Received: 12 January 2023 | Revised: 15 February 2023 | Accepted: 14 March 2023

IET Renewable Power Generation

DOI: 10.1049/rpg2.12726

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH

An efficient local multi-energy systems planning method with long-term storage

Jiahao Ma[1,2] | Ning Zhang[1] | Qingsong Wen[3] | Yi Wang[2]

[1]Department of Electrical Engineering, Tsinghua University, Beijing, China

[2]Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong SAR, China

[3]DAMO Academy, Alibaba Group (U.S.) Inc., Bellevue, USA

Correspondence
Yi Wang, Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong SAR, China.
Email: yiwang@eee.hku.hk

Funding information
Science & Technology Project of State Grid Corporation of China, Grant/Award Number: 5419-202199514A-0-5-ZN

Abstract

Long-term storage will play a crucial role in future local multi-energy systems (MES) with high penetration renewable energy integration for demand balancing. Local MES planning with long-term energy storage is essentially a very large-scale program because numerous decision variables, including binary variables, should be used to model long-term energy dependencies for accurate operational cost estimation. How to largely reduce decision variables as well as guarantee the planning model accuracy becomes one main concern. To this end, this paper proposes a novel efficient aggregation and modeling method for local MES planning. The aggregation method first decomposes input time series data (renewable energy output and energy demand) into hourly and daily components, based on which more accurate aggregation results with a few typical scenarios can be derived. By incorporating similar decomposition into the operation model of energy devices, the planning model can describe the long-term energy cycle and the hourly operation characteristic at the same time and yield accurate optimization results with limited complexity. Experimental results show that the proposed method can considerably decrease the complexity of the problem while maintaining agreement with the results based on the optimization of the full-time series.

1 | INTRODUCTION

Accommodating high penetration of renewable energy is very important for operational cost and carbon emission reduction [1]. The integration of multi-energy systems (MES) has been found to utilize the complementary characteristic of different energy carriers (i.e. electricity and heating) to provide a large amount of flexibility to accommodate renewable energy and thus guarantee a reliable, low-carbon, and economical energy supply [2]. In the complex MES, both the multi-energy demands and renewable energy outputs show significant variations, including daily and seasonal variations. [3] In this context, the collaborative planning of short-term and long-term storage has attracted more attention for compensating for these two variations respectively [4–6]. However, the optimal planning of MES with these two kinds of storage is a large-scale program problem because a massive number of decision variables should be introduced to model the operation of short-term and long-term storage for accurate operational cost estimation [7].

A popular method to simplify the problem is time series aggregation, which clusters full-time series into typical scenarios based on daily or weekly patterns of input profiles [8]. Then input time series and decision variables are reused in the same typical scenarios to decrease the redundancy of the planning problem.

Different aggregation methods for various planning problems have been proposed. Kotzur et al. compared the performance of different clustering methods (i.e. averaging, k-means, k-medoids, hierarchical aggregation etc.) in planning MES [9]. As for the length of the typical scenarios, typical days were chosen for the planning of the device capacities [10], risk-considered daily profit [11], device placement and investment year [12], and also connections [13] in MES. Besides daily patterns, works also chose typical weeks. The longer typical scenarios are suitable for the planning of systems with long-term storage devices that exchange state-of-charges (SoC) between days [14]. Based on various clustering methods, MES planning models are proposed for multi-year long-term

planning horizons[15]. Nevertheless, time series aggregation based on clustering methods smooths the original profiles, resulting in an underestimate of the total cost and a weak designed system. To tackle the problem, the error of the objective function was explicitly bounded in the framework [16]. Besides, extreme periods were added to the aggregated time series to guarantee the robustness of designed systems [17].

However, there is a drawback of MES planning with long-term storage via time series aggregation: long-term storage devices are not allowed to exchange stored energy between typical scenarios. To ensure the consistency of the SoC between typical scenarios, SoC at the beginning and the end of typical scenarios must be equal [18]. However, the true energy cycles of long-term storage devices range from days to weeks [19], so cyclic constraints result in significant errors in the designed capacity of long-term storage devices.

To overcome the drawback, several new time series aggregation methods for MES planning with long-term storage have been proposed to describe the energy exchange across typical scenarios. Kotzur et al. [19] described the SoC of long-term storage based on the superposition of inter-period and intra-period states. Inter-period states with lower time resolution describe the SoC exchanged between typical scenarios, while intra-period states with higher time resolution depict the SoC change within typical scenarios. Baumgärtner et al. [20] decomposed the optimization problem into several subproblems and recombined the solution of subproblems to generate an operation solution for long-term storage devices. Gabrielli et al. [21] reused all decision variables except SoC variables in the same typical scenarios, which was called the M1 method. Besides the M1 method, a model with higher complexity is that only integer variables are reused in the same typical scenarios, which was named after the M2 method. In both methods, the SoC of storage devices was described hour by hour for the whole year, resulting in free energy exchanges across typical scenarios. Tan et al. [22] decomposed the aggregated heating load profiles into daily average components and hourly fluctuation components, and planned the long-term and short-term storage devices to respond to daily and hourly components, respectively. Novo et al. [23] observed the SoC variation of long-term storage was unique for each typical period throughout the year. Therefore, the deviation of SoC between the start and the end of one typical period was the same, and it's feasible to model the initial SoC between two typical periods via the deviation. Tejada-Arango et al. [24] proposed two novel models to improve the planning of short-term and long-term storage, respectively. Transition matrixes were introduced to guarantee the continuity and energy exchanged between typical periods.

Despite the effectiveness of the aforementioned works, the time-series aggregation method can be investigated to further improve the performance of the optimization model. Inspired by the decomposition techniques used in these works, we try to integrate decomposition and aggregation to fill the research gap of how to improve the accuracy of the aggregated time series without importing more typical scenarios. Furthermore, to fully utilize the aggregated time series after decomposition, another research gap is how to model as many devices as possible to respond to time series with multiple components of different time resolutions. To address the aforementioned research gaps, this paper employs the decomposition to both time series aggregation and operation of devices and proposes a novel framework for the planning of local MES with long-term storage.

This paper makes the following two main contributions:

1. Propose a novel framework for optimal configuration planning of a multi-energy system with long-term storage that incorporates time series seasonal-trend decomposition into time series aggregation, and provides two major benefits: (1) the complexity of the problem is reduced significantly while the accuracy is maintained; (2) the operation of energy converters and long-term storage devices can vary throughout the year, generating a highly expressive model.
2. Develop a novel method of time series aggregation based on Multiple Seasonal-Trend decomposition using Loess (MSTL), which decouples the hourly fluctuation and the trend of the original time series, and yields hourly and daily components at once. With the help of the method, the aggregated series can depict the original time series accurately even with a low number of typical scenarios.

The rest of this paper is structured as follows. Section 2 defines the problem to be solved and introduces the proposed framework. Section 3 elaborates on the proposed novel time-series aggregation method, including the principles and procedures. Section 4 details the proposed planning model. Section 5 provides experimental results and comparative analysis. Section 6 draws conclusions.

2 | PROBLEM STATEMENT AND FRAMEWORK

2.1 | Problem statement

Local MES planning aims to determine which energy devices (including energy converters and storage) should be installed so that the overall cost, including investment and operational costs, is minimized. It is important to accurately estimate the operational cost during the planning period.

The operation of local MES should consider the constraints of energy converters and storage. For storage devices, the state-of-charges (SoC) $SOC_{g,t}$ between every two adjacent periods are coupled and can usually be formulated as follows:

$$SOC_{g,t} = (1-\gamma_g)SOC_{g,t-1} + P^{+}_{g,j,t}\eta^{+}_{g} - P^{-}_{g,j,t}/\eta^{-}_{g}, \quad \forall t, g \in \Omega_s, \tag{1}$$

where γ_g is the hourly self-discharge rate of the g-th storage device; $P^{+}_{g,j,t}$ and $P^{-}_{g,j,t}$ denote the charge and discharge power during the t-th hour, respectively, the kind of which is the j-th kind of energy carrier; η^{+}_{g} and η^{-}_{g} are the input and output efficiency, respectively; Ω_s is the set of all storage devices.

For short-term storage with a larger value of γ_g but a larger value of η^{+}_{g} and η^{-}_{g}, such as batteries, we assume the consistency

of the SoC at the beginning and the end of a day so that the operation of local MES can be decoupled on a daily basis:

$$SOC_{g,t=0,s} = SOC_{g,t=24,s}, \forall g \in \Omega_{sts}, s, \tag{2}$$

where s denotes the index of the typical scenario; Ω_{sts} is the set of all short-term storage devices.

Since short-term storage cannot charge and discharge simultaneously, binary variables are needed to represent the charging or discharging state in each hour throughout the whole year, which is formulated by:

$$x^{+}_{g,t} P^{+}_{g,\max} \geq P^{+}_{g,j,t} \geq 0, \forall g \in \Omega_{sts}, t, \tag{3a}$$

$$x^{-}_{g,t} P^{-}_{g,\max} \geq P^{-}_{g,j,t} \geq 0, \forall g \in \Omega_{sts}, t, \tag{3b}$$

$$x^{+}_{g,t} + x^{-}_{g,t} \leq 1, \forall g \in \Omega_{sts}, t, \tag{3c}$$

$$x^{+}_{g,t}, x^{-}_{g,t} \in \{0, 1\}, \forall g \in \Omega_{sts}, t, \tag{3d}$$

where $x^{+}_{g,t}$ and $x^{-}_{g,t}$ are binary variables representing whether the g-th storage device is charging or discharging in the t-th hour throughout the year, respectively. These binary variables will bring a large amount of computational complexity to the problem, turning the problem into an intractable one. A solution to the high complexity introduced by the binary variables is to implement time-series aggregation, for example, daily load profile or renewable energy output clustering. It selects typical scenarios and reuses decision variables, including the binary variables in the same scenario.

When it comes to long-term storage with a smaller value of γ_g but a smaller value of η^{+}_{g} and η^{-}_{g}, such as a hydrogen storage system, the actual energy cycles of long-term storage devices usually range from several days to several weeks. Therefore, we need to assume the consistency of the SoC at the beginning and the end of a year instead of one day in our optimal planning problem:

$$SOC_{g,t=0} = SOC_{g,t=8760}, \forall g \in \Omega_{lts}, \tag{4}$$

where Ω_{lts} is the set of all long-term storage devices.

It is unreasonable to reduce decision variables based on simple daily load profiles or renewable energy output clustering for long-term storage devices because they have larger energy charging and discharging cycles. Thus, the local MES planning model with constraints (1)-(4) is still a very large-scale mixed-integer program. How to develop new time-series aggregation and modeling methods to largely reduce decision variables as well as guarantee the accuracy of the local MES planning model is the main problem to be solved in this paper.

