier. The smart contract encodes the payment mechanism, ensuring that once the classification is complete and bills are calculated, payments are autonomously disbursed to the energy suppliers' accounts. The automated nature of these transactions reduces the latency and potential errors associated with manual processing, offering a streamlined and efficient payment process. This contract is initialized with the address of the *DataNotarization* Smart Contract, enabling it to fetch the necessary data for clustering. Specifically, it retrieves the notarized data from the blockchain to perform the clustering algorithm.

Successively, it invokes a K-means clustering algorithm `kmeans` that has three main tasks: 1) retrieving the user data stored in Smart Contract *DataNotarization*, 2) executing the K-means algorithm using such data, 3) allocating users to their respective clusters, 4) calculate the pricing through an internal function that can use a particular pricing model and 5) distribute these results to the users' accounts.

Utilizing separate smart contracts for data notarization and the K-means algorithm presents distinct advantages and disadvantages. On the positive side, separating these functionalities clearly delineates responsibilities, simplifying system management and future extensibility. Furthermore, this separation allows for easier scalability as each contract can be optimized for its specific task. The separation also offers the benefit of code reusability, particularly for the data notarization contract, which could be employed in various other contexts or projects. Conversely, the system's overall complexity could increase due to managing multiple contracts. Additionally, interacting between multiple contracts (when writing) may incur extra transaction costs in terms of gas. However, this concern is mitigated by the current implementation using an EVM-compatible environment with negligible gas costs for such operations.

Code 2 Smart Contract 2: K-Means

```
contract KmeansClustering {
  Data Notarization dataNotarization;

  // Initialize with DataNotarization
  // contract address constructor
  constructor(
    address _dataNotarizationAddr
  ) {
    dataNotarization =
    DataNotarization(_dataNotarizationAddr);
  }

  // Execute K-means algorithm and set
     prices
  function kmeans() public {
      // 1. Fetch data from
      //    DataNotarization contract
      // 2. Perform K-means clustering
      // 3. Assign users to clusters
      // 4. Calculate pricing based on
      //    clustering (internal fn)
      // 5. Distribute the pricing to
      //    accounts
  }
  // Calculate pricing
  function calculatePricing()
  internal {
      // it can use a pricing model
  }
}
```

V. District User Clustering and Class Follower Problem

In this section, the preliminary *District User Clustering* scheme based on *k-means* algorithm is formally described and the subsequent *Class Follower Problem* is introduced.

A. K-Means District User Clustering

Let us consider the set of m users $\mathcal{U} = \{u_i | i = 1, 2, .., m\}$ in the district. A *Consumption Class* $c_j \in \mathcal{C}$, with $\mathcal{C} = \{c_j | j = 1, \ldots, n\}$, is assigned to each user $u_i \in \mathcal{U}$. A *Consumption Class* c_j basically represents the user's consumption profile in a certain time frame $h \in N$, where N is the set of natural numbers, and a billing period is made of a certain number of time frames. As suggested in [10], the profile can be driven by the level of compliance of the user to the law and to the best practices in terms of energy saving and environment preservation.

To the purpose of class assignment, for each considered time frame, the set of feature vectors $\mathcal{V}(h) = \{\mathbf{v}_1(h), \mathbf{v}_2(h), \ldots, \mathbf{v}_i(h), \ldots, \mathbf{v}_m(h)\}$ is defined, where $\mathbf{v}_i(h)$ is the feature vector associated with the user u_i at time frame h. The components of $\mathbf{v}_i(h)$ are four average metrics collected by the sensors operated by i-th user over h and safely stored in the Smart Contract 1.

In this work, the following metrics for $\mathbf{v}_i(h)$ are defined:

- the average indoor temperature $\overline{T_{ih}}$ of user u_i;
- the average outdoor temperature $\overline{Te_h}$;
- the average seasonal outdoor temperature $\overline{Ts}$;
- the seasonal indoor temperature threshold enforced by the law $\overline{Tl}$;
- the number of time intervals H_{ih} in which the air conditioning system is powered on for user u_i;
- the total number of time intervals H_{tot} in h.

Now, with the defined data, the feature vector of user u_i is formally defined as $\mathbf{v}_i(h) = [\alpha_{1i}(h), \alpha_{2i}(h), \alpha_{3i}(h), \alpha_{4i}(h)]^T$ with:

- $\alpha_{1i}(h) = \overline{T_{ih}}/\overline{Te_h}$;
- $\alpha_{2i}(h) = \overline{T_{ih}}/\overline{Ts}$;
- $\alpha_{3i}(h) = \overline{T_{ih}}/\overline{Tl}$;
- $\alpha_{4i}(h) = H_{ih}/H_{tot}$.

In particular, $\alpha_{1i}(h), \alpha_{2i}(h), \alpha_{3i}(h)$ represent the average indoor temperature measured at time h compared to the average outdoor temperature, the seasonal average temperature and the threshold enforced by the law, respectively. In addition, $\alpha_{4i}(h)$ considers the level of operation of the air conditioning system by comparing the total number of hours of operation to the total number of hours H_{tot} in time frame h.

Now, the *k-means* clustering algorithm is applied to the set $\mathcal{V}(h)$ and each vector $\mathbf{v}_i(h)$ is assigned to a class $c_j \in C$. We denote by $\mathbf{k}_j(h)$ with $j = 1, \ldots, n$ the centroid of class c_j, i.e., a vector with the same dimensions of $\mathbf{v}_i(h)$ representing the center of cluster c_j at time frame h. The class assigned to the i-th user at time frame h is denoted as $P_i(h) \in C$, where $P_i(h)$ is the class with the least Euclidean distance between $\mathbf{v}_i(h)$ and the centroid $\mathbf{k}_j(h)$ for $j = 1, \ldots, n$.

For the purpose of applying discounts or penalties to each user, the centroids can be calculated by averaging the values of all time frames belonging to a specific billing period. Moreover, the classes in the set C are ordered on the basis of their centroid in an appropriate way, starting from the most virtuous class (c_1) to the least virtuous one (c_n).

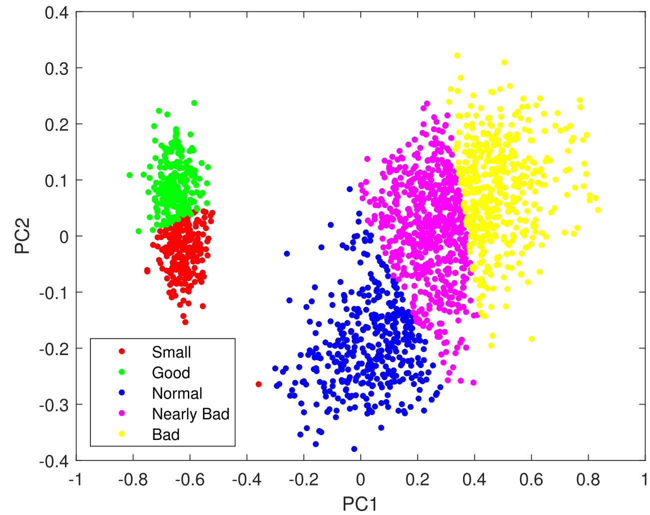


Fig. 3. K-means clustering.

