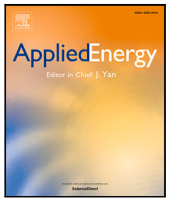
Applied Energy 326 (2022) 119942

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy



Expansion planning of active distribution networks achieving their dispatchability via energy storage systems

Ji Hyun Yi [a,*], Rachid Cherkaoui [a], Mario Paolone [a], Dmitry Shchetinin [b], Katarina Knezovic [b]

[a] Distributed Electrical Systems Laboratory, École Polytecnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
[b] Hitachi Energy Research, Baden Dättwil, Switzerland

ARTICLE INFO

Keywords:
Active distribution networks
Benders decomposition
Energy storage systems
Grid expansion
Line reinforcement
Optimal power flow
Planning dispatchability

ABSTRACT

This paper presents a combined framework for power distribution network expansion planning (DNEP) and energy storage systems (ESSs) allocation in active distribution networks (ADNs) hosting large amount of photovoltaic (PV) generations and loads. The proposed DNEP ensures the reliable operation of the targeted ADN with the objective of achieving its dispatchability while minimizing grid losses by determining the optimal grid expansion to connect new nodes, the reinforcement of existing lines, and the ESS allocation. The allocated ESSs compensate for the stochastic power flows caused by the stochastic loads and generation, allowing ADNs to follow a pre-defined power schedule at the grid connection point. The grid constraints are modeled by using a modified augmented relaxed optimal power flow (AR-OPF) model that convexifies the classical AC-OPF providing the global optimal and the exact solution of the OPF problem for radial networks. The DNEP problem's complexity is handled by employing a sequential algorithm where new nodes are added one by one, following the priorities determined by the user. In each stage of the sequential planning, the Benders decomposition algorithm determines the optimal solution for investments and ADN operation iteratively. Moreover, the siting and sizing problems associated with the ESSs and line investment are solved separately to enhance the convergence speed. Simulations are conducted on a real 55-node Swiss ADN hosting sizeable stochastic photovoltaic generation.

1. Introduction

To meet global decarbonization objectives, conventional power generation is progressively displaced by renewable energy resources (RES) such as Photovoltaic (PV) plants. However, the such energy transition has increased the stochasticity of the power generation mix, which impacts the planning and operational practices of both distribution and transmission networks. On the one hand, it causes difficulties securing a distribution network's stable and reliable operation, such as degradation in the quality of supply mainly associated with voltage control, as well as lines and transformers congestions [1]. On the other hand, the stochastic nature of the RES generation increases the prediction uncertainties of the active distribution networks (ADNs) prosumption, making the grid connecting point (GCP) power flow more difficult to comply with a pre-defined power schedule. This results in higher scheduling and activation of spinning reserves and larger expenses for grid ancillary services [2]. Several studies have suggested transmission system operator (TSO)-distribution system operator (DSO) coordination to mitigate the power imbalance within the whole system [3]. Furthermore, the authors of [4] proposed a contractual framework to delegate system balancing responsibility from TSOs to DSOs. In this regard, [5,6] proposed to integrate energy storage system (ESS) assets with renewable generation units such that the ESSs compensate for the prediction uncertainty of stochastic generation, rendering the renewable generation a *dispatchable* resource. Furthermore, [7,8] showed how ESSs can support the dispatchable operation of ADNs. More specifically, the control strategy to achieve *dispatchability-by-design* was introduced in [8] with experimental validation on a real Swiss distribution feeder. The control strategy consisted of a two-stage process following the conventional structure of system operation: *day-ahead operation* and *real-time operation*. During the day-ahead operation, a day-ahead power schedule at GCP, or *dispatch plan*, is determined based on the prosumption forecast before the beginning of the actual operation. Then, during real-time operation, the ESS power dispatch is controlled to compensate for the inevitable *dispatch errors* (*i.e.*the tracking error between the day-ahead power schedule and realized power at the GCP). In [9], the significance of ADN dispatchability is

* Corresponding author.
E-mail addresses: ji.yi@epfl.ch (J.H. Yi), rachid.cherkaoui@epfl.ch (R. Cherkaoui), mario.paolone@epfl.ch (M. Paolone), dmitry.shchetinin@hitachienergy.com (D. Shchetinin), katarina.knezovic@hitachienergy.com (K. Knezovic).

https://doi.org/10.1016/j.apenergy.2022.119942
Received 5 June 2022; Received in revised form 15 August 2022; Accepted 3 September 2022
Available online 29 September 2022

verified through its capability to reduce bulk power systems' spinning reserve requirement. In this regard, for utilities and DSOs, it is worth exploring a proper distribution network expansion planning (DNEP) strategy accounting for ADN dispatchability as well as secure operation of the local grid to seamlessly integrate RESs while minimizing the impact on ADNs planning costs.

However, according to [10], conventional DNEP is inefficient in handling the challenges mentioned above as active elements' role in managing the network's operation is limited. In the existing literature, DNEP problems are often proposed to size and site ESSs (e.g., [11–16]), distributed generation units (e.g., [13,17]), voltage regulators [18], static var compensators, and on-load-tap-changers [19]. These problems are solved with various objectives, such as minimizing power losses [11–17], assuring the reliability of power supply [11–16], improving the quality of power supply (*e.g.*, voltage) [17], and reducing operation costs [16,18,19].

However, to the best of the authors' knowledge, none of those studies mentioned above treated a planning strategy that improves the performance of the power dispatch of ADNs by tackling the prosumption stochasticity. Regarding the ADN's dispatch problem, in [8], the day-ahead operation problem is formulated as a robust optimization problem such that the possible realization of prediction error (including the worst-case scenario) can be compensated by maximizing the exploitation of given ESS assets located at the GCP. Inevitably, such a study on the control strategy of ESSs raises another crucial research question regarding how the ESS site and size should be determined to achieve ADN's dispatchability. In this context, in the authors' previous work [20], a planning strategy is proposed to allocate ESSs with a particular focus on enhancing the dispatchability of ADNs while complying with network constraints such as nodal voltages and line ampacity limits. In contrast to [8], the prosumption uncertainty is considered by formulating a stochastic optimization problem for the dispatching problem. The stochastic optimization problem can avoid over-investment compared to the robust one because it aims to achieve the economically optimal investment solution with a slight increase in risk.

In [21], the planning tool proposed in [20] was extended to consider line reinforcement on top of the ESS allocation in order to identify a better optimal investment solution even in binding operational conditions due to growing PV generation capacity within the distribution network. The co-optimization of line reinforcement and ESS allocation assured the feasible operation of the network by updating the line ampacity, while the allocated ESS compensated for the increased stochasticity stemming from the PV injection. Finally, this study extends the method proposed in [20,21] to a DNEP strategy to host newly integrated PV resources and customers, *i.e.*it takes into account the distribution network expansion as a decision variable. The proposed tool co-optimizes the investment in new line installations, line reinforcement, and ESSs allocation to achieve ADNs dispatchability.

The DNEP problem is non-linear and non-convex due to the power flow equations and the presence of binary decision variables. Due to its non-convexity, this type of problem is often solved by meta-heuristic algorithms [17,22,23]. However, these solution techniques do not guarantee optimality or even feasibility with respect to the ADN operation. Other studies used convexification schemes on the OPF model, such as linear approximation [11,16,24] or second-order cone (SOC) relaxation [19,25]. For example, the multi-stage expansion planning tool developed in [11] evaluates the operation through a probabilistic linear OPF model while embedding the ESS daily schedule identified by dynamic programming. The main drawback of the linear OPF model is that the approximation of the power flow may not be accurate enough to ensure physically feasible operation. Alternatively, the second-order cone programming (SOCP) model [19,25] can improve the modeling accuracy of the load flow, but at the cost of higher computation time compared to the linear OPF. In [19], the DNEP problem is modeled as a mixed-integer SOCP problem. The system operation is formulated as a SOCP relaxed OPF model while maneuvering a demand management scheme through the optimal load-shedding direction. Yet, the utilized OPF model did not consider the shunt element of lines. Moreover, the exactness of its solution cannot be guaranteed in the case of over-voltages and binding ampacity limits [26]. In [27], the authors investigated the solution inexactness of the general SOCP relaxation depending on different operational objectives, from which they observed the possible incompatibility between the identification of an exact solution and the efficient operation of the ADN with high penetration of distributed resources. The authors proposed to tighten the relaxation by introducing an iterative algorithm that increasingly adds cutting planes to a SOCP-relaxed OPF model. However, the formulation of operation based on such an iterative approach may not be computationally efficient to be utilized as the core algorithm of the large-scale optimization problem for ADN planning. Instead, in [28], to guarantee the exactness of the solution obtained from the SOCP model, augmented bounds are built upon the grid constraints in the so-called augmented relaxed OPF (AR-OPF) model. Furthermore, the authors proved and numerically verified the superiority of the AR-OPF model over the SOCP relaxed model proposed in [29] in terms of the exactness of the operational solution. In this regard, the ESS allocation strategy proposed in [20] was developed based on the AR-OPF model, where the model was appropriately modified to tackle the specific problem of ADNs dispatchability achieved by the ESSs. In [20], the operational solution's exactness within the system configuration defined by the optimal investment solution was compared to the standard SOCP relaxed model regarding statistically assessed errors' distribution of branch currents. The numerical result verified that the MAR-OPF model is superior to the standard SOCP-OPF model in obtaining an exact solution. Hereafter, the modified AR-OPF is referred to as MAR-OPF.

The DNEP methodology proposed in this paper likewise relies on a scenario-based stochastic MAR-OPF model. However, in contrast to the ESS allocation problem, including the network investment option in the DNEP strategy requires fundamental modification of power flow equations and operating constraints associated with changes in topology and network adjacency matrix. By employing the big-M method, the power flow equations and grid constraints are selectively activated only when the optimal connection among candidate lines is chosen, such that the system operation under the optimal connection to the new node is evaluated. The admittance matrix is adjusted along with the change of line ampacities of new line and the line to be reinforced, following the relation between the line parameters and the line ampacity presented in [21].

Moreover, the increased problem complexity caused by the simultaneous management of ESSs and lines investment may necessitate a change in the solution approach to improve the tractability of the planning problem. To cope with the computational complexity of the proposed planning framework, in [21], the reformulation of the planning problem is proposed to determine the investment decisions regarding the sites and sizes of assets in separate sub-stages. In both sub-stages, a suitably modified Benders decomposition was employed. The reformulated planning problem showed faster convergence of the Benders decomposition in the performance comparison with the original problem. In view of its advantage, particularly important when solving a planning problem consisting of numerous binary variables, the reformulation technique is fittingly applied to the DNEP problem. Moreover, a sequential algorithm is introduced to add new nodes to the existing network, making the proposed DNEP strategy capable of handling the scalability issue raised by the increasing number of new nodes. In summary, the contributions of the paper are threefold:

- The DNEP strategy relying on a scenario-based stochastic programming approach is proposed to ensure sufficient hosting capacity and achieve ADNs dispatchability when increasing stochastic renewable generation and demand.

- The ESSs and lines investment decisions are optimized based on the accurate evaluation of ADN operation via the MAR-OPF model, appropriately converted to consider the change of network topology and adjacency matrix associated with assets investment.
- The computational burden of the DNEP problem is mitigated by employing a specific sequential algorithm that consists of two sub-stages: the 1st sub-stage sequentially integrates new nodes while determining lines for reinforcement and nodes for ESS allocation, and the 2nd sub-stage determines the capacity of lines and ESSs.

The paper is organized as follows: Section 2 introduces the structure of the optimization problem and explains the key parts in detail. In Section 3, the proposed problem and the associated solution are described and discussed. Section 4 contains a detailed application example referring to the DNEP considering ESSs on real ADNs. Finally, Section 5 concludes the paper by discussing the main findings.

2. System description

The goal of the DNEP problem is to determine (i) the optimal routing to new nodes hosting additional stochastic resources (consumption/renewable generation), (ii) the reinforcement of existing lines, and (iii) the allocation of ESSs[1] to maintain a feasible and dispatchable operation of the network, while minimizing the total investment and operating costs. The inputs to the DNEP problem include characteristics of the existing networks and the new assets (ESSs and lines), locations and capacities of newly integrated load and PV generation units, the existing candidate nodes that can be connected to new nodes, and the ESS candidate nodes. The time-series prosumption profile is assumed to be known by the modeler, while its seasonal variability and prediction uncertainty is represented by operating scenarios.

