frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 05 January 2024
DOI 10.3389/fenrg.2023.1298365

OPEN ACCESS

EDITED BY
Chaojie Li,
University of New South Wales, Australia

REVIEWED BY
Yuanzheng Li,
Huazhong University of Science and Technology, China
Xiangyu Li,
University of New South Wales, Australia

*CORRESPONDENCE
Zhichun Yang,
yangzhichun36000@163.com

RECEIVED 21 September 2023
ACCEPTED 11 December 2023
PUBLISHED 05 January 2024

CITATION
Yang Z, Min H, Yang F, Liu Y, Shen Y, Zhou B, Zhou Z, Hu W and Lei Y (2024), Planning model for rural distribution networks with integrated energy stations via robust optimization method. *Front. Energy Res.* 11:1298365. doi: 10.3389/fenrg.2023.1298365

Planning model for rural distribution networks with integrated energy stations via robust optimization method

Zhichun Yang[1]*, Huaidong Min[1], Fan Yang[1], Yu Liu[1], Yu Shen[1], Bin Zhou[2], Zhiqiang Zhou[2], Wei Hu[1] and Yang Lei[1]

[1]Electric Power Research Institute of State Grid Hubei Co., Ltd., Wuhan, China, [2]State Grid Hubei Electric Power Co., Ltd., Wuhan, China

With the modernization and intelligent development of agriculture, the energy demand in rural areas continues to increase, which leads to an increased operational burden on the existing rural distribution network. Integrated energy stations (IESs), in rural areas, use renewable energy sources such as biogas, wind power, and photovoltaic as energy inputs, which can fully improve energy efficiency and help reduce the operating load and peak valley difference of rural distribution networks. In this paper, a multistage planning model is proposed for rural distribution networks with IESs based on the robust optimization method. Firstly, a rural distribution network operation framework with IESs is proposed, and a mathematical model of rural IESs is built based on the energy hub (EH). Then, the multistage robust planning model of rural distribution net-works with IESs is developed and typical scenarios of stochastic source and load are generated based on improved k-means. An iterative solution method for a two-stage robust optimization method is proposed based on the nested column constraint generation (NC&CG) algorithm. Finally, the effectiveness of the presented model and solution approach is assessed through case studies on a modified IEEE 33-node distribution network and a real 152-node distribution network.

KEYWORDS

distribution network planning, flexible control, integrated energy station, multistage robust planning, NCCG algorithm

1 Introduction

With the continuous construction of modern agriculture in rural areas, the demand for electricity among residents has increased rapidly (Han, Zhang and Li, 2022). The air conditioning load is one of the primary causes of the constant rise in power consumption. Additionally, because of the air-conditioning load's focused energy demand, the power grid's peak and valley differences are becoming more significant (International Energy Agency, 2015). In this context, the development and construction of the distribution network framework are facing new challenges. Therefore, it is crucial to help rural distribution network planning and construction through the multi-energy complementary advantages of integrated energy stations.

In response to the increasing load growth, conventional distribution network planning usually determines what type of lines to build, the capacity and location of transformer

substation (TS) based on the peak value of load forecast. As proposed in reference (Heidari and Fotuhi-Firuzabad, 2019), a distribution network planning model suitable for multiple power sources is proposed. The objective of minimizing the investment cost and operation cost is to determine the conventional lines of the distribution network. Ref (Dong et al., 2022). considers the correlation and timing of the outputs of distributed power sources in the distribution network, and conducts collaborative planning between the distribution network structure and distributed power sources, taking into account load growth. Ref (Li, Wang and Xia, 2017). proposes an active distribution network collaborative planning method considering energy storage devices and distributed power sources. Although the implementation of these planning approaches has the potential to enhance the operational adaptability of the distribution network, their practical application is hindered by the issue of substantial investment expenses. Additionally, it is worth noting that the majority of the aforementioned research focuses on urban distribution network planning, leaving the exploration of rural distribution networks in its nascent stage.

The concept of IES has been pioneered in Europe and the United States in order to address the challenges posed by the energy transition. By coordinating and integrating various energy systems, such as cooling, electricity, heat, and gas, through collaborative planning and synergistic operations, the aim is to overcome the current limitations of separate operations and enhance the overall efficiency and utilization of energy systems (Wu et al., 2016). Ref (Wu et al., 2017). utilizes the time-delay characteristics of heat networks and the thermal inertia of buildings in different regions to promote the optimal operation of multi-region electricity-heat. A multistage planning approach for IES for converged distribution networks is proposed in (Li et al., 2022a) to minimize the total costs. Ref (Ge, Liu and Ge, 2018). considered the participation of electricity, heat, and cold storage in interaction in IES planning, and compared the differences in planning capacity among the three under different objectives such as economy and energy conversion efficiency. In contrast to power storage devices, cold and heat storage devices have achieved a higher level of development and are more cost-effective, rendering them more viable for the implementation of energy storage in rural areas. However, most of the above references focuses on IES as the main body and distribution networks and other energy networks as the component elements to achieve the overall system planning. Less attention has been paid to the application value of IES as an energy coupling node for distribution network planning.

IES also faces uncertainty in the operation of multi-energy loads and renewable energy output. The different methods of handling random variable prediction errors, it is mainly divided into scenario method (Li et al., 2022b), chance-constrained programming (Huo et al., 2021), robust optimization (Gan et al., 2021), etc. A scenario-generation based approach for integrated regional energy expansion planning was proposed in (Lei et al., 2020). The method considers the temporal, autocorrelation and inter-correlation properties of renewable energy sources and multi-energy loads. Based on the spatial and temporal correlation of PV output, ref (Koutsoukis and Georgilakis, 2019). established a multi-objective distributed power grid-connected planning model by utilizing opportunity constraint planning, taking into account the total cost of PV and the cost of grid operation. A robust return on investment oriented optimal allocation method is proposed in (Jiang et al., 2020) to address the uncertainty of wind power output and multi-energy load. In which, the planning method based on robust optimization fully considers the possible worst scenarios of random variables in the planning stage, which can ensure the robustness of the planning scheme and has a wide range of application scenarios. The typical scenarios in existing scenario methods and the basic scenarios in robust planning methods are usually obtained by clustering algorithms such as k-means. However, the clustering number of existing k-means is artificially given and has a certain degree of subjectivity.

In summary, existing research on urban distribution network planning is relatively rich, but rural distribution network planning is still in its early stages. There exist notable disparities in terms of source and load characteristics between rural and urban regions, and existing plans for urban distribution networks do not apply to rural areas. In addition, it is also necessary to conduct research on distribution network planning with the distribution network as the main body and IES as the means of distribution network regulation. Therefore, the main contributions of this article are as shown below.

1) A multistage planning model for rural distribution network with integrated energy stations is proposed. The primary focus of the model is the distribution network, with the IES serving as a regulatory entity to address the increasing demand in rural areas and mitigate fluctuations in load levels. The model implements a multistage rolling plan that encompasses various aspects such as distribution line management, TS expansion, and investment and construction of the IES.
2) Utilizing an enhanced k-means clustering algorithm to derive multiple representative daily load patterns from projected annual load data. The method aims to improve clustering efficiency by calculating clustering performance evaluation indicators to ascertain the most suitable number of clusters, thereby solving the problem of traditional k-means algorithms being unable to determine the optimal number of clusters.

The subsequent organization of this article is outlined as follows. Section 2 provides an introduction to the operational framework of rural distribution networks incorporating integrated energy stations. In section 3, a multistage robust planning model for rural distribution networks with integrated energy stations is proposed. Section 4 elaborates on the solution methods for robust programming models. Section 5 presents the simulation results. Section 6 summarizes the entire text.

2 Operational framework for rural distribution networks with IES

2.1 Structure of the rural distribution network with IES

In this paper, we consider the current situation that in the development of biogas in rural, biogas projects with a variety of

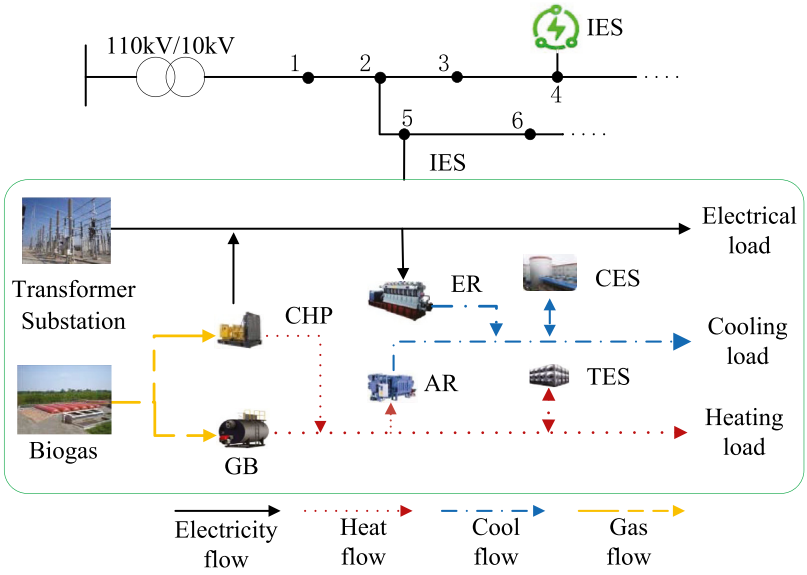

FIGURE 1
Energy flow structure of rural distribution network with integrated energy stations.

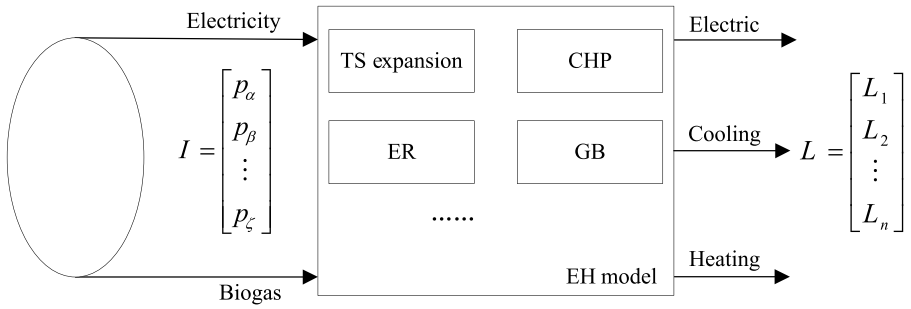

FIGURE 2
Typical EH-based RIES.

feedstock substrates are gradually formed, and the gas production capacity increases year after year (Yan et al., 2021). Therefore, biogas is used as the gas input for IES in the IES configuration. In addition, the main focus of this article is on the planning method of rural distribution networks. Therefore, non-distribution network structures such as gas networks have been weakened and set to have gas source points at the construction site of IES to meet the gas demand of IES.

