Distribution Network Expansion Planning with an Explicit Formulation for Reliability Assessment

Gregorio Muñoz-Delgado, Student Member, IEEE, Javier Contreras, Fellow, IEEE, and José M. Arroyo, Senior Member, IEEE

Abstract—This paper investigates the multistage expansion planning problem of a distribution network considering reliability. Thus, the best alternative, location, and installation time for the candidate assets are identified while jointly accounting for economic and reliability aspects. As a major salient feature, the conventional simulation-based reliability assessment is equivalently implemented through algebraic expressions whereby the effect of the network topology is explicitly represented by decision variables of the optimization process. For expository purposes, the focus is placed on the expected energy not supplied, which is a widely-used metric for reliability assessment. The resulting optimization problem is cast as an instance of mixed-integer linear programming. Hence, unlike existing heuristic and metaheuristic solution techniques for reliability-constrained distribution system planning, the proposed approach is finitely convergent to the optimal solution and can be readily implemented using commercially-available software. Simulation results show the effective performance of the proposed methodology.

Index Terms—Distribution network expansion planning, mixed-integer linear programming, multistage, reliability.

NOMENCLATURE

Acronyms

EFB Existing fixed branch.
ET Existing transformer.
ERB Existing replaceable branch.
HV High voltage.
MV Medium voltage.
NAB Newly added branch.
NRB New replacement branch.
NT New transformer.

Indices

\(b\) Index for the blocks of the annual system load-duration curve.
\(h\) Index for the blocks of the piecewise linear energy losses.
\(i, j, r, s\) Indices for nodes.
\(k, k'\) Indices for available investment alternatives.
\(l, l'\) Indices for branch types.
\(t, t'\) Indices for time stages.

Sets

\(B\) Index set of the blocks of the annual system load-duration curve.
\(K^l\) Index set of alternatives for branches of type \(l\).
\(K^{tr}\) Index set of alternatives for transformers of type \(tr\).
\(L\) Set of branch types. \(L = \{EFB, ERB, NRB, NAB\}\).
\(T\) Index set of time stages.
\(TR\) Set of transformer types. \(TR = \{ET, NT\}\).
\(\Psi_i^l, \Psi_i^{LN}, \Psi_i^N, \Psi_i^{SS}\) Index sets of nodes connected to node \(i\) by a branch of type \(l\), load nodes at stage \(t\), system nodes, and substation nodes.

Parameters

\(A_{kh}^l\) Width of block \(h\) of the piecewise linear energy losses for alternative \(k\) for branches of type \(l\) (A).
\(A_{kh}^{tr}\) Width of block \(h\) of the piecewise linear energy losses for alternative \(k\) for transformers of type \(tr\) (MVA).
\(C_{l,k}\) Investment cost coefficient for alternative \(k\) for branches of type \(l\) ($/km$).
\(C_{l,NT}\) Investment cost coefficient for alternative \(k\) for new transformers ($)$.
\(C_{i,SS}\) Investment cost coefficient for the substation at node \(i\) ($$).
\(C_{b}^{L}\) Cost coefficient for energy losses in block \(b\) of the annual system load-duration curve ($/MWh$).
\(C_{k}^{M,\theta}\) Maintenance cost coefficient for alternative \(k\) for branches of type \(\theta\) ($$).
\(C_{k}^{M,\theta}\) Maintenance cost coefficient for alternative \(k\) for transformers of type \(tr\) ($$).
\(C_{R}\) Cost coefficient for the expected energy not supplied under branch outages ($/MWh$).
\(C_{ib}^{SS}\) Cost coefficient for the energy supplied by the substation at node \(i\) in block \(b\) of the annual system load-duration curve ($/MWh$).

This work was supported in part by the Ministry of Economy and Competitiveness of Spain, under Project ENE2015-63879-R (MINECO/FEDER, UE), the Junta de Comunidades de Castilla-La Mancha, under Project POII-2014-012-P and Grant PRE2014/8064, and the Universidad de Castilla-La Mancha, under Grant GI20173944.

The authors are with the Escuela Técnica Superior de Ingenieros Industriales, Universidad de Castilla-La Mancha, 13071 Ciudad Real, Spain (e-mail: gregorio.munoz.delgado@gmail.com; Javier.Contreras@uclm.es; JoseManuel.Arroyo@uclm.es).
\(C^{U}\)  
Cost coefficient for load shedding ($/\text{MWh})

\(D_{it}\)  
Actual nodal peak demand at node \(i\) and stage \(t\) (MVA).

\(P_{ik}\)  
Upper limit for the actual current flow through alternative \(k\) for branches of type \(l\) (A).

\(G_{k}^{tr}\)  
Capacity of alternative \(k\) for transformers of type \(tr\) (MVA).

\(I_{f}\)  
Annual inflation-adjusted interest rate.

\(IB_{t}\)  
Investment budget for stage \(t\) ($).

\(\ell_{ij}\)  
Length of the branch connecting nodes \(i\) and \(j\) (km).

\(M_{kh}\)  
Slope of block \(h\) of the piecewise linear energy losses for alternative \(k\) for branches of type \(l\) (MW/km/kV).

\(M_{kh}^{r}\)  
Slope of block \(h\) of the piecewise linear energy losses for alternative \(k\) for transformers of type \(tr\) (MW/A).

\(n^{B}\)  
Number of blocks of the annual system load-duration curve.

\(n^{H}\)  
Number of blocks of the piecewise linear energy losses.

\(n_{K}^{K}\)  
Number of alternatives for branches of type \(l\).

\(n_{K}^{tr}\)  
Number of alternatives for transformers of type \(tr\).

\(n_{t}^{LN}\)  
Number of load nodes at stage \(t\).

\(n_{t}^{N}\)  
Number of system nodes.

\(n_{t}^{SS}\)  
Number of substation nodes.

\(n^{T}\)  
Number of time stages.

\(n_{t}^{k}\)  
Number of branches of type \(l\).

\(n_{t}^{k,SS}\)  
Number of branches of type \(l\) connected to a substation node.

\(p^{f}\)  
Power factor.

\(RR_{l}, RR_{NT}, RR_{SS}\)  
Capital recovery rates for investment in branches of type \(l\), new transformers, and substations.

\(V_{+}, V_{-}\)  
Lower and upper limits for nodal voltages (V).

\(W, W'\)  
Sufficiently large positive constants.

\(Z_{k}^{l}\)  
Unitary impedance magnitude for alternative \(k\) for branches of type \(l\) (\(\Omega/km\)).

\(Z_{k}^{tr}\)  
Impedance magnitude for alternative \(k\) for transformers of type \(tr\) (\(\Omega\)).

\(\Delta_{b}\)  
Duration of block \(b\) of the annual system load-duration curve (h).

\(\eta_{l}, \eta_{NT}, \eta_{SS}\)  
Lifetimes of branches of type \(l\), new transformers, and substation assets other than transformers (years).

\(\lambda_{k}\)  
Average failure rate of alternative \(k\) for branches of type \(l\) (failures/km/year).

\(\mu_{b}^{D}\)  
Load level of block \(b\) of the annual system load-duration curve, which is expressed in percent of the annual system peak demand (%).

\(\tau_{ij}^{RP}, \tau_{ij}^{SW}\)  
Durations of the repair-and-switching and switching-only interruptions associated with the failure of the branch connecting nodes \(i\) and \(j\) (h).

\(\tilde{d}_{ist}\)  
Amortized investment cost at stage \(t\) ($).

\(c_{i}^{PV,T}\)  
Costs of energy losses, maintenance, production, reliability, and load shedding at stage \(t\) ($).

\(\tilde{d}_{itb}\)  
Present value of the total cost ($).

\(\text{EENSC}_{t}\)  
Demand at node \(i\) in the fictitious system under the operating condition corresponding to node \(s\) at stage \(t\) (MVVA).

\(f_{ijktb}\)  
Cost of the expected energy not supplied associated with branch outages at stage \(t\) (V).

\(\tilde{f}_{ijktb}\)  
Current flow through alternative \(k\) for the branch of type \(l\) connecting nodes \(i\) and \(j\) of the fictitious system under the operating condition corresponding to node \(s\) at stage \(t\) (A).

\(\tilde{g}_{i}^{t}\)  
Current flow through alternative \(k\) for the branch of type \(l\) connecting nodes \(i\) and \(j\) of the fictitious system under the operating condition corresponding to node \(s\) at stage \(t\) (A).