Example: Figure 3 shows the clustering performed by k-means with $n = 5$ classes and $m = 2000$ users. At the end of the procedure, based on collected values, users are partitioned into five different classes ranked from *Small* (c_1), meaning a small power consumption, to *Bad* (c_5), in which the users with high power consumption and high relative difference between indoor and outdoor and seasonal temperatures are placed. In order to visualize the partitions in only two dimensions, the Principal Component Analysis (PCA) [37] is applied to the classified user set by projecting the data on the first two principal components (PC1 and PC2 in Fig. 3).

B. Class Follower Problem

Let us assume that user $u_i \in \mathcal{U}$ has been assigned to class c_k for the h time frame and at time h the current billing period ends. This user wants to be assigned to a different and better class c_j for time frame $h+1$, where $k \neq j$. Now, the *Class Follower Problem* (CFP) for each generic user is defined as follows:

$$\min_{\mathbf{v}(h+1) \in \mathcal{V}(h+1)} \|\mathbf{k}_j(h) - \mathbf{v}(h+1)\|. \tag{1}$$

Note that for the sake of simplicity, the index i denoting the i-th user is omitted in the following.

The aim of the objective function (1) is selecting vector $\mathbf{v}(h+1) \in \mathcal{V}(h+1)$ that exhibits the minimum Euclidean distance between vector $\mathbf{v}(h+1)$ and the centroid $\mathbf{k}_j(h)$ of the desired c_j class as calculated in previous time frame h. In that respect, it is evident that the global minimum of (1) is reached when $\mathbf{v}(h+1) = \mathbf{k}_j(h)$.

In order to minimize (1), it is necessary to determine the indoor temperatures and the HVAC actuators values during the time frame $h+1$. To this aim the time frame $h+1$ is divided in S time steps $s = 1, .., S$ and the following decision variables are defined:

$$T_{h+1}(s) \in \mathcal{R}^+ \text{ for } s = 1, 2, \ldots, S; \tag{2a}$$

$$H_{h+1}(s) \in \{0, 1\}, \text{for } s = 1, 2, \ldots, S; \tag{2b}$$

where $T_{h+1}(s)$ is the indoor average temperature collected at time step s and $H_{h+1}(s) = 1$ means that the HVAC of user u_i is powered on during the time step s, otherwise $H_{h+1}(s) = 0$ means that it is powered off.

Moreover, the feature vector elements at at time $h+1$ for each user $u_i \in \mathcal{U}$ are computed by the following variables:

$$\alpha_1(h+1) = \sum_{s=1}^{S} \frac{T_{h+1}(s)}{Te_{h+1}(s)} \tag{3a}$$

$$\alpha_2(h+1) = \frac{1}{S}\frac{1}{\overline{Ts}} \sum_{s=1}^{S} T_{h+1}(s) \tag{3b}$$

$$\alpha_3(h+1) = \frac{1}{S}\frac{1}{\overline{T_l}} \sum_{s=1}^{S} T_{h+1}(s) \tag{3c}$$

$$\alpha_4(h+1) = \frac{1}{S} \sum_{s=1}^{S} H_{h+1}(s). \tag{3d}$$

Now, the CFP (1) can be rewritten as follows:

$$\min \sum_{k=1}^{4} (k_j(h)_k - \alpha_k(h+1))^2 \tag{4a}$$

$$\text{s.t.} \tag{4b}$$

$$H_{h+1}(s) \in \{0, 1\} \ \forall s = 1, \ldots, S \tag{4c}$$

$$T_{h+1}(s) \in \mathcal{R}^+ \ \forall s = 1, \ldots, S. \tag{4d}$$

We assume that the average outdoor temperature $Te_{h+1}(s)$ at each time step s is determined by a suitable stochastic function.

To solve the CFP (4), we need to predict the values of the indoor temperature $T_{h+1}(s)$ at each time step s. However, the indoor temperature is related to intrinsic features of the building, such as the wall and floor materials and the number, position and size of the windows. Nevertheless, when an HVAC system is present, the indoor temperature is mainly driven by the thermostat set-point. We denote by $y_{h+1}(s) \in [0, 1]$ the real value of the set-point at time step s in time frame $h+1$, so that we can write:

$$T_{h+1}(s) = L(y_{h+1}(s)) \ s = 1, \ldots, S$$

where $L(y_{h+1}(s))$ denotes the building thermodynamic model and determines the indoor temperature $T_{h+1}(s)$ at time step s corresponding to the thermostat set point $y_{h+1}(s) \in [0, 1]$. Hence, the decision variables of CFP (4) are the thermostat set-point values $y_{h+1}(s)$ for $s = 1, \ldots, S$.

C. Energy Consumption Optimization

Since CFP (4) may have more optimal solutions, we have the possibility of considering a second objective function, related to the energy consumption denoted by $E(y_{h+1}(s), H_{h+1}(s))$, which is function of the set points $y_{h+1}(s)$ and the ON/OFF positions of the HVAC system at time steps $s = 1, \ldots, S$. Then, the cumulative energy consumption over the time frame $h+1$ is the following:

$$\sum_{s=1}^{S} E(y_{h+1}(s), H_{h+1}(s)). \tag{5}$$

Since equation (5) is not straightforward to ascertain, its value will be computed through an approximation in a subsequent part of this work, facilitated by the introduction of an LSTM. Now, the following lexicographic minimization problem is formulated:

$$\begin{aligned} &\text{lex min} \sum_{k=1}^{4} (k_j(h)_k - \alpha_k(h+1))^2, \ \sum_{s=1}^{S} E(y_{h+1}(s), H_{h+1}(s)) \\ &\quad \text{s.t. } y_{h+1}(s) \in [0, 1] \ \forall s = 1, \ldots, S \\ &\qquad Hy_{h+1}(s) \in \{0, 1\} \ \forall s = 1, \ldots, S. \end{aligned} \tag{6}$$

The lexicographic optimization consists of subdividing a multi-objective problem into a set of single-task optimizations that are solved in series according to their priority order [38]: the optimization with the highest priority is solved first and, then, the successive one is addressed with an additional constraint which aims at guaranteeing the optimality of the higher priority cost function. In this work, problem (6) is denoted as *Energy Consumption Optimization CFP* (OCFP).

VI. Control System Design and Simulation

In this section, a data-driven MPC strategy to solve the OCFP in real-time is introduced. Since the proposed approach is iterated over each time frame h, for the sake of simplicity, the suffix $h+1$ is omitted in the relevant notations.

A. System Architecture

The architecture of the proposed system for a specific building is depicted in Fig. 4. In more detail, the system is made of three main blocks: (i) the Plant, (ii) the State Observer and (iii) the MPC Controller. The Plant is represented by the user building and the associated HVAC system. The State Observer is basically constituted by a set of sensors installed inside and outside the building to monitor indoor and outdoor temperatures. The components of the State Observer are detailed in the subsequent Algorithm 1. The MPC Controller is composed of three sub-blocks: the outdoor temperature predictor, the LSTM-based plant thermodynamic model that predicts indoor temperatures and energy consumption based on historical thermostat set-points and HVAC operational times, and the Optimizer that solves the OCFP problem (6) over time frame $h+1$.