The structure of the rural distribution network with IES is depicted in Figure 1. The voltage level of the rural distribution network shown in Figure 1 is 10kV, with IES built on some network nodes. The Integrated Energy System comprises a diverse range of equipment, encompassing combined heat and power (CHP), gas boilers (GB), electric refrigeration (ER), absorption refrigeration (AR), cold energy storage (CES), and thermal energy storage (TES). The specific direction of energy flow is shown in Figure 1, where the CHP consumes biogas to generate electricity and purchases electricity from the distribution network to transmit to the load and ER. The thermal load and AR are sup-plied by TES heat release, CHP, and GB consumption of biogas, while the excess thermal energy is stored in TES. The cooling load is jointly supplied by CES cooling, ER, and AR cooling, and the excess cold energy is stored in CES.

2.2 Modeling of rural integrated energy stations based on EH

After the IES is connected to the rural distribution system, there are multiple energy coupling relationships within the system. Considering the energy supply perspective, the rural integrated energy station (RIES) generally includes energy supply means such as TS, GB, CHP, and ER. The EH-based model, as depicted in Figure 2, is commonly employed to describe the multi-energy attributes of RIES (Moazeni, Miragha and Defourny, 2018).

EH takes into account various forms of energy input and output, and represents its energy conversion relationship through a set of linear equations.

$$L = \begin{bmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{bmatrix} = \begin{bmatrix} \mathfrak{R}_{\alpha,1} & \mathfrak{R}_{\beta,1} & \cdots & \mathfrak{R}_{\zeta,1} \\ \mathfrak{R}_{\alpha,2} & \mathfrak{R}_{\beta,2} & \cdots & \mathfrak{R}_{\zeta,2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathfrak{R}_{\alpha,n} & \mathfrak{R}_{\beta,n} & \cdots & \mathfrak{R}_{\zeta,n} \end{bmatrix} \begin{bmatrix} p_\alpha \\ p_\beta \\ \vdots \\ p_\zeta \end{bmatrix} = UI \tag{1}$$

where $L = [L_1, L_2, \cdots, L_n]^T$ represents the output part of EH, and subscript $1, 2, \cdots, n$ represents different forms of energy demand on the load side, such as electrical load, heating load, cooling load, etc.; $I = [p_\alpha, p_\beta, \cdots, p_\zeta]^T$ represents the input part of EH, where subscript $\alpha, \beta, \cdots, \zeta$ represents different types of energy input forms, such as electricity, heat, biogas, etc.; U is the energy conversion coefficient matrix, where $\mathfrak{R}_{\alpha,1}, \mathfrak{R}_{\beta,1}, \cdots, \mathfrak{R}_{\zeta,1}$ is the conversion coefficient.

In general, EH is used to summarize the overall "energy" input and output of RIES. However, because since most devices in RIES have input and output characteristics. Therefore, this article will use the EH model to model each type of equipment separately, to construct a planning model for RIES. Unless otherwise specified in the following text, the subscripts p, s, and t respectively represent the corresponding parameters for the current planning stage, scenario, and period time; i and j represent nodes i and j respectively, and their corresponding sets are labeled in the following formulas.

The constraints of the RIES internal model based on EH for modeling various multi-functional devices are as follows.

1) Constraints for Multi-energy Load Balancing

Assuming that the type set of energy conversion equipment contained in RIES is m and the set of energy storage equipment is denoted as n, then the multi-energy load balance is

$$L_{i,t,p,s} = \sum_{m\in\Omega} U^m I^m_{i,t,p,s} + \sum_{n\in\Psi} W^n_{i,t,p,s} \tag{2}$$

where $L_{i,t,p,s}$ represents the overall multi-energy load demand of RIES at stage p, time t, and node i in operating scenario s; U^m indicates the conversion coefficient matrix of the mth energy conversion device; $I^m_{i,t,p,s}$ represents the energy input matrix of the mth energy conversion device; $W^n_{i,t,p,s}$ represents the energy storage power of the nth type of energy storage device.

2) Constraints for Energy Conversion Equipment

The energy conversion equipment considered in this article includes CHP, GB, ER, and AR, which are represented by the following general models.

$$\widehat{P}^m_{i,t,p,s} = \sum_{m\in\Omega} U^m I^m_{i,t,p,s} \tag{3}$$

where $\widehat{P}^m_{i,t,p,s}$ is the output matrix of the mth energy conversion device.

Taking CHP as an example, its energy conversion coefficient matrix is

$$U^{CHP} = \begin{bmatrix} \mathfrak{R}^{CHP}_{G-E} & \mathfrak{R}^{CHP}_{E-E} & \mathfrak{R}^{CHP}_{H-E} \\ \mathfrak{R}^{CHP}_{G-C} & \mathfrak{R}^{CHP}_{E-C} & \mathfrak{R}^{CHP}_{H-C} \\ \mathfrak{R}^{CHP}_{G-H} & \mathfrak{R}^{CHP}_{E-H} & \mathfrak{R}^{CHP}_{H-H} \end{bmatrix} \tag{4}$$

The energy input matrix is

$$I^{CHP}_{i,t,p,s} = \begin{bmatrix} P^{CHP,G}_{i,t,p,s} \\ P^{CHP,E}_{i,t,p,s} \\ P^{CHP,H}_{i,t,p,s} \end{bmatrix} \tag{5}$$

The energy output matrix is

$$\widehat{P}^{CHP}_{i,t,p,s} = \begin{bmatrix} \widehat{P}^{CHP,E}_{i,t,p,s} \\ \widehat{P}^{CHP,C}_{i,t,p,s} \\ \widehat{P}^{CHP,H}_{i,t,p,s} \end{bmatrix} \tag{6}$$

where $G-E, G-C$ and $G-H$ respectively represent the conversion coefficients of gas-to-electricity, gas-to-cooling, and gas-to-heat, and the meanings of the other parameters are analogous; CHP consumes biogas to produce heat and electricity, then only $\mathfrak{R}^{CHP}_{G-E}$ and $\mathfrak{R}^{CHP}_{G-H}$ are non-zero elements in the matrix, and all the other elements are zero; $P^{CHP,G}_{i,t,p,s}$, $P^{CHP,E}_{i,t,p,s}$ and $P^{CHP,H}_{i,t,p,s}$ respectively indicate the biogas, electricity, and heat inputs to the CHP, and only $P^{CHP,G}_{i,t,p,s}$ is a non-zero element, and all the other elements are zero; $\widehat{P}^{CHP,G}_{i,t,p,s}$, $\widehat{P}^{CHP,E}_{i,t,p,s}$ and $\widehat{P}^{CHP,H}_{i,t,p,s}$ respectively represent the biogas, electricity and heat outputs of the CHP, and only $\widehat{P}^{CHP,E}_{i,t,p,s}$ and $\widehat{P}^{CHP,H}_{i,t,p,s}$ are non-zero elements, and all the other elements are zero.

Similarly, similar models can be established for GB, ER, and AR. Other related constraints of energy conversion equipment can be expressed as

$$0 \le U^m I^m_{i,t,p,s} \le Z^{RIES}_{i,t,p,s} P^m_{rated} \tag{7}$$

where $Z^{RIES}_{i,t,p,s}$ denotes a binary variable that represents the construction status of RIES. When set to 1, it indicates construction at node i, and when set to 0, it indicates no construction at node i; P^m_{rated} represents the rated operating power of the mth type of energy transformation equipment.

3) Constraints for Energy Storage Device

The energy storage devices considered are CES and TES in this paper. In the case of CES, its energy storage properties can be described in the following manner.

$$\begin{gathered} W^n_{i,t,p,s} = 0, t = t_0 \\ W^n_{i,t,p,s} = \mu^n W^n_{i,t-1,p,s} + P^{n,ch}_{i,t,p,s} u^n_{ch} - P^{n,dis}_{i,t,p,s} \Big/ u^n_{dis} \quad i \in \Omega_{RIES} \\ 0 \le W^n_{i,t,p,s} \le Z^{RIES}_{i,t,p,s} W^n_{rated} \end{gathered} \tag{8}$$

$$\begin{gathered} 0 \le P^{n,ch}_{i,t,p,s} \le Z^{RIES}_{i,t,p,s} P^n_{rated} X^n \\ 0 \le P^{n,dis}_{i,t,p,s} \le Z^{RIES}_{i,t,p,s} P^n_{rated} Y^n \quad i \in \Omega_{RIES} \\ X^n + Y^n \le 1 \end{gathered} \tag{9}$$

where t_0 denotes the initial moment of the energy storage charging and discharging cycle, and the cycle set in this paper is 24h; μ^n is the energy dissipation coefficient; $P^{n,ch}_{i,t,p,s}$ and $P^{n,dis}_{i,t,p,s}$ mean the storing and discharging power, respectively; u^n_{ch} and u^n_{dis} respectively indicate the storing efficiency and discharging efficiency; P^n_{rated} and W^n_{rated} respectively indicate the rated operating power and installed capacity; The variables X^n and Y^n are binary variables that are

employed to impose constraints on the upper and lower boundary of the power for charging and discharging. Additionally, they serve to restrict the occurrence of simultaneous charging and discharging operations; Ω_{RIES} is the set of installation candidate nodes for RIES; The bivariate nonlinear term in Eq. 9 can be treated using the big-M method in reference (Kou and Li, 2018).