The DNEP is applied to a MV distribution network with a radial topology, where $l \in \mathcal{L}(|\mathcal{L}| = L)$ is the generic line, $i \in \mathcal{N}(|\mathcal{N}| = L{+}1)$ is the generic node, and $n' \in \mathcal{N}'(|\mathcal{N}'| = N')$ is the new node. The network is considered three-phase balanced without any coupling between the phases. The set of candidate nodes that can be connected to a new node are defined for each new node n' and indicated as $\mathcal{K}(n')$. Its element is indexed as $k \in \mathcal{K}(n')$. The binary variable for line reinforcement is denoted by $X_l \in \{0,1\}, \forall l \in \mathcal{L}$. The decision variable associated to the ampacity of the generic reinforced line is denoted by $A_l, \forall l \in \mathcal{L}$. The installation of the line connecting node $k \in \mathcal{K}(n')$ and new node $n' \in \mathcal{N}'$ is denoted by the variable $X'_k \in \{0,1\}, \forall k \in \mathcal{K}(n'), \forall n' \in \mathcal{N}'$, whereas line ampacities of the candidates for the new line are denoted by $A'_k, \forall k \in \mathcal{K}(n'), \forall n' \in \mathcal{N}'$. Line ampacities A_l and A'_k are determined for the lines where the line investment take place (i.e., $X_l = 1, \forall l \in \mathcal{L}$, and $X'_k = 1, \forall k \in \mathcal{K}(n'), \forall n' \in \mathcal{N}'$). Another set of binary variables is defined for each node $i \in \mathcal{N}$ to determine the ESS allocation (while ESS candidate nodes can be given by the user) and its array is denoted as $U_i \in \{0,1\}, \forall i \in \mathcal{N}$. The energy reservoir ($C_i$) and power rating ($R_i$) are determined at the node where the ESS is allocated ($U_i = 1, \forall i \in \mathcal{N}$).

Each node in the ADN has a non-dispatchable complex power prosumption. The prosumption profiles are classified into a group of several typical day-types indexed with d, $\forall d \in \mathcal{D}$, representing their seasonal variations. Over the planning horizon Y, the load consumption of year $y \in \{0,1,\dots,Y{-}1\} = \mathcal{Y}$ is increased with the annual rate of r_i from the previous year. The stochasticity of each day-type prosumption is modeled by a set of scenarios ($s_{i\phi t} = p_{i\phi t} + jq_{i\phi t}, \forall \phi \in \Phi_{dy}$), which is defined for $\forall d \in \mathcal{D}$ and $\forall y \in \mathcal{Y}$. The probability of each scenario is given by $\lambda_\phi, \forall \phi \in \Phi_{dy}$. The dispatching problem considering prosumption uncertainty is modeled as a scenario-based stochastic problem. The main operational objective is to achieve optimal ADN dispatchability considering the trade-off between the ESS investment cost and the penalty cost caused by the tracking error concerning the pre-defined power schedule set at the GCP. On each day, a daily dispatch plan ($DP_{tdy}, \forall t \in \{1,\dots,T\} = \mathcal{T}$) is derived day-ahead based on the aggregated prosumption prediction,[2] and the active power through GCP of the ADN (assigned with the node number $i = 1$) is dispatched in real-time to follow the dispatch plan. The active power of the ESS allocated at node i ($p^E_{i\phi t}$) is dispatched to enhance the tracking accuracy by minimizing the observed active dispatch error and, thus, the corresponding imbalance penalty. Meanwhile, ESS reactive power $q^E_{i\phi t}$ supports the reactive power flows to maintain the desired operating condition.

In summary, the DNEP problem is structured as a two-stage decision process: (1) the first stage determines the binary decision variables of the ESS location (U_i), the line to be newly constructed (X'_k), the existing lines to be reinforced (X_l), and the continuous decision variables on the ESS energy capacity (C_i), their power rating (R_i), the line ampacity of the new line (A'_k) and the reinforced lines (A_l); and (2) the second stage deals with the daily operation problem, where the decision variables on the ESSs active and reactive power are determined for all operating scenarios and the power state variables accordingly.

3. Problem formulation

The complexity of the considered DNEP problem stems mainly from three points: (i) the numerous possible connections between the existing and new nodes that introduce numerous binary decision variables; (ii) the two-layered decision-making process on planning and operation; (iii) the simultaneous decision on siting and sizing of lines and ESSs.

Regarding (i), although considering all the possible connections of new nodes at once can provide optimal investment solutions, it results in an intractable problem as the number of new nodes increases. Therefore, it is computationally more efficient to consider a sequential approach for connecting the new nodes, while sacrificing the solution's global optimality. Moreover, given that the integration of loads and distributed generation units is spread over a long-term horizon following the pre-established targets set by DSO and customers, the sequential approach following the priority order of new nodes is in accordance with the standard planning process adopted by DSOs.

(ii) can be tackled by employing the Benders decomposition algorithm [30], which is widely applied for solving large-scale optimization problems. The planning problem is decomposed into an investment decision stage (master problem) and parallel operation decision stages (subproblems). After solving the master problem, the values of the investment decision variables are fixed temporarily in the following subproblems, such that the given problem is reduced to an operational problem parameterized by the value of the investment decision variables vector. Then, the optimal value of this vector is updated by a cutting-plane approach.

(iii) is related to the decision variables of different natures (continuous and binary) constituting the investment options. The Benders cuts are generated based on the dual values associated with the continuous investment variables. They are often not efficient enough to narrow down the mixed-integer solution space. In other words, including numerous binary investment variables makes the planning algorithm even more burdensome to converge to an optimal investment solution since each solution value of binary variables has to be checked with various combinations of the continuous investment variables' values. The excessive number of Benders iterations caused by inefficient Benders cuts increases the master problem's size and slows the convergence.

[1] The considered planning problem does not consider hybrid PV-ESS systems since their behavior is user-defined and hard to predict. Future research has to focus on integrating these systems in the proposed planning tool.

[2] We assume a suitable forecasting tool to be available with known statistical distributions of prediction uncertainties.

Therefore, in the proposed method, the planning problem employing the Benders decomposition is reformulated into two sub-stages (each employing Benders decomposition) such that siting and sizing decisions of ESSs and lines are made in a decoupled way. The Benders cut for the siting problem is strengthened by calculating the optimal sizes achieved at each possible siting solution while solving parallel operation problems. When the optimal siting solution is obtained, siting solution is fixed such that the master problem of the sizing problem is reduced to a continuous programming problem.

Aside from the computational complexity issues, the last point to be addressed concerns the accurate modeling of ADN operation. The AR-OPF model is employed to model the ADN operation. However, one of the pre-requisite conditions to guarantee the exactness of the solution is to have the AR-OPF objective strictly increasing with the grid losses. Given this, the random nature of dispatch error cannot make the objective term related to the minimization of the dispatch error (which is the main operational objective of interest in this paper) comply with the necessary condition. Therefore, the exactness of the AR-OPF solution cannot be guaranteed if the objective value that corresponds to the dispatch error is significant compared to the objective term regarding the total grid losses. We propose to decompose the problem into two blocks, each consisting of an OPF problem. In this way, the dispatchability level is obtained by solving another, linearly approximated, OPF problem (the so-called 1st block problem) based on the ESSs investment cost and the penalty cost of dispatch error. Then the dispatchability level is imposed as operating constraints in the AR-OPF model-based operation problem of the 2nd block.

To summarize, the DNEP problem is solved by a sequential approach comprising several rounds of the planning problem, each tackling the addition of a single new node to the existing grid in a sequential priority order defined by the user. Each planning problem is decomposed into two blocks. In the 1st block, the economic benefits/penalties associated with the network dispatchability are obtained considering the prosumption of the new node by quantifying the optimal trade-off between the ESS allocation costs and the avoidance of the power dispatch imbalance penalties. In the 2nd block, the new lines, the reinforcement of the existing lines, and the ESSs allocation are determined to comply with the dispatch error level (obtained from the 1st block) and with the network operating constraints. The reader is referred to [20] for more details regarding the two-block structure. Compared to [20], [21] proposes to decompose the 2nd block problem into two sub-stages: 1st sub-stage where the binary investment decisions are determined, and 2nd sub-stage where values of the continuous investment variables are obtained. The Benders decomposition technique is applied to both sub-stages to tackle the computational complexity. It assigns investment-related variables in a master problem and operation-related variables in parallel subproblems representing daily dispatch problems. In each Benders iteration, the master problem optimally updates the investment solution based on evaluating the subproblem costs for the investment solutions given by all the previous iterations. The optimal planning solution is determined through a finite number of Benders iterations as the difference between the upper and lower bounds on the planning costs becomes smaller than a pre-defined threshold.

The whole algorithm of the proposed approach is illustrated in Fig. 1 and described step by step in Alg. 1. We start the planning algorithm by selecting the first new node $n' = 1$ from the pool of new N' nodes accordingly to the priority defined by the user. First, the 1st block problem is solved to determine the optimal level of dispatchability when node n' is added to the existing network. Then, candidate nodes $\mathcal{K}(n')$ are chosen based on the modeler's criteria. In this study, we selected the candidate nodes based on their distance from the new node. After solving the 1st sub-stage of the 2nd block problem, the optimal allocation of a new line connecting the new node is determined along with the possible reinforcement of existing lines and the ESSs allocation. The grid is thus expanded by having one more line (*i.e.*, $L \leftarrow L+1, \mathcal{L} \leftarrow \mathcal{L} \cup k^*$, where k^* indicates the optimally determined candidate for the connection to the new node) and node (*i.e.*, $N \leftarrow N+1, \mathcal{N} \leftarrow \mathcal{N} \cup n'$). Then, the determined line candidate's fixed cost parameter ($I^c_{k^*}$) becomes the cost parameter of line L (I^r_L). The node next in the queue (*i.e.*, $n' = 2$) is tackled in the next iteration of the planning problem. Once all new nodes are connected to the existing grid through this procedure, the 2nd sub-stage of the 2nd block problem is solved to determine the ESS sizes and the optimal line ampacities.

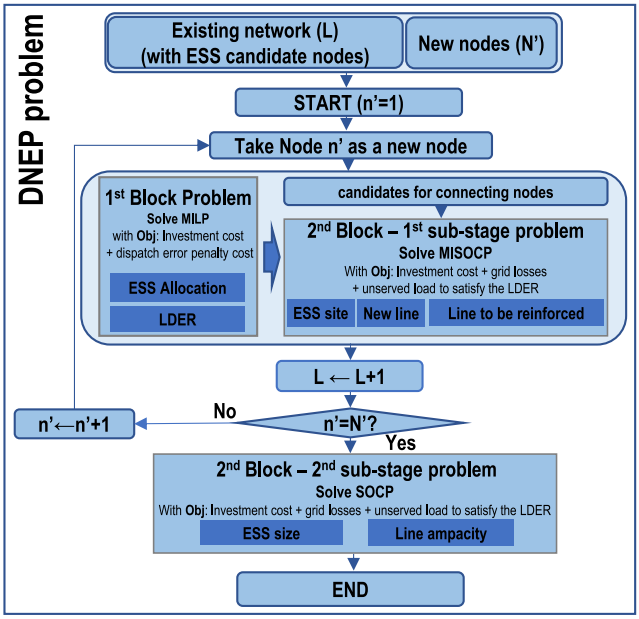


Fig. 1. Full algorithm of the proposed method.

Algorithm 1 Full algorithm of the proposed expansion planning method

Input : Existing network ($\mathcal{L}$), prosumption ($s = p + jq$), ESS candidate nodes ($\mathcal{N}_E$), new nodes ($\mathcal{N}'$)

Initialization : $n' \leftarrow 1$;
while $n' \leq N'$ **do**
  Consider Node n' as a new node
  Solve 1st block problem to determine the optimal level of dispatchability with node n' added to the existing network
  Select candidate nodes $\mathcal{K}(n')$ for connecting the existing nodes to the new node according to the modeler's criteria
  Solve 1st sub-stage of 2nd block problem to determine $U^*_i, \forall i \in \mathcal{N}$ (ESS site), $X'^*_k, \forall k \in \mathcal{K}(n')$ (connection to the new node), $X^*_l, \forall l \in \mathcal{L}$ (line to be reinforced)
  $L \leftarrow L+1, \mathcal{L} \leftarrow \mathcal{L} \cup k^*$, $N \leftarrow N+1, \mathcal{N} \leftarrow \mathcal{N} \cup n'$ {Update the network to include the new node}
  $n' \leftarrow n' + 1$
end while
return $\mathcal{N}, \mathcal{L}, U^*, X^*$
Solve 2nd sub-stage of 2nd block problem to determine ESS size ($R^*_i, C^*_i, \forall i \in \mathcal{N}_E$), line ampacities of new lines and lines to be reinforced ($A^*_l, \forall l \in \mathcal{L}$)
return R^*, C^*, A^*

3.1. 1st block problem

The 1st block problem determines the optimal ESS allocation and the grid dispatchability level by evaluating the dispatch operation for all operating scenarios with the PV generation and load consumption of a new node integrated into the existing grid. The formulation of the 1st block problem is the same as the one already proposed

in [20]. The planning problem is modeled as a mixed-integer linear programming (MILP) with the objective of minimizing the investment and total imbalance penalty costs over the planning horizon, and the set of constraints associated with the investment conditions and the linear Distflow OPF model. Shunt elements of the distribution lines are considered to appropriately model their reactive power generation and the impact on the nodal voltage profiles. In contrast, grid losses and the lines' ampacity constraints are neglected. Instead, they are suitably considered in the 2nd block problem.

