Risk–based modeling, simulation and optimization for the integration of renewable distributed generation into electric power networks

Rodrigo Mena

CentraleSupélec

THÈSE

présentée par

Rodrigo MENA

pour l'obtention du

GRADE DE DOCTEUR

Spécialité: Génie Industriel

Laboratoire d'accueil: Laboratoire de Génie Industriel

SUJET:

Risk–based Modeling, Simulation and Optimization for the Integration of Renewable Distributed Generation into Electric Power Networks

Modélisation, simulation et optimisation basée sur le risque pour l'intégration de génération distribuée renouvelable dans des réseaux de puissance électrique

soutenue le: 30 juin 2015

devant un jury composé de:

Enrique DROGUETT	University of Maryland	Reviewer
Federica FOIADELLI	Politecnico di Milano	Reviewer
Michel MINOUX	Université Pierre et Marie CURIE	Examiner
Guillaume SANDOU	CentraleSupélec	Examiner
Vera SILVA	Électricité de France R&D	Examiner
Enrico ZIO	CentraleSupélec	Supervisor
	❖	
Martin HENNEBEL	CentraleSupélec	Co–supervisor
Yan–Fu LI	CentraleSupélec	Co–supervisor

2015ECAP0034

To my brother . . .

Acknowledgements

I would like to acknowledge my thesis supervisor, Professor Enrico Zio, for having given me the opportunity to develop this Ph.D. work. Thanks for trusting me, for all the guidance and advice to improve my knowledge of how to do forefront research, for constantly spreading enthusiasm and eagerness in learning and for having encouraged me to overcome the difficulties encountered over these three years. To me, It has been a privilege to work in Prof. Zio's team and undoubtedly a real fruitful experience, both professionally and personally.

I extend my appreciation to my co–supervisors, Dr. Yan–fu Li and Dr. Martin Hennebel, and to a former collaborator of the Chair SSEC, Dr. Carlos Ruiz, whose generous support has been fundamental in the development of this work. I am grateful to you for having been always interested and attentive in following the progress of this thesis, for the willingness and patience to discuss and teach me whenever I needed technical counsel.

To all the members of the jury, Professors Federica Foiadelli, Guillaume Sandou, Michel Minoux and Enrique Droguett, and Dr. Vera Silva, my sincere gratitude for your time and consideration. It was an absolute challenge having defended my thesis in front of such excelled academics and professionals. I thank you all for the constructive remarks, comments and suggestions and, in particular, Prof. Foiadelli and Prof. Droguett that kindly accepted to be the reviewers of the manuscript.

My warmest gratitude and recognition go to Corinne Ollivier, Delphine Martin and Sylvie Guillemain, the three loving assistants (a.k.a. supreme bosses) of the Industrial Engineering Department (LGI), because of their unconditional good will to help us, the usually lost–in–translation Ph.D. students. In addition, I would like to thank my friends and colleagues from LGI, for all the unforgettable moments.

To conclude, words are not enough to express how beholden I am to my beloved parents, siblings and Elisa, they just know that I am.

Abstract

Renewable distributed generation (DG) is expected to continue playing a fundamental role in the development and operation of sustainable, efficient and reliable electric power systems, by virtue of offering a practical alternative to diversify and decentralize the overall power generation, benefiting from cleaner and safer energy sources. The integration of renewable DG in the existing electric power networks poses socio–techno–economical challenges, which have attracted substantial research and advancement.

In this context, the focus of the present thesis is the design and development of a modeling, simulation and optimization framework for the integration of renewable DG into electric power networks. The specific problem considered is that of selecting the technology, size and location of renewable generation units, under technical, operational and economic constraints. Within this problem, key research questions to be addressed are: (i) the representation and treatment of the uncertain physical variables (like the availability of diverse primary renewable energy sources, bulk–power supply, power demands and occurrence of components failures) that dynamically determine the DG–integrated network operation, (ii) the propagation of these uncertainties onto the system operational response and the control of the associated risk and (iii) the intensive computational efforts resulting from the complex combinatorial optimization problem of renewable DG integration.

For the evaluation of the system with a given plan of renewable DG, a non–sequential Monte Carlo simulation and optimal power flow (MCS–OPF) computational model has been designed and implemented, that emulates the DG–integrated network operation. Random realizations of operational scenarios are generated by sampling from the different uncertain variables distributions, and for each scenario the system performance is evaluated in terms of economics and reliability of power supply, represented by the global cost (CG) and the energy not supplied (ENS), respectively. To measure and control the risk relative to system performance, two indicators are introduced, the conditional value–at–risk (CVaR) and the CVaR deviation (DCVaR).

For the optimal technology selection, size and location of the renewable DG units, two distinct multi–objective optimization (MOO) approaches have been implemented by heuristic optimization (HO) search engines. The first approach is based on the fast non–dominated sorting genetic algorithm (NSGA–II) and aims at the concurrent minimization of the expected values of CG and ENS, then ECG and $EENS$, respectively, combined with their corresponding $CVaR(CG)$ and $CVaR(ENS)$

values; the second approach carries out a MOO differential evolution (DE) search to minimize simultaneously ECG and its associated deviation $DCVaR(CG)$. Both optimization approaches embed the MCS–OPF computational model to evaluate the performance of each DG–integrated network proposed by the HO search engine.

The challenge coming from the large computational efforts required by the proposed simulation and optimization frameworks has been addressed introducing an original technique, which nests hierarchical clustering analysis (HCA) within a DE search engine.

Examples of application of the proposed frameworks have been worked out, regarding an adaptation of the IEEE 13 bus distribution test feeder and a realistic setting of the IEEE 30 bus sub–transmission and distribution test system. The results show that these frameworks are effective in finding optimal DG–integrated networks solutions, while controlling risk from two distinct perspectives: directly through the use of CVaR and indirectly by targeting uncertainty in the form of DCVaR. Moreover, CVaR acts as an enabler of trade–offs between optimal expected performance and risk, and DCVaR integrates also uncertainty into the analysis, providing a wider spectrum of information for well–supported and confident decision making.

The main original contributions of the thesis work here presented reside in: framing the problem of optimal technology selection, size and location of renewable generation units, within an integrated simulation and optimization approach that takes into consideration multiple uncertain operational inputs through the developed MCS–OPF, allows assessing and controlling risk by introducing CVaR and DCVaR measures, and copes with computational complexity by embedding HCA into the HO search engine.

Keywords: renewable distributed generation, uncertainty, risk, simulation, optimization, conditional value–at–risk, conditional value–at–risk deviation, genetic algorithm, differential evolution, hierarchical clustering analysis

Résumé

Il est prévu que la génération distribuée par l'entremise d'énergie de sources renouvelables (DG) continuera à jouer un rôle clé dans le développement et l'exploitation des systèmes de puissance électrique durables, efficaces et fiables, en vertu de cette fournit une alternative pratique de décentralisation et diversification de la demande globale d'énergie, bénéficiant de sources d'énergie plus propres et plus sûrs. L'intégration de DG renouvelable dans les réseaux électriques existants pose des défis socio–technico–économiques, qu'ont attirés de la recherche et de progrès substantiels.

Dans ce contexte, la présente thèse a pour objet la conception et le développement d'un cadre de modélisation, simulation et optimisation pour l'intégration de DG renouvelable dans des réseaux de puissance électrique existants. Le problème spécifique à considérer est celui de la sélection de la technologie, la taille et l'emplacement de des unités de génération renouvelable d'énergie, sous des contraintes techniques, opérationnelles et économiques. Dans ce problème, les questions de recherche clés à aborder sont: (i) la représentation et le traitement des variables physiques incertains (comme la disponibilité de les diverses sources primaires d'énergie renouvelables, l'approvisionnements d'électricité en vrac, la demande de puissance et l'apparition de défaillances de composants) qui déterminent dynamiquement l'exploitation du réseau DG–intégré, (ii) la propagation de ces incertitudes sur la réponse opérationnelle du système et le suivi du risque associé et (iii) les efforts de calcul intensif résultant du problème complexe d'optimisation combinatoire associé à l'intégration de DG renouvelable.