3 Multistage robust planning model for rural distribution network with IES

The model proposed in this article takes the rural medium and low voltage distribution network as the planning object, to respond to load growth and alleviate load peak valley differences. The planning content considers the upgrading and renovation of grid lines, the increase of TS main transformer capacity, and the construction of RIES. The overall planning is divided into multiple stages and rolling.

3.1 Objective function

To achieve the above goals and fully consider the actual operation situation, this paper proposes a two-layer multistage planning model for distribution network planning and operation. The planning strategy of the distribution network is determined by the upper-level problem, as depicted in Eq. 10.

$$\begin{cases} \min f\left(Z_p^{inv}\right) = \gamma\left(F^{INV} + F^{OPE}\right) \\ s.t. G\left(Z_p^{inv}\right) \le 0 \\ \gamma = (1+\kappa)^{-Y} \end{cases} \tag{10}$$

The lower-level model solves the optimal operational mode of the distribution network based on the planning strategy of the upper-level model, as depicted in formula (11).

$$\begin{cases} \min f\left(Z_p^{inv}, X_{p,s}^{ope}\right) = F_p^{OPE} \\ s.t. \begin{cases} g_1\left(Z_p^{inv}, X_{p,s}^{ope}\right) \le 0 \\ g_2\left(Z_p^{inv}, X_{p,s}^{ope}\right) \le 0 \end{cases} \end{cases} \tag{11}$$

where F^{INV} denotes the investment planning cost; Z_p^{inv} represents the planning decision variable for stage p; $X_{p,s}^{ope}$ is the operating variable under scenario s of stage p; F_p^{OPE} represents the operating cost of stage p; γ denotes the present value coefficient; ω indicate the Bank rate, which is 5% in this paper; Y denote the total number of years of investment period; G and g_1、g_2 are constraints for the upper and lower models, respectively.

1) Investment Planning Cost

The planned investment cost includes the investment cost of line expansion, TS investment cost, and RIES construction investment cost, as shown in Eqs. 12–16.

$$F^{INV} = \sum_{p\in D}\left(F_p^L + F_p^{sub} + F_p^{RIES}\right) \tag{12}$$

$$F_p^L = \sum_{l\in\Omega_L} c_l^L z_{l,p}^L h_l \varpi^L \tag{13}$$

$$F_p^{sub} = \sum_{i\in\Omega_{SUB}} c_i^{sub} z_{i,p}^{sub} \varpi^{sub} \tag{14}$$

$$\begin{aligned} F_p^{RIES} = \sum_{i\in\Omega_{RIES}} z_{i,d}^{RIES}\varpi^{RIES}\left(c_i^{ER}P_{rated}^{ER} + c_i^{AR}P_{rated}^{AR} + c_i^{CHP}P_{rated}^{CHP}\right. \\ + c_i^{GB}P_{rated}^{GB} + c_i^{CES}P_{rated}^{CES} + c_i^{TES}P_{rated}^{TES} \\ \left. + c_i^{CES}W_{rated}^{CES} + c_i^{TES}W_{rated}^{TES}\right) \end{aligned} \tag{15}$$

$$\varpi^{\ell} = \frac{\kappa(1+\kappa)^{L^{\ell}}}{(1+\kappa)^{L^{\ell}} - 1} \quad \ell = L, sub \tag{16}$$

where ϖ^{ℓ} represents the investment recovery coefficient of equipment ℓ; L^{ℓ} denotes the life cycle of equipment ℓ; P_{rated}^{ER}, P_{rated}^{AR}, P_{rated}^{CHP}, P_{rated}^{GB}, P_{rated}^{CES} and P_{rated}^{TES} respectively represent the rated operating power of ER, AR, CHP, GB, CES, TES equipment; W_{rated}^{CES} and W_{rated}^{TES} respectively represent the Nameplate capacity of CES and TES equipment; c_l^L and c_i^{sub} respectively represent the unit construction cost of the line and TS; c_i^{ER}, c_i^{AR}, c_i^{CHP}, c_i^{GB}, c_i^{CES}, c_i^{TES} respectively represent the unit power construction costs of ER, AR, CHP, GB, CES, TES equipment; c_i^{CES} and c_i^{TES} respectively represent the unit capacity construction cost of CES and TES equipment; $z_{l,p}^L$, $z_{i,p}^{sub}$ and $z_{i,p}^{RIES}$ are binary variables, indicating the construction status of the line, TS, and RIES; h_l represents the length of line l; Ω_L is the set of candidate lines; Ω_{SUB} represent the set of installation candidate nodes for TS.

2) Operating Cost

The operating cost includes the benefits of RIES flexible regulation, the expenses associated with purchasing energy, and the operating cost of RIES, as indicated in equations Eqs. 17–20.

$$F^{OPE} = \sum_{p\in D}\sum_{s\in S}\sum_{t\in T}\left(F_{p,s}^E + F_{t,p,s}^{pur} + f_{t,p,s}^{RIES}\right) \tag{17}$$

$$F_{p,s}^E = \frac{1}{24}\tau^p \sum_{t\in T}\left[\sum_{i\in\Omega_{EL}} P_{i,t,p,t}^E + \sum_{i\in\Omega_{RIES}}\left(P_{i,t,p,s}^{EC,E} - \widehat{P}_{i,t,p,s}^{CHP,E}\right) - P^{av}\right]^2 \tag{18}$$

$$F_{t,p,s}^{pur} = \sum_{i\in\Omega_{SUB}} P_{i,t,p,s}^{sub}\rho_{i,t,p,s}^e + \sum_{i\in\Omega_{RIES}} P_{i,t,p,s}^{bio}\rho_{i,t,p,s}^g \tag{19}$$

$$\begin{aligned} f_{t,p,s}^{RIES} = \sum_{i\in\Omega_{RIES}}\left[\tau_{i,t,p,s}^c\left(1-\mu^c\right)W_{i,t,p,s}^{CES}\Big/u_{ch}^{CES} + \right. \\ \left. \tau_{i,t,p,s}^h\left(1-\mu^h\right)W_{i,t,p,s}^{TES}\Big/u_{ch}^{TES}\right] \end{aligned} \tag{20}$$

where $P_{i,t,p,s}^E$ is the active load; $\widehat{P}_{i,t,p,s}^{CHP,E}$ denotes the active output of CHP; $P_{i,t,p,s}^{EC,E}$ represents the power consumption of ER; P^{av} is the daily average electrical load; τ^p indicates the penalty factor for peak valley load difference; $P_{i,t,p,s}^{sub}$ and $P_{i,t,p,s}^{bio}$ respectively denote the electricity and biogas purchased; $\rho_{i,t,p,s}^e$ and $\rho_{i,t,p,s}^g$ indicate real-time electricity prices and biogas prices, respectively; $\tau_{i,t,p,s}^c$ and $\tau_{i,t,p,s}^h$ mean the penalty costs per unit loss of cold energy and thermal energy, respectively; μ^c and μ^h are the energy loss coefficients of CES and TES, respectively; $W_{i,t,p,s}^{CES}$ and $W_{i,t,p,s}^{TES}$ respectively represent the energy storage of CES and TES; u_{ch}^{CES} and u_{ch}^{TES} respectively represent the charging efficiency of CES and TES; Ω_{EL} is the set of electrical load nodes.

3.2 Constraints

1) Planning and Construction Constraints

$$z_{l,p}^L \le z_{l,p+1}^L \; l \in \Omega_L \tag{21}$$

$$0 \le z_{i,p}^{sub} i \in \Omega_{SUB} \tag{22}$$

$$z_{i,p}^{RIES} \le z_{i,p+1}^{RIES} i \in \Omega_{RIES} \tag{23}$$

In the above formula, RIES and routes can only be made based on the previous stage planning, and TS depends on the needs of each stage.

2) Power Flow Constraints in Distribution Networks

$$\sum_{i\in\Omega_{\mathrm{EL}}} P_{i,t,p,s}^{\mathrm{E}} = \sum_{i\in\Omega_{SUB}} P_{i,t,p,s}^{sub} + \sum_{i\in\Omega_{\mathrm{RIES}}} \left(\widehat{P}_{i,t,p,s}^{\mathrm{CHP,E}} - P_{i,t,p,s}^{\mathrm{ER,E}}\right) - \sum_{l\in Ls(i)} P_{l,t,p,s} + \sum_{l\in Le(i)} \left(P_{l,t,p,s} - r_l I_{l,t,p,s}^2\right) \tag{24}$$

$$\sum_{i\in\Omega_{\mathrm{EL}}} Q_{i,t,p,s}^{\mathrm{E}} = \sum_{i\in\Omega_{SUB}} Q_{i,t,p,s}^{sub} + \sum_{i\in\Omega_{\mathrm{RIES}}} \widehat{Q}_{i,t,p,s}^{\mathrm{CHP,E}} - \sum_{l\in Ls(i)} Q_{l,t,p,s} + \sum_{l\in Le(i)} \left(Q_{l,t,p,s} - x_l I_{l,t,p,s}^2\right) \tag{25}$$

$$U_{e(l),t,p,s}^2 = U_{s(l),t,p,s}^2 - 2\left(r_l P_{l,t,p,s} + x_l Q_{l,t,p,s}\right) + I_{l,t,p,s}^2\left(r_l^2 + x_l^2\right) \quad l \in \Omega_L \tag{26}$$

$$U_{e(l'),t,p,s}^2 + M\left(\alpha_{l',t,p,s} - 1\right) \leqslant U_{s(l'),t,p,s}^2 - 2\left(r_{l'} P_{l',t,p,s} + x_{l'} Q_{l',t,p,s}\right) + I_{l',t,p,s}^2\left(r_{l'}^2 + x_{l'}^2\right) \leqslant U_{e(l'),t,p,s}^2 + M\left(1 - \alpha_{l',t,p,s}\right) l' \in \Omega_{sw} \tag{27}$$

where r_l and x_l mean the resistance and reactance of line l, respectively; $P_{l,t,p,s}$ and $Q_{l,t,p,s}$ respectively represent the active and reactive power of line l; $Q_{i,t,p,s}^{sub}$ denotes the reactive power of transformer station; $Q_{i,t,p,s}^{\mathrm{E}}$ is the reactive load; $\widehat{Q}_{i,t,p,s}^{\mathrm{CHP,E}}$ represents the reactive output of the CHP; $Ls(i)$ and $Le(i)$ respectively represent the set of lines starting and ending with node i; $s(l)$ and $e(l)$ respectively indicate the start and end nodes of line l; $I_{l,t,p,s}$ indicates the current flowing through line l; $U_{i,t,p,s}$ means the voltage of node i; $\alpha_{l',t,p,s}$ represents a binary variable that indicates the switching state of the switching branch l'. When 0 or 1 is taken, it indicates that the branch is disconnected or connected; Ω_{sw} is the set of switch branches; M is a large number.