More specifically, the Optimizer, at each time step s, first takes the current Plant state and the target class $\mathbf{k}_j(h)$ to follow as a reference, then it determines a set of feasible inputs that minimizes the distance to $\mathbf{k}_j(h)$ and, finally, it selects the solution with the least energy consumption. At the end of the optimization procedure, the updated inputs are applied to the Plant by adjusting thermostat set-points and switching *ON* or *OFF* the HVAC system.

B. LSTM-Based Plant Thermodynamic Model

The indoor temperature variation and the HVAC energy consumption are influenced not only by the inputs, such as the thermostat set-points, but also by several intrinsic features of the building and the HVAC system. Hence, finding an analytical solution is not always feasible. Moreover, both the

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

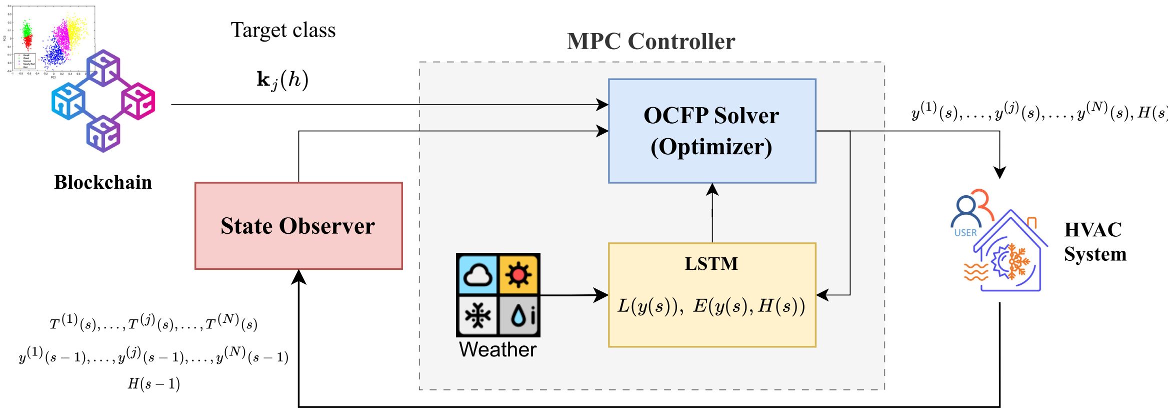


Fig. 4. MPC architecture.

TABLE I
LSTM EQUATIONS:

Input gate controller:	$i_s = \sigma(\mathbf{W}_i[h_{s-1}, \hat{x}_s] + \mathbf{b}_i)$
Forget gate controller:	$f_s = \sigma(\mathbf{W}_f[h_{s-1}, \hat{x}_s] + \mathbf{b}_f)$
Candidate vector:	$\tilde{c}_s = \tanh(\mathbf{W}_c[h_{s-1}, \hat{x}_s] + \mathbf{b}_c)$
Cell Memory:	$c_s = f_t \odot c_{s-1} + i_t \odot \tilde{c}_s$
Output gate controller:	$o_s = \sigma(\mathbf{W}_o[h_{s-1}, \hat{x}_s] + \mathbf{b}_o)$
Output:	$h_s = o_s \odot \tanh(c_s)$

temperature and the consumption depend not only on the current input but also on the past inputs applied to the HVAC system.

In this section, we propose an LSTM-based method to determine function $L(y(s))$ for $s = 1, \ldots, S$ and the energy consumption function $E(y(s), H(s))$ for $s = 1, \ldots, S$. LSTM is a variant of typical RNNs and can avoid the vanishing gradient problem existing in regular RNNs.

A data-driven solution is proposed to approximate both L and E functions using LSTM. The calculation formulas of LSTM cell units from the input to the output are obtained as shown in Table I [39], [40], [41], where i, f, $\tilde{c}$ and o, represent an input gate, forget gate, candidate vector and output gate, respectively, $\mathbf{W}_i$, $\mathbf{W}_f$, $\mathbf{W}_c$ and $\mathbf{W}_o$ denote weight matrices, $\mathbf{b}_i$, $\mathbf{b}_f$, $\mathbf{b}_c$ and $\mathbf{b}_o$ represent bias vectors, $\sigma(\cdot)$ and $\tanh(\cdot)$ denote the sigmoid and hyperbolic tangent functions, respectively, $\odot$ denotes a dotwise product, and c represents a cell state.

As shown in Fig. 5, the cell unit structure of an LSTM network consists of three gates: (i) a forget gate, (ii) an input gate, and (iii) an output gate. The forget gate determines what information of the past cell state is to be forgotten. The input gate is used to control what information of the input at the current time is to be added to the cell state. Finally, the output gate is used to determine what information of the cell state at the current time is to be used as output.

In order to train the LSTM, sample data need to be collected for a certain time from temperature sensors inside and outside the building, as well as the associated thermostat set-points and HVAC operating states. Although in OCFP problem (6) only indoor average temperature is considered, a building is composed of a set of rooms, each one is equipped with a

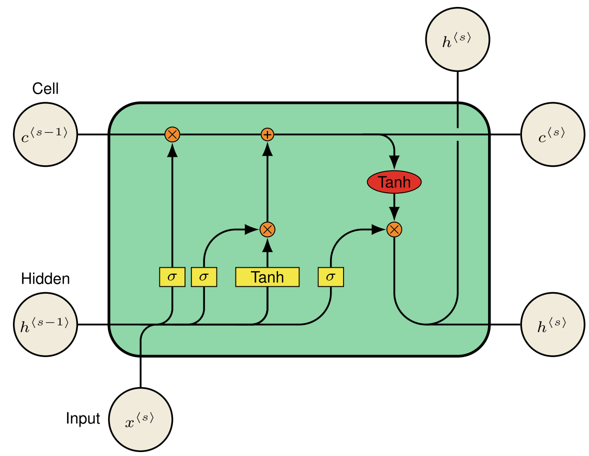


Fig. 5. LSTM cell architecture.

thermostat and indoor sensor. The sample input and output vectors collected for LSTM training at time step s are defined as follows:

$$\mathbf{X}(s) = \left[y^{(1)}, \ldots, y^{(j)}, \ldots, y^{(N)}, Te, H(s)\right]^T \tag{7}$$
$$\mathbf{T}(s) = \left[T^{(1)}, \ldots, T^{(j)}, \ldots, T^{(N)}, E\right]^T \tag{8}$$

where N is the number of rooms in the building, $y^{(j)}$ is the set-point of thermostat j, $T^{(j)}$ is the environment temperature of room j, $H(s) \in \{0, 1\}$ is the HVAC operating state and E is the cumulative energy consumption measured in Watts.

As it is shown in Fig. 6, if S consecutive time steps are considered for LSTM training, the final structure of the network is made of a series of S interconnected unit cells and the final value $\mathbf{T}(S)$ is the output prediction of indoor temperatures and energy consumption after S time steps. The mean indoor temperature value $\overline{T(S)}$ at time step S can then be easily calculated by averaging the individual rooms' predicted temperatures.