Pour l'évaluation du système avec un plan d'intégration de DG renouvelable donné, un modèle de calcul de simulation Monte Carlo non–séquentielle et des flux de puissance optimale (MCS–OPF) a été conçu et mis en œuvre, et qui émule l'exploitation du réseau DG–intégré. Réalisations aléatoires de scénarios opérationnels sont générés par échantillonnage à partir des différentes distributions des variables incertaines, et pour chaque scénario, la performance du système est évaluée en termes économiques et de la fiabilité de l'approvisionnement en électricité, représenté par le coût global (CG) et l'énergie non fournie (ENS), respectivement. Pour mesurer et contrôler le risque par rapport à la performance du système, deux indicateurs sont introduits, la valeur–à–risque conditionnelle (CVaR) et l'écart du CVaR (DCVaR).

Pour la sélection optimale de la technologie, la taille et l'emplacement des unités DG renouvelables, deux approches distinctes d'optimisation multi–objectif (MOO) ont été mis en œuvre

par moteurs de recherche d'heuristique d'optimisation (HO). La première approche est basée sur l'algorithme génétique élitiste de tri non-dominé (NSGA–II) et vise à la réduction concomitante de l'espérance mathématique de CG et de ENS, dénotés ECG et $EENS$, respectivement, combiné avec leur valeurs correspondent de $CVaR(CG)$ et $CVaR(ENS)$; la seconde approche effectue un recherche à évolution différentielle MOO (DE) pour minimiser simultanément ECG et s'écart associé $DCVaR(CG)$. Les deux approches d'optimisation intègrent la modèle de calcul MCS–OPF pour évaluer la performance de chaque réseau DG–intégré proposé par le moteur de recherche HO.

Le défi provenant de les grands efforts de calcul requises par les cadres de simulation et d'optimisation proposée a été abordée par l'introduction d'une technique originale, qui niche l'analyse de classification hiérarchique (HCA) dans un moteur de recherche de DE.

Exemples d'application des cadres proposés ont été élaborés, concernant une adaptation du réseau test de distribution électrique IEEE 13–nœuds et un cadre réaliste du système test de sous–transmission et de distribution IEEE 30–nœuds. Les résultats montrent que les cadres proposés sont efficaces dans la recherche des solutions de réseaux DG–intégrés optimales, tout en contrôlant les risques à partir de deux perspectives distinctes: directement par l'utilisation du CVaR et indirectement par en ciblant l'incertitude sous la forme du DCVaR.

Les principales contributions originales de la thèse présentée ici résident dans: encadrer le problème de la sélection optimale de la technologie, la taille et l'emplacement des unités de génération renouvelable d'énergie, dans une approche intégrée de simulation et d'optimisation qui tient compte des multiples variables opérationnelles incertains à travers du MCS–OPF développé, permet d'évaluer et contrôler le risque en introduisant les mesures CVaR et DCVaR, et fait face à la complexité de calcul en intégrant HCA dans le moteur de recherche HO.

Mots clés: génération distribuée, énergie renouvelable, incertitude, risque, simulation, optimisation, valeur–à–risque conditionnelle, écart du valeur–à–risque conditionnelle, algorithme génétique, évolution différentielle, analyse de classification hiérarchique

Table of contents

List of figures **xv**

List of tables **xvii**

Nomenclature **xix**

I Thesis **1**

1 Introduction **3**

1.1 Electric power systems 3

1.1.1 Transmission & distribution of electric power 5

1.1.2 Diversification and decentralization of electrical power generation 6

1.2 Renewable distributed generation 7

1.2.1 Potential benefits and challenges 7

1.3 Research objectives and original contributions 10

1.4 Thesis structure 12

2 Renewable DG–integrated electric power network modeling **15**

2.1 System representation 15

2.1.1 Components classification 15

2.1.2 Network topology 16

2.2 Uncertain operational inputs 18

2.2.1 Power demands and energy price 18

2.2.2 Bulk–power supply 20
2.2.3 Solar photovoltaic generation 20
2.2.4 Wind turbines generation 21
2.2.5 Electrical vehicles 22
2.2.6 Storage devices 23
2.2.7 Components availability state 24
2.3 Stochastic operation modeling 24
2.3.1 Non–sequential Monte Carlo simulation 24
2.3.2 Optimal power flow 25
2.4 Performance evaluation 27
2.4.1 Global cost 27
2.4.2 Energy not supplied 28
2.4.3 Uncertainty and risk measurement 28

3 Renewable DG–integrated electric power network planning 31
3.1 Optimal DG technologies selection, sizing and allocation 31
3.1.1 Optimization strategies formulation 31
3.2 Heuristic optimization & MCS–OPF simulation frameworks 34
3.2.1 Non–dominated sorting genetic algorithm II–based approach 35
3.2.2 MOO differential evolution–based approach 39

4 Computational Challenge 43
4.1 Clustering in heuristic optimization 43
4.2 Hierarchical clustering & differential evolution 44

5 Applications 51
5.1 A risk–based MOO and MCS–OPF simulation framework for the integration of renewable DG and storage devices 51
5.1.1 IEEE 13–bus test feeder case 51
5.1.2 Conditional value–at–risk: expected performance and risk trade–off 52

5.1.3 Optimal DG–integrated network plans 55
5.1.4 Brief summary 56
5.2 Hierarchical clustering analysis and differential evolution (HCDE) for optimal integration of renewable DG 57
5.2.1 IEEE 13–bus test feeder case 57
5.2.2 Quantification of the benefits of HCDE 58
5.2.3 Identification of time complexity conditions and limits 61
5.2.4 Optimal DG–integrated network plans 63
5.2.5 Brief summary 64
5.3 A MOO and MCS–OPF simulation framework for risk–controlled integration of DG 65
5.3.1 IEEE 30 bus sub-transmission and distribution test system 65
5.3.2 Conditional value–at–risk deviation: expected performance, uncertainty and risk trade–off 67
5.3.3 Optimal DG–integrated network plans 71
5.3.4 Brief summary 73

6 Conclusions **75**
6.1 Original contributions 75
6.2 Future work 77

Appendix A IEEE 13–bus test feeder data: Application 5.1 **79**

Appendix B IEEE 13–bus test feeder data: Application 5.2 **83**

Appendix C IEEE 30–bus sub–transmission & distribution system data: Application 5.3 **87**

References **91**

II Papers **99**

List of figures

1.1 World CO_2 emissions by sector [1] 4
1.2 Example of a passive T&D network 5
1.3 Example of an active T&D network 6
1.4 Thesis structure diagram 13

2.1 IEEE 30 bus sub–transmission and distribution test system diagram 16
2.2 Example of a nodal daily load profile, hourly normally distributed 19
2.3 Proportional correlation energy price vs aggregated load 19
2.4 Example of hourly probability distribution of EV operating states per day 22
2.5 Sampling process 25
2.6 Graphic representation of $VaR_\alpha(x)$, $CVaR_\alpha(x)$ and $DCVaR_\alpha(x)$; $x =$ loss 28

3.1 Flow chart of the proposed NSGA–II MCS–OPF framework 38
3.2 Flow chart of the proposed MOO–DE MCS–OPF framework 41

4.1 Example dendrogram for average linkage HCA 46
4.2 Example of cutoff distance calculation 47
4.3 Flow chart of the proposed HCDE framework 49

5.1 Radial 11–nodes distribution network 52
5.2 Pareto fronts for different values of β 53
5.3 Bubble plots $EENS$ v/s ECG. Diameter of bubbles proportional to $CVaR(ENS)$ (A) and $CVaR(CG)$ (B) 53