Since the inclusion of nonlinear terms $U_{i,t,p,s}^2$ and $I_{l,t,p,s}^2$ in Eqs. 24–27, the solver cannot directly solve. Firstly, phase angle relaxation is used to ignore the phase angles of voltage and current, making them $U_{i,t,p,s}^{sq}$ and $I_{l,t,p,s}^{sq}$, respectively.

$$I_{l,t,p,s}^{sq} = \frac{P_{l,t,p,s}^2 + Q_{l,t,p,s}^2}{U_{s(l),t,p,s}^{sq}} l \in \Omega_L \cup \Omega_{sw} \tag{28}$$

Then the second-order cone relaxation is used to convert the above nonconvex quadratic equation constraints to second-order cone constraints.

$$\left\| \begin{matrix} 2P_{l,t,p,s} \\ 2Q_{l,t,p,s} \\ I_{l,t,p,s}^{sq} - U_{s(l),t,p,s}^{sq} \end{matrix} \right\|_2 \le I_{l,t,p,s}^{sq} + U_{s(l),t,p,s}^{sq} l \in \Omega_L \cup \Omega_{sw} \tag{29}$$

3) Network Operation Security Constraints

The node voltage constraint is

$$\underline{U}_i \le U_{i,t,p,s} \le \bar{U}_i i \in \Omega_{EL} \tag{30}$$

The non-switching branch current constraint is

$$\left|I_{l,t,p,s}\right| \le \left(1 - z_{l,t,p,s}^L\right)\bar{I}_l + z_{l,t,p,s}^L \bar{I}_l' l \in \Omega_L \tag{31}$$

The switch branch current constraint is

$$\left|I_{l,t,p,s}\right| \le \alpha_{l',t,p,s} \bar{I}_l l' \in \Omega_{sw} \tag{32}$$

where $\bar{U}_i$ and $\underline{U}_i$ respectively denote the maximum and minimum bounds of the node i voltage; $\bar{I}_l$ indicates the upper limit of the original distribution grid branch current; $\bar{I}_l'$ represents the maximum value of the current flowing through the upgraded branch.

4) Network Reconstruction Constraints

$$\sum_{l'\in\Omega_{sw}} \alpha_{l',t,p,s} = N^{nod} - N^{sub} - N^L \tag{33}$$

where N^{nod} indicates the total number of nodes present within the network; N^{sub} denotes the aggregate quantity of TS; N^L denotes the total number of branches without contact switches in the network; Due to the presence of CHP in RIES, even under the constraints of Eq. 31, the distribution network may experience isolated operation (Li et al., 2023). Therefore, assuming that a node that is not TS has a very small power $P_{i,t,p,s}^*$ with a value of $\mathfrak{J}$, the connectivity between the non-TS node and the TS node is ensured through the following power flow constraints Eqs 34–36.

$$\sum_{l\in Ls(i)} P_{l,t,p,s}^* - \sum_{l\in Le(i)} P_{l,t,p,s}^* = P_{i,t,p,s}^* = \mathfrak{J} i \in \Omega_{\mathrm{L}} \cup \Omega_{\mathrm{sw}} \tag{34}$$

$$-\alpha_{l',t,p,s}\bar{P}_{l,t,p,s} \leqslant P_{l,t,p,s}^* \leqslant \alpha_{l',t,p,s}\bar{P}_{l,t,p,s} l' \in \Omega_{\mathrm{sw}}, l \in \Omega_L \tag{35}$$

$$-\bar{P}_{l,t,p,s} \leqslant P_{l,t,p,s}^* \leqslant \bar{P}_{l,t,p,s} l \in \Omega_{\mathrm{L}} \tag{36}$$

where $\bar{P}_{l,t,p,s}$ indicates the maximum value of the active power of line l; $P_{l,t,p,s}^*$ denotes the auxiliary power flow active power of line l, rather than the actual transmitted active power.

5) RIES Load Balance Constraints and Operational Constraints

For detailed constraints, please refer to Eqs. 2–9 above.

6) TS Operational Constraints

$$\left(P_{i,t,p,s}^{sub}\right)^2 + \left(Q_{i,t,p,s}^{sub}\right)^2 \le \left(z_{i,t,p,s}^{sub} S_i^{sub}\right)^2 i \in \Omega_{SUB} \tag{37}$$

where S_i^{sub} indicates the Nameplate capacity of TS. Since Eq. 37 is still nonlinear, the second-order cone relaxation method is also used to process it, and it will not be repeated here.

Therefore, the multistage planning model for rural distribution networks with IES can be expressed as

$$\begin{aligned} &\min F^{INV}(Z) + F^{OPE}(X) \\ &s.t.\ \ (2)-(9),\ (12)-(37) \end{aligned} \tag{38}$$

3.3 Multistage robust planning of rural distribution network with IES

Considering the uncertainty of multi-energy loads in operational scenarios, this paper proposes to establish a robust

programming model for the original problem model Eq. 38. By incorporating adjustable robust measures to characterize the variations of uncertain parameters, decision-makers in planning can modify these measures according to their risk preferences. A higher value indicates a more cautious planning approach, while a lower value signifies a riskier strategy. The uncertain parameter set is denoted as Φ, and its specific form is outlined below.

$$\begin{cases} \Phi = \left[\hat{L}_{s,t} - \Delta L_{s,t}{}^{\max}, \hat{L}_{s,t} + \Delta L_{s,t}{}^{\max}\right] \\ L_{s,t} = \hat{L}_{s,t} + B_{s,t}\Delta L_{s,t}{}^{\max} \\ \sum_{s\in S}\sum_{t\in T} B_{s,t} \le \Gamma \end{cases} \tag{39}$$

where $\hat{L}_{s,t}$ indicates the representative daily load for multiple energy loads; $\Delta L_{s,t}{}^{\max}$ is the maximum fluctuation deviation of multi-energy load; $B_{s,t}$ represents a binary variable that takes on a value of 1 when the multi-energy load of the corresponding period is brought to the boundary of the interval; Γ is an uncertainty adjustment parameter, with values ranging from 0 to T as integers.

Correspondingly, the multistage robust planning model for rural distribution networks with IES considering the uncertainty of multi-energy loads is

$$\begin{aligned} &\min_{Z_p} F^{INV} + \max_{L_p\in\Phi} \min_{X_p\in\Omega(Z_p,L_p)} F^{OPE} \\ &s.t.\ \ (2)-(9),(12)-(37) \end{aligned} \tag{40}$$

where the minimization on the left-hand side represents the first-stage issue, with the optimization variables denoted as Z_p. On the other hand, the maximum-minimization on the right-hand side corresponds to the second-stage issue, with the optimization variables L and X_p.

3.4 Scenarios generation based on improved k-means

It aims at the characteristics of large amounts of data and high similarity between data in the joint scenario of cooling/heating/electrical loads. This section proposes an improved k-means clustering algorithm to improve clustering efficiency. By calculating clustering performance evaluation indicators, the most suitable number of clusters is determined, thereby solving the limitation of traditional k-means algorithms in determining the most suitable number of clusters.

Assuming the total number of samples in the dataset is m, the most suitable number of clusters is usually defined as an integer within the range $[2, \sqrt{m}]$ (Kim and Ramakrishna, 2005). The optimal number of clusters is determined by the Calinski-Harabasz (CH) (Feng et al., 2023) index, and the definition formula is as follows.

$$CH(k) = \frac{tr(\mathbb{R}_k) \times (m-k)}{tr(\mathbb{Z}_k) \times (k-1)} \tag{41}$$

where tr represents the trace of the matrix; k denote the number of categories; $\mathbb{R}_k$ denote the Covariance matrix between different categories; $\mathbb{Z}_k$ indicates the Covariance matrix of internal data of the same category. As the k value increases, $tr(\mathbb{Z}_k)$ decreases and $tr(\mathbb{R}_k)$ increases. If there is a k value that maximizes the value of CH (k), this value is the most suitable number of clusters.

Assuming a group sample group $N = \{N_1, N_2, \ldots, N_m\}$, in the case of high sample density, the higher density area is usually selected as the initial clustering center. The density coefficient d of the sample group can be expressed as:

$$d(N_m, d_{ave}) = \sum_{m_1=1}^{m} \vartheta\left(d_{ave} - d_{\mathrm{E}}(N_m, N_{m_1})\right) \tag{42}$$

where d_{ave} denotes the average distance of the sample; ϑ represents a Sign function; d_E is a Euclidean distance function. The density coefficient d refers to the number of samples with an average distance d_{ave} in an area centered on N_m. When the density coefficients of multiple samples are equal, select the sample with higher cohesion as the clustering center.

Therefore, the improved k-means algorithm flow is explained as follows:

Step 1: Input the sample dataset based on historical data.

Step 2: Generate the optimal initial clustering center based on the sample dataset.

Step 3: Set k = 2 and call the traditional k-means algorithm subroutine to cluster the dataset.

Step 4: Determine if k is less than int. If so, make k = k+1 and skip back to step 3; If not, then call the traditional k-means algorithm subroutine again to cluster the dataset.

Step 5: Calculate the CH (k) based on the clustering results to obtain typical scenarios of cooling/heating/electrical loads.