2.2 | Proposed framework

in contrast to the SoC constraint of short-term energy storage in (2), the SoC of long-term energy storage $SOC_{g,t}$ changes by the end of a day. Thus, an intuitive idea is that: can we introduce an additional variable SOC^{d}_{g,t_d} to represent the SoC change in the d-th day, so that the "rest" SoC (i.e. $SOC_{g,t} - SOC^{d}_{g,t_d}$) can be decoupled?

Based on this idea, we decompose $SOC_{g,t}$ of long-term storage devices into two components:

$$SOC_{g,t} = SOC^{h}_{g,t_h,s} + SOC^{d}_{g,t_d}, \forall g \in \Omega_{lts}, t, \tag{5}$$

where $SOC_{g,t}$ is the actual SoC at the t-th hour in the whole year; SOC^{d}_{g,t_d} denotes the daily resolution component at the t_d-th day, which reflects the change of the SoC on a daily basis and has an annual cycle; $SOC^{h}_{g,t_h,s}$ denotes the corresponding hourly resolution component, which reflects the hourly fluctuation of the full-time series and has a daily cycle and satisfies:

$$SOC^{h}_{g,t_h=0,s} = SOC^{h}_{g,t_h=24,s}, \forall g \in \Omega_{lts}, s, \tag{6}$$

which is similar to (2). In this way, we can reuse $SOC^{h}_{g,t_h,s}$ in the same typical scenarios.

Take one-year hourly resolution data as an example, we have to introduce 8760 variables ($SOC_{g,t}$) to accurately represent the SoC of the g-th storage. However, using (5) and (6), we only need to introduce $24 \times s + 365$ variables ($24 \times s$ for $SOC^{h}_{g,t_h,s}$ and 365 for SOC^{d}_{g,t_d}) to approximate the original constraint in (4). Then, the next problem is how to implement the decomposition in (5) and (6) to reduce the approximate error as much as possible.

Inspired by the decomposition of $SOC_{g,t}$ in (5), we propose a novel framework for local MES planning as shown in Figure 1. It contains three major blocks: (1) time series decomposition-aggregation, (2) superposed operation modeling, and (3) optimization based on all components.

The operation of local MES is mainly driven by uncontrollable components such as energy demand and renewable energy output. Having a deep investigation of the behavior of these uncontrollable components helps to reduce the approximation error. Thus, we first decompose the energy demand and renewable energy output time series data into the hourly fluctuation and the daily change components before aggregation:

$$X_t = X^{d}_{t_d} + X^{h}_{t_h,s}, \forall t, \tag{7}$$

where X_t is the actual value at the t-th hour of the aggregated time series; $X^{d}_{t_d}$ represents the corresponding daily resolution component at the t_d-th day; $X^{h}_{t_h,s}$ denotes the hourly resolution component at the t_h-th hour in the s-th typical scenario, respectively. The time series decomposition-aggregation method is applied to all input time series, including all kinds of renewable generation power and energy demands.

On this basis, we implement superposed operation modeling, where the operational status of all devices except short-term storage devices can be viewed as a superposition of both $X^{d}_{t_d}$

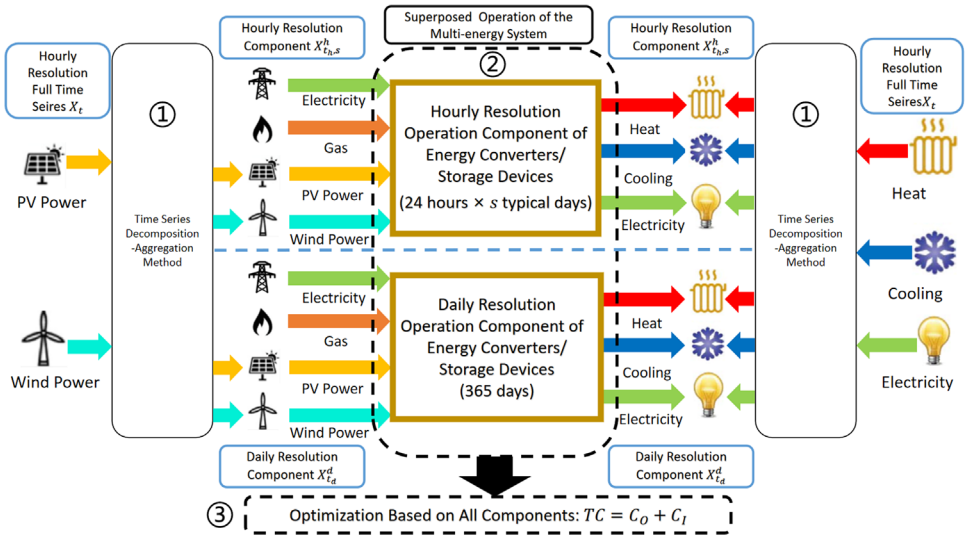

FIGURE 1 Illustration of the proposed framework for planning the local MES. Three major blocks include the time-series decomposition-aggregation method, superposed operation modeling, and optimization based on all components.

and $X^h_{t_h,s}$, which is formulated as:

$$P_{g,t} = P^d_{g,t_d} + P^h_{g,t_h,s}, \forall g \in \Omega_c, \Omega_{lls}, \Omega_{reg}, t, \tag{8}$$

where $P_{g,t}$ is the actual power of the g-th device at the t-th hour; $P^d_{t_d}$ represents the corresponding daily resolution component at the t_d-th day; $P^h_{t_h,s}$ denotes the hourly resolution component at the t_h-th hour in the s-th typical scenario, respectively; $\mathbf{\Omega}_{reg}$ is the set of all renewable energy generation devices.

Finally, we establish the new optimization model, where the total cost consisting of investment and operational costs is minimized considering both P^d_{g,t_d} and $P^h_{g,t_h,s}$, and then the optimal investment and operation strategy of each device throughout the year is derived.

The accuracy and efficiency of the proposed framework lie in the decomposition within the time series aggregation method and the superposed operation modeling. In terms of time series aggregation, the combination of different $X^d_{t_d}$ and $X^h_{t_h,s}$ can result in different X_t. Because of the multiplication principle, the aggregated time series is accurate even with a few typical scenarios. Similarly, the superposition of P^d_{g,t_d} and $P^h_{g,t_h,s}$ also guarantees a highly expressive operational model throughout the year.

3 | TIME SERIES DECOMPOSITION-AGGREGATION

In this section, a novel time series aggregation method is proposed to extract the hourly and daily variations of the full-time series. We first introduce the principle of Multiple Seasonal-Trend decomposition using Loess (MSTL) algorithm [25], which is applied in our method first and decomposes the full-time series into multiple seasonal components. On this basis, the proposed Time series Decomposition-Aggregation Method is detailed.

3.1 | Introduction to MSTL

For a full-time series X_t, such as a one-year hourly load profile, the decomposition of X_t via the MSTL algorithm is formulated as:

$$X_t = S^1_t + S^2_t + \cdots + S^n_t + T_t + R_t, \tag{9}$$

where S^i_t denotes the i-th seasonal component of X_t; T_t and R_t represent the trend and remainder components of X_t, respectively.

Since MSTL is essentially an extension of the Seasonal-Trend decomposition using Loess (STL) algorithm [26], we introduce the STL procedure first.

The additive decomposition of X_t via the STL algorithm is formulated as:

$$X_t = S_t + T_t + R_t. \tag{10}$$

Generally, the principle of STL is a series of low-pass filters, through which the slowly-changed trend component is extracted from the full-time series. As for the seasonal component, low-pass filters are applied to the cycle-subseries of the full-time series to generate seasonal filters. Noting the cycle of the seasonal component as n_p, all cycle-subseries of a series $\{X_t\}$ are defined as:

$$\{\{X_i, X_{i+n_p}, X_{i+2n_p}, \ldots\} | i = 1, 2, \ldots, n_p\}. \tag{11}$$

Therefore, the low-pass filters applied to the cycle-subseries guarantee the similarity between two nearby cycles, namely the seasonality of the output seasonal component.

In terms of the low-pass filters chosen in STL, locally-weighted regression (loess) is employed repeatedly. The principle of loess is that when calculating regressed value $g(x)$ at

$x \in R$, we only consider observed tuples (x_i, y_i) whose x_i are nearby x.

The procedure of loess is detailed as follows. Note $(x_i, y_i), i = 1, \ldots, n$ as input tuples to be regressed. To take the electricity load profile decomposition as an example, the tuples are $(t, L_{t,j=E}), t = 1, \ldots, 8760$. The loess regression curve, $g(x), x \in R$, is computed based on neighborhood weights. Let q be a positive integer parameter and $\lambda_q(x)$ be the distance of the q-th farthest x_i from x, and then the neighborhood weight $v_i(x)$ for each x_i is:

$$v_i(x) = W\left(\frac{|x_i - x|}{\lambda_q(x)}\right), \tag{12a}$$

$$W(u) = \begin{cases} (1-u^3)^3, & 0 \le u < 1 \\ 0, & u \ge 1 \end{cases}, \tag{12b}$$

which means only tuples (x_i, y_i) whose $x_i \in [x - \lambda_q(x), x + \lambda_q(x)]$ have non-zero weights when evaluating $g(x)$, and the closer x_i is to x, the bigger the neighborhood weight $v_i(x)$ is.

Next, the loess regression curve $g(x)$ at x is derived by fitting a polynomial of degree d with neighborhood weight $v_i(x)$ for each (x_i, y_i). To take $d = 2$ as an example, the loess regression curve $g(x)$ equals:

$$g(x) = a(x)x^2 + b(x)x + c(x), \tag{13}$$

where $a(x), b(x), c(x)$ are coefficients at x derived from the weighted least square method whose loss function $J(a, b, c)$ is:

$$J(a, b, c) = \frac{1}{n}\sum_{i=1}^{n} v_i(ax^2 + bx + c - y_i)^2. \tag{14}$$

To evaluate $g(x)$ for each $x \in R$, first, the coefficients of $g(x)$ at x (i.e. $a(x), b(x), c(x)$ when $d = 2$) are given by the weighted least square method based on (14). Next, $g(x)$ is calculated with given $a(x), b(x), c(x)$, and x according to the expression of $g(x)$ based on (13). If a large q is selected, the considered scope $[x - \lambda_q(x), x + \lambda_q(x)]$ becomes large when evaluating $g(x)$, which plays an effect of low-filter.