5.4 Average total DG power allocated (A) and its breakdown by type of DG: PV (B), W (C), EV (D) and ST (E) 55

5.5 Radial 11-nodes distribution network 58

5.6 $ECG_{\min}$ vs NFE for $NP \in \{10, 20, 30, 40, 50\}$ set in DE 58

5.7 $ECG_{\min}$ vs NFE for each (NP, CCC_T, p_{co}) set in HCDE 59

5.8 CCC behavior per generation g 60

5.9 Empirical NP^g *pdf* for each (NP, CCC_T, p_{co}) set in HCDE 60

5.10 Average total DG power allocated and investment cost for representative (NP, CCC_T, p_{co}) settings 63

5.11 Nodal average total DG power for representative (NP, CCC_T, p_{co}) settings 64

5.12 IEEE 30 bus sub-transmission and distribution test system diagram 66

5.13 Set of non-dominated solutions: Pareto front 67

5.14 Set of non–dominated solutions with iso–CVaR (A) and iso–DCVaR (B) curves 68

5.15 Empirical CG probability density functions 69

5.16 Total average renewable power installed. Cases $ECG^{\min}$ (A), $CVaRCG^{\min}$ (B) and $DCVaRCG^{\min}$ (C) 71

5.17 Nodal average power by type of generator 72

A.1 Mean (A) and variance (B) values of nodal power demand daily profiles 79

A.2 Hourly per day probability data of EV operating states. $\rho = -1$: charging, $\rho = 0$: disconnected, $\rho = 1$: discharging 81

B.1 Mean and standard deviation values of normally distributed nodal power demand daily profile 83

B.2 Hourly per day probability data of EV operating states 84

C.1 Accumulated mean (A) and standard deviation (B) values of nodal load daily profiles 87

List of tables

4.1 Example hierarchical structure outcome 45

5.1 Objective functions: expected and CVaR values of selected Pareto front solutions . . . 54
5.2 Average, minimum and maximum total DG power allocated per node 56
5.3 Asymptotic time complexity of the algorithms 62
5.4 Ratio κ for each (NP, CCC_T, p_{co}) 63
5.5 MOO-DE and MCS–OPF parameters 65
5.6 Ratio of power usage by type of generator 70

A.1 Feeders characteristic and technical data [2] 79
A.2 Bulk–power supply parameters 80
A.3 Parameters of PV, W, EV and ST technologies [3–5] 80
A.4 Failure rates of feeders, MG and DG units [3–6] 80
A.5 Investment, fixed O&M and variable O&M costs of MG and DG [6–8] 81

B.1 Feeders characteristic and technical data [2, 4, 9] 83
B.2 Power sources parameters and technical data [3–8, 10] 84

C.1 T&D lines characteristics and technical data [11–13] 88
C.2 Power generators technical data and uncertain model parameters [3–5, 7, 10, 14] . . 89
C.3 Power generators failure and repair rates and costs [3, 4, 7, 10, 14, 15] 89

Nomenclature

Roman Symbols

BGT	Budget	($)
$B_{(i,i')}$	Susceptance of T&D line (i, i')	(p.u.)
CCC	Cophenetic correlation coefficient	
CCC_T	Cophenetic correlation coefficient threshold	
cdf	Cumulative distribution function	
CG	Global cost	($/h)
${}^{\mathrm{G}}CG^{\Omega}$	Sample of NS realizations of ${}^{\mathrm{G}}CG_s^{\omega}$	($)
${}^{\mathrm{GTD}}CG^{\Omega}$	Sample of NS realizations of ${}^{\mathrm{GTD}}CG_s^{\omega}$	($/h)
$CI_j + COf_j$	Fixed investment and operating cost	($)
C_{LS}	Load shedding cost	($/MWh)
$\mathscr{C}$	Group (cluster) of individuals	
${}^{\mathrm{G}}CO^{\Omega}$	Sample of NS realizations of ${}^{\mathrm{G}}CO_s^{\omega}$	($)
${}^{\mathrm{GTD}}CO^{\Omega}$	Sample of NS realizations of ${}^{\mathrm{GTD}}CO_s^{\omega}$	($/h)
$COv_{(i,i')}$	Variable operating cost of the T&D line (i, i')	($/MWh)
COv_j	Variable operating cost of the power generator type j	($/MWh)
D	Set of hours of the day	
$\bar{D^1}$	Average of the Euclidean distances $d^1_{\iota,J}$	
d_C	Crowding distance	
D_L	Daylight interval	
$d_{\min}$	Minimum linkage distance $d^z_{\iota,J}$	
$d_{NC=4}$	Distance to form at least four clusters	
Δt	Time interval	(h)
D^z	Matrix of linkage distances	
D^z	Linkage distance between groups $\mathscr{C}_\iota$ and $\mathscr{C}_J$	
${}^{(\mathrm{a})}ECG$	Expected global cost (FW (a))	($)
${}^{(\mathrm{b,c})}ECG$	Expected global cost (FWs (b) & (c))	($/h)
$EENS$	Expected energy not supplied	(MW h)
ℓ	Non–dominated front index	

ENS	Energy not supplied	(MW h)
ENS^{Ω}	Sample of NS realizations of ENS_s^{ω}	(MW h)
EP_{max}	Maximum value of energy price	($/MWh)
EP^{Ω}	Sample of NS realizations of $EP_{t_s}, \forall \omega_s \in \Omega$	($/MWh)
EP_t	Energy price at hour of the day t	($/MWh)
EV	Electrical vehicles type of DG	
F	Differential variation amplification factor	
$f_{j,t}(\rho_{j,t})$	Hourly *pdf* of operational states of EV type j	
FF_j	Fill factor	
$f_{i,t}(L_{i,t}\|\mu_{i,t}, \sigma_{i,t})$	Normal pdf of nodal power load at hour of the day t	
$f_j(P_j\|\mu_j, \sigma_j)$	Normal *pdf* of bulk–power supply MG type j	
$f_i(H_i\|\alpha_i, \beta_i, p_L, H_i^*)$	Adjusted solar irradiance Beta *pdf*	
$f_j(J\|, J_{S_j}, M_{T_j})$	Uniform *pdf* of ST type j	
$f_i(U_i\|\alpha_i^*, \beta_i^*)$	Wind speed Weibull *pdf*	
G	Class of power generation components	
${}^{\mathrm{G}}CO_s^{\omega}$	Aggregated operating cost for scenario ω_s	($)
${}^{\mathrm{GTD}}CO_s^{\omega}$	Aggregated operating cost for scenario ω_s	($/h)
H_i	Solar irradiance at node i	
H_i^*	p_L%ile of the non–adjusted solar irradiance Beta *pdf* at node i	
I_{MPP_j}	Current at maximum power point of PV type j	(A)
inc	Incentive for power generation from DG sources	($/MWh)
I_{SC_j}	Short circuit current of PV type j	(A)
J	Level of charge in the storage device	(MJ)
J_{S_j}	Specific energy of the active chemical in ST type j	(MJ/kg)
f_k	Availability states stationary *pdf* of component k	
k_{I_j}	Current temperature coefficient of PV type j	(mA/°C)
L	Power demand	(MW)
$l_{k,k'}$	Linkage distance $d_{\iota,J}^{z^*}$ where X_k and $X_{k'}$ become members of $\mathscr{C}_{\iota,J}$	
$\bar{\mathscr{L}}$	Average of the linkage distances $l_{k,k'}$	
$\mathscr{L}$	Matrix of linkage distances $l_{k,k'}^{z^*} = d_{\iota,J}^{z^*}$	
L_i	Active power load at node i	(MW)
L_{max_i}	Maximum value of power load at node i	(MW)
$LS_{i,s}$	Load shedding at node i for scenario ω_s	(MW)
LS_i^{Ω}	Sample of NS realizations of $LS_{i,s}$	(MW h)
m	Total number of types of bulk–power suppliers	
MG	Subclass of bulk–power generation components	
M_{T_j}	Mass of active chemical in ST type j	(kg)
N	Set of all nodes in the electric power network	
n_c	Number of photovoltaic cells	