4 Planning model solution

This paper uses the NCCG algorithm (Zeng and Zhao, 2013; Yan et al., 2019) to solve the robust optimization model Eq. 40 established earlier. The algorithm includes an inner loop and an outer loop. Considering that this article is multistage planning, we will use stage p as an example when introducing the NCCG algorithm. To facilitate the description of the solution process, first, write Eq. 40 in the following compact form.

$$\begin{aligned} &\min_{Z_p} c^T \cdot Z_p + \max_{L_p\in\Phi} \min_{X_p\in\Omega(Z_p,L_p)} b^T \cdot X_p \\ &s.t. \qquad \varphi \cdot Z_p \le h \\ &\qquad \kappa \cdot Z_p + \sigma \cdot U_p + \upsilon \cdot X_p \le e \\ &\qquad \left\| J \cdot X_p \right\|_2 \le H \cdot X_p + w \end{aligned} \tag{43}$$

where Z_p represents the upper-level model's decision variable in stage p; X_p indicates the lower-level model's operational variable in stage p; $\varphi Z_p \le h$ represents the constraint of the upper-level issue; $\kappa Z_p + \sigma U_p + \upsilon X_p \le e$ is the linear constraint of the lower-level problem; U_p is the binary variable in the lower-level problem; $\|J \cdot X_p\|_2 \le H \cdot X_p + w$ indicates the second-order cone constraint in the lower-level model; e is a variable that reflects uncertainty.

Decompose Eq. 43 into the master problem and subproblem, represented as Eqs. 44, 45 respectively.

$$\begin{aligned} MP: \qquad &\min_{Z_p} c^T \cdot Z_p + \phi \\ &s.t. \qquad \varphi \cdot Z_p \le h \\ &\qquad\quad \phi \ge 0 \end{aligned} \tag{44}$$

$$SP: \quad \max_{L_p \in \Phi} \ \min_{X_p \in \Omega\left(Z_p^*, L_p\right)} \ b^T \cdot X_p$$
$$s.t. \ \ \kappa \cdot Z_p^* + \sigma \cdot U_p + \upsilon \cdot X_p \le e \qquad (45)$$
$$\left\| J \cdot X_p \right\|_2 \le H \cdot X_p + w$$

where ϕ means the auxiliary variable; Z_p^* indicates the optimal solution get by solving the master problem Eq. 44.

1) Outer Circulation

i: Set the maximum limit $UB_p^{MP} = +\infty$ and minimum limit $LB_p^{MP} = -\infty$ of the master problem, with initial iteration k = 0.

ii: Address Eq. 44 to get Z_p^* and ϕ, and set the lower limit $LB_p^{MP} = \phi$ of the master problem.

iii: Use the inner loop to solve the subproblem and obtain the bottom scenario $e^\star$ and the upper limit UB_p^{MP} of the master problem.

iv: Assuming k = k+1, determine whether UB_p^{MP} and LB_p^{MP} satisfy Eq. 46. If not, add constraints Eqs 44–47 and start a new cycle; If satisfied, break the cycle and obtain the most suitable solution of the planning model.

$$\left| UB_p^{MP} - LB_p^{MP} \right| \le o \qquad (46)$$

$$\kappa \cdot Z_p^* + \sigma \cdot U_p^k + \upsilon \cdot X_p^k \le e^\star$$
$$\phi \ge b^T \cdot X_p \qquad (47)$$

2) Inner Circulation

i: Set the maximum limit $UB_p^{SP} = +\infty$ and minimum limit $LB_p^{SP} = -\infty$ of the subproblem, with initial iteration a = 0 and initial binary variable $U_p = U_p^{a^*}$.

ii: Solve Eq. 48 to obtain the upper limit of the subproblem and the bottom scenario $e^\star$.

$$\max_{\Phi} \ \theta$$
$$s.t. \ \ \theta \le \left(e - \kappa \cdot Z_p^* - \sigma \cdot U_p^{a^*}\right)^T \kappa^a \qquad (48)$$
$$\upsilon^T \kappa^a + J^T \pi^a = b, \left\| \pi^a \right\|_2 \le \chi^a$$
$$\theta \ge 0$$

where a is the current number of iterations; κ represents the dual variable introduced regarding the linear constraints of the lower-level model; χ and π are dual variables introduced regarding second-order cone constraints. Update the upper bound $UB_p^{SP} = \theta$ of the subproblem.

$$\min_{X_p, U_p} \ b^T \cdot X_p$$
$$s.t. \ \ \kappa \cdot Z_p^* + \sigma \cdot U_p + \upsilon \cdot X_p \le e^\star \qquad (49)$$
$$\left\| J \cdot X_p \right\|_2 \le H \cdot X_p$$

iii: Solve Eq. 49 to get the lower limit of the subproblem.

By solving Eq. 49, the optimal solution X_p^* can be obtained and the lower bound $LB_p^{SP} = \max\left\{b^T \cdot X_p^*, LB_p^{SP}\right\}$ of the subproblem can be updated.

$$\left| UB_p^{SP} - LB_p^{SP} \right| \le o \qquad (50)$$

$$\theta \le \left(e - \kappa \cdot Z_p^* - \sigma \cdot U_p^{a^*}\right)^T \kappa^a$$
$$\upsilon^T \cdot \kappa^a + J^T \cdot \pi^a = b, \left\| \pi^a \right\|_2 \le \chi^a \qquad (51)$$
$$\phi \ge b^T \cdot X_p^a$$

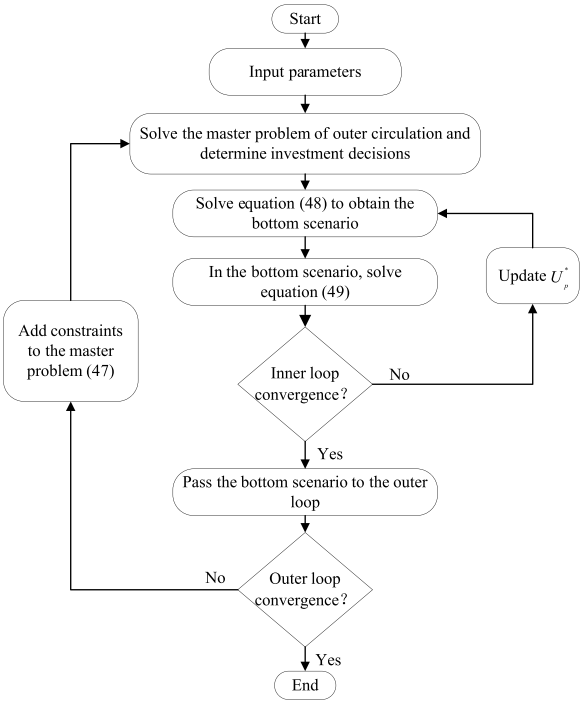

FIGURE 3
NCCG algorithm flowchart.

iv: Let a = a+1 to get a new binary variable $U_p^{a^*}$, and determine whether UB_p^{SP} and LB_p^{SP} satisfy Eq. 50. If so, break the inner loop, let $UB_p^{MP} = UB_p^{SP}$, and return $e^\star$ as the bottom scenario for the master problem; If not, add Eqs. 48–51 and return to step ii.

There is a nonlinear term $e \cdot \kappa^a$ in Eq. 51, and the big-M method is also used for linearization, which will not be repeated here.

The NCCG algorithm's flowchart is displayed in Figure 3. Firstly, solve the master problem of the external loop and determine the investment decision Z_p^*. In an iterative manner, Eqs 48, 49 are solved, subsequently, $e^\star$ and U_p^* are updated. The bottom scenario is determined by considering the convergence of the inner loop and $e^\star$ is set accordingly. Then, the bottom scenario is returned to the master problem. When the outer loop satisfies convergence, the algorithm terminates.

5 Case study

5.1 Parameter settings

In this article, it is first tested in a modified IEEE 33-node distribution system (Zhou et al., 2019) and then further applied in a 152-node 10 kV distribution system in a region. The example grid structure is shown in Figure 4 and Figure 5. The planning period is set to three stages, each lasting for 3a. The initial capacity of TS in the modified IEEE 33-node distribution

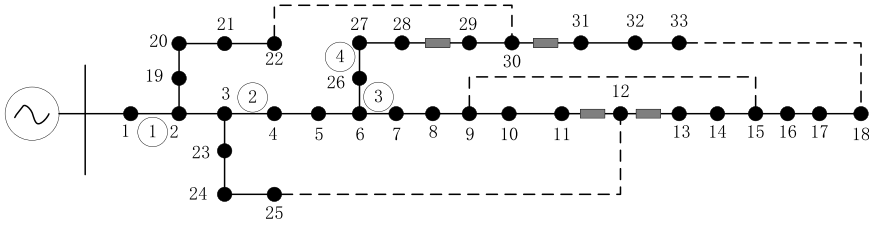

FIGURE 4
Modified IEEE 33-node system network structure.

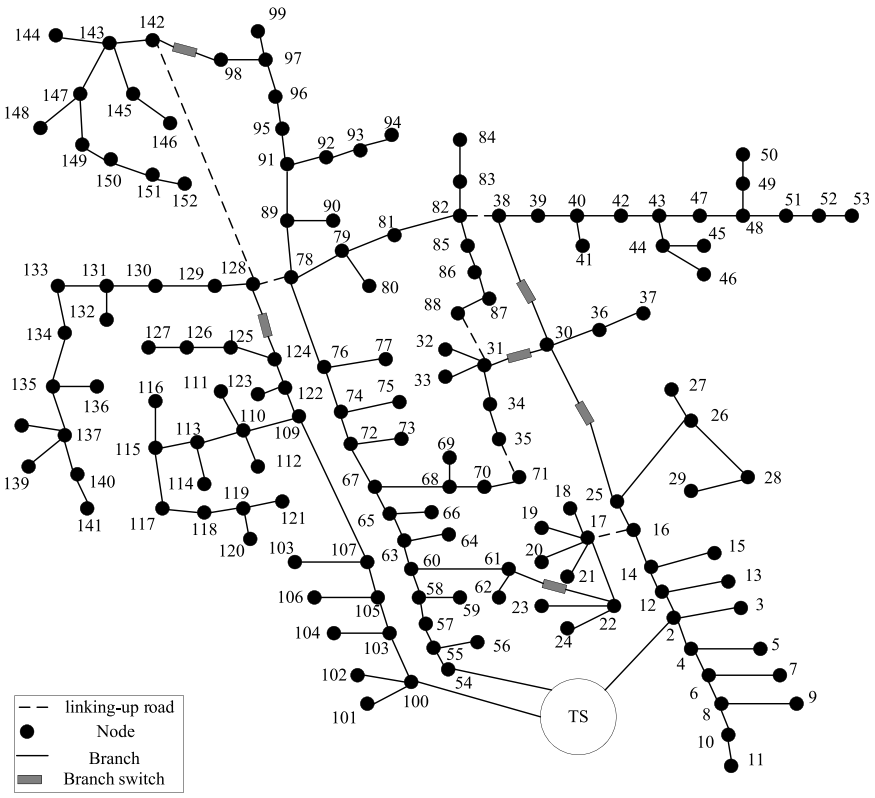

FIGURE 5
152-node system network structure.