\(\tilde{g}_{ist}\)  
Current injection at substation node \(i\) for alternative \(k\) for the transformer of type \(tr\) at stage \(t\) and block \(b\) of the annual system load-duration curve (MVVA).

\(v_{itb}\)  
Voltage magnitude at node \(i\), stage \(t\), and block \(b\) of the annual system load-duration curve (V).

\(x_{ijkt}^{l}\)  
Binary investment variable for alternative \(k\) for the branch of type \(l\) connecting nodes \(i\) and \(j\) at stage \(t\).

\(x_{ikt}^{NT}\)  
Binary investment variable for alternative \(k\) for the new transformer at substation node \(i\) and stage \(t\).

\(x_{ikt}^{SS}\)  
Binary investment variable for the substation at node \(i\) and stage \(t\).

\(y_{ijkt}^{l}\)  
Binary utilization variable for alternative \(k\) for the branch of type \(l\) connecting nodes \(i\) and \(j\) at stage \(t\).

\(y_{ikt}^{tr}\)  
Binary utilization variable for alternative \(k\) for the transformer of type \(tr\) at substation node \(i\) and stage \(t\).

\(\beta_{ijktbh}\)  
Current in block \(h\) of the piecewise linear energy losses for alternative \(k\) for the branch of type \(l\) connecting nodes \(i\) and \(j\) at stage \(t\) and block \(b\) of the annual system load-duration curve (A).
I. INTRODUCTION

MULTISTAGE distribution network expansion planning consists in determining the location, timing, and sizing of capital-intensive distribution assets such as branches, transformers, and substations to be installed in an existing distribution network so that the forecast demand growth is met over the long-term planning horizon in an economical and secure fashion [1]. The latter aspect is related to the notion of reliability, which is defined as the ability to continuously meet the electricity needs of end users with the required quality and quantity [2]. Therefore, the desired tradeoff between economics and reliability characterizing the optimal decision making requires the joint consideration of both aspects when using optimization models for distribution system planning.

Reliability is customarily incorporated in distribution system planning through the computation of standard reliability indices [2]–[10] and their associated costs. Thus, the resulting optimization process is driven by the minimization of a composite cost function comprising investment, operational, and reliability terms. Unfortunately, exact solution methodologies are currently unavailable for existing reliability-constrained distribution network expansion planning models [11]. This relevant methodological issue is a consequence of two factors, namely (i) the dependence of reliability indices on the network topology, which, in turn, is an outcome of the optimization, and (ii) the use of simulation for the computation of reliability indices in the absence of analytical expressions modeling reliability in terms of topology-related decision variables.

State-of-the-art approaches for reliability-constrained distribution network expansion planning thus rely on two broad categories of approximate methods. The first category comprises several metaheuristics [11] where a population of candidate topologies is iteratively handled. For each candidate topology, reliability is assessed through simulation. Hence, the impact of reliability on investment decisions drives the exploration of the search space without the method being able to acknowledge global optimality. An alternative two-step heuristic methodology was prescribed in [8]. First, an iterative algorithm based on standard mathematical programming was used to generate a pool of candidate topologies with no regard to reliability. Similar to the above-mentioned metaheuristics, for each candidate topology, reliability was assessed subsequently using simulation. As a result, topology-related decisions are decoupled from reliability, which may lead to suboptimal solutions. Hence, new approaches for the reliability-constrained distribution network expansion planning problem are yet to be explored.

Built upon [8] and [10], this paper presents a new approach for the precise incorporation of reliability in distribution network expansion planning. As a major salient feature over previously-reported approaches, the proposed model does not require the use of approximate solution methods. The key idea is the application of the findings of [10], where equivalent algebraic expressions for reliability assessment of a given network topology were recently developed, to a reliability-constrained distribution network expansion planning problem based on that addressed in [8]. Two recent examples of a successful application of similar, albeit simpler, algebraic expressions for reliability assessment can be found in [12] and [13], where different problems related to distribution system operation and planning were solved, respectively.

Here, the algebraic expressions of [10] are extended in a non-trivial way to account for the fact that, in a planning context, the network topology is not known a priori, being rather an outcome of the optimization process. Such an extension constitutes the main distinctive modeling aspect of this paper with respect to the reliability assessment model described in [10].

Moreover, the proposed extension of the model presented in [10] is relevant since it allows the explicit consideration of reliability assessment as part of the expansion planning problem. Thus, from a modeling perspective, the resulting reliability-constrained planning formulation significantly differs from that in [8], wherein reliability was implicitly characterized, i.e., no reliability-related mathematical expressions were included in the optimization problem. As another salient feature with respect to [8], the proposed model is suitable not only for radial distribution networks, but also for meshed and radially-operated distribution networks. For the latter, reliability indices are computed for the radial operation topology determined by the optimization process for each planning period.

In addition, the paper substantially differs methodologically from [8], wherein reliability was assessed by a simulation-based algorithm, thereby requiring the use of a heuristic solution approach for the original reliability-constrained planning model. Here, in contrast, the topology-dependent reliability assessment is explicitly formulated in terms of the decision variables of the proposed optimization problem, thereby permitting the exact solution of the resulting problem through the application of non-heuristic techniques.

The proposed model is formulated as a mixed-integer nonlinear program with nonlinearities related to (i) quadratic energy losses, and (ii) bilinear terms involving the products of decision variables. Such nonlinearities are recast as linear expressions by using well-known linearization schemes, namely a piecewise linear approximation for energy losses [14]...
and equivalent transformations for the bilinear terms based on disjunctive constraints [15] and integer algebra results [16]. The proposed solution approach yields a mixed-integer linear program for which, unlike the method applied in [8], finite convergence to the optimum is guaranteed while providing a measure of the distance to optimality along the solution process [17]. Additionally, effective off-the-shelf software based on the branch-and-cut algorithm is readily available [18].

The main contributions of this paper are twofold:

(i) From a modeling perspective, a new formulation is presented for multistage distribution network expansion planning considering reliability. Unlike previously-reported works, the cost of reliability is explicitly expressed in terms of the decision variables of the resulting optimization problem, thereby enabling the use of non-heuristic tools with well-known properties in terms of solution quality and convergence.

(ii) From a methodological perspective, mixed-integer linear programming is proposed to jointly consider economic and reliability aspects in distribution network expansion planning. To the best of the authors’ knowledge, the practical reliability-constrained planning model addressed in this paper is solved with guaranteed finite convergence to optimality for the first time. The superiority of the proposed method is backed by its computationally-effective ability to optimally solve cases for which a previously-reported heuristic required longer computing times to identify suboptimal solutions.

The remainder of this paper is organized as follows. Section II is devoted to the incorporation of reliability in distribution network expansion planning. Section III presents the formulation of the proposed model. In Section IV, numerical results are reported and analyzed. Finally, relevant conclusions are drawn in Section V.

II. Modeling Reliability in Distribution Network Expansion Planning

The reliability-constrained distribution network expansion planning problem aims to minimize the present value of the total cost \( c^{PV,T} \) over an \( n \)-dimensional vector of binary variables \( x \) and a vector of continuous variables \( y \), both modeling investment and operational decisions whose feasibility is characterized by constrained functions, \( f(x,y) \). Mathematically, this general description can be stated as follows:

Minimize \( c^{PV,T} \) subject to:

\[
\begin{align*}
\mathbf{x} &\in \{0, 1\}^n \\
f(x,y) &\leq 0 \\
\end{align*}
\]

(1) (2) (3)

The total cost minimized in (1) includes the investment cost; the operating cost, which comprises the costs of maintenance, production, energy losses, and load shedding; and the reliability cost. Investment and operating costs can be readily expressed in terms of \( x \) and \( y \) as \( c^{PV,I-O}(x,y) \). In contrast, the reliability cost is computed through several standard indices [2–9] for which no closed form in terms of decision variables \( x \) and \( y \) is currently available in the literature on distribution network expansion planning. In the absence of such analytical expressions, the reliability cost has been typically calculated through simulation, thereby preventing the application of exact methods to solve problem (1)–(3) [11].

As described in Section III, the above shortcoming featured by conventional reliability-constrained distribution network expansion planning models is overcome here based on the findings of [10], whereby reliability is explicitly characterized in terms of the decision variables of the optimization process. As a result, problem (1)–(3) is equivalently formulated in a compact way as follows:

Minimize \( c^{PV,T} = c^{PV,I-O}(x,y) + c^{PV,R}(z) \) subject to:

\[
\begin{align*}
\mathbf{x} &\in \{0, 1\}^n \\
f(x,y) &\leq 0 \\
g(x,z) &\leq 0 \\
\end{align*}
\]

(4) (5) (6) (7)

where \( z \) is a new vector of variables relating topology decisions and reliability, \( c^{PV,R}(z) \) is a function explicitly representing the present value of the cost of reliability in terms of the newly added decision variables included in \( z \), and \( g(x,z) \) is a new set of constrained functions modeling the relationship between \( x \) and \( z \).