Based on loess, the procedure of the STL algorithm is introduced in Algorithm 1. Generally, the procedure is divided into the inner loop, which updates both the trend component and seasonal components, and the outer loop, which guarantees the robustness of the algorithm. n_o and n_i are the outer loop and inner loop times, respectively.

As for the k-th inner loop, first, the latest trend component T_t^k is subtracted from the full-time series X_t, yielding the detrended component C_t^{k+1}. Next, loess is used to smooth all cycle-subseries of the detrended series, which strengthens the seasonality of C_t^{k+1}. Then a low-pass filter consisting of two moving averaging procedures with a length of n_p extracts more trend components from the C_t^{k+1} and yields L_t^{k+1} which changes slowly. Finally, the latest seasonal component S_t^{k+1} is defined as $C_t^{k+1} - L_t^{k+1}$, and the the latest trend component T_t^{k+1} is derived after doing loess to $X_t - S_t^{k+1}$.

ALGORITHM 1 STL Algorithm

Input: Full time series X_t, cycle of the seasonal component n_p, number of outer loop and inner loop n_o, n_i
Output: The seasonal decomposition: $[T_t, S_t, R_t]$

Initialize: $T_t = 0, S_t = 0, \rho_t = 1$;
for *j in 1:*n_o **do**
  Initialize the inner loop: $T_t^0 = T_t, S_t^0 = S_t$;
  for *k in 1:*n_i **do**
    Detrending of X_t: $C_t^k = X_t - T_t^{k-1}$;
    Cycle-subseries(with the cycle of n_p) loess Smoothing of C_t^k;
    L_t^k = low-pass filter (C_t^k) (two moving averaging procedures with a length of n_p);
    Detrending of C_t^k: $S_t^k = C_t^k - L_t^k$;
    Updating Trend: $T_t^k = loess(X_t - S_t^k)$;
  end
  Updating seasonal and trend component: $T_t = T_t^{n_i}, S_t = S_t^{n_i}$;
  Updating remainder: $R_t = X_t - T_t - S_t$;
  Updating robustness weights: $\rho_t = B(|R_t|/(6 \times \text{median}(|R_t|)))$;
end
Return $[T_t, S_t, R_t]$;

In the outer loop, the remainder R_t is calculated based on the latest seasonal component $S_t^{n_i}$ and the trend component $T_t^{n_i}$. On this basis, STL defines robustness weights ρ_t for each observed tuples (t, X_t) to reflect how extreme R_t is, where ρ_t and $B(u)$ are defined as:

$$\rho_t = B(|R_t|/(6 \times \text{median}(|R_t|))), \tag{15a}$$

$$B(u) = \begin{cases} (1-u^2)^2, & 0 \le u < 1 \\ 0, & u \ge 1 \end{cases}, \tag{15b}$$

which means the bigger $|R_t|$ is, the smaller the robustness weight ρ_t is.

In the next inner loop, the neighborhood weights v_t are multiplied by the robustness weights ρ_t in each loess procedure. The larger $|R_t|$ is, the smaller weight the value at t-th period has in the inner loop afterward, which weakens the effect of outliers and strengthens the robustness of the algorithm.

Besides tackling outliers, the robustness of the STL algorithm is also reflected in the tolerance of the missing values, which owes to the utilization of loess. During the procedure of loess, observed tuples (x_i, y_i) are not required to have evenly distributed x_i. If some data is lost in the full-time series X_t, in the following procedures the missing value will be replaced with the latest loess regressed value $g(x)$, which is calculated based on the observations close to the missing time.

For more details of parameters in the STL algorithm and their suggested values, please refer to [26].

The MSTL algorithm uses STL iteratively to estimate each seasonal component in the full-time series. Because the seasonal components with lower seasonal cycles might be absorbed by the ones with higher cycles, the MSTL algorithm extracts multiple seasonal components in ascending order of seasonal cycles.

ALGORITHM 2 MSTL algorithm

Input: Full time series X_t, number of iteration n_{iter}, cycle of the seasonal component $N_p = [n_{p_1}, \cdots, n_{p_n}]$
Output: The seasonal decomposition: $[T_t, S_t = [S_t^1, S_t^2, \cdots, S_t^n], R_t]$

Sort $N_p = [n_{p_1}, \cdots, n_{p_n}]$ in the ascending order;
Initialize $S_t = [S_t^1, S_t^2, \cdots, S_t^n]$ into all zero series;
for *i in 1:*n_{iter} **do**
  for *j in 1:*n **do**
    $X_t = X_t + S_t^j$;
    $[T_t, S_t^j, R_t^j]$ = STL$(X_t, n_{p_j}, ...)$;
    $X_t = X_t - S_t^j$;
  end
end
$R_t = X_t - T_t$;
Return $[T_t, S_t, R_t]$

As shown in Algorithm 2, at the very beginning the input list of cycles $N_p = [n_{p_1}, \ldots, n_{p_n}]$ is sorted in ascending order. After initializing the seasonal components, in each iteration loop, the MSTL algorithm first adds the latest seasonal component S_t^j to the full-time series X_t to strengthen the seasonality of the cycle n_{p_j}. Next, the STL algorithm is implemented to extract the latest seasonal component with seasonal cycle n_{p_j}. Then S_t^j is subtracted from X_t to avoid being absorbed by seasonal components with the higher seasonal cycle. Finally, the trend component T_t is updated using the results from the last implementation of STL, and the remainder component R_t is derived after subtracting T_t from X_t.

Figure 2 shows the decomposition results of the full-time series of electricity load $L_{t,j=E}$ in one year via the MSTL algorithm. Since the length of all time series considered in this paper is one year, the yearly component of the full-time series is reflected in the trend component.

3.2 | Framework of time series decomposition-aggregation

The novel Time Series Decomposition-Aggregation method incorporates the MSTL algorithm and time series aggregation. As shown in Figure 3, the aggregation strategy contains three major blocks: (1) decomposition of the full-time series X_t via the MSTL algorithm, (2) the selection of typical scenarios, and (3) upsample of the remaining components.

First, the full-time series X_t of hourly resolution is decomposed into the trend T_t, daily component S_t^d, weekly component S_t^w, and the remainder R_t via the MSTL algorithm [25]. Note that our framework is general and other advanced seasonal-trend decomposition algorithms [27–29] can also be adopted here.

Second, the typical scenarios are selected from the sum of all components reflecting hourly fluctuation, including S_t^d, S_t^w, and R_t, via clustering methods such as k-means. The output is the hourly resolution component $X_{t_h,s}^h$.

Finally, the trend component T_t and the remainder of the clustering procedure R_t' are added for upsampling. Because $X_{t_h,s}^h$ has extracted hourly fluctuation of the full-time series, the components left mainly reflect the slow change of the full-time series. Therefore, the sum of T_t and R_t' are upsampled to daily resolution, yielding the daily resolution component $X_{t_d}^d$.

4 | PROPOSED PLANNING MODEL

In this paper, the aim of local MES planning is to optimally invest in energy converters and storage to satisfy the electricity, heating, and cooling demand with a given penetration rate of renewable energy generation. In this section, the detail of the superposed optimization is described.

4.1 | Objective function

The objective is to minimize the total cost, including the investment cost C_I and the operation cost C_O:

$$TC = C_I + C_O. \tag{16}$$

4.1.1 | Investment Cost

The investment cost is converted to the annual cost considering the payback period of each device and the interest rate:

$$C_I = \sum_{g \in \Omega_c, \Omega_s} \frac{i}{1-(1+i)^{-Y_g}} C_g, \tag{17}$$

where i is the interest rate; Y_g denotes the payback period of the g-th device; C_g is the investment cost of the g-th energy converter device or storage device, which is detailed separately below.

For each energy converter device, the investment cost is mainly determined by its power capacity:

$$C_g = C_{g,\text{unit}} P_{g,\max}^{\text{in}}, \forall g \in \Omega_c, \tag{18}$$

where $C_{g,\text{unit}}$ and $P_{g,\max}^{\text{in}}$ denote the unit input power capacity cost and capacity of the g-th energy converter device, respectively.

For a short-term storage device, the investment cost is mainly determined by its storage capacity:

$$C_g = C_{g,\text{unit}}^E E_{g,\max}, \forall g \in \Omega_{sts}, \tag{19}$$

where $C_{g,\text{unit}}^E$ and $E_{g,\max}$ denote the unit SoC capacity cost and the maximum SoC of the g-th storage device, respectively.

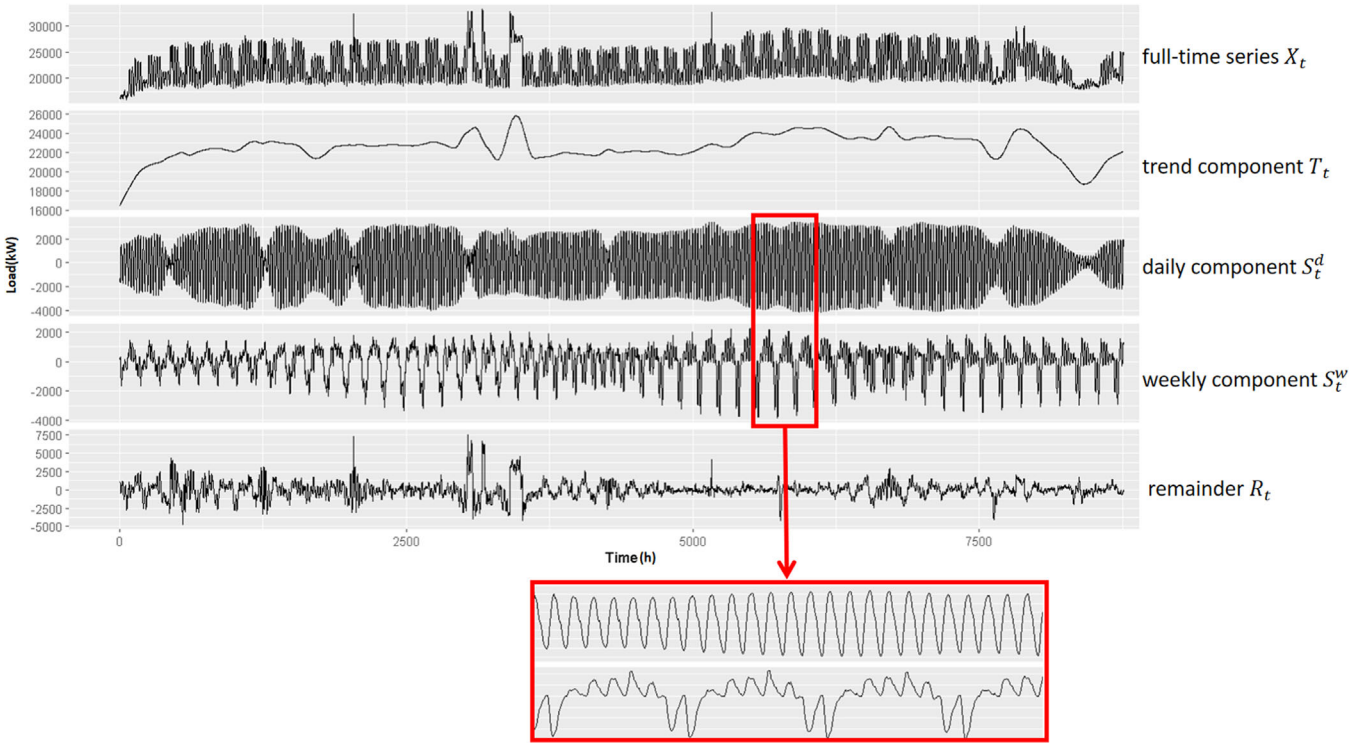

FIGURE 2 The results of the decomposition of electricity load via MSTL algorithm [25] in a year. Parts of the daily and the weekly components are zoomed in to represent the seasonality.