system is 3.15 MVA, and the candidate main transformer capacities are 3.15MVA, 6.3 MVA, and 10 MVA; The initial capacity of TS in the 152-node distribution system is 10 MVA, and the candidate main transformer capacities are 6.3 MVA, 10 MVA, and 20 MVA; The cost of TS capacity increase is 500,000$/(MVA). The parameters of RIES candidate nodes and candidate equipment (Ehsan and Yang, 2019) can be found in Table 1 and Table 2. There are already biogas supply points built at the candidate nodes. The candidate lines and the parameters of the selected models are shown in Table 3 and Table 4, with a lifespan of 25a. Procurement of electricity from the power grid through the utilization of time-of-use tariffs can be observed in detail in Table 5 (Subramanian and Das, 2019) The constant price of biogas is set at 0.193 $/(kWh).

5.2 Scenario generation results

The setting of load parameters needs to take into account the impact of load development, seasonal factors, and other factors at each planning stage. Firstly, obtain the 8760h cold/hot/electrical load of a certain region through DesT software (DeST introduction, 2023). Then, a total of 1000 sets of data are generated through the implementation of Monte Carlo simulation. Subsequently, the enhanced k-means clustering methodology is employed to get the cooling, heating, and electricity loads for four representative days. The curves are depicted in Figure 6. Where Figure 6B and Figure 6D show significant changes in cold and heat loads, with high peaks and load-saving characteristics in summer and winter, respectively; The changes in cooling and heating loads in Figure 6A and Figure 6C are

TABLE 1 Parameters of energy conversion equipment in RIES.

Device	Type	Rated capacity (MW)	Energy efficiency coefficient	Unit cost (M$)
CHP	I	1.0	Electricity:0.28 Heat:0.55	0.4936
	II	0.85	Electricity:0.24 Heat:0.52	0.4420
	III	0.66	Electricity:0.22 Heat:0.48	0.4031
	IV	0.42	Electricity:0.20 Heat:0.45	0.3425
GB	I	1.7	0.88	0.2108
	II	1.5	0.72	0.1824
	III	0.84	0.75	0.1680
	IV	0.6	0.69	0.1425
ER	I	3	3.5	0.5027
	II	2	3.1	0.4025
	III	1.5	2.7	0.3745
	IV	0.8	2.4	0.3256
AR	I	2	1.45	0.1546
	II	1.75	1.44	0.1229
	III	1	1.40	0.0987
	IV	0.85	1.41	0.0824

RIES, Candidate nodes: 2、6、25、43、55、74、109、124、142.

TABLE 2 Parameter of energy storage equipment in RIES.

Device type	CES			TES		
	I	II	III	I	II	III
Rated power (MW)	1.0	0.75	0.5	1.0	0.75	0.5
Rated capacity (MWh)	16	12	8	4.5	3.0	2.0
Loss coefficient	0.002	0.0015	0.001	0.005	0.004	0.003
Charging efficiency	0.65	0.65	0.65	0.65	0.65	0.65
Discharging efficiency	0.65	0.65	0.65	0.65	0.65	0.65
Unit power construction cost (M$)	0.067	0.043	0.0306	0.0104	0.009	0.0074
Unit capacity construction cost (M$)	0.067	0.043	0.0306	0.0104	0.009	0.0074

relatively gentle, with characteristics of spring and autumn loads. In terms of describing load uncertainty, this article considers that the maximum deviation of various load fluctuations is 10%, and the parameter Γ is set to 6.

TABLE 3 Information of candidate lines.

Candidate line number	1	2	3	4	5	6
Starting node	1	1	1	2	54	100
End node	2	54	100	12	55	103

5.3 Planning results

5.3.1 Modified IEEE 33-node example

This article sets up two comparative schemes based on whether to consider RIES or not. Among them, without considering RIES, the cooling and heating loads are converted into electrical loads using a thermoelectric energy efficiency coefficient of 2.5; When considering RIES, the optimized configuration combination of RIES devices is shown in Table 6. Based on four typical daily load scenarios, the proposed rural distribution network planning method with RIES was first tested in the modified IEEE 33-node system. The planning and operation outcomes of each stage are presented in Table 7. Where the numbers inside and outside the parentheses of the RIES column represent the configuration combination number and installation location of the RIES

TABLE 4 Parameter of a line candidate model.

Type	Unit length cost (M$·km^{-1})	Impedance/(Ω·km^{-1})	Maximum current (A)
1	0.1753	0.8 + j0.4	150
2	0.2491	0.65 + j0.4	170
3	0.2825	0.45 + j0.4	215

TABLE 5 Peak Valley electricity prices.

Time periods(h)	Electricity price ($)
01:00–06:00	0.41
07:00–20:00	1.33
21:00–24:00	0.41

TABLE 6 Planning results of RIES equipment combination configuration.

Type	Configure capacity (MW)					
	CHP	GB	AR	ER	CES	TES
1	1	1.5	1.2	2.8	1	1
2	0.66	0.84	1	2	0.75	0.75
3	0.42	0.6	0.85	1.5	0.5	0.5

equipment; The numbers inside and outside the parentheses of the line represent the combination number and installation position of the line configuration; /represents no investment, the same applies later.

When considering RIES, there is no need to expand or renovate the line and TS, and only two RIES were invested in planning phase 1; Without considering RIES, TS was increased by 3.15MVA in planning stage 1, and both the line and TS were upgraded and renovated in planning stage 3. After considering RIES, the total annualized planning cost has slightly decreased by approximately 13.2%. In terms of operational efficiency, considering RIES, the annual operating costs of each planning stage have decreased by approximately 35.35% compared to not considering RIES.

5.3.2 152-Node system calculation example

Based on four typical scenarios, further application is carried out in a 152-node system. The distribution network planning and operation results of each planning stage are depicted in Table 8.

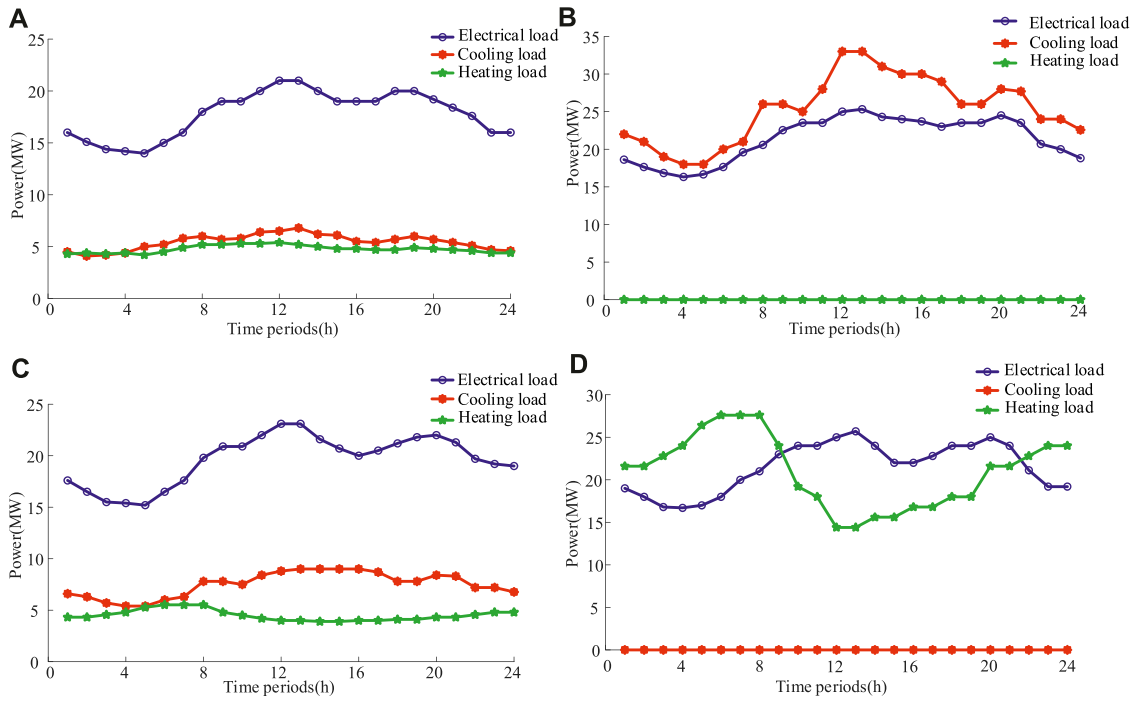

FIGURE 6
(A) Typical scenario 1; **(B)** Typical scenario 2; **(C)** Typical scenario 3; **(D)** Typical scenario 4.

TABLE 7 Comparison of planning results for modified IEEE 33-node system.

Program	Stage	Planning results			Annualized planning cost (M$)	Annualized operating cost (M$)	Total cost (M$)
		RIES	Line	TS expansion			
Consider RIES	1	2(1),6(2)	—	—	0.2556	3.9799	4.2355
	2	—	—	—	0	5.2449	5.2449
	3	—	—	—	0	7.1855	7.1855
Not considering RIES	1	—	—	3.15MVA	0.1102	6.3477	6.4579
	2	—	—	—	0	8.0407	8.0407
	3	—	1 (2)	6.3MVA	0.2478	10.8821	11.1299

TABLE 8 Comparison of planning results for the 152-node system.