III. Problem Formulation

This section presents the detailed formulation of problem (4)–(7) for a reliability-constrained distribution network expansion planning model based on that described in [8]. Thus, a radially-operated distribution system is considered, a multistage planning framework is adopted, an approximate lossless network model is used, the cost of losses is included in the objective function, and several investment alternatives exist for each asset. Moreover, as is customary in distribution system planning [8], [9], [13], [19]–[21], the seasonal or chronological aspect of demand is modeled by the discretization of the annual system load-duration curve into a set of blocks each characterized by its load level and its duration. Furthermore, for the sake of simplicity, such a system-wide demand characterization is applied on a nodal basis. As a consequence, the investment- and operation-related expressions described in the next sections are based on those presented in [8], which also form the basis of recently-reported distribution planning models [9], [21].

As for reliability, the assessment discussed in [8] is equivalently implemented. Thus, the reliability assessment relies on the following practical assumptions, which are customarily adopted for the sake of tractability [8]–[10], [22]–[26]:

(i) Only sustained interruptions due to single branch outages are considered. Branch outages are characterized by failure rates and interruption durations.

(ii) A radially-operated distribution system is considered wherein each branch connected to a substation is equipped with a circuit breaker without a recloser at the output of the substation. Moreover, all branches are equipped with a switch that enables partial restoration...
[7], i.e., the isolation of the part of the system downstream of the fault in order to meet the demand of the healthy portion of the system. Thus, once a sustained fault has occurred, the first circuit breaker upstream of the fault trips, thereby curtailing all downstream load demands. Subsequently, the system topology is reconfigured by operating switches and circuit breakers to reduce the non-supplied energy. To that end, the first switch upstream of the fault is opened in order to isolate the fault. Then the circuit breaker is closed so that the supply to all load demands between the circuit breaker and the switch is restored. Finally, once the isolated fault is cleared, the corresponding switch is closed and complete service is reestablished. Note that additional fault-clearing operations are included in the investment cost coefficients for branches.

In addition, it is worth emphasizing that, as a salient modeling aspect with respect to [8], reliability is equivalently assessed through a novel formulation. Based on [10], the proposed model comprises a set of algebraic expressions relying on the characterization of the operation of a fictitious system under different operating conditions. As highlighted in Section I, this distinctive modeling feature is relevant because it enables the application of mixed-integer linear programming, which is advantageous in terms of solution quality and computational behavior.

### A. Objective Function and Cost-Related Terms

The aim of the proposed model is the minimization of the present value of the total cost, which, based on [8], [9], and [21], is formulated as the following summation of three terms:

\[
c_{PV,T} = \sum_{t \in T} \frac{(1 + I_f)^{-t}}{I_f} c^P_t + \sum_{t \in T} \left[ (1 + I_f)^{-t} \left( c^M_t + c^P_t + c^L_t + c^U_t + c^R_t \right) \right] + \frac{(1 + I_f)^{-n_T}}{I_f} \left( c^M_{n_T} + c^P_{n_T} + c^L_{n_T} + c^U_{n_T} + c^R_{n_T} \right).
\]  (8)

In the first term, the present worth value of the investment cost is represented under the assumption of a perpetual or infinite planning horizon [27]. The second term characterizes the present value of the sum of the operating cost and the reliability cost. Finally, in the third term, the present value of the sum of the operating cost and the reliability cost incurred after the last time stage is modeled. Note that such a term depends on the values of those costs at the last time stage while also assuming a perpetual planning horizon. The cost terms in (8) are formulated as follows:

\[
c_i = \sum_{l \in \{NRB, NAB\}} RR^l \sum_{k \in \mathcal{K} \setminus \{i, j\}} C_{ik}^l \epsilon_{lj} x_{ijkl} + RR^{SS} \sum_{i \in \Psi^{SS}} C_i^M \eta_{SS} \sum_{k \in K^{NT}} C_k^M N^T x_{ijkl} ; \forall t \in T
\]  (9)

\[
c_i = \sum_{l \in \mathcal{L}} \sum_{k \in \mathcal{K} \setminus \{i, j\}} \sum_{(i, j) \in T^l} C_i^M \left( y_{ijkl} + y_{ijkl}^b \right)
\]  (10)

\[
c_i = \sum_{b \in B} \Delta_b p^f \left( \sum_{t \in T} \sum_{k \in K^{TR}} \sum_{i \in \Psi^{SS}} C_{ib}^s \eta_{SS} \right) \sum_{k \in K^{TR}} \eta_{SS} \sum_{b \in B} y_{ijkl} ; \forall t \in T
\]  (11)

\[
c_i = \sum_{b \in B} \Delta_b p^f \left( \sum_{t \in T} \sum_{k \in K^{TR}} \sum_{i \in \Psi^{SS}} C_{ib}^s \eta_{SS} \right) \sum_{k \in K^{TR}} \eta_{SS} \sum_{b \in B} y_{ijkl} ; \forall t \in T
\]  (12)

\[
c_i = EENS_C ; \forall t \in T
\]  (13)

\[
c_i = \sum_{b \in B} \Delta_b p^f \left( \sum_{t \in T} \sum_{k \in K^{TR}} \sum_{i \in \Psi^{SS}} C_{ik}^M \eta_{SS} \sum_{k \in K^{TR}} \eta_{SS} \right) \sum_{k \in K^{TR}} \eta_{SS} \sum_{b \in B} y_{ijkl} ; \forall t \in T
\]  (14)

where capital recovery rates are computed as \( RR^t = \frac{I_t (1 + I_f)^{n_T}}{(1 + I_f)^{n_T} - 1} \forall t \in \{NRB, NAB\} ; RR^{NT} = \frac{I_t (1 + I_f)^{n_T}}{(1 + I_f)^{n_T} - 1} ; \) and \( RR^{SS} = \frac{I_t (1 + I_f)^{m_T}}{(1 + I_f)^{m_T} - 1} \).

It is worth pointing out that, for each time stage, one binary variable per branch, \( x_{ijkl} \), is used to model the associated investment decision. In contrast, two continuous variables, \( f_{ijklb}^1 \) and \( f_{ijklb}^2 \), as well as two binary variables, \( y_{ijkl} \) and \( y_{ijkl}^b \), are associated with each branch in order to model the magnitude and direction, respectively, of the corresponding current flow. Note that \( f_{ijklb}^1 \) is positive and equal to the branch current flow between nodes \( i \) and \( j \) measured at node \( i \) only when the current flows from \( i \) to \( j \), being 0 otherwise.

In (9), the amortized cost of investment at each stage is formulated as the sum of the costs associated with the replacement and addition of branches, the reinforcement and construction of substations, and the installation of new HV/MV transformers. In the absence of decision variables modeling the investment in protection devices, their installation costs are included in the investment cost coefficients for branches. Expressions (10) model the maintenance costs of branches and transformers at each stage. In (11), the total production cost at each stage is characterized as the sum of the production costs across substation nodes.

In (12) and (13), the costs of load shedding and reliability are modeled, respectively. As done in [8], [9], and [26], such costs are jointly considered in the objective function (8).

Expressions (12) correspond to the penalty costs for load shedding under normal operation. Load shedding under normal operation, which is represented by decision variables \( c_{ib}^l \), refers to the violations of the nodal balance equations for the intact expanded system, i.e., with no component out of service.
The economic quantification of such levels of nodal imbalance is implemented by the cost coefficient \( C^U \).

In contrast, expressions (13) model the reliability cost at each stage. As customarily done in the related literature [8], [9], [26], reliability is represented by the expected energy not supplied, which is a well-known metric [3], [5]–[8] that refers to the expected energy curtailment due to branch outages. It is worth mentioning that the economic quantification of this reliability metric is typically implemented by a regulation-dependent cost coefficient, \( C^R \), that may be different from the cost coefficient for load shedding under normal operation, \( C^U \). Note also that the proposed modeling framework is general enough to accommodate other standard reliability metrics [2]–[10]. For the sake of conciseness, details on such modeling capability have been omitted.