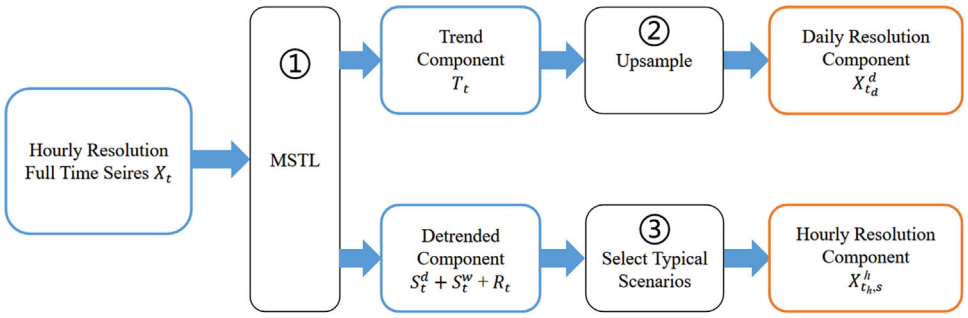

FIGURE 3 Framework of time series decomposition-aggregation.

For a long-term storage device such as a hydrogen storage system, the devices of the charging, storage, and discharging procedure are PEMEC, gas tanks, and PEMFC, respectively. Therefore, the investment cost is determined by these three aspects:

$$C_g = C^+_{g,\text{unit}} P^+_{g,\max} + C^-_{g,\text{unit}} P^-_{g,\max} + C^{SOC}_{g,\text{unit}} SOC_{g,\max}, \forall g \in \Omega_{lls}, \tag{20}$$

where $C^+_{g,\text{unit}}$ and $C^-_{g,\text{unit}}$ are the unit charging power capacity cost and unit electric discharging power capacity cost of the g-th storage device; $P^+_{g,\max}$ and $P^-_{g,\max}$ are the maximum charging power and electric discharging power.

4.1.2 | Operation cost

The operation cost includes the cost of imported electricity and gas and the penalty of the curtailed renewable energy generation power, containing both the daily resolution component and the hourly resolution component:

$$C_O = \sum_{t_d=1}^{T_d} \left(\sum_{j \in J_{\text{in}}} p^d_{j,t_d} I^d_{j,t_d} + \sum_{g \in \Omega_{reg}} p^d_{j=E,t_d} P^{c,d}_{g,j=E,t_d} \right) + \sum_{s=1}^{S} \sum_{t_h=1}^{T_h} \left(\sum_{j \in J_{\text{in}}} n_s p^h_{j,t_h,s} I^h_{j,t_h,s} + \sum_{g \in \Omega_{reg}} p^h_{j=E,t_h,s} P^{c,h}_{g,j=E,t_h,s} \right), \tag{21}$$

where J_{in} is the set of all kinds of imported energy carriers, which is $\{E, G\}$ in this framework, where E, G represent electricity and gas, respectively. p^{d}_{j,t_d} and I^{d}_{j,t_d} are the price and the imported power of the j-th kind of energy on the t_d-th day. Because the imported power remains the same during the t_d-th day, the daily resolution price is the sum of all hourly resolution prices on the t_d-th day. $p^{h}_{j,t_h,s}$ and $I^{h}_{j,t_h,s}$ are the price and the imported power of the j-th kind of energy of the s-th scenario at the t-th hour. $P^{c,d}_{g,j=E,t_d}$ and $P^{c,h}_{g,j=E,t_h,s}$ are the daily and hourly resolution curtailed power of the g-th renewable energy generation devices, which is punished at the cost of the real-time electricity price. n_s denotes the number of the s-th scenario in the entire year.

4.2 | Constraints of devices

The constraints of each device are decomposed into the daily resolution operation constraints, the hourly resolution operation constraints.

4.2.1 | Energy converter devices

For each energy converter device, we assume the relationship between its input and output power can be described by linear constraints:

$$\boldsymbol{F}\left(P^{\text{in}}_{g,j_{\text{in}},t}, P^{\text{out}}_{g,j,t}\right) = \boldsymbol{0}, \forall g \in \Omega_c, j, t, \tag{22a}$$

$$\boldsymbol{G}\left(P^{\text{in}}_{g,j_{\text{in}},t}, P^{\text{out}}_{g,j,t}\right) \geq \boldsymbol{0}, \forall g \in \Omega_c, j, t, \tag{22b}$$

where $\boldsymbol{F}(\cdot), \boldsymbol{G}(\cdot)$ represent linear equations and inequalities, respectively; $P^{\text{in}}_{g,j_{\text{in}},t}$ denotes the input power of the g-th energy converter at the t-th hour; j_{in} denotes the index of the kind of input energy; $P^{\text{out}}_{g,j,t}$ represents the j-th kind of output power at the t-th hour.

Take an energy converter device whose output power is proportional to input power as an example, we have the following constraint:

$$\eta_{g,j} P^{\text{in}}_{g,j_{\text{in}},t} = P^{\text{out}}_{g,j,t}, \forall t, \tag{23}$$

where $\eta_{g,j}$ represents the conversion efficiency.

After decomposition, the daily and hourly resolution components of the operation are formulated as:

$$\eta_{g,j} P^{\text{in},d}_{g,j_{\text{in}},t_d} = P^{\text{out},d}_{g,j,t_d}, \forall g \in \Omega_c, j, t_d, \tag{24a}$$

$$\eta_{g,j} P^{\text{in},h}_{g,j_{\text{in}},t_h,s} = P^{\text{out},h}_{g,j,t_h,s}, \forall g \in \Omega_c, j, t_h, s, \tag{24b}$$

where $P^{\text{in},d}_{g,j_{\text{in}},t_d}$ and $P^{\text{out},d}_{g,j,t_d}$ denote the daily resolution input power and j-th kind of output power of the g-th energy converter on the t_d-th day; $P^{\text{in},h}_{g,j_{\text{in}},t_h,s}$ and $P^{\text{out},h}_{g,j,t_h,s}$ denote the hourly resolution input and j-th kind of output power of the g-th energy converter at the t_h-th hour in the s-th scenario.

The upper bound of the input power is derived from the sum of the maximum of the hourly and daily resolution input power components. Similarly, the lower bound is given by the sum of the minimum of the two components to ensure its non-negativity:

$$P^{\text{in},d}_{g,\min,s} \leq P^{\text{in},d}_{g,j_{\text{in}},t_d} \leq P^{\text{in},d}_{g,\max,s}, \forall g \in \Omega_c, \sigma(t_d) = s, t_d, \tag{25a}$$

$$P^{\text{in},h}_{g,\min,s} \leq P^{\text{in},h}_{g,j_{\text{in}},t_h,s} \leq P^{\text{in},h}_{g,\max,s}, \forall g \in \Omega_c, s, t_h, \tag{25b}$$

$$P^{\text{in},d}_{g,\max,s} + P^{\text{in},h}_{g,\max,s} = P^{\text{in}}_{g,\max} \geq 0, \forall g \in \Omega_c, s, \tag{25c}$$

$$P^{\text{in},d}_{g,\min,s} + P^{\text{in},h}_{g,\min,s} = 0, \forall g \in \Omega_c, s, \tag{25d}$$

where $P^{\text{in},d}_{g,\max,s}$ and $P^{\text{in},d}_{g,\min,s}$ denote the maximum and minimum of the daily resolution input power during the s-th scenario in the whole year, respectively; $P^{\text{in},h}_{g,\max,s}$ and $P^{\text{in},h}_{g,\min,s}$ denote the maximum and minimum of the hourly resolution input power in the s-th scenario, respectively; $\sigma(\cdot)$ denotes the mapping function between the t_d-th day and its corresponding typical scenario.

4.2.2 | Short-term storage devices

The hourly resolution power and SoC variables of short-term storage devices are formulated as:

$$\begin{aligned} SOC^{h}_{g,t_h+1,s} &= (1-\gamma_g) SOC^{h}_{g,t_h,s} + P^{+,h}_{g,j,t_h,s}\eta^{+}_{g} \\ &- P^{-,h}_{g,j,t_h,s}/\eta^{-}_{g}, \forall g \in \Omega_{sts}, t_h, s, \end{aligned} \tag{26}$$

$$0 \leq SOC^{h}_{g,t_h,s} \leq SOC_{g,\max}, \forall g \in \Omega_{sts}, t_h, s. \tag{27}$$

The daily energy cycle is formulated as:

$$SOC^{h}_{g,t_h=0,s} = SOC^{h}_{g,t_h=24,s}, \forall g \in \Omega_{sts}, t_h, s. \tag{28}$$

The constraints that restrict short-term storage devices from simultaneous charging and discharging are formulated as:

$$0 \leq P^{+,h}_{g,j,t_h,s} \leq x^{+,h}_{g,t_h,s} P^{+}_{g,\max}, \forall g \in \Omega_{sts}, t_h, s, \tag{29a}$$

$$0 \leq P^{-,h}_{g,j,t_h,s} \leq x^{-,h}_{g,t_h,s} P^{-}_{g,\max}, \forall g \in \Omega_{sts}, t_h, s, \tag{29b}$$

$$x^{+,h}_{g,t_h,s} + x^{-,h}_{g,t_h,s} \leq 1, \forall g \in \Omega_{sts}, t_h, s, \tag{29c}$$

$$x^{+,h}_{g,t_h,s}, x^{-,h}_{g,t_h,s} \in \{0,1\}, \forall g \in \Omega_{sts}, t_h, s. \tag{29d}$$

The relationship between the power and energy capacity is:

$$P^{+}_{g,\max} = P^{-}_{g,\max} = SOC_{g,\max}/\tau_g, \forall g \in \Omega_{sts}, \tag{30}$$

where τ_g denotes the time to fully charge or discharge the g-th short-term storage device.