Program	Stage	Planning results			Annualized planning cost (M$)	Annualized operating cost (M$)	Total cost (M$)
		RIES	Line	TS expansion			
Consider RIES	1	2(1),6(1),25(1),55(1),109(1),142(2)	—	—	0.9088	21.2285	22.1373
	2	—	—	—	0	27.6814	27.6814
	3	43(2),73(3)	—	—	0.1529	36.1457	36.2986
Not considering RIES	1	—	2(2),4(1)	10MVA	0.4048	39.2517	39.6565
	2	—	3 (1)	10MVA	0.3217	46.1826	46.5043
	3	—	1 (3)	6.3MVA	0.2852	57.8327	58.1179

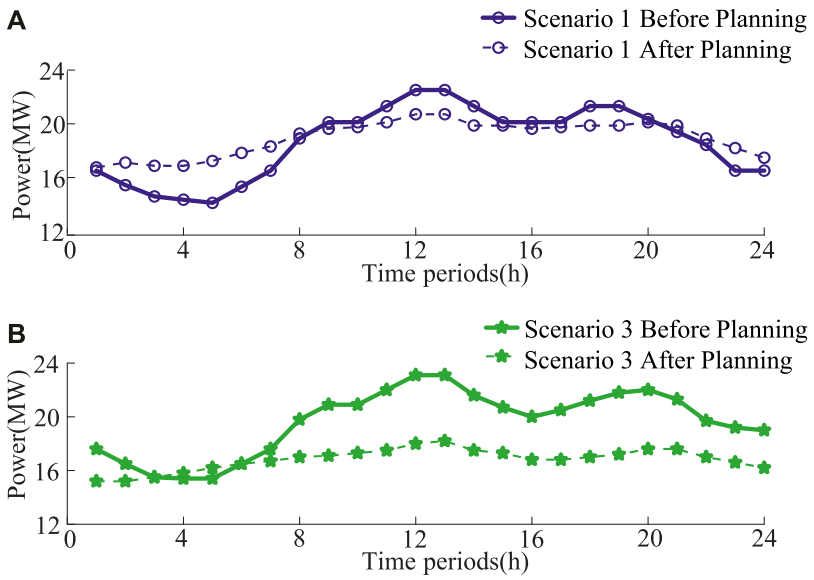

FIGURE 7
(A) Scenario 1 load comparison results in a 152-node system; **(B)** Scenario 3 load comparison results in a 152-node system.

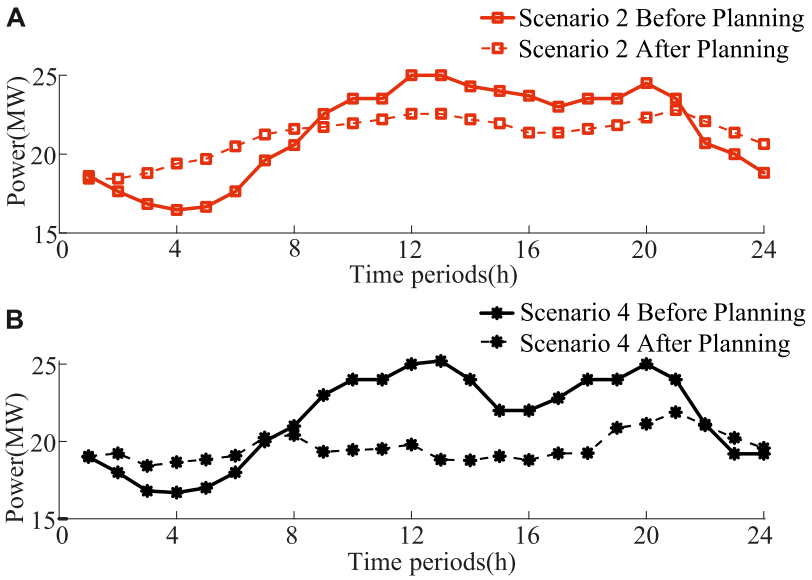

FIGURE 8
(A) Scenario 2 load comparison results in a 152-node system; **(B)** Scenario 4 load comparison results in a 152-node system.

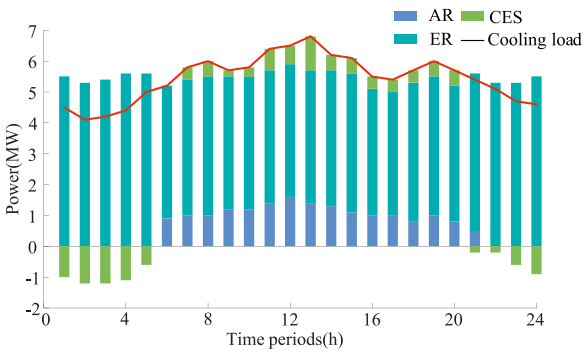

FIGURE 9
The cooling output of RIES after planning.

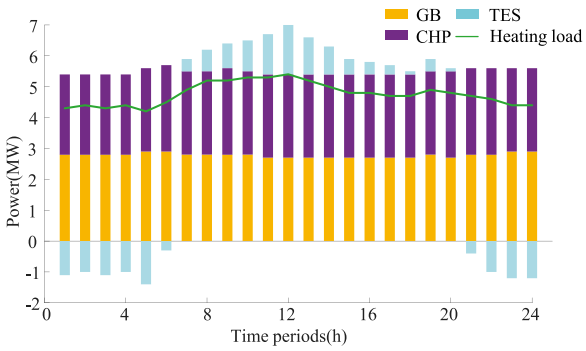

FIGURE 10
Heating output of RIES after planning.

As can be seen from the planning results in Table 8, when considering RIES, it is only planned in stages 1 and 3, and the investment is mainly in stage 1, accounting for 85.59%; Without considering RIES, there are investments in line and TS expansion at all stages. The total annualized planning cost when considering RIES is approximately 1.05 times that when not considering RIES. The annualized cost during the initial planning stage (stage 1) is approximately 2.245 times that without considering RIES. In the sub-sequent planning stages (planning stages 2 and 3), under the same load growth rate, the annualized planning cost is about 25.2% of that without considering RIES, and the in-vestment period is delayed by one stage. It can be seen that when considering RIES, there is only a significant investment in the initial stage, and the subsequent investment reduction effect is significant. There are two main reasons for this. 1) Before the construction of RIES, the cooling and heating loads in rural areas were mainly supplied by electricity, and the absence of energy storage devices was observed in the system. Therefore, after considering the construction of RIES, the initial investment is large, but it also effectively transfers some of the power supply pressure of the distribution net-work; Without considering RIES, the initial stage is based on the planning of the existing power grid, so the investment is relatively low; 2) After the distribution network planning, due to the regulatory role of RIES, the energy supply curve has been effectively smoothed. And the supply pressure of various loads was allocated to corresponding equipment, alleviating the pressure of electricity load growth.

As can be noticed from the operational results in Table 8, the distribution network operation economy after considering RIES has shown good results. The total annualized operating cost decreased by approximately 41.16% compared to when RIES was not considered. RIES can store energy during electricity price valleys and release energy during peak hours while using

TABLE 9 Comparison of results between robust planning schemes considering RIES and traditional planning schemes.

Category	Robust model	Deterministic model
TS expansion	—	—
line	—	—
RIES	2(1),6(2)	2 (1)
Energy purchasing (M$)	9.8791	6.8393
RIES operating cost (M$)	1.8872	0.9945
RIES flexible regulation (M$)	4.8996	6.3118
Total investment cost (M$)	0.2556	0.1632
Total operating cost (M$)	16.6659	14.1456

biogas to replace a portion of electricity during peak hours can further reduce operating costs; When not taking into account the RIES, the operating cost of the system is solely dependent on the load and electricity price, resulting in reduced economic efficiency.

Furthermore, with regard to the regulation of the disparity between peak and valley levels in the power grid, a comparison is made between the electricity load curves before and after the planning. The results are shown in Figure 7 and Figure 8. The peak-to-valley ratio of Scenarios 1 and 3 has increased from 66.67% to 65.80% before planning to 83.07% and 83.51% after planning; The peak to valley ratio of scenarios 2 and 4 has increased from 64.98% to 64.85% before planning to 84.14% and 80.87% after planning. The peak-to-valley ratio of four typical scenarios has been improved to varying degrees.

Take Scenario 1 as an example to conduct a detailed analysis of the operating mechanism of RIES. Figure 9 and Figure 10 depict the cooling and heating output of RIES subsequent to the implementation of the planning phase. From Figure 9, it can be seen that 01:00 to 05:00 are the low periods of cooling load, during which AR is not activated and ER output is relatively high; Starting from 06:00, CES stops cold storage and ER output decreases to cooperate with AR cooling; CES starts cooling from 07:00, and after reaching a peak at 19:00, the cooling load begins to decrease. At 22:00, CES stops cooling and starts cooling storage again, resulting in a decrease in AR output and an increase in ER output.

As can be noticed from Figure 10, the heat load curve is relatively flat, and due to the fixed price of biogas. Therefore, there is almost no significant fluctuation in the output of GB and CHP. TES stores heat from 01:00 to 06:00 and 21:00 to 24:00 and begin to release heat from 07:00 to 20:00 to cooperate with GB, CHP for heating load, and AR heating.

5.4 Comparison with traditional solutions

Table 9 presents a comparison of the results between the robust planning scheme considering RIES and the traditional planning scheme in the IEEE 33-node scenario. From Table 9, it can be seen that: 1) The robust model incurs higher investment and operational costs in comparison to the deterministic model. This is because the robust model takes into account the uncertainty of the cooling/heating/electricity loads in the planning process and the planning decisions made are more conservative than the deterministic model. Similarly, there is an increase in operational expenses. The increase in operating cost is mainly in the area of energy purchase; 2) The robust model reduces the cost of regulation of the distribution network compared to the deterministic model. This is because there are more RIES inputs, and the operational cost of the grid can be effectively reduced by energy storage and peak shifting, and energy substitution. Therefore, the robust planning model presented in this article can not only play a good effect on peak and valley difference regulation but also has a certain resistance to fluctuation risk.