Finally, as done in [8], [9], and [21], the costs of energy losses in branches and transformers are modeled in (14) by quadratic functions. Based on [14], such nonlinearities are approximated by tangent lines, as also done in [8], [9], and [21]. Thus, expressions (14) are replaced with:

\[
\begin{align*}
\alpha_t^L &= \sum_{b \in B} \Delta t C^L_{tb} p f \left[ \sum_{tr \in T_R} \sum_{k \in K^{tr}} \sum_{i \in \Psi^{SS}} \sum_{h=1}^{n^H} M^H_{khh} \beta^{tr}_{iktbb} + \sum_{\ell \in L} \sum_{\ell \in K^{tr}} \sum_{i \in \Psi^{SS}} \sum_{h=1}^{n^H} M^H_{khh} \beta^{tr}_{iktbb} \right] ; \\
g^{tr}_{iktbb} &= \sum_{h=1}^{n^H} \beta^{tr}_{iktbb} ; \forall tr \in T_R, \forall i \in \Psi^{SS}, \forall k \in K^{tr}, \\
0 &\leq g^{tr}_{iktbb} \leq A^{tr}_{iktbb} ; \forall h = 1 \ldots n^H, \forall tr \in T_R, \forall i \in \Psi^{SS}, \forall k \in K^{tr}, \forall t \in T, \forall b \in B, \\
f^{l}_{ijkb} &= \sum_{h=1}^{n^H} \beta^{l}_{ijkbh} ; \forall l \in L, \forall i \in \Psi^{l}, \forall j \in \Psi^{N}, \forall k \in K^{l}, \\
0 &\leq \beta^{l}_{ijkbh} \leq A^{l}_{ijkbh} ; \forall h = 1 \ldots n^H, \forall l \in L, \forall i \in \Psi^{l}, \forall j \in \Psi^{N}, \forall k \in K^{l}, \forall t \in T, \forall b \in B.
\end{align*}
\]

This approximation yields piecewise linear functions, which, for practical purposes, are indistinguishable from nonlinear models if enough segments are used. Expressions (15) represent the linearized costs of energy losses, while (16)–(17) and (18)–(19) model the piecewise linearization for energy losses in transformers and branches, respectively.

**B. Kirchhoff’s Laws and Operational Limits**

System operation is modeled by the following constraints:

\[
\begin{align*}
\nabla \leq v_{ibt} &\leq \nabla ; \forall i \in \Psi^N, \forall t \in T, \forall b \in B, \\
0 &\leq f_{ijkb}^l = \sum_{k \in K^{l}} F_{ijk}^l ; \forall l \in L, \forall i \in \Psi^{l}, \forall j \in \Psi^{N}, \forall k \in K^{l}, \forall t \in T, \forall b \in B, \\
0 &\leq g^{tr}_{iktbb} = \sum_{k \in K^{tr}} C^{tr}_{iktbb} ; \forall tr \in T_R, \forall i \in \Psi^{SS}, \forall k \in K^{tr}, \\
0 &\leq \beta^{l}_{ijkbh} \leq A^{l}_{ijkbh} ; \forall h = 1 \ldots n^H, \forall l \in L, \forall i \in \Psi^{l}, \forall j \in \Psi^{N}, \forall k \in K^{l}, \forall t \in T, \forall b \in B.
\end{align*}
\]

Therefore, constraints (20) limit nodal voltage magnitudes. Similarly, constraints (21)–(23) set the upper and lower bounds on flows and injections. Finally, the effect of the distribution network is characterized in (24)–(25) by a lossless linearized network model that was first proposed by Haffner et al. [19] and successfully applied in [8], [9], [20], and [21]. As described in [19], the linearized network model is an adapted version of the per-unit dc model used for the transmission network that is based on three assumptions: (i) nodal voltages lie in a narrow range around the rated value used as base voltage, (ii) all per-unit branch current flows and nodal power injections have the same power factor, and (iii) the per-unit voltage drop across a branch is equal to the difference between the per-unit magnitudes of the nodal voltages at both ends of the branch.

As per assumption (i), per-unit values of nodal apparent power injections and nodal current injections are identical. Assumption (ii) allows representing per-unit branch current flows and nodal power injections, which are complex numbers, through their magnitudes. Hence, Kirchhoff’s current law can be expressed as a set of linear scalar equalities in terms of per-unit magnitudes of branch current flows and nodal power injections, giving rise to nodal balance equations (24). In addition, assumption (iii) allows formulating Kirchhoff’s voltage law for each branch in use as a linear expression relating the per-unit magnitudes of branch current flows, nodal voltages, and branch impedances. Constraints (25) extend this result to account for the utilization state of all branches.

Bearing in mind that nodal voltage magnitudes are decision variables of the optimization process, the proposed model provides an approximate characterization of the dependence of the actual nodal current demands on nodal voltages.

Note that constraints (25) include nonlinearities involving the products of binary variables, \( y_{ijk}^l \), and continuous variables, \( f_{ijkb}^l \) and \( v_{ibt} \), for which the following linear equivalent is formulated using the disjunctive-constraint-based transformation described in [15]:

\[
- W (1 - y_{ijk}^l)^{-} \leq Z_{ijkb}^l - f_{ijkb}^l - (v_{ibt} - v_{ijb}) \leq W (1 - y_{ijk}^l) ; \forall l \in L, \forall i \in \Psi^{l}, \forall j \in \Psi^{N}, \forall k \in K^{l}, \forall t \in T, \forall b \in B.
\]

If \( y_{ijk}^l \) is equal to 1, the corresponding constraint (26) becomes 0 \( \leq Z_{ijkb}^l - (v_{ibt} - v_{ijb}) \leq 0 \), which is identical to the condition \( Z_{ijkb}^l - (v_{ibt} - v_{ijb}) = 0 \) resulting from (25). Conversely, if \( y_{ijk}^l \) is equal to 0, the corresponding constraint (26) yields \( -W \leq (v_{ibt} - v_{ijb}) \leq W \), i.e., \( |v_{ibt} - v_{ijb}| \leq W \), since \( f_{ijkb}^l \) is equal to 0 as per (21). Thus, for a sufficiently large positive value for parameter \( W \), no relation between nodal voltage magnitudes \( v_{ibt} \) and \( v_{ijb} \) is imposed, as modeled in (25) for \( y_{ijk}^l \) equal to 0. Since nodal voltage magnitudes are bounded by \( \nabla \) and \( \nabla \) in (20),
the largest possible value for \(|v_{i-thb} - v_{j-thb}|\) is \(\overline{V} - \underline{V}\), which is thus the minimum value for \(W\).

C. Investment and Utilization Constraints

The constraints associated with investment and utilization decisions are formulated as:

\[
\begin{align*}
x_{ijk}^t &\in \{0, 1\}; \forall l \in \{NRB, NAB\}, \forall (i, j) \in \mathcal{T}^l, \forall k \in K^l, \forall t \in T \quad (27) 

x_{it}^{kSS} &\in \{0, 1\}; \forall i \in \Psi^{kSS}, \forall t \in T 

N_{NT}^{i} &\in \{0, 1\}; \forall i \in \Psi^{kSS}, \forall k \in K^{NT}, \forall t \in T 

y_{ijkt}^l &\in \{0, 1\}; \forall l \in L, \forall i \in \Psi_j^l, \forall j \in \Psi^N, \forall k \in K^l, \forall t \in T 

y_{ijkt}^{kTR} &\in \{0, 1\}; \forall t \in T, \forall i \in \Psi^{kSS}, \forall k \in K^{TR}, \forall t \in T 

\sum_{t \in T} \sum_{k \in K^l} x_{ijk}^t \leq 1; \forall l \in \{NRB, NAB\}, \forall (i, j) \in \mathcal{T}^l 