C. Model Predictive Control Scheme

The real-time control of the HVAC system is implemented by the MPC approach. In detail, at each time step s, the MPC solves the OCFP (6), in which the functions $L(y(s))$ and $E(y(s), H(s))$ are approximated by the LSTM pre-trained

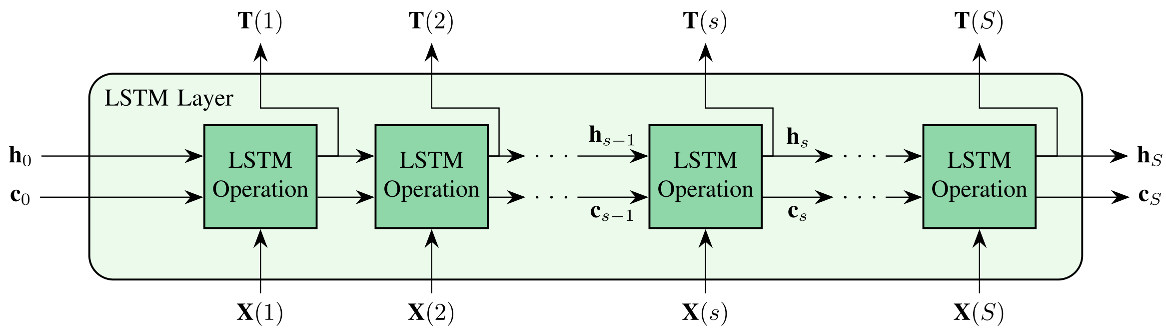


Fig. 6. LSTM final structure.

network. The time horizon of the MPC scheme is equal to S time steps, i.e., the duration of the time frame $h+1$.

Algorithm 1 Proposed MPC Scheme

1: Set reference as $\mathbf{k}_j(h)$
2: Collect:
- Indoor temperatures: $T^{(1)}(s), \dots, T^{(j)}(s), \dots, T^{(N)}(s)$
- Thermostat set-points: $y^{(1)}(s-1), \dots, y^{(j)}(s-1), \dots, y^{(N)}(s-1)$
- HVAC state: $H(s-1)$

3: Average indoor temperatures
4: Perform lexicographic optimization (6) over S steps:
5: **for** each time step s **do**
6: Compute candidate input vector $\hat{\mathbf{X}}(s)$
7: Select input vector with least energy consumption
8: Apply thermostat set-points $y^{(1)}(s), \dots, y^{(j)}(s), \dots, y^{(N)}(s)$ and $H(s)$ to the HVAC system
9: **end for**

In Algorithm 1, the proposed MPC scheme is shown. At step 1, the system takes the followed class $\mathbf{k}_j(h)$ as reference. At step 2, the State Observer collects current indoor temperatures: $T^{(1)}(s), \dots, T^{(j)}(s), \dots, T^{(N)}(s)$, last applied room thermostat set-points $y^{(1)}(s-1), \dots, y^{(j)}(s-1), \dots, y^{(N)}(s-1)$, as well as last HVAC operating state $H(s-1)$.

At step 4, the Optimizer pre-processes the data by averaging the indoor temperatures and performs the lexicographic optimization (6) over the next S time steps by minimizing the two cost functions. The non-linear problem can be solved by iterative methods, such as Sequential Least SQuares Programming (SLQP) [42]. Moreover, for each time step s, the Optimizer calculates a candidate optimal input vector $\hat{\mathbf{X}}(s)$ for the LSTM model as follows:

$$\hat{\mathbf{X}}(s) = \big[y^{(1)}(s+i), \dots, y^{(j)}(s+i), \dots, y^{(N)}(s+i), \\ H(s+i), \dots, y^{(N)}(s+S-1), H(s+S-1)\big]^T \quad (9)$$

for $i = 0, 1, 2, \dots, S-1$.

Here, the outdoor temperature $Te(s)$, required by the LSTM network as part of the input, is not considered as a decision variable as it is assumed as predicted by a stochastic function and, therefore, is not included in (9).

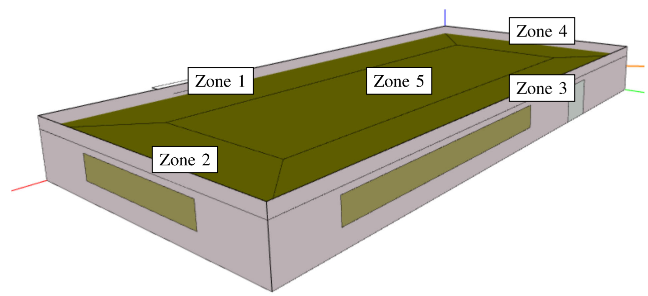


Fig. 7. 5-zones air-conditioned building.

The input vector spans from time step s to $s+S-1$ and allows the Optimizer to fit to the MPC time horizon. At the end of the optimization procedure, the input vector with the least cumulative energy consumption over the considered future time horizon is selected as optimal solution and the first set of thermostat set-points: $y^{(1)}(s), \dots, y^{(j)}(s), \dots, y^{(N)}(s)$, as well as the operating state $H(s)$ are applied to the HVAC system.

The whole control scheme reiterates for time step $s+1$ up to the end of the billing period $h+1$.

VII. Case Study

In this section, the results of the numerical simulations performed to validate the proposed approach are described and discussed. The well-known building energy simulation *EnergyPlus* is used with a standard 5-zone building to collect data and simulate the plant in the MPC scheme.

A. The Studied Building

We use a single-floor rectangular building of $30.5m \times 15.2m$ made of five different air-conditioned zones with individual thermostats, named Zone 1 to Zone 5. The building is located in the city of Bari, in Italy. There are windows on all four exterior facades and glass doors on the south and north facades. The air conditioning system is made of centralized HVAC equipment with an electric chiller, air-cooled condenser, and a return plenum. The total floor area is 463.6 m^2.

The building is shown in Fig. 7 and is included in the standard *EnergyPlus* package as *5ZoneAirCooled.idf* file.

B. The Used Dataset

The dataset used to train the LSTM network is collected by running multiple Monte Carlo simulations with *EnergyPlus* and the 5-zones building file. To get the outdoor temperature

TABLE II
SAMPLES FROM THE DATASET FOR JULY 7

H	Zone 1 Thermostat	Zone 2 Thermostat	Zone 3 Thermostat	Zone 4 Thermostat	Zone 5 Thermostat	System Status	Air Temp. Zone 1	Air Temp. Zone 2	Air Temp. Zone 3	Air Temp. Zone 4	Air Temp. Zone 5	Outdoor Temperature	Chiller Electr. Rate
1	21.3	24.4	19.9	20.1	21.7	ON	22.2	19.2	20.6	21.5	20.6	22.0	2649.8
2	19.9	19.8	21.1	21.7	21.9	ON	21.3	21.7	19.9	20.1	21.4	21.6	2555.9
3	24.3	20.6	23.5	22.7	19.5	OFF	20.1	19.9	21.1	20.9	21.2	21.2	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...
24	24.9	22.2	22.6	24.3	21.4	OFF	21.9	21.8	22.0	22.2	22.1	19.2	0

throughout the simulations, we use a weather file in *EnergyPlus Weather File* (EPW) format containing one year of historical weather data of Bari city, in Italy.