3.1.1. Modeling of the ESS investment and operation

Eqs. (1a)–(1c) model the possible investment range for the ESS allocation. Power ratings and energy capacity ranges are given by (1a) and (1b). $\underline{C}_i/\overline{C}_i$ is the minimum/maximum possible ESS energy reservoir capacity at node i; $\underline{R}_i/\overline{R}_i$ represents minimum/maximum possible ESS power rating capacity at node i; and $\underline{CR}$ represents the limit of the rate at which the ESS is discharged with respect to its maximum energy capacity. The power rating and energy reservoir are determined according to (1c). $\mathcal{IC}_e$ represents the ESS investment cost defined by (1d). $\mathcal{I}_i^f$, $\mathcal{I}^p$, and $\mathcal{I}^e$ are cost parameters for the ESS installation fixed cost at node i, power rating, and energy reservoir, respectively. All variables with subscript i are defined for $i \in \mathcal{N}$.

$$\underline{R}_i U_i \le R_i \le \overline{R}_i U_i, \forall i \tag{1a}$$

$$\underline{C}_i U_i \le C_i \le \overline{C}_i U_i, \forall i \tag{1b}$$

$$\Delta t R_i \le \frac{C_i}{\underline{CR}}, \forall i \tag{1c}$$

$$\mathcal{IC}_e = \sum_{i\in\mathcal{N}} (\mathcal{I}_i^f U_i + \mathcal{I}^p R_i + \mathcal{I}^e C_i) \tag{1d}$$

The ESSs operational characteristics are given by (2a)–(2c) assuming their behavior being ideal in the 1st block problem. The circular capability curve defined by the maximum complex power of the given ESS is piecewise-linearized by the array of coefficient parameters α_i, β_i, κ_i associated with ESS active power, reactive power, and the power rating, respectively. $E^E_{i\phi t}$ represents state-of-energy (SoE) of ESS installed at node i for time t of scenario ϕ, which is governed by the SOE lower/upper bound ($\overline{E}/\underline{E}$), as shown in (2c). The SoE changes as the ESS (dis)charges by $p^E_{i\phi t}$ at each time interval t of scenario ϕ, as described in (2b). As given in (2d), the ESS is operated such that the final SoE is within $\pm\nu\%$ of margin from the given initial SoE to make the continuous operation over consecutive days feasible.

$$\alpha_i p^E_{i\phi t} + \beta_i q^E_{i\phi t} \leqslant \kappa_i R_i, \ \forall i, \forall \phi, \forall t \tag{2a}$$

$$E^E_{i\phi(t+1)} = E^E_{i\phi t} + \Delta t \cdot p^E_{i\phi t}, \ \forall i, \forall \phi, \forall t \tag{2b}$$

$$\underline{E} C_i \le E^E_{i\phi t} \le \overline{E} C_i, \ \forall i, \forall \phi, \forall t \tag{2c}$$

$$E^E_{i\phi(1)} - \nu C_i \le E^E_{i\phi(T+1)} \le E^E_{i\phi(1)} + \nu C_i, \ \forall i, \forall \phi, \forall t \tag{2d}$$

(2a)–(2d) are collectively referred to as $\Xi(\eta) \ge 0$ where $\eta := \{p^E, q^E, E^E, R, C\}$ is the set of variables. The notation without subscript corresponds to the vectors of variables and parameters for all nodes, all timesteps, and all scenarios.

3.1.2. Modeling the ADN operation

As the ESS investment and the optimal dispatchability level are the planning objectives of the 1st block problem, we assume that the new node n' may be connected to any existing node among the candidate nodes.

The set of all nodes including the new node is denoted by $\mathcal{O}$ (*i.e.*, $\mathcal{O} = \mathcal{N} \cup \{n'\}$). The line connected upstream to node i is also indexed by i. $p_{i\phi t}$ and $q_{i\phi t}$ are active and reactive prosumption for node $i \in \mathcal{O}$, scenario $\phi \in \Phi_{dy}$, and time $t \in \mathcal{T}$. $\tilde{p}_{itdy}$ is the active prosumption prediction for node $i \in \mathcal{O}$, time $t \in \mathcal{T}$, day $d \in \mathcal{D}$, and year $y \in \mathcal{Y}$. The prosumption scenarios are generated based on the assumption that the prosumption follows a normal distribution.[3] Therefore, the prosumption prediction at node i ($\tilde{p}_{itdy}$) is given as the average prosumption over the scenario set. The dispatch plan (DP_{tdy}) is equivalent to the aggregated prosumption prediction, as shown in (3a). Eq. (3b) states that the gap between the aggregated prosumption scenario and the dispatch plan is reduced by the sum of ESS active power. The resulting leftover prediction error at node i, scenario ϕ, and time t is indicated by $\epsilon_{i\phi t}$. The algebraic sum of errors over all nodes is equivalent to the observed dispatch error at GCP, which we aim to minimize (see (4a).)

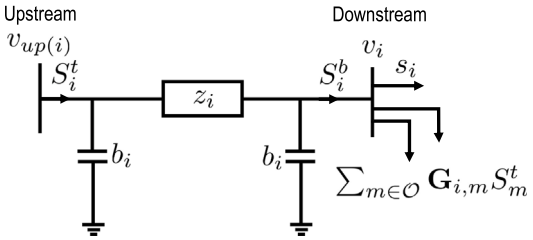


Fig. 2. Illustration of the adopted nomenclature with respect to the classic two-port II model of a transmission line adopted from [28].

The power flow equations are derived by applying the Kirchhoff's law to the two-port II branch model shown in Fig. 2. The node connected upstream of node i is denoted by $up(i)$. Take any connection among the candidate connections to the new node n' as the new line to node n'. Then, the network with the new connected node becomes a radial system where only single line is connected upstream to each node. Therefore, node i and line connected upstream of it are both denoted by index $i \in \mathcal{O}$. $\mathbf{G}$ is the network adjacency matrix, where $\mathbf{G}_{ki}$ is defined for $k, i \in \mathcal{O}$ and $\mathbf{G}_{ki} = 1$ if $k = up(i)$, or $\mathbf{G}_{ki} = 0$ otherwise. $S^t_{i\phi t} = P^t_{i\phi t} + jQ^t_{i\phi t}/S^b_{i\phi t} = P^b_{i\phi t} + jQ^b_{i\phi t}$ are the complex power injected from node $up(i)$ to line i /from line i to node i, respectively. z_i is longitudinal impedance of line i, while $\bar{z}$ represents the complex conjugate of z. b_i is the shunt susceptance of line i. $\Re(.)$ represents the real parts of a complex number. The active power balance equations at the upper and lower sides of line i is given by (3c), while the reactive power balance equations at the upper and lower sides of line i is given by (3d) and (3e). Squared nodal voltages of upperstream/lowerstream nodes ($v_{up(i)\phi t}/v_{i\phi t}$) of line i are calculated based on (3f), while being constrained by squared voltage upper limit (v^M) and lower limit (v^m) (see (3g).) All the variables with index i are defined for $i \in \mathcal{O}$.

$$DP_{tdy} = \sum_{i\in\mathcal{O}} \tilde{p}_{itdy}, \forall t, \forall d, \forall y \tag{3a}$$

$$DP_{tdy} - \sum_{i\in\mathcal{O}} p_{i\phi t} = \sum_{i\in\mathcal{O}} (\epsilon_{i\phi t} + p^E_{i\phi t}), \forall \phi, \forall t, \forall d, \forall y \tag{3b}$$

$$P^t_{i\phi t} = P^b_{i\phi t} = p_{i\phi t} + p^E_{i\phi t} + \sum_{m\in\mathcal{O}} \mathbf{G}_{im} P^t_{i\phi t}, \forall i, \forall \phi, \forall t \tag{3c}$$

$$Q^t_{i\phi t} = q_{i\phi t} + q^E_{i\phi t} + \sum_{m\in\mathcal{O}} \mathbf{G}_{im} Q^t_{i\phi t} - (v_{up(i)\phi t} + v_{i\phi t}) b_i, \forall i, \forall \phi, \forall t \tag{3d}$$

$$Q^b_{i\phi t} = q_{i\phi t} + q^E_{i\phi t} + \sum_{m\in\mathcal{O}} \mathbf{G}_{im} Q^t_{i\phi t}, \forall i, \forall \phi, \forall t \tag{3e}$$

$$v_{i\phi t} = v_{up(i)\phi t} - 2\Re\left(\bar{z}_i (S^t_{i\phi t} + j v_{up(i)\phi t} b_i)\right), \forall i, \forall \phi, \forall t \tag{3f}$$

$$v^m \le v_{i\phi t} \le v^M, \forall i, \forall \phi, \forall t \tag{3g}$$

The objective is to minimize the ESS investment cost ($\mathcal{IC}_e$) and the total penalty cost of the dispatch error over the planning horizon. ω_D is the weight coefficient for the imbalance penalty while N_{dy} stands

[3] The user can, however, generate these scenarios according to other parametric or non-parametric prosumption models that maybe available.

for the number of days in year y classified in the group of typical day-type d. r_d is the discount rate. Ω_1 and Ω_2 represent the set of control variables in the first and second stage decision process, respectively.

$$\min_{\Omega_1,\Omega_2} \mathcal{IC}_e + \sum_{y\in\mathcal{Y}}\sum_{d\in D}\sum_{t\in\mathcal{T}}\sum_{\phi\in\Phi_{dy}}\sum_{i\in\mathcal{O}} \frac{1}{(1+r_d)^y} N_{dy} w_D \lambda_\phi |\epsilon_{i\phi t}| \tag{4a}$$

subject to: (1), (3),

$$\Xi_{dy}(\eta) \geq 0, \forall d, \forall y \tag{4b}$$

By solving the above optimization problem, the optimal ESS allocation is determined, along with the optimal dispatchability level which is quantified by "leftover dispatch error rate" (LDER). LDER ($\theta_{\phi t}$) is defined for each scenario ϕ and time-step t and computed as the ratio of the optimal dispatch error and original aggregated prosumption prediction error.

$$\theta_{\phi t} = \frac{|\sum_{i\in\mathcal{O}} \epsilon^*_{i\phi t}|}{|\sum_{i\in\mathcal{O}} \Delta p_{i\phi t}|}, \forall t, \forall \phi \tag{5}$$

Then, the optimally defined LDER is incorporated into the 2nd block problem as a parameter for defining dispatchability level constraints.

3.2. 2nd block problem

The objective of the 2nd block problem is to determine the optimal connection between the existing network and node n', line reinforcement, and ESS allocation to minimize the grid losses and load curtailment while satisfying LDER constraints. The investment cost consists of the construction cost of the new line, the reinforced lines, and the ESSs allocation cost (6). The new line investment and the line reinforcement cost consist of two parts: fixed cost, which is invariant with the ampacity of the line and accounts for the construction, labor, etc., and the line conductor cost, which varies with line ampacity. $\mathcal{I}^c_k$ denotes the fixed cost parameter of the new line investment for line candidate $k \in \mathcal{K}(n')$, whereas $\mathcal{I}^r_l$ is the fixed cost parameter of the line reinforcement for line l. ρ_l is the length of line l. The conductor cost is linearly dependent on the line length. Based on the line cost data from [31], we fit a quadratic curve to the conductor cost as a function of line ampacity, where $\delta_2, \delta_1, \delta_0$ are the coefficients for the squared, linear and constant terms of the quadratic curve.

$$\begin{aligned}\mathcal{IC} = \ & \mathcal{IC}_e + \sum_{k\in\mathcal{K}(n')} \mathcal{I}^c_k X'_k + \sum_{l\in\mathcal{L}} \mathcal{I}^r_l X_l \\ & + \sum_{k\in\mathcal{K}(n')} \rho_k(\delta_2 (A'_k)^2 + \delta_1 A'_k + \delta_0 X'_k) \\ & + \sum_{l\in\mathcal{L}} \rho_l(\delta_2 (A_l)^2 + \delta_1 A_l + \delta_0 X_l)\end{aligned} \tag{6}$$

The system state during the operation horizon with each investment decision set is evaluated by solving the daily convexified AC-OPF problem, named the MAR-OPF problem. Therefore, the 2nd block problem is formulated as a MISOCP problem. Regarding the investment decisions, we have binary (U, X, X') and continuous investment decisions (R, C, A, A') separately in the 1st and the 2nd sub-stage problems, making them MISOCP problem and SOCP problem, respectively. The structure of the 2nd block problem is illustrated by the diagram shown in Fig. 3. We apply the Benders decomposition technique to solve both sub-stage problems.