$\mathscr{F}$	Non–dominated front	
NF	Number of objective functions	
NFE	Number of objective function evaluation	
NG_{max}	Maximum number of generations	
NP	Population size	
NS	Number of operational scenarios realizations	
NZ	Number of steps	
$p^{-}_{j,t}$	Hourly *pdf* of charging operational state of EV type j	
p_C	Crossover probability	
p_{co}	Cutoff level coefficient	
$p^{+}_{j,t}$	Hourly *pdf* of discharging operational state of EV type j	
pdf	Probability density function	
PF^{DG}	DG penetration factor	
$PF^{DG}_{max_j}$	Maximum allowed penetration factor of DG technology type j	
P_{G_i}	Active power injected or generated at node i	(MW)
P_i	Active power leaving node i	(MW)
$P_{i,j}$	Power output of generator type j at node i	(MW)
$P_{j,t}$	Power output of generator type j at hour of the day t	(MW)
P_j	Available bulk–power supply MG type j	(MW)
p_L	Probability that the hour of the day t is in the interval D_L	
p_M	Mutation probability	
$P_{max_{(i,i')}}$	Power rating of the T&D line (i, i')	(MW)
P_{max_j}	Maximum power capacity of generator type j	(MW)
$P^{\Omega}_{U_{i,j}}$	Sample of NS realizations of $P_{U_{i,j,s}}, \forall \omega_s \in \Omega$	(MW)
POP	Population	
P_{R_j}	Rated power of generator type j	(MW)
$P_{U_{i,j,s}}$	Used power from the generator type j	(MW)
PV	Solar photovoltaic type of DG	
$p^{0}_{j,t}$	Hourly *pdf* of disconnected operational state of EV type j	
Q	Matrix of location and capacity size of power generation units	
$q_{i,j}$	Number of power generation units type j allocated at node i	
Q^M	Matrix of location and capacity size of MG	
q'	Non–negative integer	
Q^R	Matrix of location and capacity size of renewable DG	
r	Total number of types of renewable technologies	
RG	Subclass of renewable DG components	
S_{ref}	Reference apparent power in the network	(MVA)
ST	Storage devices type of DG	
t	Hour of the day	

T_{A_i}	Ambient temperature at node i	(°C)
t^{Δ}	Operational scenarios of duration i	(h)
TD	Class of transmission and distribution lines components	
t^{H}	Lifetime of the project	(h)
t_j^J	Upper bound of discharging time interval of ST type j	(h)
T_{No_j}	Nominal cell operation temperature of PV type j	(°C)
t^{ρ}	Time of residence in the operating state $\rho_{j,t}$	(h)
U_{A_j}	Average wind speed of W type j	(m/s)
U_{CI_j}	Cut–in wind speed of W type j	(m/s)
U_{CO_j}	Cut–out wind speed of W type j	(m/s)
U_i	Wind speed at node i	(m/s)
k_{V_j}	Voltage temperature coefficient of PV type j	(mV/°C)
V_{MPP_j}	Voltage at maximum power point of PV type j	(V)
V_{OC_j}	Open circuit voltage of PV type j	(V)
W	Wind turbines type of DG	
Y	Set of transmission and/or distribution lines	
$\mathbb{Z}^*$	Set of non–negative integers	
Greek Symbols		
α_i, β_i	Shape parameters of Beta distribution at node i	
α_i^*, β_i^*	Scale and Shape parameters of Weibull distribution at node i	
α	Confidence level	
Γ	Gamma function	
η_k	Availability state of component k	
λ_{F_k}	Failure rate of component k	(n/h)
λ_{R_k}	Repair rate of component k	(n/h)
$\mu_{i,t}$	Mean of nodal power load at hour of the day t	(MW)
μ_j	Mean of bulk–power supply MG type j	(MW)
Ω	Set of all the NS realizations of operational scenarios ω	
ω	Set of sampled variables or operational scenarios	
Φ	Standard Normal *cdf*	
ϕ	Standard Normal *pdf*	
$\rho_{j,t}$	Possible operational states of EV type j	
$\sigma_{i,t}$	Standard deviation of nodal power load at hour of the day t	(MW)
σ_j	Standard deviation of bulk–power supply MG type j	(MW)
τ_j	Maximum number of units of DG technology j to allocate	
δ_i	Voltage angle at node i	
Subscripts		
g	Generation index	
i	Node index	

$\imath$	Group (cluster) of individuals index
j	Power generator type index
$\jmath$	Group (cluster) of individuals index
k	Component index
k	Individual index
k'	Individual index
s	Scenario index
t	Hour of the day index
z	Steps index

Acronyms / Abbreviations

CHP	Combined heat and power
CLLI	Contingency load loss index
CO_2	Carbon dioxide
CVaR	Conditional value–at–risk
DCVaR	Conditional value–at–risk deviation
DE	Differential evolution
DG	Distributed generation
ECOST	Expected value of non–distributed energy cost
EENS	Expected energy not supplied
FW	Framework
HCA	Hierarchical clustering analysis
HCDE	Hierarchical clustering differential evolution
HO	Heuristic optimization
LHS	Latin Hypercube Sampling
MCS	Monte Carlo simulation
MINLP	Mixed–integer non–linear problem
MOO	Multi–objective optimization
NSGA–II	Fast non–dominated sorting genetic algorithm
O&M	operation and maintenance
OPF	Optimal power flow
SAIDI	System average interruption duration index
SAIFI	System average interruption frequency index
SO	Single objective
T&D	Transmission and distribution
TPES	Total primary energy supply
TVD	Total voltage deviation
VaR	Value–at–risk

I Thesis

1 Introduction

The present thesis describes the works done in the design and development of a modeling, simulation and optimization framework for the integration of renewable distributed generation (DG) into electric power networks. Specifically, the problem considered is that of technology selection, sizing and allocation of renewable DG units, taking into account uncertainty and risk.

This introductory chapter is organized as follows. Section 1.1 acquaints the reader with the research context, presenting the motivations for advancements in sustainable electric power systems and the role that renewable DG plays in the diversification and decentralization of power generation. Section 1.2 discusses the potential benefits of DG and the major challenges involved in its integration into existing power networks. In Section 1.3, key research questions and ensuing objectives are formulated in the effort of contributing to shape frameworks for well–supported decision–making in optimal DG integration. Finally, a schematic representation of the overall structure of the thesis and general descriptions of the developed frameworks are provided in Section 1.4.

1.1 Electric power systems

The adverse environmental effects accompanying the intensive and prolonged use of fossil fuels are not anymore a midterm conjecture but an ongoing reality. Even though initiatives with considerable participation, like the Kyoto Protocol and the Copenhagen Accord, have set targets to limit carbon dioxide (CO_2) emissions in order to constrain the global temperature rise to less than 2°C relative to the pre–industrial level between 2008–12, no overall mitigation has been achieved [1]. Indeed, greenhouse gas emissions from fuel combustion have been steadily increasing during the last decade, mainly by cause of the accelerated economic and population growths of industrialized countries, that do not face emission targets and, consequently, demand progressively larger amounts of energy from fossil fuels.

Given the above, the development of environmental sustainability has become a worldwide imperative for all sectors of highly compromised human activities such as industry, transport, residential and electricity and heat generation. In particular, the electric power generation sector is

considered an important point of concern, since it subscribes 42% of the global CO_2 emissions, by far being the largest contribution, as shown in Figure 1.1 [1].