We set the time step s to one hour and we draw each thermostat set-point value, for each sample, from a uniform distribution between 19°C and 25°C. Similarly, we set the HVAC operating state to *ON* or *OFF* from a random distribution between 0 and 1, by rounding each sampled value to the nearest integer.

We collect 43,800 one-hour samples over five consecutive yearly simulations. Each sample includes the individual thermostat set-points at time step s and the HVAC operating state as the input $\mathbf{X}(s)$, and the individual indoor temperatures and the energy consumption in Watts as the output $\mathbf{T}(s)$. Table II shows some sample rows of the dataset.

C. The LSTM Training

We use the collected dataset to train the LSTM network and predict the thermodynamic behavior of the building. We present the samples to the training procedure in groups of $S = 6$ hourly time steps. More specifically, each input sample $\mathbf{X}(s)$ is grouped with the previous $\mathbf{X}(s-1)$, $\mathbf{X}(s-2), \ldots \mathbf{X}(s-5)$ samples and is paired with the output sample $\mathbf{T}(s)$, in order for the resulting LSTM network to be able to predict the indoor temperatures and energy consumption after six consecutive thermostats set-points and six HVAC operating state settings.

We run the training on a server equipped with a 14 cores *Intel Core i7* CPU and 32GB RAM. *Python Keras* library and *TensorFlow* are used to implement the LSTM network with *Adam* optimizer. The dataset is normalized and split into 80% of values for training and 20% for test. *Mean Square Error* (MSE) metric is adopted to evaluate the performance of the training procedure.

In Fig. 8, the results after 60 training epochs are shown. At the end of the procedure, the MSE for training and testing converges to 5e−3, allowing good predictions for the MPC Controller.

D. The OCFP Solution

In order to assess the performance of the proposed approach, we consider for the case study a district $\mathcal{U}$ of $m = 2000$ users located in Bari (Italy) for a billing period of one week in July. We set $\overline{T_s} = 25°C$, which is the average temperature in July and $\overline{T_l} = 26°C$, that we assume as the minimum temperature prescribed by the law for public offices and institutions to save energy.

We assume that at the end of the current billing period, at time frame h, the user set $\mathcal{U}$ has been partitioned in

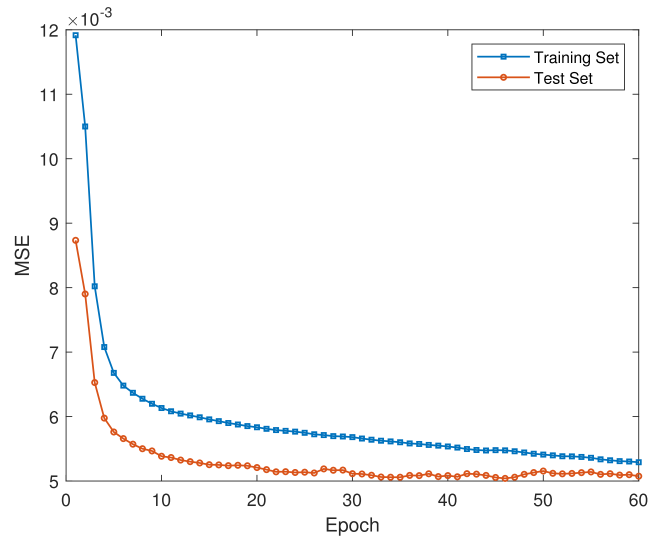


Fig. 8. LSTM training.

TABLE III
PARTIONED USERS SET

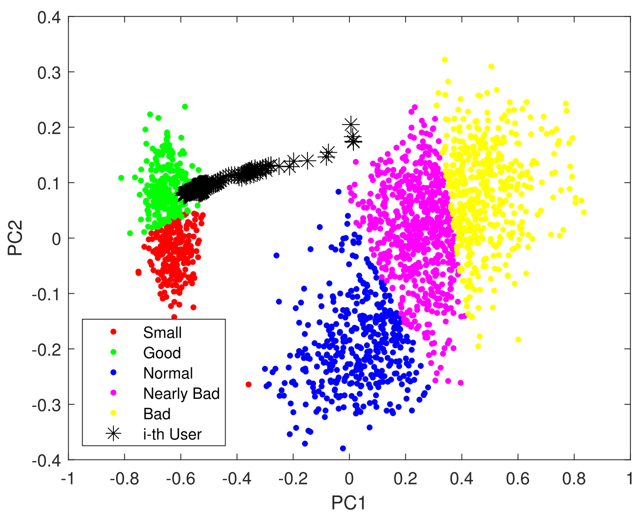
	Class	No.	Temperature [°C] Indoor	Temperature [°C] Outdoor	Daily Usage [Hrs]
1	Small	501	27.00	27.00	7.01
2	Good	317	26.00	27.00	10.76
3	Normal	291	22.05	27.50	12.48
4	Nearly Bad	516	21.53	28.00	13.71
5	Bad	375	21.17	27.87	14.10
	Total	2000			

$n = 5$ behavior classes by the *k-means* algorithm described in Section V-A. Table III shows the classes resulting from applying the *k-means* clustering algorithm.

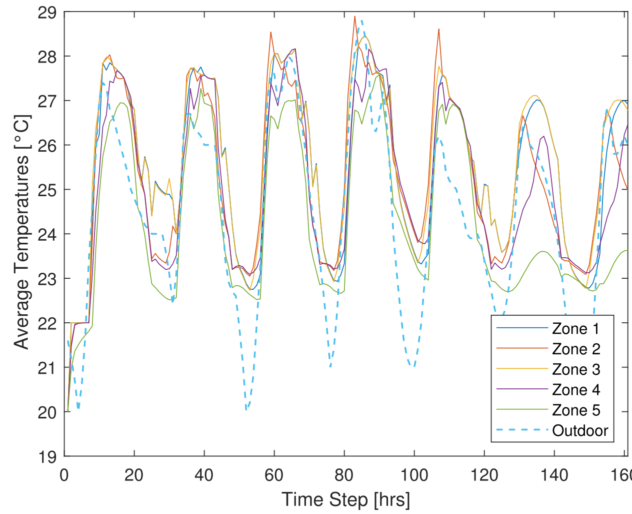
The classes are ranked based on the average indoor temperature compared to the average outdoor temperature, the seasonal average temperature and the temperature prescribed by the law. In addition, the total operational time of the HVAC is considered. The closer the indoor temperature is to the reference values and the less operational time, the better the class is ranked. Based on the above evaluations, we labeled the resulting classes from *Small* (c_1) to *Bad* (c_5).

Now, two *OCFP* scenarios for time frame $h + 1$ are proposed. In that respect, we assume that user u_i has been ranked in class *Nearly Bad* (c_4) for the current time frame h and that he wishes to scale up to class *Good* (c_2) and *Normal* (c_3) for the first and second scenario respectively.