In the 1st sub-stage, the Benders master problem determines the route connecting the existing network and the new node, the line for reinforcement, and the site of the ESSs. Then, each subproblem, which solves a daily OPF modeled by the AR-OPF model, determines the ESSs capacity and the line ampacity of the new and the reinforced lines for each day-type and year to minimize the unserved load and comply with the dispatchability constraint. The change of network topology associated with the choice of the new line is modeled by employing the big-M method [32], as it enables the selective activation of power

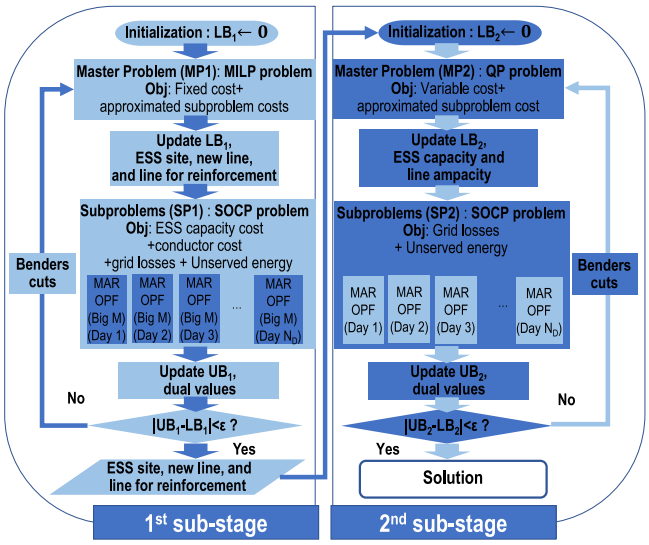


Fig. 3. Full structure of the 2nd block problem.

flow equations and the network constraints depending on the choice of the new line. The unserved load takes value to ensure the feasibility of the subproblem regardless of the investment. When the convergence of the Benders decomposition is reached, the binary solutions regarding the site of ESSs and lines investment are determined. The 1st sub-stage problem of the 2nd block (along with the 1st block) is solved repetitively till all the connections to the new nodes are decided. Then, the siting solutions are passed to the 2nd sub-stage problem. In the 2nd sub-stage problem, the master problem considers the ESS capacity and the line ampacity for the nodes and lines to be invested. The subproblem evaluates the fitness of the determined allocations in terms of the same operational requirements and objectives as the 1st sub-stage subproblem.

3.2.1. Modified augmented relaxed optimal power flow

3.2.1.1. AR-OPF model. In this section, the AR-OPF model, including ESS assets, is described. As only a summary of power flow equations is provided, readers are referred to [28]. All variables and parameters shared between the Distflow and AR-OPF models have been previously defined in Section 3.1.2. In a radial network, line l and the node connected downstream of the line l can be both denoted by index $l \in \mathcal{L}$.[4] The node connected upstream of node l is denoted by $up(l)$. The power flow equations are derived based on the two-port II branch model shown in Fig. 4.

One essential difference between the AR-OPF model and the linear model used in the 1st block problem (see Section 3.1.2) is the modeling of current and line losses. Let f_l be the square of the longitudinal current through line l, which produces grid losses through the line longitudinal impedance $z_l = r_l + jx_l$. I^t_l and I^b_l represent the square of the line current injected from the sending and the receiving end of line l, respectively. Moreover, the AR-OPF model employs auxiliary state variables that correspond to upper-bound/lower-bound line power flows at the sending end ($\bar{S}^t_l = \bar{P}^t_l + j\bar{Q}^t_l$ / $\hat{S}^t_l = \hat{P}^t_l + j\hat{Q}^t_l$) and the receiving end ($\bar{S}^b_l = \bar{P}^b_l + j\bar{Q}^b_l$ / $\hat{S}^b_l = \hat{P}^b_l + j\hat{Q}^b_l$), upper-bound squared longitudinal current ($\bar{f}_l$), upper-bound squared line current from the sending end

[4] In a radial network, the line can be also denoted following the index of node which is connected downstream of it (*i.e.*, both the line and the node are denoted by index $i \in \mathcal{N}$) as described in Section 3.1.2 and illustrated in Fig. 2. However, the notation according to line index is preferred in this section to better describe the consideration of different line candidates in the AR-OPF model.

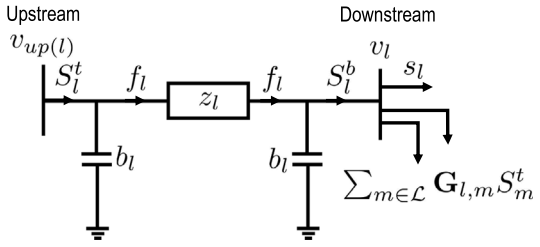


Fig. 4. Illustration of the adopted nomenclature with respect to the classic two-port II model of a transmission line adopted from [28].

and the receiving end of line l ($\bar{I}_l^t$ and $\bar{I}_l^b$) and, upper-bound squared nodal voltage ($\bar{v}_{up(l)}$, $\bar{v}_l$). The power flow equations and grid constraints employing these auxiliary variables impose conservative bounds on voltage limits and line ampacities for the nodal voltages and line currents.

$$s'_l = p'_l + jq'_l = (p_l + up_l^+ - up_l^-) + j(q_l + uq_l^+ - uq_l^-),\ \forall l \tag{7a}$$

$$S_l^t = s'_l + s_l^E + \sum_{m\in\mathcal{L}} \mathbf{G}_{lm} S_l^t + z_l f_l - j(v_{up(l)} + v_l) b_l, \forall l \tag{7b}$$

$$S_l^b = s'_l + s_l^E + \sum_{m\in\mathcal{L}} \mathbf{G}_{lm} S_l^t, \forall l \tag{7c}$$

$$v_l = v_{up(l)} - 2\Re\left(\bar{z}_l\left(S_l^t + jv_{up(l)} b_l\right)\right) + |z_l|^2 f_l, \forall l \tag{7d}$$

$$f_l v_{up(l)} \geq |S_l^t + jv_{up(l)} b_l|^2, \forall l \tag{7e}$$

Eqs. (7b), (7c), and (7d) represent the power balance at the sending end and the receiving end of line l (in the presence of ESSs) and the nodal voltage equation, respectively. Note that the complex prosumption of each node is indicated by (7a) by introducing variables related to active/reactive unserved load on both positive and negative sides (up^+/uq^+ and up^-/uq^-). We aim to reduce the values of these variables to ensure the feasibility of the ADN operation. As shown in (7e), the equation defining the squared longitudinal line current is relaxed to a SOC constraint.

$$\bar{S}_l^t = s'_l + s_l^E + \sum_{m\in\mathcal{L}} \mathbf{G}_{lm} \bar{S}_l^t + z_l f_l - j(v_{up(l)} + v_l) b_l, \forall l \tag{7f}$$

$$\bar{S}_l^b = s'_l + s_l^E + \sum_{m\in\mathcal{L}} \mathbf{G}_{lm} \bar{S}_l^t, \forall l \tag{7g}$$

$$\hat{S}_l^t = s'_l + s_l^E + \sum_{m\in\mathcal{L}} \mathbf{G}_{lm} \hat{S}_l^t - j(\bar{v}_{up(l)} + \bar{v}_l) b_l, \forall l \tag{7h}$$

$$\hat{S}_l^b = s'_l + s_l^E + \sum_{m\in\mathcal{L}} \mathbf{G}_{lm} \hat{S}_l^t, \forall l \tag{7i}$$

$$\bar{v}_l = \bar{v}_{up(l)} - 2\Re\left(\bar{z}_l(\hat{S}_l^t + j\bar{v}_{up(l)} b_l)\right), \forall l \tag{7j}$$

$$v^m \leq v_l,\ \ \bar{v}_l \leq v^M, \forall l \tag{7k}$$

$$\bar{P}_l^t \leq P_l^M,\qquad \bar{Q}_l^t \leq Q_l^M, \forall l \tag{7l}$$

$$\bar{f}_l v_{up(l)} \geq |\max\left\{|\hat{P}_l^t|, |\bar{P}_l^t|\right\}|^2 + |\max\left\{|\hat{Q}_l^t + j\bar{v}_{up(l)} b_l|, |\bar{Q}_l^t + jv_{up(l)} b_l|\right\}|^2, \forall l \tag{7m}$$

$$\bar{f}_l v_l \geq |\max\left\{|\hat{P}_l^b|, |\bar{P}_l^b|\right\}|^2 + |\max\left\{|\hat{Q}_l^b - j\bar{v}_l b_l|, |\bar{Q}_l^b - jv_l b_l|\right\}|^2, \forall l \tag{7n}$$

$$I_l^t v_{up(l)} \geq |\max\left\{|\hat{P}_l^t|, |\bar{P}_l^t|\right\}|^2 + |\max\left\{|\hat{Q}_l^t|, |\bar{Q}_l^t|\right\}|^2, \forall l \tag{7o}$$

$$I_l^b v_l \geq |\max\left\{|\hat{P}_l^b|, |\bar{P}_l^b|\right\}|^2 + |\max\left\{|\hat{Q}_l^b|, |\bar{Q}_l^b|\right\}|^2, \forall l \tag{7p}$$

$$I_l^t \leq (I_l^M)^2,\ I_l^b \leq (I_l^M)^2, \forall l \tag{7q}$$

The same set of power balance equations are given for the upper-/lower-bound power flow variables ((7f),(7g)/(7h),(7i)). The nodal voltage equation using the upper-bound squared voltage variable is (7j). The squared voltage and upper-bound squared voltage variables are bounded by squared voltage lower limit (v^m) and squared voltage upper limit (v^M), respectively, as in (7k). The upper-bound active/reactive power flow variables are respectively bounded by the active/reactive power flow limits defined for line l as given by (7l). As given in (7m)–(7p), the squared longitudinal current and the squared line current of sending/receiving end are defined by SOC constraints by employing the maximum value among upper-bound and lower-bound power flow variables. The line ampacity limits given by (7q), where I_l^M is the line ampacity of line l.

3.2.1.2. Definition of the dispatch error. The active prosumption (including the unserved load) at node l, and the longitudinal line losses of line l are expressed by their predicted values ($\tilde{p'}_{lt}$ and $r_l \tilde{f}_{lt}$) and the deviation from each prediction point ($\Delta p'_{l\phi t}$, $r_l \Delta f_{l\phi t}$) as shown in (8a) and (8b). The aggregated deviation of both prosumption and the losses over all lines (left-hand side of (8c)) is compensated by the ESSs active power. The residual dispatch error observed at the GCP is denoted as $\sum_l \epsilon_{l\phi t}$, and it should be smaller than, or equal to, the optimal dispatch error defined by multiplying LDER by the aggregated prosumption deviation (see (8e)). Two slack variables ($\gamma_{\phi t}^s$ and $\zeta_{\phi t}^s$) are introduced as shown in (8c) and (8d) to approximate the grid losses deviation accurately through an iterative algorithm proposed in [20]. The two values are updated at each iteration of solving the modified AR-OPF until $\gamma_{\phi t}^s$ becomes smaller than predefined tolerance after s iterations.[5]

$$p'_{l\phi t} = \tilde{p'}_{lt} - \Delta p'_{l\phi t}, \forall l, \forall \phi, \forall t \tag{8a}$$

$$r_l f_{l\phi t} = r_l \tilde{f}_{lt} - r_l \Delta f_{l\phi t}, \forall l, \forall \phi, \forall t \tag{8b}$$

$$\sum_{l\in\mathcal{L}} (\Delta p'_{l\phi t} + r_l \Delta f_{l\phi t}) = \sum_{l\in\mathcal{L}} (\epsilon_{l\phi t} + p_{l\phi t}^E) + \zeta_{\phi t}^s,\ \forall \phi, \forall t \tag{8c}$$

$$\sum_{l\in\mathcal{L}} \Delta p'_{l\phi t} + (\gamma_{\phi t}^s + \zeta_{\phi t}^{*s-1}) = \sum_{l\in\mathcal{L}} (\epsilon_{l\phi t} + p_{l\phi t}^E),\ \forall \phi, \forall t \tag{8d}$$

$$|\sum_{l\in\mathcal{L}} \epsilon_{l\phi t}| \leq \theta_{\phi t} |\sum_{l\in\mathcal{L}} \Delta p_{l\phi t}|,\ \forall \phi, \forall t \tag{8e}$$

Eqs. (7) and (8) compose the MAR-OPF model. For the sake of readability, they are collectively denoted by $\Theta(\varphi) \geq 0$ where $\varphi := \{S^t, v, f, \hat{S}^t, \bar{v}, \bar{f}, \bar{S}^t, s, s^E, \epsilon, up^+, up^-, uq^+, uq^-\}$ is the set of variables. The notation without subscript corresponds to the vector of variables for all nodes/lines.