\sum_{l \in T} x_{it}^{kSS} \leq 1; \forall i \in \Psi^{kSS} 

\sum_{l \in T} \sum_{k \in K^{NT}} x_{ijkt}^{kNT} \leq 1; \forall i \in \Psi^{kSS} 

x_{ijkt}^{kNT} \leq \sum_{l' = 1}^t x_{ijkt}^{kNRB}; \forall i \in \Psi^{kSS}, \forall k \in K^{NT}, \forall t \in T 

y_{ijkt}^l + y_{ijkt}^{kERB} \leq 1; \forall (i, j) \in \mathcal{T}^{SW,ERB}, \forall k \in K^{ERB}, \forall t \in T 

y_{ijkt}^l + y_{ijkt}^{kERB} = 1; \forall (i, j) \in (\mathcal{T}^{ERB} \setminus \mathcal{T}^{SW,ERB}), \forall k \in K^{ERB}, \forall t \in T 

y_{ijkt}^l + y_{ijkt}^{kNRB} = \sum_{t' = 1}^t x_{ijkt}^{kNRB}; \forall l \in \{NRB, NAB\}, \forall (i, j) \in (\mathcal{T}^l \setminus \mathcal{T}^{SW,l}), \forall k \in K^l, \forall t \in T 

y_{ijkt}^{ERB} + y_{ijkt}^{ERO} = 1 - \sum_{t' = 1}^t \sum_{k \in K^{NRB}} x_{ijkt}^{kNRB}; \forall (i, j) \in (\mathcal{T}^{ERB} \setminus \mathcal{T}^{SW,ERB}), \forall k \in K^{ERB}, \forall t \in T 

y_{ijkt}^{NRB} \leq \sum_{l' = 1}^t x_{ijkt}^{kNRB}; \forall i \in \Psi^{kSS}, \forall k \in K^{NT}, \forall t \in T 

\sum_{l \in \{NRB, NAB\}} \sum_{k \in K^l} C_{ik}^{l,m} l_{ik}^{l,m} + \sum_{i \in \Psi^{N}} C_{ik}^{l,SS} x_{it}^{kSS} 

\sum_{l \in L} \sum_{i \in \Psi^{l}} \sum_{k \in K^l} y_{ijkt}^{l} = 1; \forall j \in \Psi^L, \forall t \in T
\end{align*}
\]

\[
\sum_{l \in L} \sum_{i \in \Psi^{l}} \sum_{k \in K^l} y_{ijkt}^{l} \leq 1; \forall j \notin \Psi^L, \forall t \in T. \quad (45)
\]

Constraints (27)–(31) set the binary nature of investment and utilization variables. As per (32)–(34), each system component and location can only undergo expansion once along the planning horizon. Constraints (35) impose that the installation of a new transformer require the expansion or construction of the corresponding substation. Unlike [8], meshed topologies are allowed through network reconfiguration constraints (36)–(38), which model the utilization of branches while explicitly characterizing the direction of current flows. Constraints (39)–(41) are analogous to (36)–(38) for the branches that are not reconfigurable under normal operation. Constraints (42) formulate the utilization of new transformers. A budgetary limit for investments at each stage is imposed in (43). Finally, radial operation is modeled through constraints (44)–(45). Constraints (44) impose the existence of an incoming flow for load nodes while expressions (45) set a maximum of one incoming flow for the remaining nodes.

D. Reliability Constraints

In [10], we showed that standard reliability indices can be expressed in terms of the durations of the so-called repair-and-switching and switching-only interruptions affecting system nodes. Moreover, we also showed that, for a given radially-operated network topology, such durations can be equivalently computed through a set of binary-valued continuous variables representing the operation of a fictitious, topologically identical system. Here, we present the application of those findings to explicitly incorporate reliability in the mathematical formulation of reliability-constrained distribution network expansion planning. It should be noted that such an application requires not only extending the notation to represent investment alternatives and a multistage setting, but also the non-trivial modification of the expressions described in [10] to handle the lack of knowledge of the network topology along the planning horizon, which is an outcome of the optimization process. The proposed reliability constraints are formulated as follows:

\[
EENS_C = C^R \left[ \sum_{s \in \Psi_1^N} (\delta_{st}^{RP} + \delta_{st}^{SW}) \sum_{b \in B} \Delta_b \frac{S_{b0}}{I_{b}} \rho f_{b} D_{st} \right]; \forall t \in T \quad (46)
\]

\[
\delta_{st}^{RP} = \sum_{l \in L} \sum_{k \in K^l} \sum_{(i,j) \in T} \left( p_{ijkt}^{l} + \tilde{p}_{ijkt}^{l} \right) \lambda_{k} f_{ij} \tau_{ij}^{RP}; \forall s \in \Psi_1^{L_{N}}, \forall t \in T \quad (47)
\]

\[
\delta_{st}^{SW} = \delta_{st}^{T} - \delta_{st}^{SW}; \forall s \in \Psi_1^{L_{N}}, \forall t \in T; \quad (48)
\]

\[
\delta_{st}^{T} = \sum_{l \in L} \sum_{i \in \Psi^{SS}} \sum_{j \in \Psi^{N}} \sum_{k \in K^l} \delta_{ijt}^{CB,l} \sum_{s' \in \Psi} \left( \tilde{p}_{ijkt}^{l} + \tilde{p}_{ijkt}^{l} \right) \sum_{k \in K^l} \tilde{p}_{ijkt}^{l}; \forall s \in \Psi_1^{L_{N}}, \forall t \in T \quad (49)
\]

\[
\delta_{ijt}^{CB,l} = \sum_{s \in \Psi^{N}} \left( \sum_{l' \in L} \sum_{i \in \Psi^{l'}} \sum_{k \in K^{l'}} \tilde{p}_{rskt}^{l'} \lambda_{k} f_{rs} \tau_{rs} \right) \sum_{k \in K^l} \tilde{p}_{ijkt}^{l}; \forall l \in L, \forall i \in \Psi^{SS}, \forall j \in \Psi_1^{l}, \forall t \in T \quad (50)
\]
\[
\delta_{st}^{SW} = \sum_{i \in L} \sum_{k \in K^l} \sum_{t \in T^l} \left( \bar{j}_{ijskt}^l + \bar{f}_{ijskt}^l \right) \lambda_i^l b_i^l \tau_{ij}^{SW};
\]
\[
\forall s \in \Psi_s^{LN}, \forall t \in T
\]
\[
\tilde{d}_{ist} = \sum_{l \in L} \sum_{k \in K^l} \sum_{t \in T^l} \left( \bar{j}_{ijskt}^l + \bar{f}_{ijskt}^l \right) \gamma^l_i \delta_{ist}^{SW};
\]
\[
\forall t \in T
\]
\[
\tilde{d}_{ist} = 0; \forall i \in \Psi_i^{N}, \forall s \in \Psi_s^{N} | i = s, \forall t \in T
\]
\[
\sum_{l \in L} \sum_{k \in K^l} \sum_{t \in T^l} \left( \bar{f}_{ijskt}^l - \bar{j}_{ijskt}^l \right) = g_{ist}^{SS} - \tilde{d}_{ist};
\]
\[
\forall i \in \Psi_i^{N}, \forall s \in \Psi_s^{N}, \forall t \in T
\]
\[
0 \leq \bar{f}_{ijskt}^l \leq y_{ijskt}^l; \forall l \in L, \forall i \in \Psi_i^l, \forall j \in \Psi_j^N, \forall s \in \Psi_s^N,
\]
\[
\forall k \in K^l, \forall t \in T
\]
\[
0 \leq g_{ist}^{SS} \leq 1; \forall i \in \Psi_i^{SS}, \forall s \in \Psi_s^{N}, \forall t \in T.
\]

According to [10], the cost of the expected energy not supplied at each stage is modeled in (46) in terms of the expected durations of the repair-and-switching and switching-only interruptions affecting system nodes, namely \( \delta_{st}^{RP} \) and \( \delta_{st}^{SW} \), respectively. Thus, from (8), (13), and (46), the general term \( c_{PV,R}(z) \) in (4) corresponds to:
\[
\sum_{t \in T} \left[ (1 + I_f)^{-t} C_R \left( \sum_{s \in \Psi_s^{LN}} \left( \delta_{st}^{RP} + \delta_{st}^{SW} \right) \sum_{b \in B} \frac{\Delta_b}{8760} f_{ijskt}^l b_i^l D_{st} \right) \right]
\]
\[
+ \frac{(1 + I_f)^{-n^T}}{I_f} C_R \left( \sum_{s \in \Psi_s^{LN}} \left( \delta_{st}^{RP} + \delta_{st}^{SW} \right) \sum_{b \in B} \frac{\Delta_b}{8760} f_{ijskt}^l b_i^l D_{sn^T} \right)
\]
\[
\text{where vector } z \text{ comprises variables } \delta_{st}^{RP} \text{ and } \delta_{st}^{SW}. \text{ In (47)–(51), such expected durations are cast in terms of variables } \bar{f}_{ijskt}^l, \text{ which are binary-valued continuous variables that allow characterizing the dependence of reliability on the topology of the network along the planning horizon. A detailed explanation of (47)–(51) can be found in [10], where essentially identical expressions were described, albeit with a simplified notation disregarding indices } l, k, \text{ and } t.
\]

Based on [10], variables \( \bar{f}_{ijskt}^l \) model the per-unit branch current flows in a fictitious system with the same topology as the expanded network under a set of operating conditions, one per system node \( s \), and considering a constant power factor across the network, as formulated in (52)–(56). Similar to the approximate network model used for normal operation, two continuous variables, \( \bar{f}_{ijskt}^l \) and \( \bar{j}_{ijskt}^l \), are associated with each branch. Thus, variables \( \bar{f}_{ijskt}^l \) are different from 0 only when the fictitious current flows from \( i \) to \( j \), being 0 otherwise.