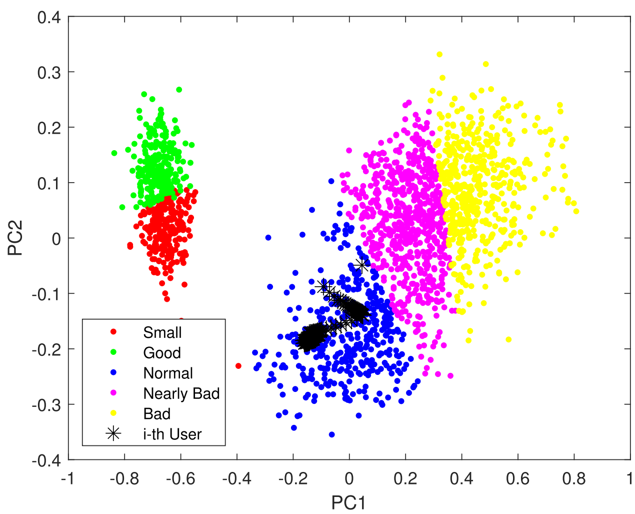
We implement the MPC scheme by using the pre-trained LSTM to model the thermodynamics of the i-th building and we use *EnergyPlus* to simulate the plant. We set the MPC time



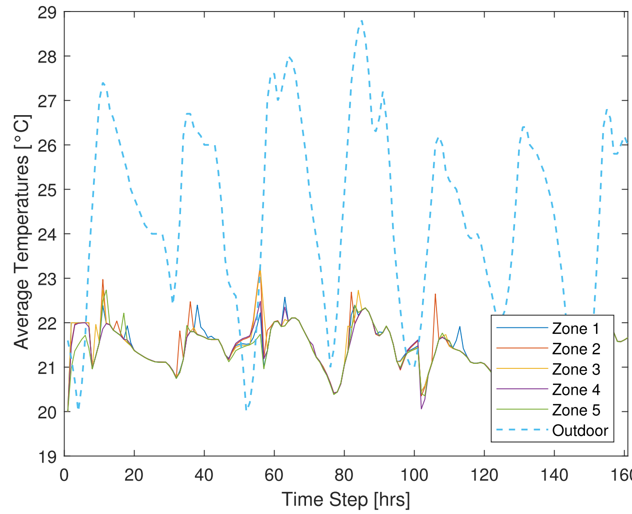
(a) Scenario 1 - From *Nearly Bad* to *Good*



(b) Scenario 1 - Average Temperatures



(c) Scenario 2 - From *Nearly Bad* to *Normal*



(d) Scenario 2 - Average Temperatures

Fig. 9. Optimized Class Follower MPC scheme progress over the billing period.

horizon to $S = 6$ hours and we reiterate the control scheme up to the end of the next billing period. To implement and solve the lexicographic problem (6), we use *Python* and the SLQP optimization algorithm provided by the *SciPy* library.

The results of the class following procedure for the time frame $h + 1$ in the two considered scenarios are shown in Fig. 9.

In Fig. 9(a) and 9(c), black star points represent the average centroids calculated over the whole billing period for the i-th user at each time step s for Scenario 1 and 2 respectively. We observe that, in both cases, the points move from the *Nearly Bad* (c_4) cluster to the *Good* (c_2) and *Normal* (c_3) cluster, respectively, by gradually approaching the relative cluster center. Since the *Good* (c_2) class is far more distant from *Nearly Bad* (c_4) than the *Normal* (c_3) class, the initial error for Scenario 1 is much greater than for Scenario 2. As a result, the path of Scenario 1 looks much more extended than the one of Scenario 2.

The above difference is also observed in Fig. 10, which shows the class tracking relative error of Scenario 1 compared to Scenario 2 and confirms that the convergence to a stable

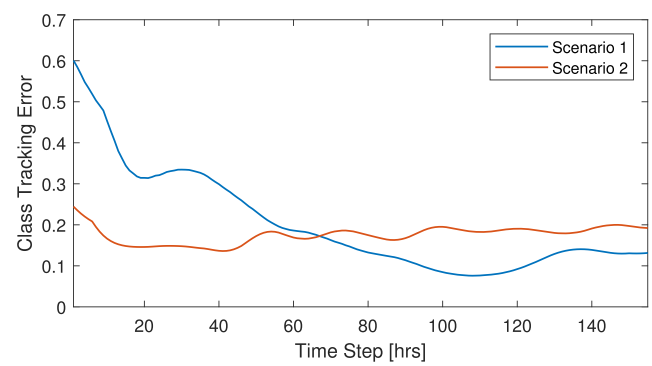


Fig. 10. Class follower tracking error.

residual error of around 0.14, in the first case, takes more time steps than of Scenario 2. However, the residual error of Scenario 2 stabilizes around 0.20, which is higher than that of Scenario 1. As a result, the numerical simulations show that the proposed scheme works better when the initial error is large and the followed class is relatively distant.

In addition, in Fig. 9(b) and 9(d) the average individual indoor temperatures and the average outdoor temperature are

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

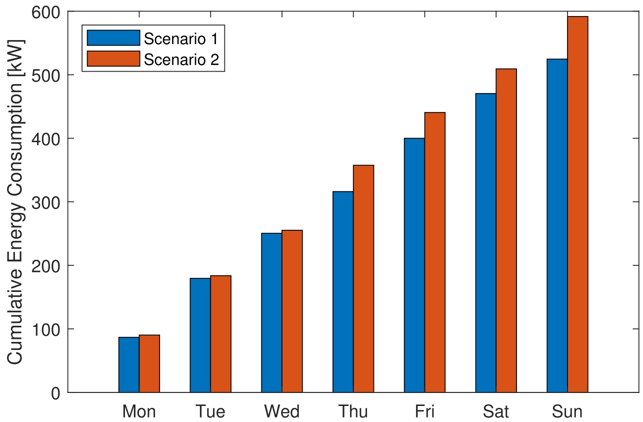


Fig. 11. Cumulative energy consumption.

shown. In Scenario 1, the indoor temperatures are much closer to the outdoor temperature. On the contrary, in Scenario 2, the indoor temperatures are lower than the outdoor temperature. As observed in Table III, the values obtained by the MPC in both scenarios are in line with the respective followed classes *Good* and *Normal*.

Finally, in Fig. 11, the cumulative energy consumption in kW over the considered billing period is shown. It is evident that, in Scenario 2, at the end of the week, the HVAC consumed more electricity than in Scenario 1. In Scenario 2, the gap between the indoor and outdoor temperature and the higher operational time leads to higher energy consumption by the HVAC to fit the following class requirements.

VIII. Conclusion

This paper introduces a novel approach to managing the HVAC control system at the DEMS level to optimize the user HVAC consumption. The HVAC control system includes the synergistic use of blockchains, MPC strategies and LSTM networks. In particular, the district users are divided into consumption classes, and each class is associated with a suitable energy price based on consumption behaviors. A blockchain platform ensures data integrity and facilitates transparent transactions, enables dynamic categorization of users based on their energy consumption behaviors, and processes payments on-chain. Moreover, the users can modify their behaviors and pass to a less expensive class. To this aim, the MPC strategy is proposed to provide the thermostat set points and the time intervals in which the HVAC must be switched off or switched on.