3.2.2. Use of the big-m method into the MAR-OPF model

The big-M method is employed to selectively activate the OPF constraints for the new line and node only when a new line is chosen among the candidates. In this regard, let arrays Π and A'' be the investment status of existing lines and the candidate routes represented by binary variables, and line ampacities, respectively. Ψ represents the connection status of existing lines and the candidate routes. The size of $\mathcal{L}$ is increased by adding the new node/line, and thus X, A in the array Π and A''.

$$\Pi = [X^\top\ X'^\top]^\top,\ A'' = [A^\top\ A'^\top]^\top,\ \Psi = [\mathbb{1}_{|\mathcal{L}|}^\top\ X'^\top]^\top \tag{9}$$

All the power flow variables of nodes and lines take values only when the corresponding nodes and lines are connected within the system. In other words, when a line from $k \in \mathcal{K}(n')$ is not connected to the new node n', the state variables of the new node and the line are fixed to zero. Otherwise, the state variables are imposed by the given operating constraints. In the following equations, all the variables with index j and i are defined for $j \in (\mathcal{L} \cup \{k \mid k \in \mathcal{K}(n')\})$ and $i \in \mathcal{O}$, respectively. P_j^M, Q_j^M, and I_j^M take a value of maximum active/reactive power flow and line current, respectively. By employing the big-M method, (10a) and (10b) replace (7l) to redefine the ranges for the upper-bound power flow variables at line j to updated value when line j is determined to be invested (both for line reinforcement and for new line installation.) Likewise, (10c) imposes the possible ranges on the values of squared longitudinal current, its upper-bound variable,

[5] Further explanation on the formulation can be found in [20].

and squared currents at sending/receiving end of line j considering the investment status of line j.

$$0 \le \bar{P}_j^t \le P_j^M \cdot (1-\Pi_j) + A_j'' \sqrt{v^M/2},\ \forall j \tag{10a}$$

$$0 \le \bar{Q}_j^t \le Q_j^M \cdot (1-\Pi_j) + A_j'' \sqrt{v^M/2},\ \forall j \tag{10b}$$

$$0 \le \left\{ f_j, \bar{f}_j, I_j^t, I_j^b \right\} \le (I_j^M)^2 \cdot (1-\Pi_j) + (A''_j)^2,\ \forall j \tag{10c}$$

The formulation of the branch power flow equations for both power flow variables and auxiliary bound variables does not change from (7b)–(7g) due to the equations above that selectively activate the variables according to the investment status of the lines. The voltage equations (7d) and (7j) are replaced by (11a) and (11b) by employing the big-M method. The voltages of the nodes connected to line j take values only when line j is connected.

$$\left| -v_j + v_{up(j)} - 2\Re(\bar{z}_j(\hat{S}_j^t + j v_{up(j)} b_j)) + |z_j|^2 f_j \right| \le M(1-\Psi_j),\ \forall j \tag{11a}$$

$$\left| -\bar{v}_j + \bar{v}_{up(j)} - 2\Re(\bar{z}_j(\hat{S}_j^t + j \bar{v}_{up(j)} b_j)) \right| \le M(1-\Psi_j),\ \forall j \tag{11b}$$

For the sake of readability, the equations regarding the MAR-OPF employing the big-M method are grouped as $\Theta^B(\varphi^B) \ge 0$, where $\varphi^B := \{S^t, v, f, \hat{S}^t, \bar{v}, \bar{f}, \bar{S}^t, s, s^E, \epsilon, \Pi, A'', \Psi\}$ is the set of variables. The notation without subscript corresponds to the vector of variables for all buses/lines.

3.2.3. 1st sub-stage—Determination of the site

Algorithm 2 1st sub-stage of the 2nd block problem

Input : Existing network ($\mathcal{L}$), prosumption ($s_i = p_i + jq_i, \forall i \in \mathcal{N} \cup n'$), ESS candidate nodes, new node (n'), candidate nodes $\mathcal{K}(n')$

1: **Initialization :** $LB_1 \leftarrow 0, n \leftarrow 1$
2: Solve master problem (MP1)
3: $LB_1 \leftarrow \mathcal{MC}^{1st*}$
4: Update U^*, X'^*, X^* to feed the solution to subproblems
5: Solve parallel subproblems (SP1)
6: $UB_1 \leftarrow \sum_i I_i^c U_i^* + \sum_l (I_l^r + \rho_l \delta_0) X_l^* + \sum_k (I_k^n + \rho_k \delta_0) X_k'^* + \sum_y \sum_d SC_{dy}^{1st*}$
7: Update the dual values (τ, ξ)
8: **if** $|UB_1 - LB_1| \le \epsilon$ **then**
9: **return** U^*, X'^*, X^*
10: **else**
11: $n \leftarrow n+1$
12: Update Benders cut with dual values
13: Go to step 2
14: **end if**

The problem structure of the 1st sub-stage problem is illustrated by the left-hand side of Fig. 3 and Alg. 2. The new line that connects the new node to the existing node, the nodes for ESS, and the lines for reinforcement are determined in the master problem. The optimal ESS sizes, line ampacity for the invested lines specific to each day type, and resulting operational benefit are evaluated within each subproblem. As this stage aims to determine the best solution for the binary variables, the sizes of different ESSs and conductors determined for each day type are not the final optimal solutions. The Benders iteration terminates when the upper-bound and the lower-bound of the total cost become sufficiently close to each other. As the new line connecting to the new node is determined at the end of the Benders iteration, the adjacency matrix and the array of line parameters get modified considering the new line.

3.2.3.1. Master problem. The formulation of the master problem is given in (12). Its objective is the sum of the fixed investment cost for ESS installation, new line installation, line reinforcement, and the constant part of the conductor cost for both of the new lines and the reinforced lines. In addition, the lower approximation of the total expected subproblem cost is added. As the targeted networks are supposed to operate with a radial configuration, only a single line must be connected to the new node (see (12b).) Eq. (12c) indicates that if any new node is already added to the existing grid (*i.e.*, $n' \ge 2$), the investment statuses for the added lines are 1 to update their required line ampacity while determining a new connection.

α_{dy} stands for the lower approximation of the subproblem cost associated to the investment solution determined by the master problem. Starting from the lower-bound for the operation cost ($\underline{\alpha}$), as shown in (12d), α_{dy} gets updated by the set of Benders cut corresponding to all day-types and years accumulated in each iteration indexed by $n \in \mathcal{N}_1$, as given in (12e) (see (14) for the formulation of the cuts). The lower-bound of the total cost LB_1 is the optimal objective value of the master problem (i.e., $LB_1 = \mathcal{MC}^{1st*}$).

$$\min_{U,\Pi,\alpha} :\ \mathcal{MC}^{1st} = \sum_{i\in\mathcal{N}} I_i^f U_i + \sum_{k\in\mathcal{K}(n')} (I_k^c + \rho_k \delta_0) X'_k + \sum_{l\in\mathcal{L}} (I_l^r + \rho_l \delta_0) X_l + \sum_{\in\mathcal{Y}} \sum_{d\in D} \alpha_{dy} \tag{12a}$$

subject to: $$\sum_{k\in\mathcal{K}(n')} X'_k = 1 \tag{12b}$$

$$X_{(L-n'+1):L} = 1, \forall n' \ge 2 \tag{12c}$$

$$\alpha_{dy} \ge \underline{\alpha}, \forall d, \forall y \tag{12d}$$

$$\alpha_{dy} \ge \Gamma 1_{dy}^{(n)}, \forall d, \forall y, \forall n, \tag{12e}$$

3.2.3.2. Subproblem. In the subproblem associated with day-type d and year y, a daily MAR-OPF model employing the big-M method sizes the ESS capacity and line ampacity based on the investment constraints, and evaluates its operational advantages on the system conditions. The line ampacities are optimally sized not only for the new line candidates (connecting the grid to the new node n') but also for the lines added by the previous rounds of 1st sub-stage optimization problem and for the existing to-be-reinforced lines. The problem is modeled based on the MAR-OPF with the big-M method as discussed in Section 3.2.2. Moreover, the equations modeling the investment constraints on the conductor size range and the ESSs capacity range are included. Conductor's possible ampacity range is modeled by (13b), where A^m/A^M represent minimum/maximum possible line ampacity, respectively.

By incorporating the constraints related to the line ampacity range and the ESS investment and operation-related constraints (see (1) and (13d)) into the OPF constraints explained in Section 3.2.2, the subproblem of the 1st sub-stage is defined as follows. The objective function of the subproblem is the combination of the capacity cost of the ESSs energy reservoir, power rating, the conductor cost for new lines and the to-be-reinforced lines, as well as the operational cost, given by (13a). The operational cost consists of grid losses and unserved energy costs. Eqs. (13e) and (13f) show that the ESS location and the lines to be invested are fixed to the optimal solution values of the master problem. τ_{idy} and χ_{jdy} are the duals of constraints related to the fixed ESS location and the line for investment.

$$\begin{aligned}\min_{\forall\varphi^B,\eta} :\ SC_{dy}^{1st} = {} & \frac{N_d}{(1+r_d)^y} \sum_{t\in\mathcal{T}} \sum_{\phi\in\Phi_{dy}} \lambda_\phi (w_l \sum_j r_j f_{j\phi t} \\ & + w_u \sum_j (up_{j\phi t}^+ + up_{j\phi t}^- + uq_{j\phi t}^+ + uq_{j\phi t}^-)) \\ & + \frac{N_d}{365*Y} \sum_i (I^e C_i + I^p R_i) \\ & + \frac{N_d}{365*Y} \sum_j \rho_j (\delta_2 (A''_j)^2 + \delta_1 A''_j)\end{aligned} \tag{13a}$$

subject to: (1), (8),

$$A^m \Pi_j \le A''_j \le A^M \Pi_j \tag{13b}$$

$$\Theta^B(\varphi_{\phi t}^B) \ge 0, \forall t, \forall \phi \tag{13c}$$

$$\Xi(\eta_{\phi t}) \geq 0, \forall t, \forall \phi \tag{13d}$$

$$U_{idy} = U_i^* \ : \ \tau_{idy}, \ \forall i, \forall d, \forall y \tag{13e}$$

$$\Pi_{jdy} = \Pi_j^* \ : \ \chi_{jdy}, \ \forall j, \forall d, \forall y \tag{13f}$$

w_l and w_u are the weight coefficients associated with grid losses and unserved load, respectively.

$$\Gamma 1_{dy}^{(n)} = \Big[\ SC_{dy}^{1st*} - \sum_i (\tau_{idy}(U_i - U_i^*)) - \sum_j (\chi_{jdy}(\Pi_j - \Pi_j^*)) \ \Big], \forall d, \forall y, \forall n \tag{14}$$

As shown in (14), the bound used for the Benders cut is obtained for each day-type and year by employing the dual values linked to the location of ESS and line investment. The upper-bound (UB_1) of the planning cost is the sum of the fixed cost from the master problem and the total subproblem cost.

3.2.4. 2nd sub-stage—Determination of ESSs and lines sizes

Algorithm 3 2nd sub-stage of the 2nd block problem

Input : Expanded network ($\mathcal{L}$), prosumption ($s = p + jq$), ESS location (U^*), Line for investment (X'^*, X^*)

1: **Initialization :** $LB_2 \leftarrow 0, n \leftarrow 1$
2: Solve master problem (MP2)
3: $LB_2 \leftarrow \mathcal{MC}^{2nd*}$
4: Update R^*, C^*, A^* to feed the solution to subproblems
5: Solve parallel subproblems (SP2)
6: $UB_2 \leftarrow \sum_l (\mathcal{I}^c C_l^* + \mathcal{I}^r R_l^*) + \sum_l \rho_l (\delta_2 (\Delta A_l^*)^2 + (2\delta_2 I_l^M + \delta_1)\Delta A_l^*) + \sum_y \sum_d SC_{dy}^{2nd*}$
7: Update the dual values (μ, ϑ, ι)
8: **if** $|UB_2 - LB_2| \leq \epsilon$ **then**
9: **return** R^*, C^*, A^*
10: **else**
11: $n \leftarrow n + 1$
12: Update Benders cut with dual values
13: Go to step 2
14: **end if**

As all the new nodes are connected after the N' runs of the 1st sub-stage problem, the lines for reinforcement and the nodes for hosting ESSs are also determined. The binary solutions are passed to the 2nd sub-stage, where the size decisions take place. In the 2nd sub-stage, the set and the number of existing lines (*i.e.*, $\mathcal{L}$ and L, respectively) are equivalent to the set and the number of lines in the network with the new nodes connected after solving the final round of the 1st sub-stage problem. Therefore, lines (both existing and new lines determined to invest in the 1st sub-stage problems) are indicated with index $l \in \mathcal{L}$. Moreover, given that the network of interest is radial, the node (both existing and new nodes connected in the 1st sub-stage problems) connected downstream of line l is also indexed as l. The 2nd sub-stage problem is solved by employing the Benders decomposition as described in the right side of Fig. 3 and Alg. 3. It should be noted that the formulation of 2nd sub-stage problem is identical to the 2nd sub-stage problem developed in [21].