The operating conditions for the fictitious system are characterized in (52)–(53) by decision variables \( \bar{d}_{ist} \). Thus, such conditions are different from those related to load levels under the normal state or contingency conditions, which are typically represented by parameters. Rather, the per-unit demand at node \( i \) and stage \( t \) under the operating condition associated with node \( s \), \( \bar{d}_{ist} \), is equal to 1 if \( i = s \) and node \( i \) is connected to other nodes, being 0 otherwise. Note that the connection of node \( i \) to other nodes is modeled in (52) through binary utilization variables \( y_{ijskt}^l \). Constraints (54) represent the nodal balance equations for the fictitious system. Constraints (55) bound fictitious branch current flows using the corresponding binary utilization variables. Hence, as per (30) and (36)–(41), variables \( \bar{f}_{ijskt}^l \) lie in the interval [0,1], \( \bar{f}_{ijskt}^l \) and \( \bar{j}_{ijskt}^l \) cannot be simultaneously different from 0, and fictitious current flows through non-existing branches are set to 0. Finally, the fictitious injections by the substations are limited in (56). It should be pointed out that the constraint set (54)–(56) features a unimodular matrix structure, and consequently continuous variables \( \bar{d}_{ist}, \bar{f}_{ijskt}^l, \) and \( g_{ist}^{SS} \) can only take two values, namely 0 and 1. Thus, similar to the fictitious current flows described in [10], variables \( \bar{f}_{ijskt}^l \) are useful for the purposes of reliability assessment. As a matter of fact, for a given network topology, i.e., for fixed values of \( y_{ijskt}^l \), constraints (54)–(56) correspond to the expressions presented in [10]. Therefore, for further details on how fictitious branch current flows are used to model reliability, interested readers should refer to [10].

The presence of variables \( y_{ijskt}^l \) in (54), through variables \( \bar{d}_{ist} \), and (55) allows modeling the relationship between the operating conditions for the fictitious system and the network topology resulting from the expansion planning optimization process. In other words, the dependence of reliability-related variables, \( \bar{f}_{ijskt}^l, \bar{g}_{ist}^{SS}, \) and \( \bar{d}_{ist} \), on the network topology is modeled in (52)–(56) through operation-related variables, \( y_{ijskt}^l \), which, in turn, are dependent on expansion decisions, \( x_{ijskt} \), as per (37)–(38) and (40)–(41). Thus, the required relationship between reliability- and expansion-related decisions is mathematically characterized. This aspect, aside from the increased notational complexity, represents a salient feature with respect to the model described in [10], which was devised for a given network topology, thereby relying on operating conditions for the fictitious system that were characterized by parameters.

As another relevant difference with respect to the optimization problems addressed in [8] and [10], the explicit incorporation of algebraic expressions for reliability in the planning model gives rise to a mixed-integer nonlinear program. It is worth mentioning that expressions (49) include nonlinearities involving products of continuous variables, \( \delta_{ijskt}^{CB,l} \), and binary-valued continuous variables, \( \bar{f}_{ijskt}^l \). Additionally, nonlinearities comprising the cross products of binary-valued variables, \( \bar{f}_{ijskt}^l \), are present in (50). Based on [16], such nonlinearities can be reformulated as linear equivalents.

Thus, expressions (49) are linearized as follows:
\[
\delta_{st}^T = \sum_{l \in L} \sum_{i \in \Psi_i^{LN}} \sum_{j \in \Psi_i^l} \sum_{k \in K^l} \delta_{ij, st}^{T,l}; \forall s \in \Psi_s^{LN}, \forall t \in T
\]
\[
\delta_{ij, st}^{T,l} \leq W^T \left( \sum_{k \in K^l} \left( \bar{f}_{ijskt}^l + \bar{j}_{ijskt}^l \right) \right); \forall l \in L, \forall i \in \Psi_i^{SS},
\]
\[
\forall j \in \Psi_j^l, \forall s \in \Psi_s^{LN}, \forall t \in T
\]
\[
\delta_{ij, st}^{CB,l} - \delta_{ij, st}^{T,l} \leq W^T \left[ 1 - \sum_{k \in K^l} \left( \bar{f}_{ijskt}^l + \bar{j}_{ijskt}^l \right) \right];
\]
\[
\forall l \in L, \forall i \in \Psi_i^{SS}, \forall j \in \Psi_j^l, \forall s \in \Psi_s^{LN}, \forall t \in T
\]
where $\delta_{ijst}^{CB,1}$ represents $\delta_{ijst}^{CB,1} \sum_{k \in K^1} (\tilde{f}_{ijskt} + \tilde{f}_{jiskt})$. Note that a suitable value for the big-M parameter $W'$ is $\sum_{l \in L} \sum_{k \in K^1} \sum_{(i,j) \in T^1} \lambda_{ij}^{SW} \tilde{f}_{ij}^{SW}$.

Finally, expressions (50) can be linearized through the following set of equations:

$$\delta_{ijst}^{CB,1} = \sum_{\psi_i} \sum_{\psi_{i'}} \sum_{\psi_{i''}} \delta_{ijst}^{CB,1'} ; \forall \psi_i \in L, \forall \psi_{i'} \in \Psi^{SS},$$

$$\forall j \in \psi_i, \forall \psi_{i'} \in T$$

$$\delta_{ijst}^{CB,1'} \leq \sum_{k \in K^1} \tilde{f}_{ijskt} \lambda_{k}^{SW} \tilde{f}_{rs}^{SW} ; \forall \psi_i \in L, \forall \psi_{i'} \in L,$$

$$\forall \psi_{i'} \in L, \forall \psi_{i''} \in \Psi^{N}$$

$$\delta_{ijst}^{CB,1'} \leq \sum_{k \in K^1} \tilde{f}_{ijskt} ; \forall \psi_i \in L, \forall \psi_{i'} \in L, \forall \psi_{i''} \in \Psi^{SS},$$

$$\forall j \in \psi_i, \forall \psi'_{i} \in \Psi^{N}, \forall \psi_{i'} \in \Psi^{N}_{s}, \forall \psi_{i''} \in \Psi^{SS},$$

$$\forall \psi_{i'} \in L,$$

$$\delta_{ijst}^{CB,1'} \geq \sum_{k \in K^1} \tilde{f}_{ijskt} \lambda_{k}^{SW} \tilde{f}_{rs}^{SW} + \sum_{k \in K^1} \tilde{f}_{ijskt} - 1 ; \forall \psi_i \in L,$$

$$\forall \psi_{i'} \in L, \forall \psi_{i''} \in \Psi^{SS}, \forall j \in \psi_i, \forall \psi'_{i} \in \Psi^{N}, \forall \psi_{i'} \in \Psi^{N}_{s}, \forall \psi_{i''} \in \Psi^{SS},$$

$$\forall \psi_{i'} \in L,$$

$$\delta_{ijst}^{CB,1'} \leq \sum_{k \in K^1} \tilde{f}_{ijskt} \lambda_{k}^{SW} \tilde{f}_{rs}^{SW} + \sum_{k \in K^1} \tilde{f}_{ijskt} - 1 ; \forall \psi_i \in L,$$

$$\forall \psi_{i'} \in L, \forall \psi_{i''} \in \Psi^{SS}, \forall j \in \psi_i, \forall \psi'_{i} \in \Psi^{N}, \forall \psi_{i'} \in \Psi^{N}_{s}, \forall \psi_{i''} \in \Psi^{SS},$$

where $\delta_{ijst}^{CB,1'}$ represents $\left( \sum_{k \in K^1} \tilde{f}_{ijskt} \lambda_{k}^{SW} \tilde{f}_{rs}^{SW} \right) \sum_{k \in K^1} \tilde{f}_{ijskt}$.

Both linearizations thus constitute additional relevant distinctive features with respect to [8] and [10].