4.2.3 | Long-term hydrogen storage system

the hydrogen storage system is allowed to participate in both daily and hourly resolution responses, so the SoC transitions are formulated as:

$$SOC^{h}_{g,t_h+1,s} = SOC^{h}_{g,t_h,s} + P^{+,h}_{g,j=E,t_h,s}\eta^{+}_{g} - P^{-,h}_{g,j=E,t_h,s,}/\eta^{-}_{g,j=E}, \forall g \in \Omega_{lts}, t_h, s, \quad (31a)$$

$$SOC^{d}_{g,t_d+1} = SOC^{d}_{g,t_d} + P^{+,d}_{g,j=E,t_d}\eta^{+}_{g} - P^{-,d}_{g,j=E,t_d}/\eta^{-}_{g,j=E}, \forall g \in \Omega_{lts}, t_d, s, \quad (31b)$$

where $P^{+,d}_{g,j,t_d}$ and $P^{-,d}_{g,j=E,t_d}$ denote the daily resolution charge and discharge electric power of the g-th storage device during the t_d period, respectively.

The energy cycles of SOC^{d}_{g,t_d} and $SOC^{h}_{g,t_h,s}$ are set to be annual and daily:

$$SOC^{h}_{g,t_h=0,s} = SOC^{h}_{g,t_h=24,s}, \forall g \in \Omega_{lts}, t_h, s, \quad (32a)$$

$$SOC^{d}_{g,t_d=0} = SOC^{d}_{g,t_d=365}, \forall g \in \Omega_{lts}, t_d. \quad (32b)$$

The relationship between electricity and heating output of the PEMFC is:

$$P^{\text{out},h}_{g,t_h,s,j=E}/\eta_{g,j=E} = P^{\text{in},h}_{g,t_h,s,j=H}/\eta_{g,j=H}, \forall g \in \Omega_c, t_h, s, \quad (33a)$$

$$P^{\text{out},d}_{g,t_d,j=E}/\eta_{g,j=E} = P^{\text{in},d}_{g,t_d,j=H}/\eta_{g,j=H}, \forall g \in \Omega_c, t_h, s. \quad (33b)$$

Similar to (25), the charging constraints are:

$$P^{+,h}_{g,\min,s} \le P^{+,h}_{g,j,t_h,s} \le P^{+,h}_{g,\max,s}, \forall g \in \Omega_{lts}, s, t_h, \quad (34a)$$

$$P^{+,d}_{g,\min,s} \le P^{+,d}_{g,j,t_d} \le P^{+,d}_{g,\max,s}, \forall g \in \Omega_{lts}, \sigma(t_d) = s, t_d, \quad (34b)$$

$$P^{+,d}_{g,\max,s} + P^{+,h}_{g,\max,s} = P^{+}_{g,\max} \ge 0, \forall g \in \Omega_{lts}, s, \quad (34c)$$

$$P^{+,d}_{g,\min,s} + P^{+,h}_{g,\min,s} = 0, \forall g \in \Omega_{lts}, s. \quad (34d)$$

The discharging constraints are:

$$P^{-,h}_{g,\min,s} \le P^{-,h}_{g,j=E,t_h,s} \le P^{-,h}_{g,\max,s}, \forall g \in \Omega_{lts}, s, t_h, \quad (35a)$$

$$P^{-,d}_{g,\min,s} \le P^{-,d}_{g,j=E,t_d} \le P^{-,d}_{g,\max,s}, \forall g \in \Omega_{lts}, \sigma(t_d) = s, t_d, \quad (35b)$$

$$P^{-,d}_{g,\max,s} + P^{-,h}_{g,\max,s} = P^{-}_{g,\max} \ge 0, \forall g \in \Omega_{lts}, s, \quad (35c)$$

$$P^{-,d}_{g,\min,s} + P^{-,h}_{g,\min,s} = 0, \forall g \in \Omega_{lts}, s. \quad (35d)$$

As for the capacity of the SoC, note that the daily resolution charge power $P^{+,d}_{g,j_{\text{in}},t_d}$ and discharge power $P^{-,d}_{g,j_{\text{in}},t_d}$ remain unchanged during each scenario. Therefore, the actual daily resolution SoC component SOC^{d}_{g,t_d} changes linearly to SOC^{d}_{g,t_d+1} on an hourly basis in the corresponding typical scenario. Therefore, the estimation of the upper bound and lower bound is formulated as:

$$((24 - t_h)SOC^{d}_{g,t_d} + t_h SOC^{d}_{g,t_d+1})/24 + SOC^{h}_{g,t_h,s} \le SOC_{g,\max}, \forall g \in \Omega_{lts}, t_h, \sigma(t_d) = s, \quad (36a)$$

$$((24 - t_h)SOC^{d}_{g,t_d} + t_h SOC^{d}_{g,t_d+1})/24 + SOC^{h}_{g,t_h,s} \ge 0, \forall g \in \Omega_{lts}, t_h, \sigma(t_d) = s. \quad (36b)$$

4.2.4 | Renewable energy generation

Considering the curtailed power, the output of renewable energy generation satisfies:

$$P^{\text{out},h}_{g,j=E,t_h,s} + P^{c,h}_{g,j=E,t_h,s} = P^{\max,h}_{g,t_h,s}, \forall g \in \Omega_{reg}, t_h, s, \quad (37a)$$

$$P^{\text{out},d}_{g,j=E,t_d} + P^{c,d}_{g,j=E,t_d} = P^{\max,d}_{g,t_d}, \forall g \in \Omega_{reg}, t_d, \quad (37b)$$

where $P^{\text{out},h}_{g,j=E,t_h,s}$ and $P^{\text{out},d}_{g,j=E,t_d}$ are the hourly and daily resolution output power of the g-th renewable generation. $P^{c,h}_{g,j=E,t_h,s}$ and $P^{c,d}_{g,j=E,t_d}$ are the curtailed hourly and daily resolution power.

Because the actual output and curtailed power are non-negative, we have:

$$P^{\text{out},d}_{g,j=E,t_d} \ge P^{\text{out},d}_{g,\min,s}, \forall g \in \Omega_{reg}, \sigma(t_d) = s, t_d, \quad (38a)$$

$$P^{\text{out},h}_{g,j=E,t_h,s} \ge P^{\text{out},h}_{g,\min,s}, \forall g \in \Omega_{reg}, s, t_h, \quad (38b)$$

$$P^{\text{out},d}_{g,\min,s} + P^{\text{out},h}_{g,\min,s} = 0, \forall g \in \Omega_{reg}, s, \quad (38c)$$

$$P^{c,d}_{g,j=E,t_d} \ge P^{c,d}_{g,\min,s}, \forall g \in \Omega_{reg}, \sigma(t_d) = s, t_d, \quad (38d)$$

$$P^{c,h}_{g,j=E,t_h,s} \ge P^{c,h}_{g,\min,s}, \forall g \in \Omega_{reg}, s, t_h, \quad (38e)$$

$$P^{c,d}_{g,\min,s} + P^{c,h}_{g,\min,s} = 0, \forall g \in \Omega_{reg}, s. \quad (38f)$$

4.2.5 | Energy balance

For all kinds of considered energy carriers, namely electricity (E), heating (H), cooling (C), and natural gas (G), the hourly resolution power balance is formulated as:

$$I^{h}_{j,t_h,s} + \Delta L_{j,s} + \sum_{g \in \Omega_c, \Omega_{reg}} P^{\text{out},h}_{g,j,t_h,s} - \sum_{g \in \Omega_c} P^{\text{in},h}_{g,j,t_h,s}$$

$$+\sum_{g\in\Omega_s}\left(P^{-,h}_{g,j,t_h,s}-P^{+,h}_{g,j,t_h,s}\right)=L^{h}_{j,t_h,s},\forall j\in J,t_h,s, \quad (39)$$

where $L^{h}_{j,t_h,s}$ denotes the j-th kind of energy demand in the t_h-th hour of the s-th scenario; $\Delta L_{j,s}$ represents the daily resolution load transferred to the daily resolution balance in the s-th scenario; $J=\{E,G,H,C\}$ denotes the set of all energy carriers including electricity, gas, heating, and cooling.

The daily resolution power balance is formulated as:

$$I^{d}_{j,t_d}-\Delta L_{j,s}+\sum_{g\in\Omega_c,\Omega_{reg}}P^{\text{out},d}_{g,j,t_d}-\sum_{g\in\Omega_c}P^{\text{in},d}_{g,j,t_d}+\sum_{g\in\Omega_s}\left(P^{-,d}_{g,j,t_d}-P^{+,d}_{g,j,t_d}\right)=L^{d}_{j,t_d},\forall j\in J,t_d,\sigma(t_d)=s, \quad (40)$$

where L^{d}_{j,t_d} denotes the j-th kind of energy demand on the t_d-th day.

4.3 | Model linearization

Because there exist products of binary variable and continuous variable in (29), this model leads to a mixed-integer linear program (MILP). In this section, the big M method is used to linearize the model.

For each period t_h, big M is introduced to indicate whether the short-term storage is charging or discharging:

$$0\le P^{+,h}_{g,j,t_h,s}\le x^{+,h}_{g,t_h,s}M,\forall g\in\Omega_{sts},t_h,s, \quad (41a)$$

$$0\le P^{-,h}_{g,j,t_h,s}\le x^{-,h}_{g,t_h,s}M,\forall g\in\Omega_{sts},t_h,s, \quad (41b)$$

$$x^{+,h}_{g,t_h,s}+x^{-,h}_{g,t_h,s}\le 1,\forall g\in\Omega_{sts},t_h,s, \quad (41c)$$

$$x^{+,h}_{g,t_h,s},x^{-,h}_{g,t_h,s}\in\{0,1\},\forall g\in\Omega_{sts},t_h,s. \quad (41d)$$

Besides, the power capacity of short-term storage devices is estimated without importing binary variables:

$$0\le P^{+,h}_{g,j,t_h,s}\le P^{+}_{g,\max},\forall g\in\Omega_{sts},t_h,s, \quad (42a)$$

$$0\le P^{-,h}_{g,j,t_h,s}\le P^{-}_{g,\max},\forall g\in\Omega_{sts},t_h,s. \quad (42b)$$

Finally, the proposed MILP model is derived:

$$\begin{aligned}&\text{Obj:}\quad \text{Eq. (16)}\\ &\text{s.t.}\quad \text{Eqs. (17)}-\text{(28)},\text{(31)}-\text{(42)}.\end{aligned} \quad (43)$$

5 | CASE STUDIES

In this section, the experimental setup is first introduced. Next, the basic results derived from the proposed framework are discussed and compared with the ones given by full-time series optimization (FSO). Finally, two comparative analyses verify the effectiveness of the proposed time-series decomposition-aggregation method and decomposition-based planning method.