The application of the proposed model to a case study of a smart district shows promising results. The system effectively guides users from less virtuous energy consumption classes towards more energy-efficient ones while maintaining thermal comfort. Moreover, it is observed that the control scheme performs better with larger initial tracking errors, suggesting its robustness in scenarios requiring significant user behavioral adjustments.

In conclusion, the presented research highlighted the potential of combining blockchain with advanced control techniques to improve energy systems management within smart urban districts.

Future work will study the system's scalability to larger districts, the integration with other smart grid components and the extension of extensive tests under different seasonal conditions with a broader set of user profiles. Additionally, there is an opportunity to refine the predictive algorithms, considering the impact of integrating renewable energy sources into the proposed model.

References

[1] X. Deng, Y. Zhang, Y. Zhang, and H. Qi, "Toward smart multi-zone HVAC control by combining context-aware system and deep reinforcement learning," *IEEE Internet Things J.*, vol. 9, no. 21, pp. 21010–21024, Nov. 2022.

[2] O. Van Cutsem, D. Ho Dac, P. Boudou, and M. Kayal, "Cooperative energy management of a community of smart-buildings: A blockchain approach," *Int. J. Electr. Power Energy Syst.*, vol. 117, May 2020, Art. no. 105643.

[3] O. Erdinç, A. Tascikaraoglu, N. G. Paterakis, and J. P. S. Catal ao, "Novel incentive mechanism for end-users enrolled in DLC-based demand response programs within stochastic planning context," *IEEE Trans. Ind. Electron.*, vol. 66, no. 2, pp. 1476–1487, Feb. 2019.

[4] Q. Shi, C.-F. Chen, A. Mammoli, and F. Li, "Estimating the profile of incentive-based demand response (IBDR) by integrating technical models and social-behavioral factors," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 171–183, Jan. 2020.

[5] A. Mnatsakanyan, A. H. AlMazrooqi, E. B. Muruaga, and P. A. Banda, "Electricity market structure with individual pricing mechanism," in *Proc. 9th IEEE Int. Symp. Power Electron. Distrib. Gener. Syst. (PEDG)*, Jun. 2018, pp. 1–6.

[6] I. Brahmia, J. Wang, H. Xu, H. Wang, and L. D. O. Turci, "Robust data predictive control framework for smart multi-microgrid energy dispatch considering electricity market uncertainty," *IEEE Access*, vol. 9, pp. 32390–32404, 2021.

[7] M. Biemann, P. A. Gunkel, F. Scheller, L. Huang, and X. Liu, "Data center HVAC control harnessing flexibility potential via real-time pricing cost optimization using reinforcement learning," *IEEE Internet Things J.*, vol. 10, no. 15, pp. 13876–13894, Aug. 2023.

[8] J. P. Tan, A. L. A. Ramos, M. V. Abante, R. L. Tadeo, and R. R. Lansigan, "A performance review of recurrent neural networks long short-term memory (LSTM)," in *Proc. 3rd Int. Conf. Emerg. Technol. (INCET)*, May 2022, pp. 1–5.

[9] R. Kwadzogah, M. Zhou, and S. Li, "Model predictive control for HVAC systems—A review," in *Proc. IEEE Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2013, pp. 442–447.

[10] G. Olivieri, G. Volpe, A. M. Mangini, and M. Pia Fanti, "A district energy management approach based on Internet of Things and blockchain," in *Proc. IEEE Int. Conf Dependable, Autonomic Secure Comput., Int. Conf Pervasive Intell. Comput., Int. Conf Cloud Big Data Comput., Int. Conf Cyber Sci. Technol. Congr. (DASC/PiCom/CBDCom/-CyberSciTech)*, Sep. 2022, pp. 1–6.

[11] F. Casino, T. K. Dasaklis, and C. Patsakis, "A systematic literature review of blockchain-based applications: Current status, classification and open issues," *Telematics Informat.*, vol. 36, pp. 55–81, Mar. 2019.

[12] M. J. M. Chowdhury, A. Colman, M. A. Kabir, J. Han, and P. Sarda, "Blockchain versus database: A critical analysis," in *Proc. 17th IEEE Int. Conf. Trust, Secur. Privacy Comput. Commun. 12th IEEE Int. Conf. Big Data Sci. Eng. (TrustCom/BigDataSE)*, Aug. 2018, pp. 1348–1353.

[13] P. Haji Mirzaee, M. Shojafar, H. Cruickshank, and R. Tafazolli, "Smart grid security and privacy: From conventional to machine learning issues (threats and countermeasures)," *IEEE Access*, vol. 10, pp. 52922–52954, 2022.

[14] M. B. Mollah et al., "Blockchain for future smart grid: A comprehensive survey," *IEEE Internet Things J.*, vol. 8, no. 1, pp. 18–43, Jan. 2021.

[15] C. Lazaroiu and M. Roscia, "Smart district through IoT and blockchain," in *Proc. IEEE 6th Int. Conf. Renew. Energy Res. Appl. (ICRERA)*, Nov. 2017, pp. 454–461.

[16] K. Christidis and M. Devetsikiotis, "Blockchains and smart contracts for the Internet of Things," *IEEE Access*, vol. 4, pp. 2292–2303, 2016.

[17] S. Benedict, P. Rumaise, and J. Kaur, "IoT blockchain solution for air quality monitoring in SmartCities," in *Proc. IEEE Int. Conf. Adv. Netw. Telecommun. Syst.*, Dec. 2019, pp. 1–6.

[18] P. Zhuang, T. Zamir, and H. Liang, "Blockchain for cybersecurity in smart grid: A comprehensive survey," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 3–19, Jan. 2021.

[19] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, "Hawk: The blockchain model of cryptography and privacy-preserving smart contracts," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2016, pp. 839–858.

[20] X. Li et al., "A quantitative and qualitative review of blockchain research from 2015 to 2021," *Sustainability*, vol. 15, no. 6, p. 5067, Mar. 2023.

[21] T. B. Malla, A. Bhattarai, A. Parajuli, A. Shrestha, B. B. Chhetri, and K. Chapagain, "Status, challenges and future directions of blockchain technology in power system: A state of art review," *Energies*, vol. 15, no. 22, p. 8571, Nov. 2022.

[22] E. Androulaki et al., "Hyperledger fabric: A distributed operating system for permissioned blockchains," in *Proc. 13th EuroSys Conf.*, Porto, Portugal, Apr. 23/26, 2018, pp. 1–15.

[23] Z. Wu and W. Chu, "Sampling strategy analysis of machine learning models for energy consumption prediction," in *Proc. IEEE 9th Int. Conf. Smart Energy Grid Eng. (SEGE)*, Aug. 2021, pp. 77–81.

[24] M. Roccotelli, A. M. Mangini, and M. P. Fanti, "Smart district energy management with cooperative microgrids," *IEEE Access*, vol. 10, pp. 36311–36326, 2022.

[25] D. Muralidar, "A review about energy management techniques in industrial buildings," in *Proc. Innov. Power Adv. Comput. Technol.*, Apr. 2017, pp. 1–9.