5.5 Impact of uncertain parameters

To evaluate the flexibility of the proposed approach in relation to the cautiousness of planning schemes, five different sets of uncertain adjustment parameters were chosen for comparative simulation. The specific configurations of these parameters and the corresponding outcomes of the simulations are displayed in Table 10.

The data presented in Table 10 demonstrates that when the uncertainty adjustment parameter is set to 0, the robust planning model is equal to the Deterministic system. As the uncertainty parameters increase, the corresponding planning and operating costs increase. That is to say, when formulating the planning plan, more uncertainty is taken into account, resulting in a more conservative plan. With the increase of RIES in the planning scheme, the corresponding operating costs are also increasing.

TABLE 10 Comparison of planning scheme results under different uncertainty adjustment parameters.

Uncertainty adjustment parameters	Planning cost (M$)	Operating cost (M$)
$\Gamma = 0$	0.1632	14.1456
$\Gamma = 6$	0.2556	16.6659
$\Gamma = 12$	0.3162	18.6271
$\Gamma = 18$	0.3769	21.7882
$\Gamma = 24$	0.4402	24.4814

6 Conclusion

In this article, a multistage robust planning approach is introduced for rural distribution networks that incorporate integrated energy stations. The method takes the distribution network as the main body, and the integrated energy station plays the role of regulation to alleviate the pressure of power supply in rural areas; the improved k-means clustering algorithm is employed to obtain multiple typical scenarios to improve the clustering efficiency. The method's feasibility and effectiveness were evaluated using a modified version of the IEEE 33-node system, and then further based on a 152-node distribution system in a local area for practical application, the following conclusions are obtained.

1) The distribution network planning method proposed in this article, which takes into account RIES, has the advantage of reducing investment cycle and planning costs, although it has a relatively large investment in the initial planning stage. And as the planning period increases, its investment economy will also become more apparent. In terms of load regulation, RIES can effectively improve the peak-to-valley ratio of loads through the strategy of energy storage equipment storing energy during valley hours and releasing energy during peak hours.
2) The proposed model incorporates the consideration of load uncertainty. By continuously changing uncertainty parameters, the conservatism of the entire planning scheme can be flexibly adjusted. It is beneficial for planners to make a reasonable choice between operating costs and operational risks.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

ZY: Conceptualization, Investigation, Resources, Software, Writing–original draft. HM: Conceptualization, Formal Analysis, Investigation, Resources, Visualization, Writing–original draft. FY: Methodology, Resources, Validation, Writing–review and editing. YLi: Software, Validation, Writing–original draft. YS: Data curation, Project administration, Writing–review and editing. BZ: Funding acquisition, Investigation, Writing–review and editing. ZZ: Investigation, Supervision, Writing–original draft. WH: Software, Writing–original draft. YLe: Software, Validation, Writing–original draft.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

Authors ZY, HM, FY, YLi, YS, WH, and YLe were employed by Electric Power Research Institute of State Grid Hubei Co., Ltd. Authors BZ and ZZ were employed by State Grid Hubei Electric Power Co., Ltd.

Publisher's note

References

DeST introduction (2023). DeST introduction. Available at: http://dest.tsinghua.edu.cn/(Accessed July 4, 2023).

Dong, F., Hou, Y., Li, W., and Wang, Y. (2022). Intelligent decision-making of distribution network planning scheme with distributed wind power generations. *Int. J. Electr. Power and Energy Syst.* 136, 107673. doi:10.1016/j.ijepes.2021.107673

Ehsan, A., and Yang, Q. (2019). Scenario-based investment planning of isolated multi-energy microgrids considering electricity, heating and cooling demand. *Appl. energy* 235, 1277–1288. doi:10.1016/j.apenergy.2018.11.058

Feng, Y., Zhang, W., Yin, S., Tang, H., Xiang, Y., and Zhang, Y. (2023). A collaborative stealthy DDoS detection method based on reinforcement learning at the edge of internet of things. *IEEE Internet Things J.* 10, 17934–17948. doi:10.1109/JIOT.2023.3279615

Gan, W., Yan, M., Yao, W., Guo, J., Ai, X., Fang, J., et al. (2021). Decentralized computation method for robust operation of multi-area joint regional-district integrated energy systems with uncertain wind power. *Appl. Energy* 298, 117280. doi:10.1016/j.apenergy.2021.117280

Ge, S., Liu, X., and Ge, L. (2018). Planning method of regional integrated energy system considering demand side uncertainty. *Int. J. Emerg. Electr. Power Syst.* 20 (1), 20180195. doi:10.1515/ijeeps-2018-0195

Han, J., Zhang, L., and Li, Y. (2022). Spatiotemporal analysis of rural energy transition and upgrading in developing countries: the case of China. *Appl. Energy* 307, 118225. doi:10.1016/j.apenergy.2021.118225

Heidari, S., and Fotuhi-Firuzabad, M. (2019). Integrated planning for distribution automation and network capacity expansion. *IEEE Trans. Smart Grid* 10 (4), 4279–4288. doi:10.1109/tsg.2018.2855218

Huo, D., Gu, C., Greenwood, D., Wang, Z., Zhao, P., and Li, J. (2021). Chance-constrained optimization for integrated local energy systems operation considering correlated wind generation. *Int. J. Electr. Power and Energy Syst.* 132, 107153. doi:10.1016/j.ijepes.2021.107153

International Energy Agency (2015). Air conditioning use emerges as one of the key drivers of global electricity-demand growth. Available at: https://www.iea.org/news/air-conditioning-use-emerges-as-one-of-the-key-drivers-ofglobal-electricity-demand-growth (Accessed July 4, 2023).

Jiang, M., Liu, T., Goh, H., Zhang, D., Dai, W., Liu, J., et al. (2020). Estimation of operation cost of residential multiple energy system considering uncertainty of loads and renewable energies. *IEEE Access* 9, 4874–4885. doi:10.1109/access.2020.3046621

Kim, M., and Ramakrishna, R. S. (2005). New indices for cluster validity assessment. *Pattern Recognit. Lett.* 26 (15), 2353–2363. doi:10.1016/j.patrec.2005.04.007

Kou, X., and Li, F. (2018). Interval optimization for available transfer capability evaluation considering wind power uncertainty. *IEEE Trans. Sustain. Energy* 11 (1), 250–259. doi:10.1109/tste.2018.2890125

Koutsoukis, N., and Georgilakis, P. (2019). A chance-constrained multistage planning method for active distribution networks. *Energies* 12 (21), 4154. doi:10.3390/en12214154

Lei, Y., Wang, D., Jia, H., Chen, J., Li, J., Song, Y., et al. (2020). Multi-objective stochastic expansion planning based on multi-dimensional correlation scenario generation method for regional integrated energy system integrated renewable energy. *Appl. energy* 276, 115395. doi:10.1016/j.apenergy.2020.115395

Li, H., Ren, Z., Fan, M., Li, W., Xu, Y., Jiang, Y., et al. (2022b). A review of scenario analysis methods in planning and operation of modern power systems: methodologies, applications, and challenges. *Electr. Power Syst. Res.* 205, 107722. doi:10.1016/j.epsr.2021.107722

Li, H., Ren, Z., Trivedi, A., Srinivasan, D., and Liu, P. (2023). Optimal planning of dual-zero microgrid on an island towards net-zero carbon emission. *IEEE Trans. Smart Grid*, 1. doi:10.1109/TSG.2023.3299639

Li, H., Ren, Z., Trivedi, A., Verma, P., Srinivasan, D., and Li, W. (2022a). A noncooperative game-based approach for microgrid planning considering existing interconnected and clustered microgrids on an island. *IEEE Trans. Sustain. Energy* 13 (4), 2064–2078. doi:10.1109/tste.2022.3180842

Li, R., Wang, W., and Xia, M. (2017). Cooperative planning of active distribution system with renewable energy sources and energy storage systems. *IEEE Access* 6, 5916–5926. doi:10.1109/access.2017.2785263

Moazeni, S., Miragha, A. H., and Defourny, B. (2018). A risk-averse stochastic dynamic programming approach to energy hub optimal dispatch. *IEEE Trans. Power Syst.* 34 (3), 2169–2178. doi:10.1109/tpwrs.2018.2882549

Subramanian, V., and Das, T. K. (2019). A two-layer model for dynamic pricing of electricity and optimal charging of electric vehicles under price spikes. *Energy* 167, 1266–1277. doi:10.1016/j.energy.2018.10.171

Wu, C., Gu, W., Jiang, P., Li, Z., Cai, H., and Li, B. (2017). Combined economic dispatch considering the time-delay of district heating network and multi-regional indoor temperature control. *IEEE Trans. Sustain. Energy* 9 (1), 118–127. doi:10.1109/tste.2017.2718031

Wu, J., Yan, J., Jia, H., Hatziargyriou, N., Djilali, N., and Sun, H. (2016). Integrated energy systems. *Appl. Energy* 167, 155–157. doi:10.1016/j.apenergy.2016.02.075

Yan, B., Yan, J., Li, Y., Qin, Y., and Yang, L. (2021). Spatial distribution of biogas potential, utilization ratio and development potential of biogas from agricultural waste in China. *J. Clean. Prod.* 292, 126077. doi:10.1016/j.jclepro.2021.126077

Yan, M., Ai, X., Shahidehpour, M., Li, Z., Wen, J., Bahramira, S., et al. (2019). Enhancing the transmission grid resilience in ice storms by optimal coordination of power system schedule with pre-positioning and routing of mobile DC de-icing devices. *IEEE Trans. Power Syst.* 34 (4), 2663–2674. doi:10.1109/tpwrs.2019.2899496

Zeng, B., and Zhao, L. (2013). Solving two-stage robust optimization problems using a column-and-constraint generation method. *Operations Res. Lett.* 41 (5), 457–461. doi:10.1016/j.orl.2013.05.003

Zhou, Q., Shahidehpour, M., Yan, M., Wu, X., Alabdulwahab, A., and Abusorrah, A. (2019). Distributed secondary control for islanded microgrids with mobile emergency resources. *IEEE Trans. Power Syst.* 35 (2), 1389–1399. doi:10.1109/tpwrs.2019.2942269