5.1 | Experimental setup

We are planning a local MES, whose input energy sources include imported electricity and gas and renewable energy consisting of solar and wind generation. The normalized output of solar and wind generation is derived from [31]. The installed power capacity is determined by the renewable energy generation r, which is defined as the ratio of all annual renewable energy generation to annual electricity demand.

To embody the utility of collaboration between short-term and long-term storage in consuming renewable generation, we investigate the performance of the multi-energy system with a very high penetration of renewable energy generation, that is, $r=2$.

The multi-energy demand consists of electricity, heating, and cooling. The tested full-time series of these loads is from The Austin station in Building Data Genome Project 2 [30]. Figure 4 illustrates the demands for electricity, heating, and cooling throughout the year.

5.2 | Basic results

Figure 5 illustrates the results of the aggregated cooling load profiles via the proposed Time Series Decomposition-Aggregation method. To verify its accuracy, Figure 6 compares the RMS error of aggregated time series via the k-means method without decomposition and the proposed Time Series Decomposition-Aggregation method. The aggregated time series include electricity load, heating load, cooling load, solar power, and wind power profiles, and the number of typical scenarios ranges from 1 to 30 days. It can be shown that with the help of decomposition, the aggregated time series describes the original one more accurately compared with direct k-means clustering.

Table 1 compares the total cost and its composition, and the number of variables and constraints of FSO and the proposed model. It can be noted that the variables and constraints decrease by nearly two orders of magnitude in the proposed model while maintaining a fine agreement in total cost (+6.78%), investment cost (+10.84%), and operation cost (−4.55%) with the results given by FSO.

The comparison of the operation of the hot water sensible thermal storage (HWTS) and the hydrogen storage is illustrated in Figure 7. The pixel in the i-th row of the j-th column represents the SoC at the i-th hour of the j-th day. The yellower the pixel is, the higher the state of charge (SOC) is at this time.

Generally, the main operation characteristic of these two storage systems differs from each other in both models. As a short-term storage device, HWTS focuses on the daily energy

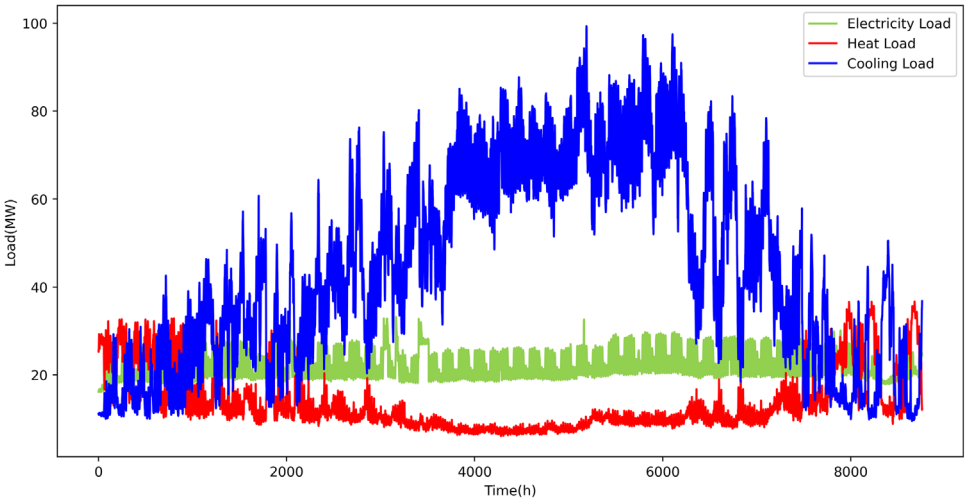

FIGURE 4 Electricity, heating, and cooling demands [30].

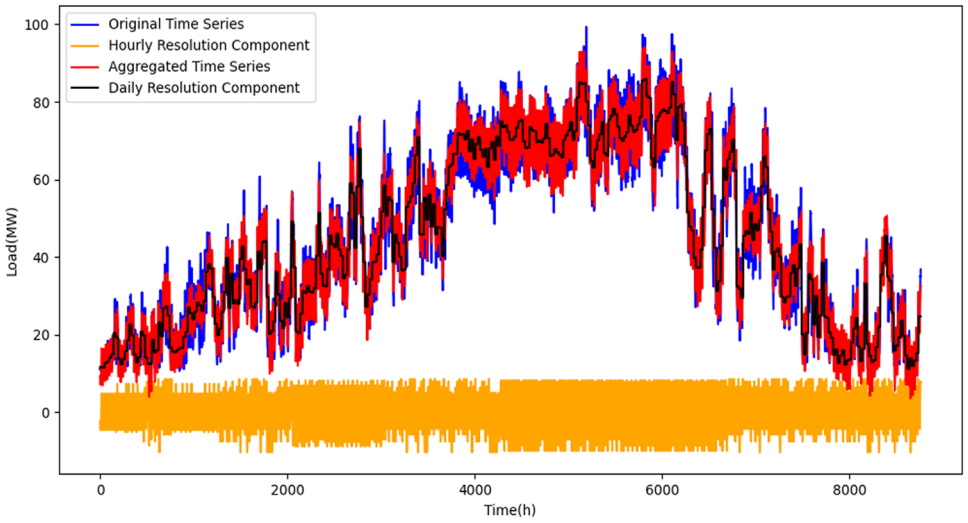

FIGURE 5 The original full-time cooling load and aggregation result using the proposed time series aggregation method. The sum of the hourly resolution component and the daily resolution component yields the aggregated time series.

TABLE 1 Costs and complexity of the proposed model with 10 typical scenarios and the full-time series optimization (FSO).

	FSO	Proposed model	Relative difference (%)
Total cost (million RMB)	68.866	73.536	+6.78
Investment cost (million RMB)	50.721	56.217	+10.84
Operation cost (million RMB)	18.145	17.319	−4.55
Binary variables	35040	960	−97.26
Continuous variables	236540	15843	−93.30
Constraints	332893	44118	−86.74

imbalance, getting charged from 10:00 to 16:00 and discharged during the rest of the time. The daily energy cycle given by FSO supports (28) in our model. Because the aggregation reduces the variation of the load profiles, the fluctuation of the SoC is reduced in the proposed model. However, within limited complexity, the proposed model describes a highly similar operation pattern to the SoC of HWTS.

On the other hand, the operation pattern of the long-term storage device is more complicated. It can be shown that the energy cycle of hydrogen storage ranges from days to weeks both in FSO and the proposed framework. However, the proposed model can follow a complicated pattern and generate a similar SoC series, guaranteeing the accuracy of the proposed model.

Besides, it can be noted from Figure 7 that the seasonality of the operation pattern of the long-term storage device

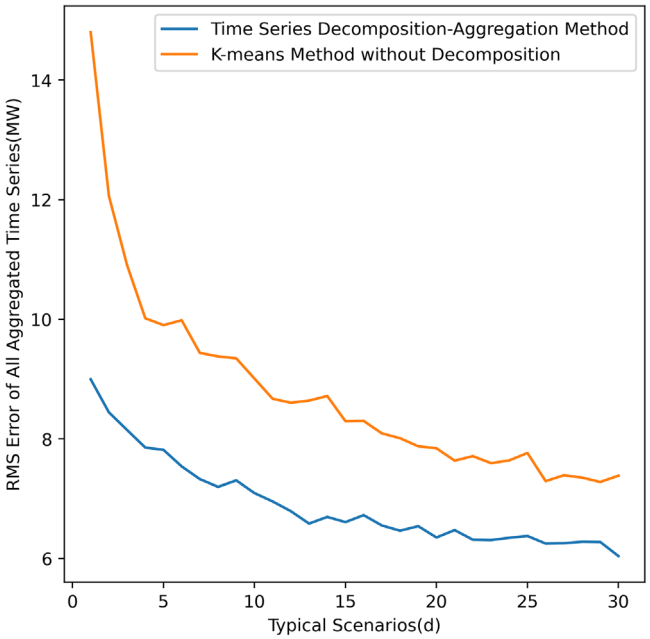

FIGURE 6 The comparison of the RMS error of all aggregated time series via the k-means method without decomposition and the proposed Time Series Decomposition-Aggregation method. The aggregated time series include electricity load, heating load, cooling load, solar power, and wind power profiles, and the number of typical scenarios ranges from 1 to 30 days.

is not obvious. Instead, it exchanges energy across weeks. There are two reasons for this phenomenon. First, cooling and heating load profiles have obvious but contemporary seasonality. Because different energy carriers are coupled through energy converters, the cooling and heating loads can be transformed into equivalent electricity loads. Therefore, the total equivalent electricity loads are high in winter because of heating loads, and high in summer because of cooling loads, weakening the seasonality of the total equivalent electricity loads. Second, renewable energy generation, especially wind power, usually has an imbalance across weeks. In this case, long-term storage also acts as a more economical solution for tackling the imbalance than short-term storage because of its low self-discharge rate and low unit SoC capacity cost.

5.3 | Comparative analyses

5.3.1 | Benchmarks

To further verify the accuracy and efficiency of the proposed framework, we compare our proposed method with two benchmarks to show the benefits of the decomposition of time series and the operation of devices.