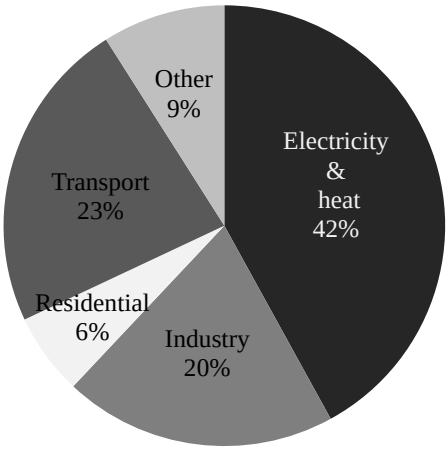

Figure 1.1 World CO_2 emissions by sector [1]

In the pursuit of sustainable energy systems, the heightening of regulatory targets and the engagement of leading–role countries to participate in international agreements certainly set a challenging framework for advancement, but these measures do not constitute solutions by themselves. Nevertheless, the awareness on the global environmental problem has triggered a revolution towards cleaner, safer and more reliable systems along all the energy value chain [10, 15–17].

Electric power systems are the 'spinal cord' of the energy value chain and they are facing a stimulating transition across all their three main components, generation, transmission and distribution, led by both technological development of new equipment and devices and enhanced actions in planning, operation and management strategies and driven by the opportunity of generating electrical power by making use of low–carbon and, in preference, renewable energy sources. This offers a great opportunity to overcome the growing energy demand, mitigating greenhouse emissions and alleviating the energy market instability associated to the depletion of fossil fuels [18].

The integration of renewable generation into power systems is implemented mainly in two ways, depending on the scale of the available primary energy sources: large–scale renewable generation is predominantly connected upstream sub–transmission and distribution networks, acting as a conventional bulk–power plant, while small–scale renewable generation units are allocated close or directly on the 'customers site of the meter', which is known as distributed generation (DG) [19–21]. Both plan of actions offer the advantage of generating clean electric power, but it is particularly DG that is playing a crucial role because it provides the possibility of concurrently diversifying and decentralizing the overall power generation.

The connection of diverse renewable DG units onto sub–transmission and distribution networks implies conceptual and operational transformations which are explained in the next section.

1.1.1 Transmission & distribution of electric power

Transmission and distribution (T&D) of electric power has been traditionally a passive 'fit–and–forget' strategy, characterized by unidirectional power flows supplied by centralized generation systems [16, 19, 22, 23]. Moreover, the restrictive structure of conventional T&D settings makes it difficult to supply power to remote areas, due to the extra T&D expenses associated and the risk of compromising a large portion of power supply due to the occurrence of outages upstream sub–transmission and distribution networks [24, 25].

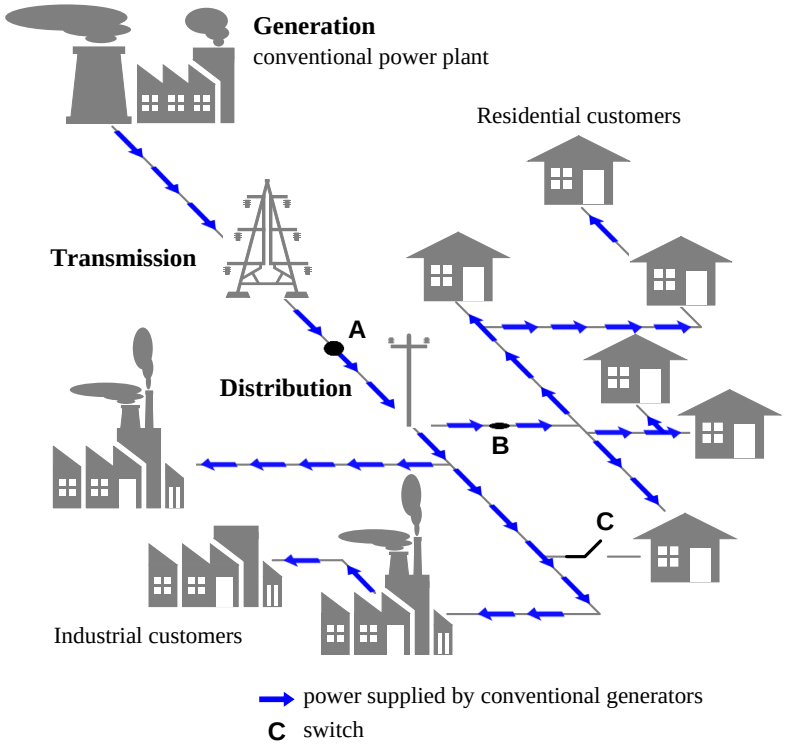

Figure 1.2 Example of a passive T&D network

Figure 1.2 illustrates the constricted configuration of a typical passive T&D network. It can be noticed that any contingency upstream point **A** may produce a large loss of power supply, in this example, affecting mainly the industrial customers given the radial topology of that portion of the network. For residential customers, the impact of an outage level at points **A** or **B** can be deaden by closing the switch at point **C**, which provides a meshed character to that portion of the network and, therefore, some level of redundancy. Notwithstanding mesh–structured T&D networks are more flexible in terms of reliability of power supply, their operation is more expensive and complicated since infrastructure and operation and maintenance costs (O&M) of T&D lines increase, more control and protection devices are needed with the respective synchronization in their operation [9, 26, 27].

1.1.2 Diversification and decentralization of electrical power generation

On a large scale, hydro–power technology is the major contributor to carbon–free generation, representing the 2.4% of the global total primary energy supply (TPES) in 2012 [28]. Over decades, it has played an important environmental role but current expansions of hydro–power capacity are negligible with respect to other pollutant technologies like oil, coal, gas, bio–fuels and waste, which state the 31.4, 29.0, 21.3, 10.0% of the TPES in 2012 [28], respectively, and account for the 99% of the global CO_2 emissions [1]. Indeed, global hydro–power additions decreased in 2013 [29], indicating a loss of priority in mitigating CO_2 emissions by use of this technology. Moreover, large scale hydro–power plants operate under the conventional passive logic, then, no further decentralization can be achieved from them.