[26] A. Rajith, S. Soki, and M. Hiroshi, "Real-time optimized HVAC control system on top of an IoT framework," in *Proc. 3rd Int. Conf. Fog Mobile Edge Comput. (FMEC)*, Apr. 2018, pp. 181–186.

[27] W.-H. Chen, S. Yang, and F. You, "Thermal comfort control on sustainable building via data-driven robust model predictive control," in *Proc. Amer. Control Conf. (ACC)*, May 2023, pp. 591–596.

[28] R. Eini and S. Abdelwahed, "Learning-based model predictive control for smart building thermal management," in *Proc. IEEE 16th Int. Conf. Smart Cities, Improving Quality Life Using ICT IoT AI (HONET-ICT)*, Oct. 2019, pp. 038–042.

[29] R. Z. Homod, K. S. Mohamed Sahari, H. A. F. Almurib, and F. H. Nagi, "RLF and TS fuzzy model identification of indoor thermal comfort based on PMV/PPD," *Building Environ.*, vol. 49, pp. 141–153, Mar. 2012.

[30] S. Taheri, P. Hosseini, and A. Razban, "Model predictive control of heating, ventilation, and air conditioning (HVAC) systems: A state-of-the-art review," *J. Building Eng.*, vol. 60, Nov. 2022, Art. no. 105067.

[31] K. Friansa, J. Pradipta, I. N. Haq, P. H. K. Utama, M. Wasesa, and E. Leksono, "Enhancing HVAC electricity load prediction accuracy using bi-LSTM method based on daily dataset," in *Proc. 8th Int. Conf. Instrum., Control, Autom. (ICA)*, Aug. 2023, pp. 92–96.

[32] X. Wang, X. Wang, L. Wang, L. Jiang, and Y. Zhan, "A distributed fusion LSTM model to forecast temperature and relative humidity in smart buildings," in *Proc. IEEE 16th Conf. Ind. Electron. Appl. (ICIEA)*, Aug. 2021, pp. 1–6.

[33] R. E. Alden, H. Gong, E. S. Jones, C. Ababei, and D. M. Ionel, "Artificial intelligence method for the forecast and separation of total and HVAC loads with application to energy management of smart and NZE homes," *IEEE Access*, vol. 9, pp. 160497–160509, 2021.

[34] K. Sekniqi, D. Laine, S. Buttolph, and E. G. Sirer, "Avalanche platform," Netw. Distrib. Ledgers, Jun. 2020.

[35] M. Westerkamp, "Verifiable smart contract portability," in *Proc. IEEE Int. Conf. Blockchain Cryptocurrency (ICBC)*, May 2019, pp. 1–9.

[36] C. Wu, J. Xiong, H. Xiong, Y. Zhao, and W. Yi, "A review on recent progress of smart contract in blockchain," *IEEE Access*, vol. 10, pp. 50839–50863, 2022.

[37] I. Jolliffe, *Principal Component Analysis* (Springer Series in Statistics). Berlin, Germany: Springer, 2002.

[38] 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.

[39] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[40] J. Bi, X. Zhang, H. Yuan, J. Zhang, and M. Zhou, "A hybrid prediction method for realistic network traffic with temporal convolutional network and LSTM," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 3, pp. 1869–1879, Jul. 2022.

[41] H. Wang et al., "Risk assessment and mitigation in local path planning for autonomous vehicles with LSTM based predictive model," *IEEE Trans. Autom. Sci. Eng.*, vol. 19, no. 4, pp. 2738–2749, Oct. 2022.

[42] S. J. W. J. Nocedal, *Sequential Quadratic Programming*. New York, NY, USA: Springer, 2006, pp. 529–562.

Giuseppe Olivieri received the M.Sc. degree in management engineering and blockchain and digital currencies. He is currently pursuing the Ph.D. degree with the Laboratory of Control and Automation, Polytechnic University of Bari. He is also the Senior Official of Italian Ministry of Economy and Finance. His research interests include blockchain technologies, optimization techniques, and artificial intelligence. These areas fuel his academic and professional endeavors and have also led to several academic and journalistic contributions, blending educational insights with effective public communication.

Gaetano Volpe received the Ph.D. degree in electrical and information engineering from the Polytechnic University of Bari, Italy, in 2024. Since 2020, he has been associated with the Laboratory of Control and Automation, Polytechnic University of Bari. He is currently an Adjunct Professor of Data Driven Industrial Automation with the Polytechnic University of Bari. His research interests include traditional and machine learning-based control of complex systems, optimization techniques for industrial processes, blockchain, and cybersecurity.

Agostino Marcello Mangini (Senior Member, IEEE) received the Laurea degree in electronics engineering and the Ph.D. degree in electrical engineering from the Polytechnic University of Bari, Bari, Italy, in 2003 and 2008, respectively. He has been a Visiting Scholar with the University of Zaragoza, Zaragoza, Spain. He is currently an Associate Professor with the Department of Electrical and Information Engineering, Polytechnic University of Bari. He has authored or coauthored over 140 printed publications. His current research interests include modeling, simulation, control of discrete-event systems, Petri nets, deep reinforcement learning, supply chains and urban traffic networks, distribution, and internal logistics. He was on the program committees of the 2007–2023 IEEE International SMC Conference on Systems, Man, and Cybernetics. He was on the Editorial Board of the 2017–2023 IEEE Conference on Automation Science and Engineering. Moreover, he was the Publication Chair of the 2017 IEEE SOLI and the 2019 IEEE SMC conferences, the Workshop Co-Chair of the 2023 IEEE SMC conference, the Workshop and Tutorial Chair of the 2023 IEEE CASE Conference, and the Special Session Chair of the 3rd IEEE International Conference on Automation in Manufacturing, Transportation and Logistics (iCaMaL2023). He is an Associate Editor of IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING and *IEEE Systems, Man, and Cybernetics Magazine*.

Maria Pia Fanti (Fellow, IEEE) received the Laurea degree in electronic engineering from the University of Pisa, Pisa, Italy, in 1983. She was a Visiting Researcher with Rensselaer Polytechnic Institute, Troy, NY, USA, in 1999. Since 1983, she has been with the Department of Electrical and Information Engineering, Polytechnic University of Bari, Italy, where she is currently a Full Professor of System and Control Engineering and the Chair of the Laboratory of Automation and Control. Her research interests include management and modeling of complex systems, such as transportation, logistics, and manufacturing systems; discrete event systems; Petri nets; consensus protocols; and fault detection. She has published more than 345 articles and two textbooks on her research topics. She received the Award for Best Associate Editor of IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS from the IEEE SMCS in 2008 and the Award for the Best Technical Committee Chair. She was a member of the Board of Governors of the IEEE Systems, Man, and Cybernetics Society and the Administrative Committee of the IEEE Robotics and Automation Society and the Founder and the Chair of the Technical Committee on Automation in Logistics of the IEEE Robotics and Automation Society. She was the General Chair of the 2011 IEEE Conference on Automation Science and Engineering, the 2017 IEEE International Conference on Service Operations and Logistics, and the Informatics and the 2019 IEEE Systems, Man, and Cybernetics Conference. She is a Senior Editor of IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING and an Associate Editor of IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS and IEEE TRANSACTIONS ON INTELLIGENT VEHICLES.