3.2.4.1. Master problem. The ESSs energy reservoirs and power ratings, ampacities for the lines to be reinforced and the new line are determined in this problem, while the ESSs locations and the lines for investment are all fixed to the optimal binary solution given by the 1st sub-stage. Among the total investment cost (see (6)), we take only the variable cost regarding the ESSs sizes and lines ampacities. Particularly, the conductor cost is split into the cost related to the original line ampacity and the required change ($A_l = I_l^M + \Delta A_l$). Note that the original ampacities for the newly connected lines are 0. In this way, the variable part of the investment cost is included in the objective function of the master problem, as shown in (15a). The lower approximation of the subproblem cost is represented by β_{dy}. The approximated value starts from the lower-bound subproblem cost value $\underline{\beta}_{dy}$, and it is updated at each Benders iteration by accumulated Benders cut over the iterations ($\forall n \in \mathcal{N}_2$) (15c). The problem takes into account the possible range of changes of conductor's sizes (15b). Moreover, the investment conditions for the ESSs are also included as constraints.

$$\min_{\forall C,R,\Delta A,\beta} \ : \ \mathcal{MC}^{2nd} = \sum_{l\in\mathcal{L}} (\mathcal{I}^e C_l + \mathcal{I}^p R_l) + \sum_{l\in\mathcal{L}} \rho_l (\delta_2 (\Delta A_l)^2 + (2\delta_2 I_l^M + \delta_1)\Delta A_l) + \sum_y \sum_d \beta_{dy} \tag{15a}$$

subject to: (1a) – (1c),

$$0 \leq \Delta A_l \leq \Delta A^M \tag{15b}$$

$$\beta_{dy} \geq \underline{\beta}, \ \beta_{dy} \geq \Gamma 2_{dy}^{(n)}, \forall d, \forall y, \forall n \tag{15c}$$

3.2.4.2. Subproblem. The subproblem of the 2nd sub-stage evaluates the system operational condition for each investment choice made in the master problem. The operational objective is identical to that of the subproblem in the 1st sub-stage (16a). As the grid topology is already fixed, the modified AR-OPF model described in Section 3.2.1 is employed (16b). The ampacity limits take into account the changed ampacity values (16d). Moreover, the parameters in the admittance matrix, such as line resistance, reactance, and susceptance, are adjusted according to the change of line ampacity to enhance the accuracy of the operation model.[6] The power flow through ESS, which was considered as an ideal battery in the subproblem of the 1st sub-stage problem, is modeled in a realistic way based on the internal-resistance model introduced in [33]. The set of virtual nodes added to the network for modeling the ESS power injection is defined as $\mathcal{N}_E$. Note that the ESS operational constraints (16c) are only imposed on the ESS power injections from/to virtual nodes $\forall i \in \mathcal{N}_E$. The ESSs power ratings, energy reservoirs, and the changes of lines ampacities are fixed by the master problem solution (16e). The corresponding dual variables are μ_{ldy}, ϑ_{ldy}, ι_{ldy}. The Benders cut at nth iteration is constructed in the same way of the 1st sub-stage, using the subproblem objective cost and dual values. The upper-bound of the 2nd sub-stage problem, UB_2, is calculated by summing the variable investment cost and subproblem costs.

$$\min_{\forall \varphi,\eta} \ : \ SC_{dy}^{2nd} = \frac{N_d}{(1+r_d)^y} \sum_{t\in\mathcal{T}} \sum_{\phi\in\Phi_{dy}} \sum_{l\in\mathcal{L}} \lambda_\phi (w_l r_l f_{l\phi t} + w_u \sum_{l\in\mathcal{L}} (up_{l\phi t}^+ + up_{l\phi t}^- + uq_{l\phi t}^+ + uq_{l\phi t}^-)) \tag{16a}$$

subject to: (8), (13d),

$$\Theta(\varphi_{\phi t}) \geq 0, \ \forall t, \forall \phi \tag{16b}$$

$$\Xi(\eta_{\phi t}) \geq 0, \forall t, \forall \phi \tag{16c}$$

$$0 \leq \left\{ I_l^t, I_l^b \right\} \leq (I_l^M)^2 + 2 I_l^M \Delta A_l + (\Delta A_l)^2, \ \forall l \tag{16d}$$

$$R_{ldy} = R_l^* \ : \ \mu_{ldy}, C_{ldy} = C_l^* \ : \ \vartheta_{ldy}, \quad \Delta A_{ldy} = \Delta A_l^* \ : \ \iota_{ldy}, \forall l, \forall d, \forall y \tag{16e}$$

$$\Gamma 2_{dy}^{(n)} = \Big[\ SC_{dy}^{2nd*} - \sum_{i\in\mathcal{N}} (\mu_{ldy}(R_l - R_l^*) - \vartheta_{ldy}(C_l - C_l^*)) + \sum_{l\in\mathcal{L}} (\iota_{ldy}(\Delta A_l - \Delta A_l^*)) \ \Big], \forall d, \forall y, \forall n \tag{17}$$

[6] Based on the line data available in industry, the reactance and susceptance values are fitted to a linear model. In contrast, the line resistance showed a better fit to a hyperbolic model.

Table 1
ESS parameter and candidate nodes.

Max. power rating	10 MVA	Max. energy reservoir	10 MWh
Cost for power rating	$200/kVA	Cost for energy reservoir	$300/kWh
Capital investment cost		$0.1 M	
ESS candidate nodes		4, 16, 27, 41, 45	

Table 2
Parameters related to line investment.

New line installation		Line reinforcement	
Fixed cost (OL)	$0.2 M/km	Fixed cost (OL)	$0.12 M/km
Fixed cost (UC)	$0.8 M/km	Fixed cost (UC)	$0.4 8 M/km
Coefficients for conductor cost function		δ_2 : 0.011, δ_1 : 0.065, δ_0 : 95.22	

Table 3
Investment result.

ESS	Node #	Energy reservoir (kWh)	Power rating (kVA)
	16	4657	1196
	27	5733	1431
Line	Investment type	Line #	Ampacity (A)
	Reinforcement	3–10	134
	New lines	10–56	51
		35–57	38
		56–58	34
		43–59	35
		16–60	34
		14–61	51
		61–62	43
		28–63	39

Table 4
Cost and operational benefits.

Investment cost ($ Million)	6.86
Dispatch error penalty cost ($ Million)	3.49
Grid losses (GWh)	2.16
EENS (kWh)	14.99
Total energy served (GWh)	241.2

4. Simulations

4.1. The simulation setup

The proposed DNEP is tested on a real Swiss 21 kV distribution network with 55 nodes (see Fig. 5(a)). Initially, 2.7 MWp of PV generation and 805 kVA of hydropower generation are installed. The initial network topology and parameters can be found in [34]. The planning horizon is set at 10 years, and the annual load consumption growth is set at 7% (*i.e.*, $r_i = 0.07$). The ESS cost parameters and possible ranges of energy reservoirs and power ratings are shown in Table 1. The candidate nodes for ESS installation are set according to the indications of the operator of this specific grid. The new nodes are given on the grid topology with their indices indicating their priorities for integration. Both load consumption and 1MWp PV system are located on each new node. Each new node can be connected to the ones of the existing grid among 4 candidates given by the user. In Table 2, the fixed costs associated with the line investment are given depending on the type of investment and connection. The estimated cost for constructing underground cables (UC) is set four times higher than the one for overhead lines (OL), considering the cost range reported in [35]. The coefficient values for the conductor cost function's quadratic, linear, and constant terms are given in Table 2 (see Section 3.2.) The penalty cost for the dispatch error is assumed as $897/MWh, which corresponds to the 99.9th percentile of the actual imbalance price settled from 2018 to 2019 in the Swiss energy market [36]. The chosen price coefficient is notably higher than a typical price settled in the energy markets to secure a sufficient level of ADN dispatchability.

The operating scenarios for the planning problem are generated to model the prediction uncertainty following the scenario generation algorithm described in Alg. 4 in Appendix. 1000 prosumption scenarios are generated with equal probabilities based on the assumption that both the load and PV irradiation follow the normal distribution [8]. Then, the number of operating scenarios for each day type is reduced to 10 by the K-medoids clustering technique [37].

The simulations are carried out on a laptop PC with an Intel® Core(TM) i7-8750H CPU at 2.2 GHz and a physical system memory of 32 GB.

4.2. Grid expansion assessment on a real Swiss power distribution grid

This subsection shows the solution of the planning problem using the determined set of operating scenarios to achieve the optimal level of dispatchability of the targeted ADNs by allocating ESSs. Moreover, the line investment regarding new line installation and existing line reinforcement is co-optimized to minimize the grid losses and the expected energy not served (EENS) while complying with the grid constraints regarding nodal voltage and line ampacity.

Fig. 5(a) illustrates the grid topology with 8 new nodes (green circle) and the candidate nodes for ESS allocation (purple circle). The new nodes are integrated into the grid in ascending order of the node index number. In Fig. 5(b), the line candidates to the new nodes are indicated (green dotted line). Fig. 5(c) shows the result in terms of the location of the new lines (green line), the line to be reinforced (thick red line), and the nodes to host ESSs (orange circle) determined by the proposed planning tool. Table 3 shows the capacities of ESSs and the invested lines' ampacities.

The connection to each new node is optimally determined based on the conductor cost (which relies on the line length and the type of line installation) and the evaluation of the system operation. For example, the optimal connecting node for Node 56 is decided to be Node 10, considering the relatively low fixed cost for overhead lines and the shorter distance to Node 56 than other candidate nodes. In this way, Node 58 is connected to Node 56, making the two new nodes connected in series to Node 10. The line ampacity between Node 56 and 58 is optimally decided to host the prosumption injection from Node 58. The ampacity of Line 10–56 is larger than Line 56–58 to host the total prosumption of Nodes 56 and 58. 5 new nodes are determined to be connected downstream of Node 10. These connections result in investment in upgrading the line between Node 10 and Node 3 because its initial line ampacity (90 A) was insufficient to host the additional PV injection from the 5 new nodes. The line is replaced to have an ampacity of 134 A. The ESS was installed at Node 27 with a 4.66 MWh energy reservoir and 1.20 MVA power rating to compensate for the increased prosumption uncertainty from the 5 new nodes. Another ESS is installed at Node 16 to mainly compensate for the uncertainty of the additional PV injection of Node 60 and the existing PV generation at Node 16 (as it hosts the biggest PV plant of 1.6 MWp).

Fig. 6(a) shows the aggregated prosumption scenarios along with the aggregated prosumption prediction, whereas Fig. 6(b) represents the dispatch result after investment on ESS and lines. Thanks to the proposed planning strategy, the dispatch result shows that the power flow at the GCP of all scenarios follows the dispatch plan with the dispatch error determined by the optimal dispatchability level of ADN. Furthermore, Fig. 7 illustrates the optimal ESSs investment and the resulting dispatch error penalty cost determined from the 1st block problem. 'EO' and 'EX' represent the case with and without the ESS allocation. The 'EO' graphs show the increase in ESSs investment cost

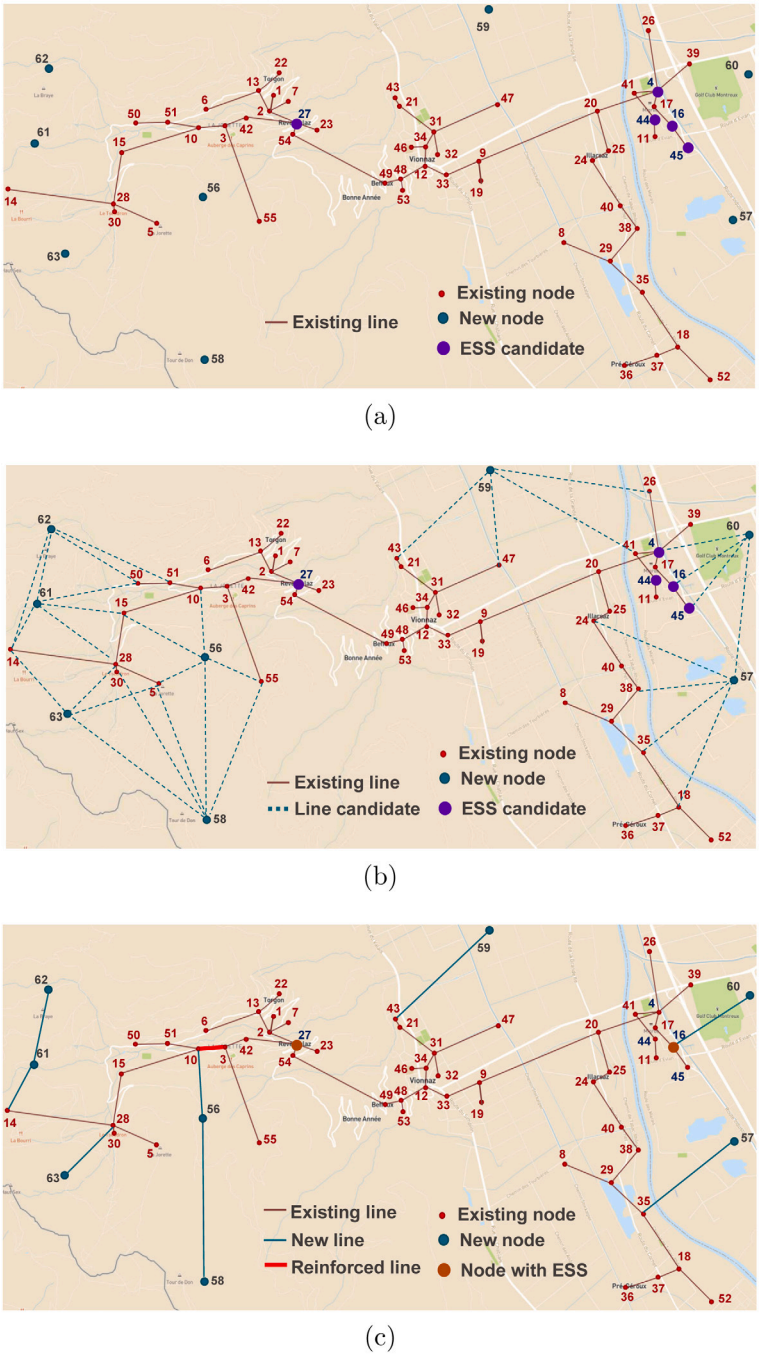


Fig. 5. Grid topology with: (a) candidate nodes for ESS allocation and new nodes, (b) with line candidates, (c) with ESSs, new lines and line reinforcement.

with the integration of new nodes hosting PV generation and load. The resulting benefit concerning the imbalance penalty cost corresponds to an 82% reduction of dispatch error compared to the case without ESS.