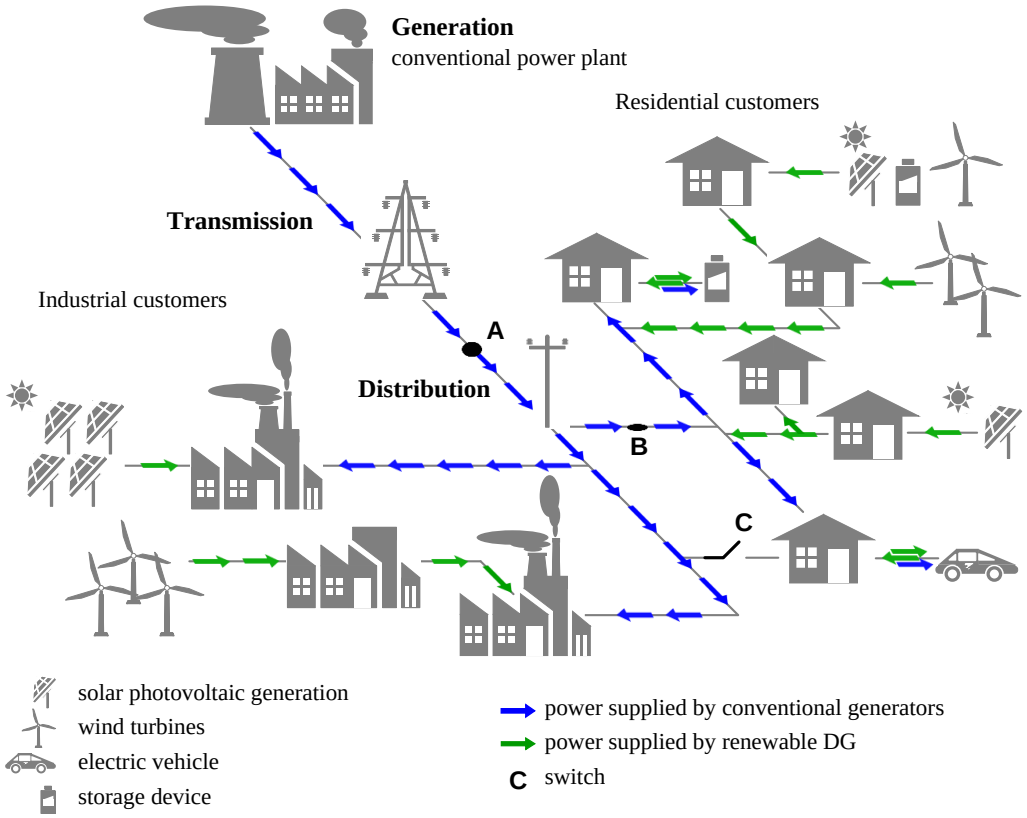

Figure 1.3 Example of an active T&D network

It is the rapid expansion of non–hydro renewable capacity, that in 2014 rose globally by approx. 7% (350 (TW h))[29], that is making a difference. With the integration of renewable DG technologies, such as small–scale hydro, solar photovoltaic, on and off–shore wind and storage devices, electric power systems are evolving towards an active operational strategy, with possibly bidirectional power flows and, thus, more reliable, decentralized and diversified sources [16, 19],

as shown in Figure 1.3. In the Figure, the impact on the power supply of outages at points **A** and/or **B** is smoothed by the deployment of different renewable DG units throughout the T&D network.

Finally, renewable DG technologies make use of local/regional renewable energy sources, which boosts them in view of the requested environmental sustainability, while engaging the participation of households, small businesses and specialized power companies and offering interesting techno–economical benefits.

1.2 Renewable distributed generation

In the literature, DG is commonly defined as modular (independent) generation units located at or in the neighborhood of power demand spots, connected to sub–transmission and/or distribution networks rather than high voltage transmission [19–21, 27, 30]. The modular characteristic refers to their smaller scale power capacity, generally in the range of generation 1 (kW)–5 (MW) [25]. It is important to mention that some DG applications can reach a considerable power capacity, between 50 and 300 (MW); however, there is disagreement among authors about at what extent these can be considered as DG.

The technological spectrum of DG is formed by two main groups: carbon–free and low–carbon 'efficient' generation devices. The first group includes small hydro and tidal turbines, solar photovoltaic panels, wind turbines, geothermal steam–turbines, etc., whereas the second one adds biomass, waste and combined heat and power (CHP) devices and installations, based on reciprocating engines and combustion gas turbines [19]. Moreover, storage devices and electric vehicles are also feasible technologies accompanying DG integration, to increase the overall efficiency of activated electric power networks [6, 19, 31, 32].

Without loss of generality, in the present thesis work the focus of attention is oriented to carbon–free DG given their superior contribution to environmental sustainability. Below, the potential technical and economical benefits on one side and the possible operational complications on the other are discussed.

1.2.1 Potential benefits and challenges

Under the assumption that DG power is 'dispatchable', i.e., DG units are able to provide and sell energy in parallel to or as competitors of the centralized bulk–power supply to satisfy the system demand, and given the fact that by integrating DG into an existing network the power flows through shorter and possibly bidirectional paths, the main technical benefits that can be achieved are [3, 4, 9, 10, 16, 17, 19, 22, 33–44]:

- Improvement of reliability of power supply.

- Reduction of power losses.
- Voltage stability.
- Enhancement of power quality.
- Alleviation of T&D lines congestion.
- Power supply autonomy of rural/isolated areas.

The most representative economical benefits expected from the integration of renewable DG are the following [4, 10, 16, 17, 19, 23, 34, 35, 38–41, 43, 45, 46]:

- Deferral of investments for conventional power plants and T&D upgrades.
- Reduction of T&D O&M costs.
- Decrement of generation costs and energy price.
- Reduction of fossil fuel costs.
- Diminution of price volatility associated to fossil fuels.
- Reduction of investment risks.

However, planners and operators have to face complex economic, regulatory, technical and operational constraints, within which the dynamics introduced by renewable DG may generate complications that can counteract the potential benefits [25, 26, 35]. Indeed, the traditional passive power systems structures are designed to operate with unidirectional power flows, and in order to incorporate renewable DG, the need for more protection and control devices may be significant. Reactive power may also result compromised since many renewable DG technologies supply only active power, affecting also the voltage stability and the introduction of harmonics [19].

The above–mentioned challenges have prompted substantial research and development. In particular, DG planning has been a fundamental baseline of advancement to properly seize its potential advantages. The specific problem associated to DG planning consists in selecting the technology, size and location of renewable generation units while respecting the techno–economical, regional and regulatory constraints imposed by the existing system.

One of the main difficulties associated to DG planning is the proper modeling of the intrinsic uncertain behavior of primary renewable energy sources (e.g. solar irradiance, wind speed and water inflow, in the case of solar photovoltaic, wind turbines and hydro–power technologies, respectively) and of the stochastic occurrence of unexpected events on the DG units, such as failures and stoppages, that may interrupt or curtail the power generation capacity. Indeed, these sources of uncertainty come on top of those already present in the operating power systems, outage events due to failures (or stoppages) of T&D lines and/or conventional power generators, variability and growth of power demand, volatility in the energy price and fluctuations in the bulk–power supply, among others. As

a consequence, for any proposed DG–integrated network plan, the uncertain operational conditions need to be considered in the system response as calculated by solving power flow equations under a number of representative scenarios [19, 42, 43, 47].

The decision making for DG–integrated network planning is usually framed as an optimization problem, which can be especially complex depending on the size and topology of the network, the number of load nodes, number of available DG technologies and the aforementioned uncertainty and the non–linear conditions arising from technical constraints [16, 20, 42]. Optimality of the renewable DG plan is customarily sought with regards to cost–based, operational or technical targets. Among cost–based objectives are the costs of energy and fuel for generation, investments, O&M, energy purchased from conventional power plants, energy losses, CO_2 emissions, taxes, incentives, incomes, etc. [3, 4, 7, 9, 10, 15, 16, 19, 22, 38, 48–57]. Operational objective functions consider the performance and reliability of power supply in terms of indicators like the expected energy not supplied (EENS) [37, 58], contingency load loss index (CLLI) [9], expected value of non–distributed energy cost (ECOST) , system average interruption frequency index (SAIFI), system average interruption duration index (SAIDI) [16, 22, 50], among others. Concerning technical targets, the most commonly used indexes are the total voltage deviation (TVD) [51] and energy losses [10, 44]. Obviously, the optimal solutions must also comply with the system technical constraints like generation capacities, T&D lines rating and voltage drops [19, 25].

In the search for optimal DG–integrated network configurations under uncertainty, the use of only expected or cumulative indicators as objective function(s) hinders the possibility of controlling the risk associated to the optimal solution(s): the expected performance of an optimal DG–integrated network may be satisfactory but be exposed to high variability or to risky scenarios with non–negligible probabilities.

One way to account for uncertainty and risk is to frame DG planning as a portfolio optimization problem, in which the different types of DG technologies are treated analogously to financial assets [42, 46, 59–65].

In portfolio optimization theory, the mean–variance approach is the most common [66] and has been applied to DG planning also [42, 46, 59, 62, 65, 67]. However, from a risk–perspective it entails a drawback that cannot be ignored: the variance measure includes the values of performance that symmetrically fall short of or exceed the expected or mean value; in the search for optimal DG technologies portfolios, lower levels of uncertainty (variance) in the performance function can be obtained with portfolios that lead to rarer occurrences of both beneficial and/or non–desired (risky) scenarios. Then, for controlling the risk side, it is necessary to introduce additional indicators that provide information on the extent of asymmetry of the performance function, weighting accordingly the risky part of it, e.g., by skewness and kurtosis indicators that estimate the asymmetry and peakedness of a probabilistic performance function, respectively [68].