### E. Resulting Mixed-Integer Linear Program

The proposed mixed-integer linear program consists in the minimization of (8) subject to (9)–(13), (15)–(24), (26)–(48), (51)–(56), and (58)–(64). Table I provides the size of the resulting optimization problem in terms of the numbers of constraints, binary variables, and real variables.

Bearing in mind the main limitation featured by the approximate network model, namely the consideration of a constant power factor across the system, and the practical assumptions adopted for reliability assessment, we recognize that the use of the proposed model leads to results that may be optimistic and that a complete study would require the consideration of an accurate network model and a more sophisticated reliability characterization including additional post-fault network reconfiguration and multiple outages. This generalization would, however, render the problem essentially intractable through optimization and would have to be solved by heuristics or repeated simulations. These modeling limitations notwithstanding, the solution of the proposed multistage reliability-constrained expansion planning problem is acceptable for distribution planning purposes and provides the planner with a first estimate of a cost-effective and reliable expansion scheme.

### IV. Numerical Results

The proposed model has been applied to a 54-node test system based on that described in [9]. As shown in Fig. 1, the system comprises 50 load nodes, 4 substation nodes, and 63 branches. For reproducibility purposes, system data can be downloaded from [28], where base values for power and voltage are 1 MVA and 13.5 kV, respectively, and the currency used is U.S. dollars, hereinafter denoted by $. It should be noted that investment and maintenance cost coefficients...
are based on those provided in [8] and [9]. Analogously, the values of the cost coefficients for load shedding, energy losses, and the energy supplied by substations rely on those used in [8]. Moreover, according to [1], the cost coefficient for the expected energy not supplied under branch outages, $C_{R_k}$, is equal to $11200$/MWh. As done in [8], [13], [20], and [21], three load levels, namely high, medium, and low, are considered. In addition, three blocks are used in the piecewise linearization for energy losses. Investment decisions are made over a 10-year planning horizon divided into yearly stages considering a 10% inflation-adjusted interest rate and an annual investment budget equal to $2$ million.

The performance of the proposed approach has been assessed with the heuristic presented in [8]. This method has been implemented to yield a pool of six inexpensive candidate expansion plans in terms of investment and operating costs. For each solution, reliability is computed ex post through a simulation-based algorithm equivalent to the proposed algebraic expressions. Simulations have been run on a Dell PowerEdge R920X64 with four Intel® Xeon® E7-4820 processors at 2.00 GHz and 768 GB of RAM using CPLEX 12.6 [18] and GAMS 24.7 [29]. For this case study, the proposed model, which comprises 241779 constraints, 3820 binary variables, and 226632 real variables, was solved to optimality, i.e., the optimality gap of CPLEX was set at 0%. The same stopping criterion was adopted for the executions of CPLEX required by the heuristic approach presented in [8].

Table II summarizes the economic results attained by both methods. The second column of this table lists the present values of the costs associated with the optimal solution to the proposed model. The same information is reported in columns 3–8 for the six expansion plans identified by the heuristic approach presented in [8]. As can be seen in Table II, jointly optimizing all costs yields a significant 9.0% cost reduction over the best solution determined by the heuristic approach, namely Plan 5. Such a cost savings is a consequence of the larger investment cost featured by the optimal solution, which substantially reduces the cost of reliability while slightly increasing the production cost. More specifically, the optimal solution involves the construction of two new substations, whereas each solution of the heuristic is characterized by the construction of one substation at most.

Figs. 2 and 3 respectively show the investment decisions and the topology associated with the optimal solution provided by the proposed approach and Plan 5 identified by the heuristic.

### Table II

<table>
<thead>
<tr>
<th>Costs</th>
<th>Proposed approach</th>
<th>Heuristic Plan 1</th>
<th>Heuristic Plan 2</th>
<th>Heuristic Plan 3</th>
<th>Heuristic Plan 4</th>
<th>Heuristic Plan 5</th>
<th>Heuristic Plan 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investment</td>
<td>1.90</td>
<td>0.41</td>
<td>0.42</td>
<td>0.87</td>
<td>0.98</td>
<td>0.95</td>
<td>1.20</td>
</tr>
<tr>
<td>Maintenance</td>
<td>0.22</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>Production</td>
<td>27.76</td>
<td>26.67</td>
<td>26.74</td>
<td>26.84</td>
<td>26.85</td>
<td>26.95</td>
<td>26.75</td>
</tr>
<tr>
<td>Losses</td>
<td>1.08</td>
<td>1.71</td>
<td>1.67</td>
<td>1.60</td>
<td>1.63</td>
<td>1.67</td>
<td>1.67</td>
</tr>
<tr>
<td>Load shedding</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Reliability</td>
<td>6.82</td>
<td>13.25</td>
<td>12.58</td>
<td>12.02</td>
<td>11.90</td>
<td>11.78</td>
<td>12.35</td>
</tr>
<tr>
<td>Total</td>
<td>37.78</td>
<td>42.24</td>
<td>41.61</td>
<td>41.59</td>
<td>41.53</td>
<td>41.51</td>
<td>42.17</td>
</tr>
</tbody>
</table>

As can be observed, the main difference arises at node 54 where the construction of a new substation is planned by the proposed approach as opposed to the heuristic. As a consequence, while new load nodes are fed by the newly built substation at node 54 at the optimal solution, such nodes are supplied by the existing substations at nodes 51 and 52 under the heuristic solution.

Table III presents cost information for each stage. As can be noted, for the proposed approach, the largest investment costs are incurred at stages 1 and 5, when the construction of two new substations is planned at nodes 54 and 53, respectively, including the installation of two new transformers at those nodes. For Plan 5, the largest investment is made at stage 6.
when a substation and a transformer are built at node 53. In addition, for both solutions, the costs related to maintenance, production, losses, and reliability rise along the planning horizon due to the demand growth. As can also be seen, no load shedding is required by any solution.

The relevance of explicitly incorporating the reliability cost in the optimization process can be quantified by comparing the optimal solution provided by the proposed approach with the most inexpensive expansion plan disregarding reliability. It is worth mentioning that such a solution featuring the minimum value of the sum of investment and operating costs corresponds to Plan 1 identified by the heuristic. As can be observed in Table II, the optimal solution provided by the proposed approach is 6.8% more expensive than Plan 1 in terms of investment and operating costs while yielding a substantial 48.5% reduction in the reliability cost. Overall, a 10.6% reduction in the total cost is achieved, thereby clearly substantiating the advantage of the proposed approach.

Fig. 4 shows the investment decisions and the topology for Plan 1. As can be seen, the absence of newly built substations in Plan 1 yields significant topological differences with respect to the optimal solution. Thus, new load nodes, which are fed by the two newly installed substations at the optimal solution, are supplied by the existing substations at nodes 51 and 52 in Plan 1. As another distinctive investment-related feature, no branch undergoes replacement at the optimal solution, whereas, for Plan 1, two branches are subject to replacement, namely branches 1-9 and 1-51.

Detailed cost information for Plan 1 is also presented in Table III. For this solution, the largest investment cost is incurred at stage 5 when a relatively larger number of branch-related expansion decisions are made, as shown in Fig. 4. Similar to the optimal solution attained by the proposed approach and Plan 5, the costs related to maintenance, production, losses, and reliability rise along stages due to the demand growth, whereas no load shedding is required.

As for computational burden, the proposed approach attained optimality in 46.97 min. In contrast, the heuristic took 107.40 min to identify the pool of candidate solutions with individual computing times ranging between 0.45 min for Plan 1 and 51.64 min for Plan 6. Thus, the proposed approach substantially reduced the overall computational effort required by the heuristic by 56.3%.