The total planning cost, consisting of investment and operational objectives costs, is shown in Table 4, along with total grid losses and active power served in the grid for the planning horizon (10 years). The final penalty cost for 10 years regarding the dispatch error is $3.49 Million (which corresponds to 3.89 GWh). The total investment cost is $6.86 Million, where the ESS investment cost is $3.85 Million, and the line investment cost is $3.01 Million. The EENS is only 15 kWh, which is negligible enough to conclude that the planning result successfully assures the reliability of the power supply. Table 5 shows the computation time of the 1st block problem, 1st and 2nd sub-stages, and the total time for solving the given planning problem. Note that there are 8 rounds of the 1st block problem and the 1st sub-stage of the 2nd block problem. The 2nd block problem takes more computation time than the 1st block problem due to its exact system operation evaluation based on the modified AR-OPF model (subproblem of the 2nd block, which is indicated as 'SP' in the table). Based on the reported total computation time for solving master problems (indicated as 'MP') and subproblems along with the number of Benders iterations (indicated as 'BI'), the average computation times of MP and SP (4 (day-types) × 10 (years) parallel subproblems) per each iteration are calculated as 7.78 s and 2541 s, respectively. It is observed that the computation times for determining the connection to each new node are within the same order of magnitude, implying that the running time of the proposed framework increases linearly with the size of the new nodes. Lastly, the master problem and subproblem computation time of 2nd sub-stage are

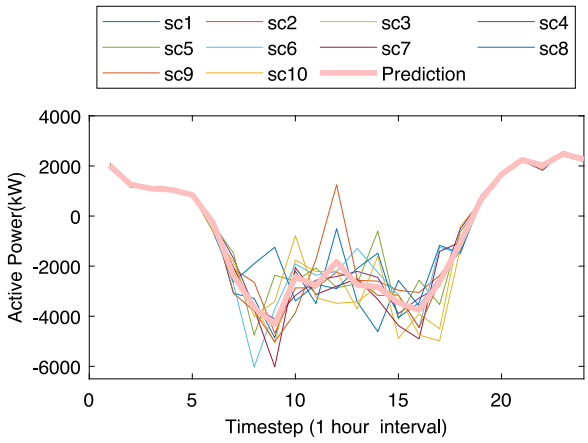


Fig. 6. Example of dispatch result for Day-type 4, year 10: (a) Aggregated prosumption scenarios and prosumption prediction, (b) Power flow through GCP of scenarios, prosumption prediction and dispatch plan (DP).

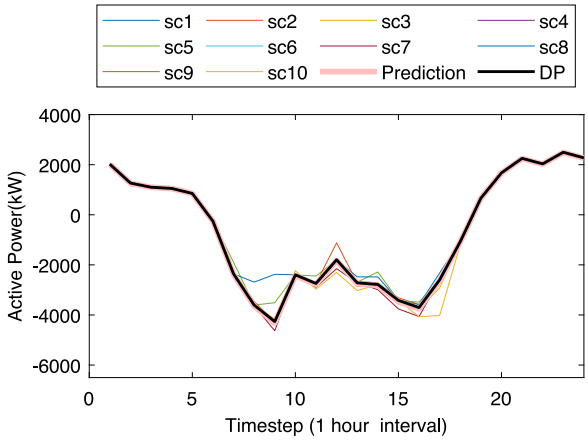


Fig. 7. The 1st block cost result.

Table 5
Computation time (seconds) and the number of Benders iterations.

Sub-stage	New Node ID	MP	SP	BI#
1st	56	8.47	2484	48
	57	8.86	2478	54
	58	6.54	2836	44
	59	9.44	2563	53
	60	8.46	2704	49
	61	7.11	2816	42
	62	9.26	2945	51
	63	6.40	2964	37
2nd		2.66	2300	25
Total		1.08E+6 (≈300 h)		

shown. Thanks to the fixed binary investment decision obtained in the 1st sub-stage, the 2nd sub-stage problem converges swiftly to the final investment solution in 25 Benders iterations.

4.3. Scalability analysis regarding varying investment candidates

In this section, we analyze the computation time for solving a planning problem considering various candidates for ESS nodes and connections to a new node from the existing grid. We tested the proposed planning framework on a 28-node benchmark distribution network [38,39]. We solved the planning problem for 16 cases accounting for the numbers of candidate nodes for ESS allocation and

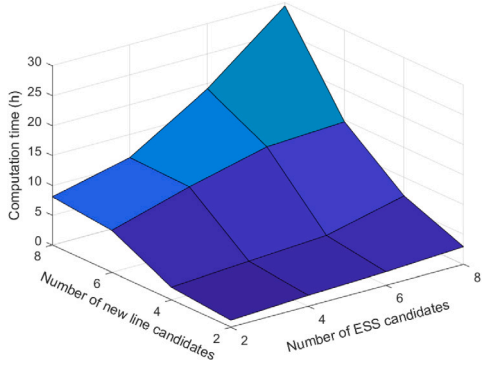


Fig. 8. Computation time with different number of candidates.

Table 6
Computation time (hours) with various number of candidates (28-node system).

New lines #	ESSs #			
	2	4	6	8
2	1.18	1.84	2.25	2.92
4	2.13	2.96	3.75	6.86
6	7.05	10.83	14.02	14.69
8	8.09	11.15	19.13	29.50

candidate connections to a new node ranging from 2 to 8 as shown in Table 6. The number of considered operating scenarios for each day type was 10, while the number of day types was 4. The local consumption and generation are considered fixed over the planning horizon. Fig. 8 illustrates the trend of computation times associated with the number of candidate nodes for ESS (indicated as 'ESSs #' in Table 6) and the number of candidates of new lines (indicated as 'New lines #' in Table 6), which ranges from 1.18 h to 29.50 h. Regarding the given system, the computation time increases more rapidly with the increase of the number of new line candidates than the number of ESS candidates.

4.4. Scalability analysis regarding varying system size

The planning methodology is also tested on distribution systems of different sizes ranging from 13 nodes to 123 nodes (13-node and 28-node systems are from [38,39], 55-node system is from [34], and 123-node IEEE benchmark system can be found in [40]). 5 candidate nodes for ESS allocation and the 4 candidate connections for a new node are considered, while the number of scenarios, day-types and the growth of prosumption is set to be the same as in the previous analysis. Table 7 reports the average computation time for the master problem and the subproblem for 1st and 2nd sub-stage problems. Note that only the computation times for the 2nd block problems are reported in this table. It is observed that the subproblems (i.e., operation problem) take most of the total computation time. Its computation time increases with the system size, suggesting that the subproblem size is determined by the number of constraints and variables varying with the system size. Solving the planning problem for the system with 28 nodes takes 1.8 times longer than the 13-node system. However, the 55-node system took a shorter computation time than the 28-node system despite the longer computation time for the subproblems thanks to the smaller number of Benders iterations. The lower average computation time to solve the master problem for the 55-node system implies a smaller number of accumulated Benders cuts in the master problem. On the other hand, the total computation time for solving the case of the 123-node system was more than 20 times higher than that of the 55-node system, significantly exceeding the proportion of the subproblem

Table 7
Computation time (seconds) of the 2nd block with respect to system size.

System size (Number of nodes)	Average 1st MP	Average 1st SP	Average 2nd MP	Average 2nd SP	Total
13	1.7	72.5	0.5	31.6	7579
28	1.3	74.8	1.3	158.9	13 488
55	0.9	98.4	0.7	172.5	13 299
123	7.2	816.7	2.1	425.0	297 000

computation time for the 123-node system compared to the case of the 55-node system. It suggests that the convergence speed of the proposed method depends not only on the size of the network but also on how the Benders cuts pin down the solution space to identify the optimal solution. This optimal solution is affected mainly by the structure of the solution space, which is subjected to the network's operation condition.

5. Conclusion

This study presents a DNEP tool for ADNs to achieve their dispatchability. The connections to the new nodes and line reinforcement are determined based on a proper evaluation of the network operation modeled employing the AR-OPF model. ESSs are employed to compensate for the prosumption uncertainty such that the realized power flow at the GCP can track a day-ahead computed dispatch plan. The proposed DNEP utilizes a sequential algorithm to add the new nodes according to the pre-defined priorities. We decompose each planning problem integrating a new node into two blocks. In the 1st block, the allocation of ESS is determined along with the corresponding LDER by implementing the linearly approximated OPF model. In the 2nd block, the modified AR-OPF model is used to check the compatibility of the allocated ESS capacity for the feasible operation of ADNs to satisfy the LDER. This block determines the optimal locations and sizes of the ESSs, the new line to connect the new node, the lines for reinforcement, and their ampacities to minimize the unserved energy and the grid losses while satisfying the LDER. We tested the proposed method on a real Swiss ADN with 8 new nodes hosting 8 MWp PV capacity. The results show that the DNEP can successfully determine the optimal level of dispatchability while securing the required hosting capacity of the ADN under increasing stochastic prosumption. Moreover, the extended computation time analysis for solving the planning problem indicates that the running time of the proposed method increases linearly with the number of new nodes, avoiding a scalability issue associated with the size of network expansion. The scalability analysis on various candidates and different system sizes is also carried out to demonstrate that the proposed planning strategy can apply to distribution networks with generic sizes considering multiple investment options. The post-analysis of the planning problem's objective values with different operating scenarios shows the trade-off between the computation time and the reliability of the planning solutions, providing helpful insight into the choice of an operating scenario set to the modeler.

Future studies may not only address the ESS allocation but also the configuration of hybrid ESS integrated to PV systems. The modeling of these hybrid PV-ESS installations is not trivial; thus its integration into the DNEP tool requires a dedicated study to be carried out by future research.

Data availability

Data will be made available on request.

Acknowledgments

This project is carried out within the frame of the Swiss Centre for Competence in Energy Research on the Future Swiss Electrical Infrastructure (SCCER-FURIES) with the financial support of the Swiss Innovation Agency (Innosuisse - SCCER program), Hitachi Energy, Switzerland, and ABB, Switzerland.

Appendix. Scenario generation

The scenarios used in this work are generated by following Alg. 4. First, we classified the historical load data at each node based on the season and weekdays/weekends, which grouped them into different day-types representing the seasonal variability of the prosumption. The mean and standard deviation values of the magnitude of prosumptions at each timestep are extracted from the consumption profiles classified in each day type. A similar process is conducted for PV generation, and the historic PV irradiation data is grouped by seasons and separated into sunny/cloudy days. With the extracted mean and standard deviation value of the load and PV irradiation, 1000 prosumption scenarios are generated with equal probabilities to model the prediction uncertainty based on the assumption that both the load and PV irradiation follow the normal distribution [8]. Then, the number of operating scenarios for each day type is reduced through the K-medoids clustering technique [37].

It should be noted that the adopted scenario generation process has some limitations in modeling accurate prosumption profiles. Firstly, using representative day types instead of complete 8760-h operation profiles may not be enough to picture the total variation of the prosumption profiles over a year. Therefore, the number of day-types should be chosen carefully, considering the trade-off between the computation burden and the accuracy of seasonal variability modeling. The criteria for categorizing day-types can largely affect the number of day-types as well as the reliability of the planning solution. Moreover, operating scenarios are modeled as discontinuous representative days. Therefore, continuous ESS operation over consecutive days is not modeled[7] Secondly, the prosumption scenarios are generated while the prosumption day-ahead prediction uncertainty is modeled as a unidimensional normal distribution. The user can, however, generate these scenarios according to other parametric or non-parametric prosumption models that may be available. Moreover, one can improve the quality of prosumption modeling by considering of temporal auto-correlation of prosumption uncertainty. Given that the ESS investment decisions are influenced by how accurately the scenarios model the prosumption uncertainty, in practice, the modeler should have a reliable forecasting

[7] Instead, the final SoE of each representative day is constrained to be 10% higher or lower than the initial SOE of each representative day to ensure the feasibility of consecutive day operation.