Likewise derived from portfolio optimization theory, direct risk–based frameworks have been formulated and applied to address DG planning problems under uncertain conditions. The most widely used risk measures are the value–at–risk (VaR) and conditional value–at–risk (CVaR) [61, 63, 64, 69–73], which focus on non–desirable performance outcomes given by a portfolio relative to a specific confidence level or percentile.

Even though integrated portfolio optimization approaches, for instance mean–variance–skewness or mean–variance–CVaR, present robustness advantages by conjointly controlling the level of uncertainty and risk associated to the expected value of performance of different portfolios, they can considerably increase the complexity of the concurrent optimization problem. Considering the expected or mean value and the necessary deviation and risk measures as representative objectives of a portfolio increases significantly the number of objective functions to be simultaneously optimized, further constraining the feasible space and, eventually, hindering the understanding of the information delivered to the decision–makers.

Moreover, optimization problems associated to DG planning are, in general, non–linear, non–convex and combinatorial in nature. Non–linearity can be given by the power flow equality constraints and/or objectives functions involving power losses. The non–convex and combinatorial characteristics are mainly due to the decision variables representing discrete (integer) locations, number of units and type of technology of DG. Non–convex mixed–integer non–linear problems (MINLP) are difficult to solve by conventional mathematical models, with multiple local optima and at least non–deterministic polynomial–time hard (NP–hard) computational complexity [7, 20]. This calls for alternative methods of solutions, like heuristic optimization techniques (HO) belonging to the class of evolutionary algorithms (EAs), which have been proposed as a most effective way of solution. These methods are suited to cope straightforwardly with non–convex combinatorial problems, discontinuous feasible spaces, non–linear and non–differentiable objective functions [20, 42].

1.3 Research objectives and original contributions

The objectives of the present thesis focus on the design and development of a modeling, simulation and optimization framework for the integration of renewable DG into electric power networks. Specifically, the problem considered is the selection of the technology, size and location of multiple DG units, under technical, operational and economic constraints. The following key research questions are addressed:

(i) Representation and treatment of the uncertain operational inputs, like the availability of diverse primary renewable energy sources, bulk–power supply, power demands and occurrence of components failures.

(ii) Propagation of uncertainties onto the model of system operational response and control of the associated risk.

(iii) Computational efforts resulting from the combinatorial optimization problem associated to renewable DG integration.

In answer to the key research questions formulated, the main original contributions of this thesis work are:

(a) Design and implementation of a non–sequential Monte Carlo simulation (MCS) and optimal power flow (OPF) computational model, denoted MCS–OPF, that emulates the T&D network operation integrating a given renewable DG plan. Random realizations of operational scenarios are generated by sampling from the different uncertain variables models, evaluating for each scenario the performance of the DG–integrated network in terms of economics and reliability of power supply, represented by the global cost (CG) and the energy not supplied (ENS), respectively.

(b) Integration of two indicators to measure and control uncertainty and risk, namely conditional value–at–risk (CVaR) and conditional value–at–risk deviation (DCVaR), respectively.

(c) With respect to the optimal technology selection, size and location of the renewable DG units, two distinct multi–objective optimization (MOO) strategies have been implemented by heuristic optimization (HO) search engines, in which the MCS–OPF model is nested to assess the performance of each DG–integrated network proposed along the evolutionary searching process:

- In the first approach, the fast non–dominated sorting genetic algorithm (NSGA–II) is used for simultaneous minimization of the expected values of CG and ENS (ECG and $EENS$, respectively) combined with their respective $CVaR(CG)$ and $CVaR(ENS)$ values.
- The second approach performs a MOO differential evolution (DE) search to minimize concurrently ECG and its associated deviation $DCVaR(CG)$.

(d) To cope with the large computational efforts required by the developed MOO frameworks with nested MCS–OPF, an original technique is introduced which embeds hierarchical clustering analysis (HCA) within a DE search engine. The technique identifies, in a controlled manner, groups of similar individuals (DG plans) in the DE population and, then, evaluates ECG performing MCS–OPF on selected representative individuals of the groups only, thus reducing the number of objective function evaluations in each iteration of the DE evolution loop.

1.4 Thesis structure

This thesis manuscript is divided into two parts. The first part is composed by six Chapters which introduce the readers to the research context, the motivations and objectives, and present in detail the methodological developments performed to address the specific problem of optimal DG planning under uncertainty. In particular, Chapter 2 provides the basics of the modeling of a DG–integrated electric power network, focusing on the system representation, the different models use to treat the uncertain operational inputs considered, the construction of the MCS–OPF computational model and the respective evaluation of the system performance based on economic, and uncertainty and risk measures. In Chapter 3, the formulation of diverse optimization strategies to address the optimal DG planning problem and the HO with nested MCS–OPF frameworks developed are presented. Chapter 4 briefly summarizes the use of clustering techniques in HO to increase the computational performance and, then, gives a complete description of the development of the DE search engine with embedded HCA. In Chapter 5, examples of applications of the proposed frameworks are illustrated, with reference to an adaptation of the IEEE 13 bus distribution test feeder [2] and a realistic setting of the IEEE 30 bus sub–transmission and distribution test system of literature [11]. Finally, Chapter 6 draws the conclusions of the thesis work carried out and proposes some opportunities for future research advancements. Figure 1.4 shows a diagram with the overall vision of the thesis structure.

The second part, includes the collection of papers published or under revision, which are the result of the research work here performed, and that the reader can refer to for further details. Paper (i) presents a direct risk–based simulation and MOO framework for the integration of renewable DG and storage based on NSGA–II, introducing the CVaR to find optimal DG plans, trading–off expected performance and risk. Paper (ii) addresses the challenge of reducing the computational efforts required to implement HO search engines with nested MCS–OPF: the technique developed integrates HCA and DE for optimal integration of renewable DG and, by defining control parameters, adapts itself along the evolutionary search determining whether it is convenient to perform clustering of the decision variables or not and at what scale to, then, reduce the number of objective function evaluations. Paper (iii) introduces a MOO framework for risk–controlled integration of renewable DG into electric power systems, which is based on DE search and MCS–OPF. This framework measures uncertainty in the system performance by the use of DCVaR that, due to its axiomatic relation to the CVaR, allows the conjoint control of risk.