Benchmark 1 is inspired by the M1 method in [21]. The full-time series are aggregated in the benchmark via clustering methods such as k-means without decomposition. All power decision variables $P_{g,t_h,s}$ are reused in the same typical scenarios, while all SoC variables SOC_{g,t_d,t_h} describe the SoC on an hourly basis throughout the year to embody energy cycles higher than one day. The transition constraints of SoC are formulated as:

$$SOC_{g,t_d,t_h+1} = SOC_{g,t_d,t_h} + P^{+}_{g,t_h,s}\eta^{+}_{g} - P^{-}_{g,t_h,s}/\eta^{-}_{g}, \forall g \in \Omega_s, t_h, \sigma(t_d) = s, s, \tag{44a}$$

$$SOC_{g,t_d,t_h=0} = SOC_{g,t_d-1,t_h=23} + P^{+}_{g,t_h=1,s}\eta^{+}_{g} - P^{-}_{g,t_h=1,s}/\eta^{-}_{g}, \forall g \in \Omega_s, \sigma(t_d) = s, s, \tag{44b}$$

$$SOC_{g,t_d=0,t_h=0} = SOC_{g,t_d=364,t_h=23}, \forall g \in \Omega_s. \tag{44c}$$

Benchmark 2 is based on our proposed framework, but the hydrogen storage system has no hourly response, namely:

$$SOC^{h}_{g,t_h,s} = 0, P^{+,h}_{g,j,t_h,s} = P^{-,h}_{g,j=E,t_h,s} = 0, \forall g \in \Omega_{hs}, j, t_h, s. \tag{45}$$

To verify the importance of the hourly response in long-term storage devices and the decomposition of the operation, the results of the proposed framework whose hydrogen storage system is with and without hourly response are compared.

TABLE 2 Total cost and its composition, and the complexity of the proposed model and Benchmark 1 with 10 typical scenarios. Benchmark 1 is inspired by the M1 method in [21].

	Proposed Model	Benchmark 1	Improvement
Total cost (million RMB)	73.536(+6.78%)	61.810(−10.25%)	3.47%
Investment cost (million RMB)	56.217(+10.84%)	41.779(−17.63%)	6.79%
Operation cost (million RMB)	17.319(−4.55%)	20.031(+10.4%)	5.85%
Binary variables	960(−97.26%)	960(−97.26%)	0%
Continuous variables	15843(−93.30%)	32540(−86.24%)	7.06%
Constraints	44118(−86.74%)	60733(−81.76)	4.98%

5.3.2 | Comparative results

For the comparison with Benchmark 1, Table 2 gives the total cost and its composition, and the number of variables and constraints of the proposed model and Benchmark 1 with the same number of typical scenarios, namely 10 typical scenarios.

It can be noted that both the proposed model and Benchmark 1 decrease the complexity of FSO by nearly two orders of magnitude. However, the proposed model achieves higher accuracy both in the total cost and its composition, namely 3.47%, 6.79%, and 5.85% improvement in total cost, investment cost, and operation cost, respectively. Furthermore, the complexity of the proposed model is lower than Benchmark 1, with 7.06% and 4.98% improvement in the number of continuous variables

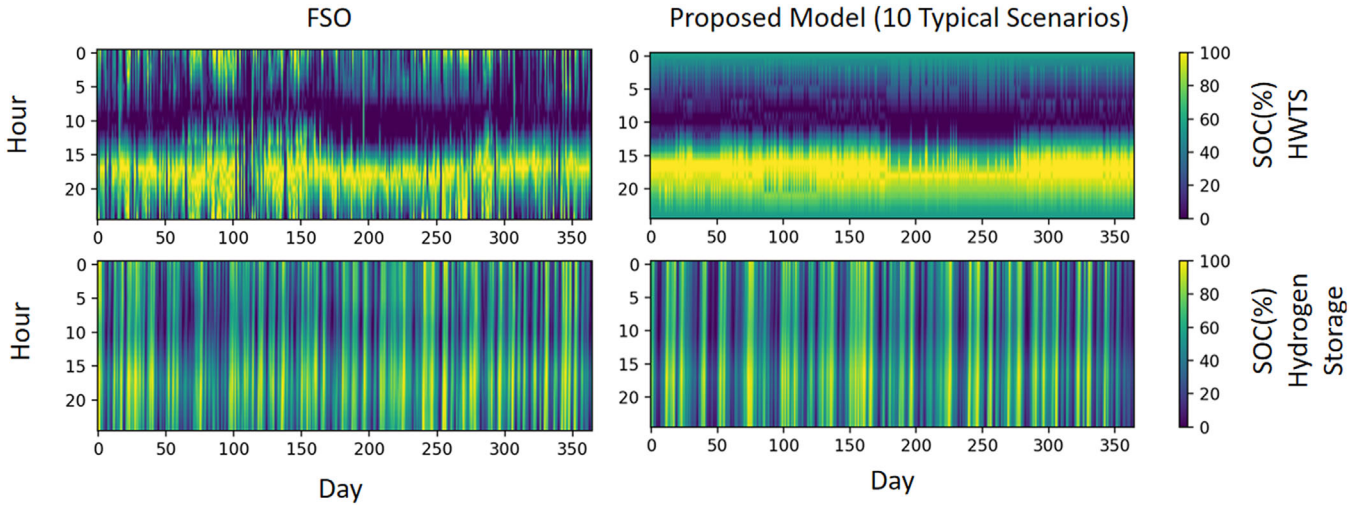

FIGURE 7 Comparison of the SoC of the hot water sensible thermal storage (HWTS) and the hydrogen storage throughout the year using full-time series optimization and the proposed framework with 10 typical scenarios.

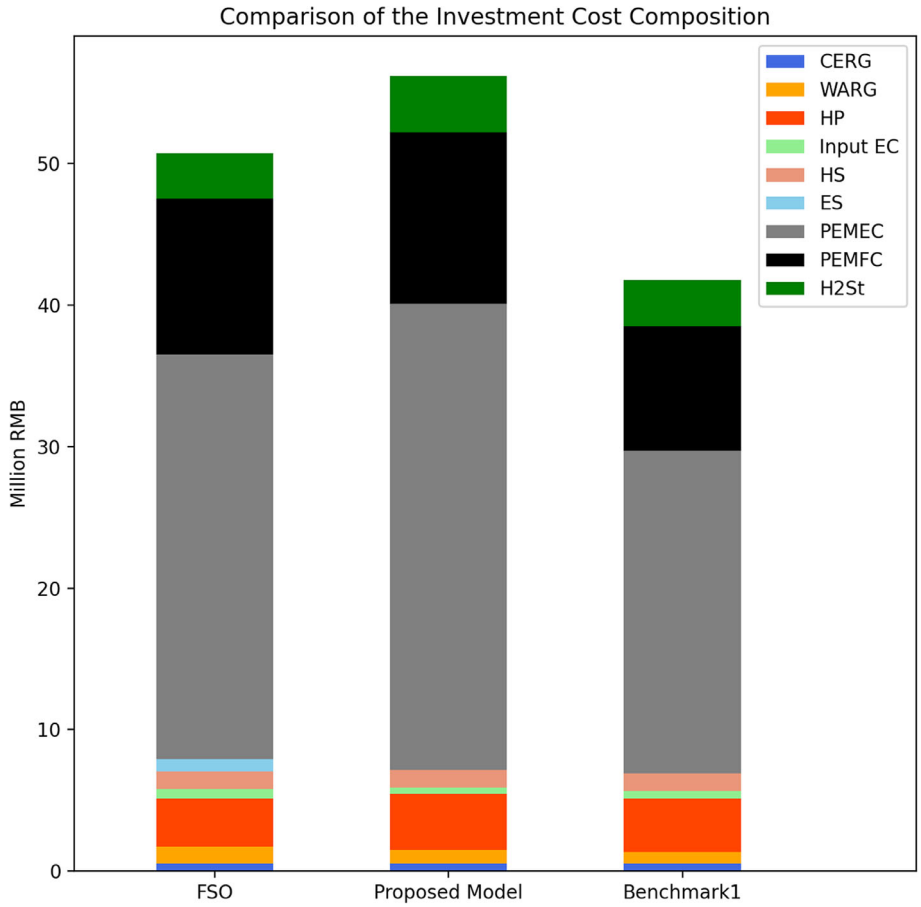

FIGURE 8 Comparison of the investment cost composition derived from the full-time series optimization, the proposed framework, and Benchmark 1.

and constraints. The improvement in all these aspects verifies the accuracy and efficiency of the proposed framework.

To further verify the accuracy of the proposed framework, Figure 8 compares the composition of the investment cost derived from FSO, the proposed framework, and Benchmark 1. It can be observed that with a very high penetration rate of renewable energy and punishment of shedding renewable energy, the cost of the hydrogen storage system accounts for the main part of the investment cost. While both the proposed model and Benchmark 1 yield similar investment cost structures

TABLE 3 Total cost and its composition of the proposed model and Benchmark 2 with 10 typical scenarios. Benchmark 2 is still the proposed framework, but the hydrogen storage system has no hourly response.

Cost(million RMB)	Proposed model	Benchmark 2	Improvement
Total cost	73.536(+6.78%)	117.024(+69.93%)	+63.15%
Investment cost	56.217(+10.84%)	36.177(−28.67%)	+17.83%
Operation cost	17.319(−4.55%)	80.847(+345.56%)	+341.01%
Operation cost in importing	5.244(−29.46%)	15.296(+105.76%)	+76.30%
Operation cost in curtailng	12.075(+12.73%)	65.550(+511.99%)	+499.26%

with the one given by FSO, the proposed model results in a more accurate investment suggestion, especially in PEMEC and PEMFC.

The improvement is mainly credited to the decomposition of the time series. The decomposition of the time series guarantees a higher accuracy of the aggregated time series, enhancing the efficiency and accuracy of the proposed model.

For the comparison with Benchmark 2, Table 3 gives the total cost and its composition of the proposed model and Benchmark 2 with the same number of typical scenarios, namely 10 typical scenarios.

It can be noted that the results of Benchmark 2 do not agree with the ones given by FSO, especially in the operation cost of curtailing renewable energy (+511.99% relative difference). In comparison, the relative difference is much smaller in the proposed model (+12.73%), which means higher consumption of renewable energy generation.

The improvement is mainly credited to the decomposition of the operation of long-term storage devices, which greatly strengthens the ability of long-term storage devices to consume the highly-fluctuated renewable energy generation.

6 | CONCLUSIONS

This paper proposes a novel framework for the efficient planning of MES with long-term storage. By incorporating decomposition into time series aggregation methods and the operation of energy devices, the planning model is able to describe the energy exchange of long-term storage devices throughout the year and also the hourly operation characteristic at the same time. Experiments show a considerable decrease in the complexity of the problem while guaranteeing the accuracy of the results. The variables and constraints decrease by nearly two orders of magnitude in the proposed model while a fine agreement is maintained with the results given by full-time series optimization in total cost (+6.78%), investment cost (+10.84%), and operation cost (−4.55%). In addition, we find that short-term responses of long-term storage devices are crucial for the MES to accommodate a high penetration of renewable energy. Comparison between the results of the proposed framework and Benchmark 2 shows a significantly higher operation cost of curtailing renewable energy (+511.99% in relative difference) when long-term storage devices have no short-term responses, while the relative difference is much smaller in the proposed model (+12.73%) where long-term storage devices have short-term responses.