This case study is also useful to substantiate the use of the approximate network model. In this regard, in Table IV, results for the injections at substations and the minimum nodal voltage magnitudes at the last time stage provided by the proposed approach have been compared with those achieved with a full ac load flow model. Moreover, Table V presents statistical information on the errors in the values of branch current flows, injections at substations, and nodal voltage

<table>
<thead>
<tr>
<th>Stage</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Investment</td>
<td>0.102</td>
<td>0.005</td>
<td>0.011</td>
<td>0.006</td>
<td>0.112</td>
<td>0.000</td>
<td>0.002</td>
<td>0.008</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>Maintenance</td>
<td>0.013</td>
<td>0.014</td>
<td>0.017</td>
<td>0.018</td>
<td>0.022</td>
<td>0.022</td>
<td>0.022</td>
<td>0.023</td>
<td>0.025</td>
</tr>
<tr>
<td></td>
<td>Production</td>
<td>1.247</td>
<td>1.444</td>
<td>1.711</td>
<td>1.965</td>
<td>2.231</td>
<td>2.633</td>
<td>2.850</td>
<td>3.193</td>
<td>3.431</td>
</tr>
<tr>
<td></td>
<td>Losses</td>
<td>0.049</td>
<td>0.055</td>
<td>0.069</td>
<td>0.090</td>
<td>0.085</td>
<td>0.097</td>
<td>0.108</td>
<td>0.120</td>
<td>0.130</td>
</tr>
<tr>
<td></td>
<td>Load shedding</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Reliability</td>
<td>0.301</td>
<td>0.337</td>
<td>0.417</td>
<td>0.561</td>
<td>0.518</td>
<td>0.595</td>
<td>0.665</td>
<td>0.735</td>
<td>0.819</td>
</tr>
<tr>
<td>2</td>
<td>Investment</td>
<td>0.007</td>
<td>0.007</td>
<td>0.003</td>
<td>0.006</td>
<td>0.016</td>
<td>0.003</td>
<td>0.002</td>
<td>0.012</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>Maintenance</td>
<td>0.012</td>
<td>0.013</td>
<td>0.014</td>
<td>0.015</td>
<td>0.018</td>
<td>0.020</td>
<td>0.020</td>
<td>0.022</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>Production</td>
<td>1.231</td>
<td>1.416</td>
<td>1.676</td>
<td>1.928</td>
<td>2.151</td>
<td>2.518</td>
<td>2.719</td>
<td>3.044</td>
<td>3.278</td>
</tr>
<tr>
<td></td>
<td>Losses</td>
<td>0.057</td>
<td>0.069</td>
<td>0.083</td>
<td>0.103</td>
<td>0.124</td>
<td>0.159</td>
<td>0.181</td>
<td>0.209</td>
<td>0.223</td>
</tr>
<tr>
<td></td>
<td>Load shedding</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Reliability</td>
<td>0.372</td>
<td>0.486</td>
<td>0.618</td>
<td>0.762</td>
<td>1.001</td>
<td>1.259</td>
<td>1.411</td>
<td>1.619</td>
<td>1.722</td>
</tr>
<tr>
<td>3</td>
<td>Investment</td>
<td>0.007</td>
<td>0.007</td>
<td>0.003</td>
<td>0.006</td>
<td>0.016</td>
<td>0.101</td>
<td>0.002</td>
<td>0.008</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>Maintenance</td>
<td>0.012</td>
<td>0.013</td>
<td>0.014</td>
<td>0.015</td>
<td>0.018</td>
<td>0.021</td>
<td>0.021</td>
<td>0.022</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>Production</td>
<td>1.231</td>
<td>1.416</td>
<td>1.676</td>
<td>1.928</td>
<td>2.151</td>
<td>2.518</td>
<td>2.719</td>
<td>3.044</td>
<td>3.278</td>
</tr>
<tr>
<td></td>
<td>Losses</td>
<td>0.057</td>
<td>0.069</td>
<td>0.088</td>
<td>0.108</td>
<td>0.126</td>
<td>0.148</td>
<td>0.167</td>
<td>0.184</td>
<td>0.209</td>
</tr>
<tr>
<td></td>
<td>Load shedding</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Reliability</td>
<td>0.372</td>
<td>0.486</td>
<td>0.611</td>
<td>0.771</td>
<td>0.945</td>
<td>1.097</td>
<td>1.235</td>
<td>1.344</td>
<td>1.519</td>
</tr>
</tbody>
</table>
magnitudes at the last time stage resulting from using the proposed approximate network model instead of the full ac load flow model. As can be seen, the results presented in Tables IV and V validate the accuracy of the approximate network model and are thus consistent with the experience reported by Haffner et al. [19], [20], who stated on page 922 of [19]: “In the networks that have been analyzed, mean errors in the magnitudes of nodal voltages obtained using this approximate model were adequate for the study objectives, when compared with the solution for load flow.”.

Moreover, the quality of the piecewise linear approximation for energy losses has also been assessed by solving the problem with quadratic losses wherein binary investment and utilization variables were fixed to the values attained by the proposed mixed-integer linear program with a three-piece linearization. For this particular case study, the total cost only differs by 0.21%, thereby corroborating the appropriateness of the piecewise linear approximation in terms of solution quality.

Table VI presents the results from a sensitivity analysis corresponding to the application of a scaling factor to nodal peak demands. The second column of this table lists the present value of the total cost obtained by the proposed approach for each value of the scaling factor. The same information is reported in columns 3–8 for the six expansion plans identified by the heuristic approach presented in [8], where the best solution is highlighted. The last column reports the cost reductions attained by the proposed approach over the best solutions obtained by the heuristic. Note that the proposed approach yields significant cost savings ranging between 5.5% for a 50% scaling factor and 9.0% for a scaling factor equal to 100%. These results further corroborate the superiority of the proposed approach over the heuristic described in [8].

Finally, Fig. 5 depicts the impact of the reliability cost coefficient, $C^R$, on the investment cost. As can be seen, for values ranging between $0/MWh and $2800/MWh, the investment cost is constant, thereby revealing that the reliability cost has no influence on the expansion plan. For values of the cost coefficient between $2800/MWh and $4480/MWh, the investment cost abruptly grows. In contrast, the increase rate of this cost significantly declines for larger values of $C^R$. This stabilization of the investment cost shows that such values of $C^R$ have a similar effect on the solution of the expansion planning problem.

V. CONCLUSION

This paper has presented a new model based on mixed-integer linear programming to incorporate reliability in the
multistage distribution network expansion planning problem. Unlike existing methods based on metaheuristics and heuristics, the proposed tool allows the distribution planner to devise the optimal investment plan while jointly accounting for investment, operating, and reliability costs. To that end, the conventional simulation-based reliability assessment is equivalently characterized through a set of nonlinear equations based on the operation of a fictitious system under a set of conditions associated with each system node. The subsequent use of some well-known linearization schemes allows us to recast such equations as linear expressions, thereby giving rise to a mixed-integer linear program that is suitable for efficient off-the-shelf software.

The economic and computational benefits of co-optimizing the reliability cost with investment and operating costs have been illustrated with a case study based on a 54-node benchmark over a 10-year planning horizon. Moreover, numerical results reveal the superiority of the proposed approach over a previously-reported heuristic.

The proposed modeling framework paves the way for the consideration of practical aspects such as distributed generation, which is currently underway, and multiple branch outages, which requires further research effort and numerical studies that will constitute part of our future work. Further work will also address the uncertainty associated with demand and renewable-based generation. Other interesting avenues of research are (i) the consideration of the recently-proposed convexifications of the ac network model, (ii) the extension of the model to account for additional post-fault network reconfiguration, the effect of the nonlinear dependence of the cost of the expected energy not supplied on interruption durations, and the installation of protection devices, and (iii) the calculation of reliability indices for meshed networks.

REFERENCES


Fig. 5. Impact of $C^R$ on the investment cost.
Gregorio Muñoz-Delgado (S’14) received the Ingeniero Industrial degree and the M.Sc. degree from the Universidad de Castilla-La Mancha, Ciudad Real, Spain, in 2012 and 2013, respectively. He is currently pursuing the Ph.D. degree at the Universidad de Castilla-La Mancha.

His research interests are in the fields of power systems planning, operations, and economics.

Javier Contreras (SM’05–F’15) received the B.S. degree in electrical engineering from the University of Zaragoza, Zaragoza, Spain, in 1989, the M.Sc. degree from the University of Southern California, Los Angeles, CA, USA, in 1992, and the Ph.D. degree from the University of California, Berkeley, CA, USA, in 1997.

He is a Professor at the Universidad de Castilla-La Mancha, Ciudad Real, Spain. His research interests include power systems planning, operations, and economics, as well as electricity markets.

José M. Arroyo (S’96–M’01–SM’06) received the Ingeniero Industrial degree from the Universidad de Málaga, Málaga, Spain, in 1995, and the Ph.D. degree in power systems operations planning from the Universidad de Castilla-La Mancha, Ciudad Real, Spain, in 2000.

From June 2003 through July 2004, he held a Richard H. Tomlinson Postdoctoral Fellowship at the Department of Electrical and Computer Engineering of McGill University, Montreal, QC, Canada. He is currently a Full Professor of electrical engineering at the Universidad de Castilla-La Mancha. His research interests include operations, planning, and economics of power systems, as well as optimization.