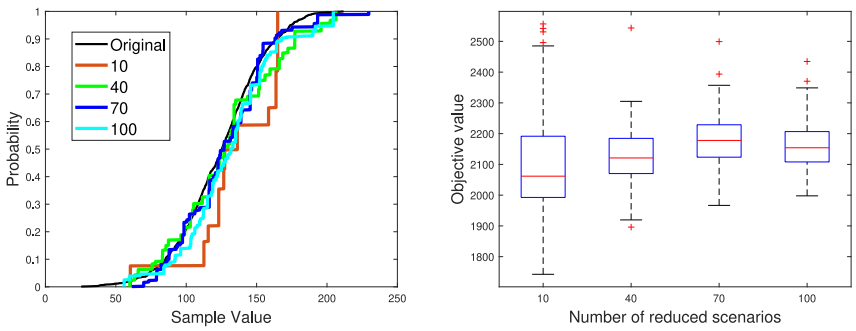


Fig. A.9. (a) CDF of initial scenario set and reduced scenario sets with different number of reduced scenarios regarding the value of a single random variable (PV irradiation value at timestep 75, day-type 1, and year 1), (b) Boxplot of objective values of planning problem with different number of prosumption scenarios.

and scenario generation tool. It is finally worth pointing out that the modeling of prosumption uncertainty is beyond the scope of this paper.

Algorithm 4 Scenario generation

1: Group the historical load data at each node based on the seasons and weekdays/weekends.
2: Calculate the mean and standard deviation value of the historical load profiles in each group
3: Group the historical solar irradiation data based on the seasons and sunny/cloudy days.
4: Calculate the mean and standard deviation values of the historical solar irradiation profiles in each group
5: Generate 1000 prosumption scenarios (Φ_{dy}) assuming that the prosumption follows the normal distribution.
6: Set the number of desired clusters (i.e., the number of reduced scenarios) for scenario reduction.
7: Run the K-medoids clustering function.
8: The members within each cluster are counted to calculate the probability set (λ_{Φ}) of the reduced scenario set.
9: Save the medoids as the reduced scenario set Φ'_{dy}
10: **return** $\Phi'_{dy}, \lambda_{\Phi_{dy}}, d \in D, y \in \mathcal{Y}$

A.1. Statistic analysis on varying number of operating scenarios

We analyzed the relationship between the number of operating scenarios and the trade-off between the computation time and the solution quality. Note that the operating scenario set considered for the planning problem is the outcome of the K-medoids clustering technique (see Alg. 4 in Section 4.1) applied to the initial scenario set. Therefore, the number of clusters (i.e., operating scenarios to be used in the planning problem) determines how the reduced scenario subset can represent the stochastic characteristic of the initial scenario set. Fig. A.9.(a) shows the cumulative distribution function (CDF) of the initial scenario set and different reduced scenario sets regarding the value of a single random variable (PV irradiation value at timestep 15, day-type 1, and year 1). It is evident that the stochastic similarity between the reduced scenario set and the initial scenario set improves with the number of operating scenarios, as shown in Fig. A.9.(a).

In this regard, it is worthwhile a-posteriori investigating the impact of the number of operating scenarios on the computation time and the optimal solution quality of the planning problem by observing the distribution of the objective values throughout the repetitive runs of planning exercises with the different number of operating scenarios. 100 runs of simulation are conducted for each case considering 10, 40, 70, and 100 reduced scenarios, respectively, to observe the average computation time, the average expected objective value, and the normalized standard deviation as shown in Table A.8. Fig. A.9.(b) indicates the median of the objective values and the 25% and 75% percentile of the objective value, along with outliers (shown as red 'x's). The number of reduced scenarios 40 is obtained from the algorithm introduced in [20], which determines the number of reduced scenarios by evaluating the statistical similarities between the initial and reduced scenarios set. The reader can find the details about the algorithm in [20]. The cases with 70 and 100 scenarios are chosen as the reference cases to compare the solution quality with the 10 and 40 scenarios cases.

Table A.8
Post analysis with respect to different number of scenarios (tested on 13-node system, 4 day-types, 24 timesteps, 2 ESS node candidates, 1 new line candidate).

Number of scenarios	Average Computation time (h)	Average objective value	Normalized standard deviation
10	0.29	2124.7	0.0810
40	1.08	2130.6	0.0454
70	4.02	2180.6	0.0413
100	6.15	2158.7	0.0375

All the metrics (i.e., the average expected objective value, the normalized standard deviation values, and the range between 25%–75% percentile) show a relatively small difference between the case with 40 scenarios and those with a larger number of scenarios in contrast to the significant gap in the average computation time between the 40 scenarios case and other cases. The result demonstrates that 40 scenarios can yield acceptable solution reliability while maintaining a reasonable computation time. Then, the solution's reliability in the case of 10 operating scenarios (used for the previous simulation exercises) is evaluated with reference to the 40 scenarios case. While the average computation time for the 10 scenarios case is nearly a quarter of the time for solving the 40 scenarios case, the average expected objective value difference is merely 0.3%. The difference in the average objective value of the 10 scenarios case with other cases (the 40, 70, and 100 scenarios cases) is at a maximum of 2.6%, which is an acceptable magnitude of error considering the merit of the 10 scenarios case in terms of significantly short computation time. However, the normalized standard deviation, and the 25%–75% percentile interval remain larger compared to the other number of scenarios cases. Given this, the modeler can choose to consider 40 scenarios to improve solution reliability in realistic planning practices, but at the cost of significantly higher computation time.

References

[1] Karimi M, Mokhlis H, Naidu K, Uddin S, Bakar AA. Photovoltaic penetration issues and impacts in distribution network–A review. Renew Sustain Energy Rev 2016;53:594–605.

[2] Murugananthan B, Gnanadass R, Padhy N. Challenges with renewable energy sources and storage in practical distribution systems. Renew Sustain Energy Rev 2017;73:125–34.

[3] Silva J, Sumaili J, Bessa RJ, Seca L, Matos MA, Miranda V, Caujolle M, Goncer B, Sebastian-Viana M. Estimating the active and reactive power flexibility area at the TSO-DSO interface. IEEE Trans Power Syst 2018;33(5):4741–50.

[4] Kim S, Pollitt M, Jin Y, Kim J, Yoon Y. Contractual framework for the devolution of system balancing responsibility from the transmission system operator to distribution system operators. 2017.

[5] Perez E, Beltran H, Aparicio N, Rodriguez P. Predictive power control for PV plants with energy storage. IEEE Trans Sustain Energy 2012;4(2):482–90.

[6] Appino RR, Ordiano JÁG, Mikut R, Faulwasser T, Hagenmeyer V. On the use of probabilistic forecasts in scheduling of renewable energy sources coupled to storages. Appl Energy 2018;210:1207–18.

[7] Koeppel G, Korpås M. Improving the network infeed accuracy of non-dispatchable generators with energy storage devices. Electr Power Syst Res 2008;78(12):2024–36.

[8] Sossan F, Namor E, Cherkaoui R, Paolone M. Achieving the dispatchability of distribution feeders through prosumers data driven forecasting and model predictive control of electrochemical storage. IEEE Trans Sustain Energy 2016;7(4):1762–77. http://dx.doi.org/10.1109/TSTE.2016.2600103.

[9] Bozorg M, Sossan F, Le Boudec J-Y, Paolone M. Influencing the bulk power system reserve by dispatching power distribution networks using local energy storage. Electr Power Syst Res 2018;163:270–9.

[10] Abbey C, Baitch A, Bak-Jensen B, Carter C, Celli G, El Bakari K, Fan M, Georgilakis P, Hearne T, Ochoa LN, et al. Planning and optimization methods for active distribution systems. CIGRE; 2014.

[11] Celli G, et al. A comparison of distribution network planning solutions: Traditional reinforcement versus integration of distributed energy storage. In: 2013 IEEE grenoble conference. IEEE; 2013, p. 1–6.

[12] Asensio M, et al. Joint distribution network and renewable energy expansion planning considering demand response and energy storage—Part I: Stochastic programming model. IEEE Trans Smart Grid 2016;9(2):655–66.

[13] Muñoz-Delgado G, et al. Multistage generation and network expansion planning in distribution systems considering uncertainty and reliability. IEEE Trans Power Syst 2015;31(5):3715–28.

[14] Canizes B, et al. Optimal expansion planning considering storage investment and seasonal effect of demand and renewable generation. Renew Energy 2019;138:937–54.

[15] Saboori H, Hemmati R, Abbasi V. Multistage distribution network expansion planning considering the emerging energy storage systems. Energy Convers Manage 2015;105:938–45.

[16] Shen X, Shahidehpour M, et al. Expansion planning of active distribution networks with centralized and distributed energy storage systems. IEEE Trans Sustain Enery 2016;8(1):126–34.

[17] Samper ME, Vargas A. Investment decisions in distribution networks under uncertainty with distributed generation—Part I: Model formulation. IEEE Trans Power Syst 2013;28(3):2331–40.

[18] Tabares A, Franco JF, Lavorato M, Rider MJ. Multistage long-term expansion planning of electrical distribution systems considering multiple alternatives. IEEE Trans Power Syst 2015;31(3).

[19] Xie S, et al. Multi-objective active distribution networks expansion planning by scenario-based stochastic programming considering uncertain and random weight of network. Appl Energy 2018;219:207–25.

[20] Yi JH, Cherkaoui R, Paolone M. Optimal allocation of ESSs in active distribution networks to achieve their dispatchability. IEEE Trans Power Syst 2020.

[21] Yi JH, Cherkaoui R, Paolone M, Shchetinin D, Knezovic K. Optimal co-planning of ESSs and line reinforcement considering the dispatchability of active distribution networks. IEEE Trans Power Syst 2022.

[22] Koutsoukis NC, Georgilakis PS, Hatziargyriou ND. Multistage coordinated planning of active distribution networks. IEEE Trans Power Syst 2017;33(1):32–44.

[23] Moradijoz M, Moghaddam MP, Haghifam M-R. A flexible distribution system expansion planning model: a dynamic bi-level approach. IEEE Trans Smart Grid 2017;9(6):5867–77.

[24] Shen X, Shahidehpour M, et al. Multi-stage planning of active distribution networks considering the co-optimization of operation strategies. IEEE Trans Smart Grid 2016;9(2).

[25] Haghighat H, Zeng B. Stochastic and chance-constrained conic distribution system expansion planning using bilinear benders decomposition. IEEE Trans Power Syst 2017;33(3):2696–705.

[26] Christakou K, Tomozei D-C, Le Boudec J-Y, Paolone M. AC OPF in radial distribution networks–Part I: On the limits of the branch flow convexification and the alternating direction method of multipliers. Electr Power Syst Res 2017;143:438–50.

[27] Abdelouadoud S, Girard R, Neirac F-P, Guiot T. Optimal power flow of a distribution system based on increasingly tight cutting planes added to a second order cone relaxation. Int J Electr Power Energy Syst 2015;69:9–17.

[28] Nick M, Cherkaoui R, Boudec JL, Paolone M. An exact convex formulation of the optimal power flow in radial distribution networks including transverse components. IEEE Trans Automat Control 2018;63(3):682–97. http://dx.doi.org/10.1109/TAC.2017.2722100.

[29] Low SH. Convex relaxation of optimal power flow—Part I: Formulations and equivalence. IEEE Trans Control Netw Syst 2014;1(1):15–27.

[30] Geoffrion AM. Generalized benders decomposition. J Optim Theory Appl 1972;10(4):237–60.

[31] Zhao Z, Mutale J. Optimal conductor size selection in distribution networks with high penetration of distributed generation using adaptive genetic algorithm. Energies 2019;12(11):2065.

[32] Cococcioni M, Fiaschi L. The Big-M method with the numerical infinite M. Optim Lett 2021;15(7):2455–68.

[33] Stai E, Reyes-Chamorro L, Sossan F, Le Boudec J-Y, Paolone M. Dispatching stochastic heterogeneous resources accounting for grid and battery losses. IEEE Trans Smart Grid 2017;9(6):6522–39.

[34] Yi JH. 55Nodes-swiss-distribution-grid. 2022, URL https://github.com/DESL-EPFL/55nodes-Swiss-distribution-grid.

[35] Fenrick SA, et al. Cost and reliability comparisons of underground and overhead power lines. Util Policy 2012;20(1).

[36] The prices for balance energy in Swissgrid. 2020, URL https://www.swissgrid.ch/en/home/customers/topics/bgm/balance-energy.html.

[37] Jain AK, Dubes RC, et al. Algorithms for clustering data, Vol. 6. Prentice hall Englewood Cliffs, NJ; 1988.

[38] Gupta RK. Estimated medium voltage distribution network models for Switzerland. 2020, URL https://go.epfl.ch/SwissMVNetworkDB.

[39] Gupta R, Sossan F, Paolone M. Countrywide pv hosting capacity and energy storage requirements for distribution networks: the case of switzerland. Applied Energy 2021;281:116010.

[40] Yi JH. IEEE-benchmark-distribution-grids. 2022, URL https://github.com/DESL-EPFL/IEEE_benchmark_distribution_grids.