The main limitation of the proposed framework is that the network of each energy carrier is not considered. The large number of nodes in the network can bring a curse of dimensionality when selecting typical scenarios, which becomes one of the obstacles when promoting the proposed framework to the cases with networks. In this case, more characteristics of full-time series such as the correlation between each two series should be mined and utilized.

Future works on the planning of MES with long-term storage should include two aspects. First, the networks of each energy carrier should be considered. More characteristics should be exploited to overcome the curse of dimensionality. Second, the robustness of the designed system against climate risk and uncertainty should be considered. Because the penetration rate of renewable generation is high in the test case, it's necessary to consider the extreme weather and the uncertainty of renewable generation power.

NOMENCLATURE

Indexes and Sets

Ω_c	set of energy converter devices
Ω_s	set of storage devices
Ω_{sts}	set of short-term storage devices
Ω_{lts}	set of long-term storage devices
Ω_{reg}	set of renewable energy generation devices
g	index of devices
J	set of all kinds of energy carriers ($\{E, H, C, G\}$)
J_{in}	set of all kinds of imported energy ($\{E, G\}$)
j	index of energy carriers
S	number of typical scenarios (day)
s	index of typical scenarios (day)
T_d	number of days of a year
t_d	time index (day of the year)
T_h	number of hours in a scenario(day)
t_h	time index (hour of the day)

Decision Variables

I	imported power
x	binary variables of charging or discharging state
P	input or output power
SOC	state of charge
L	energy load

Parameters and Constants

p	the price of the imported power of the j-th energy
η	conversion or storage efficiency

γ self-discharge rate
τ the time to fully charge or discharge a storage device

Supercripts

h	hourly resolution component
d	daily resolution component
in	input
out	output
+	charge
−	discharge
c	curtailed power
initial	the initial time
final	the final time

Acronym

max	maximum
min	minimum
PEMEC	proton exchange membrane electrolyzer
PEMFC	proton exchange membrane fuel cell
CHP	combined heat and power
AB	auxiliary boiler
CERG	compression electric refrigerator group
WARG	water absorption refrigerator group
HP	heat pump
HWTS	hot water sensible thermal storage
LiB	lithium battery
HS	hydrogen storage
ES	electrical storage
Input EC	input electricity capacity
Input GC	input gas capacity

AUTHOR CONTRIBUTIONS

Jiahao Ma: Methodology; software; writing – Original – Draft; Ning Zhang: Data Curation; Funding Acquisition; Resources; writing – Original – Draft; writing – Review – Editing; Qingsong Wen: Project Administration; validation; writing – Original – draft; writing – Review – Editing; Yi Wang: Conceptualization; supervision; writing – Original – draft; writing – Review – Editing.

ACKNOWLEDGEMENTS

The work was jointly supported by the National Key R&D Program of China (2022YFB2403300), the National Natural Science Foundation of China (72242105), and the Seed Fund for Basic Research for New Staff of The University of Hong Kong (202107185032).

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Ning Zhang https://orcid.org/0000-0003-0366-4657
Yi Wang https://orcid.org/0000-0003-1143-0666

REFERENCES

1. Petkov, I., Paolo, G.: Power-to-hydrogen as seasonal energy storage: an uncertainty analysis for optimal design of low-carbon multi-energy systems. Appl. Energy. 274, 115197 (2020)
2. Wang, Yi., Zhang, N., Zhuo, Z., Kang, C., Kirschen, D.: Mixed-integer linear programming-based optimal configuration planning for energy hub: Starting from scratch. Appl. Energy. 210, 1141–1150 (2018)
3. Gabrielli, P., Poluzzi, A., Kramer, G.J., Spiers, C., Mazzotti, M., Gazzani, M.: Seasonal energy storage for zero-emissions multi-energy systems via underground hydrogen storage. Renewable Sustainable Energy Rev. 121, 109629 (2020)
4. Murray, P., Orehounig, K., Grosspietsch, D., Carmeliet, J.: A comparison of storage systems in neighbourhood decentralized energy system applications from 2015 to 2050. Appl. Energy. 231, 1285–1306 (2018)
5. Jiang, H., Qi, B., Du, E., Zhang, N., Yang, X., Yang, F., Zhaoyuan, Wu.: Modeling hydrogen supply chain in renewable electric energy system planning. IEEE Trans. Ind. Appl. 58(2), 2780–2791 (2021)
6. Leng, R., Li, Z., Xu, Y.: A comprehensive literature review for optimal planning of distributed energy resources in distribution grids. 2022 IEEE PES Innovative Smart Grid Technologies-Asia (ISGT Asia), pp. 369–373 (2022)
7. Teichgraeber, H., Brandt, A.R.: Time-series aggregation for the optimization of energy systems: Goals, challenges, approaches, and opportunities. Renewable Sustainable Energy Rev. 157, 111984 (2022)
8. Schütz, T., Schraven, M.S., Fuchs, M., Remmen, P., Müller, D.: Comparison of clustering algorithms for the selection of typical demand days for energy system synthesis. Renewable Energy. 129, 570–582 (2018)
9. Hoffmann, M., Kotzur, L., Stolten, D., Robinius, M.: A review on time series aggregation methods for energy system models. Energies 13(3), 641 (2020)
10. Liu, Z., Yang, P., Peng, J., Lin, W., Ji, C.: Capacity allocation for regional integrated energy system considering typical day economic operation. In: 2018 IEEE International Conference on Energy Internet (ICEI), pp. 60–65. IEEE, (2018)
11. Li, Z., Xu, Y., Feng, X., Wu, Q.: Optimal stochastic deployment of heterogeneous energy storage in a residential multienergy microgrid with demand-side management. IEEE Trans. Ind. Inf. 17(2), 991–1004 (2020)
12. Li, Z., Xu, Y., Fang, S., Mazzoni, S.: Optimal placement of heterogeneous distributed generators in a grid-connected multi-energy microgrid under uncertainties. IET Renew. Power Gener. 13(14), 2623–2633 (2019)
13. Ma, T., Wu, J., Hao, L., Lee, W.-J., Yan, H., Li, D.: The optimal structure planning and energy management strategies of smart multi energy systems. Energy. 160, 122–141 (2018)
14. Kotzur, L., Markewitz, P., Robinius, M., Detlef S.,: Impact of different time series aggregation methods on optimal energy system design. Renewable Energy. 117, 474–487 (2018)
15. Mavromatidis, G., Petkov, I.: Mango: A. novel optimization model for the long-term, multi-stage planning of decentralized multi-energy systems. Appl. Energy. 288, 116585 (2021)
16. Baumgürtner, N., Bahl, B., Hennen, M., Bardow A.,: Rises3: Rigorous synthesis of energy supply and storage systems via time-series relaxation and aggregation. Comput. Chem. Eng. 127, 127–139 (2019)
17. Teichgraeber, H., Lindenmeyer, C.P., Baumgärtner, N., Kotzur, L., Stolten, D., Robinius, M., Bardow, A., Brandt, A.R.: Extreme events in time series aggregation: A. case study for optimal residential energy supply systems. Appl. Energy. 275, 115223 (2020)
18. Hilbers, A.P., Brayshaw, D.J., Gandy, A.: Reducing climate risk in energy system planning: a posteriori time series aggregation for models with storage. Appl. Energy. 334, 120624 (2023)
19. Kotzur, L., Markewitz, P., Robinius, M., Stolten, D.: Time series aggregation for energy system design: Modeling seasonal storage. Appl. Energy. 213, 123–135 (2018)

20. Baumgärtner, N., Shu, D., Björn B., Hennen, M., Hollermann, D.E., Bardow, A.: Deloop: Decomposition-based long-term operational optimization of energy systems with time-coupling constraints. Energy. 198, 117272 (2020)
21. Gabrielli, P., Gazzani, M., Martelli, E., Mazzotti, M.: Optimal design of multi-energy systems with seasonal storage. Appl. Energy. 219, 408–424 (2018)
22. Tan, J., Wu, Q., Zhang, X.: Optimal planning of integrated electricity and heat system considering seasonal and short-term thermal energy storage. IEEE Trans. Smart Grid (2022)
23. Novo, R., Marocco, P., Giorgi, G., Lanzini, A., Santarelli, M., Mattiazzo, G.: Planning the decarbonisation of energy sys-3 tems: The importance of applying time series clustering to long-term models. Energy Convers. Manage: X. 15, 100274 (2022)
24. Tejada-Arango, D.A., Domeshek, M., Wogrin, S., Centeno, E.: Enhanced representative days and system states modeling for energy storage investment analysis. IEEE Trans. Power Syst. 33(6), 6534–6544 (2018)
25. Bandara, K., Hyndman, R.J., Bergmeir, C.: Mstl: A. seasonaltrend decomposition algorithm for time series with multiple seasonal patterns. arXiv preprint arXiv:2107.13462, (2021)
26. Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I.: STL: A. seasonal-trend decomposition. J. Off. Stat 6(1), 3–73 (1990)
27. Wen, Q., Zhang, Z., Li, Y., Sun, L.: Fast RobustSTL: Efficient and robust seasonal-trend decomposition for time series with complex patterns. In: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD), p. 2203–2213. ACM, New York (2020)
28. Wen, Q., Gao, J., Song, X., Sun, L., Xu, H., Zhu, S.: RobustSTL: A. robust seasonal-trend decomposition algorithm for long time series. In: AAAI Conference on Artificial Intelligence (AAAI), pp. 5409–5416. AAAI Press, Palo Alto, CA (2019)
29. Dokumentov, A., Hyndman, R.J.: STR: Seasonal-trend decomposition using regression. INFORMS J. Data Sci. 1(1), 50–62 (2022)
30. Miller, C., Kathirgamanathan, A., Picchetti, B., Arjunan, P., Park, J.Y., Nagy, Z., Raftery, P., Hobson, B.W., Shi, Z., Meggers, F.: The building data genome project 2, energy meter data from the ashrae great energy predictor iii competition. Sci. Data. 7(1), 1–13 (2020)
31. Zhuo, Z., Zhang, N., Yang, J., Kang, C., Smith, C., O'Malley, M.J., Kroposki, B.: Transmission expansion planning test system for ac/dc hybrid grid with high variable renewable energy penetration. IEEE Trans. Power Syst. 35(4), 2597–2608 (2019)

How to cite this article: Ma, J., Zhang, N., Wen, Q., Wang, Y.: An efficient local multi-energy systems planning method with long-term storage. IET Renew. Power Gener. 18, 426–441 (2024). https://doi.org/10.1049/rpg2.12726