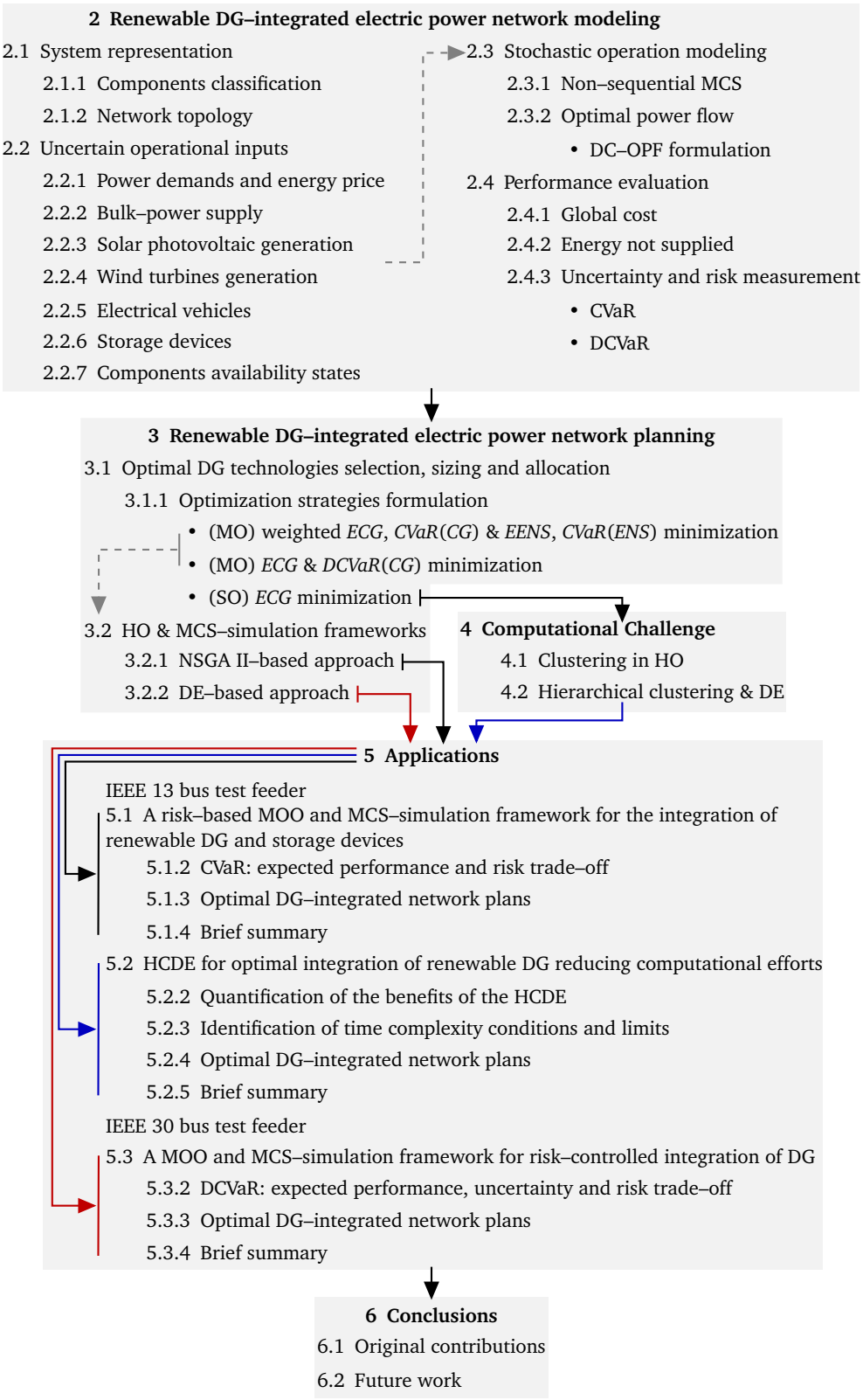

Figure 1.4 Thesis structure diagram

2 Renewable DG–integrated electric power network modeling

In this Chapter, the MCS–OPF is presented, including representation of the DG–integrated network, the modeling of the different uncertain operational inputs considered, the process of generating the random operational scenarios and the OPF problem formulation. The definition of the performance evaluation function, based on economics, reliability of power supply and uncertainty–risk measurement, is introduced as well.

2.1 System representation

2.1.1 Components classification

The starting point for modeling a DG–integrated network is the definition of the type of components involved. For example in Figure 2.1 which illustrates the IEEE 30 bus sub–transmission and distribution test system, three main classes of components can be identified: power generators, T&D lines and loads.

The power generator class G considers the subclasses containing all the different types of conventional bulk power suppliers and renewable technologies to be integrated, denoted by MG and RG, respectively. In this study, four different types of DG technologies are considered: solar photovoltaics, wind turbines, electric vehicles and storage devices. The nomenclature denoting the classes, subclasses and types of components is summarized as follows:

- G: Class of power generation components, $G = MG \cup RG$.
  - MG: Subclass of bulk–power generation components.
  - RG: Subclass of renewable DG components, $RG = PV \cup W \cup EV \cup ST$.
    - PV: Solar photovoltaics.
    - W: Wind turbines.
    - EV: Electrical vehicles.

- ST: Storage devices.

- TD: Transmission and distribution lines components.
- L: Power demand components.

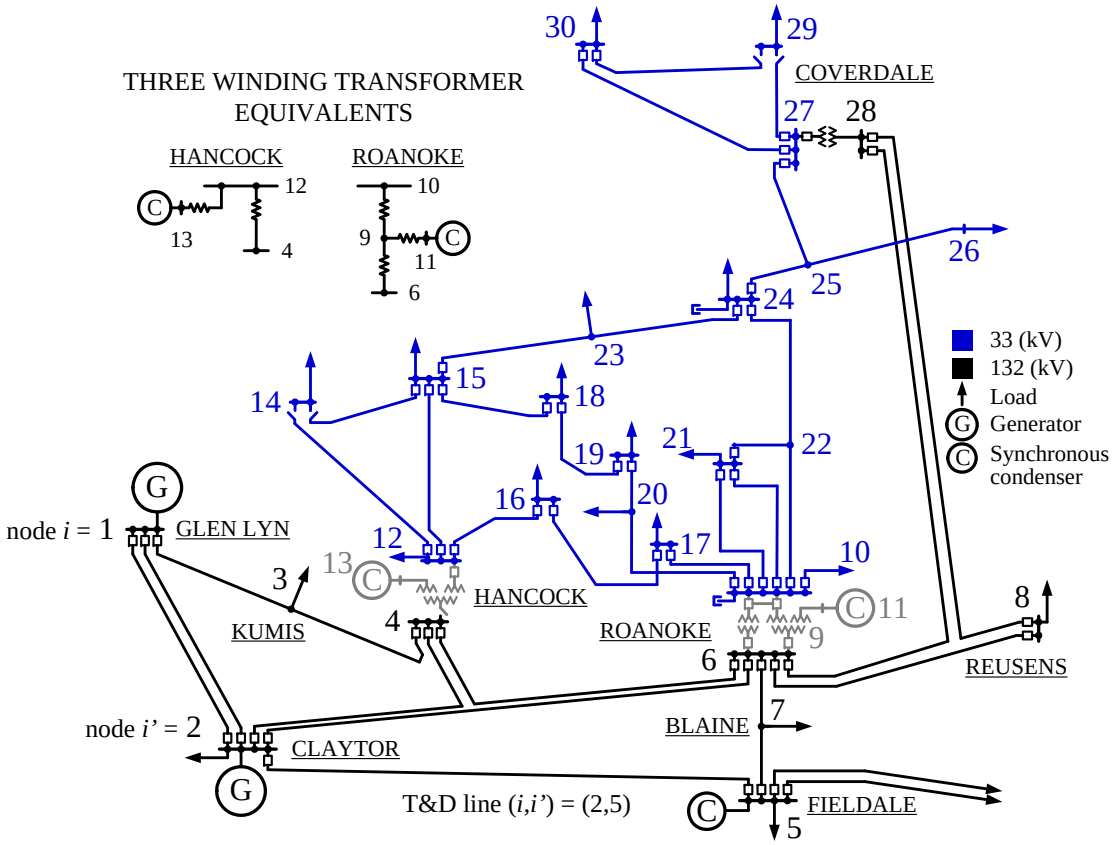

Figure 2.1 IEEE 30 bus sub–transmission and distribution test system diagram

2.1.2 Network topology

The network topology is described as a graph, i.e., a set of nodes and the various connections that link them. The nodes represent spatial points at which generation components (MG and RG) and loads are located or can be allocated, whereas the connections between nodes are the transmission and distribution lines. A single node and the set of all nodes are indicated by the index i and $N = \{i : i \in \{1, 2, \ldots, n\}\}$, respectively, where n – total number of nodes in the network.

Consequently, the links connecting nodes define the set of transmission and distribution lines as follows:

$$Y = \{(i, i') : \text{nodes } i \text{ and } i' \text{ are connected}, \forall i, i' \in N\} \tag{2.1}$$

Assuming stationary operational conditions, the network performance is considered to be dictated by the locations and magnitudes of the power available in each generation unit, the loads and the technical limits of the T&D lines. To indicate the location and capacity size of the different types of generation units present in the network, the matrix Q, $\forall i \in N$ is defined by the following expression:

$$
\begin{array}{c}
\overbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}^{(m+r)\text{ types of power generators } G} \\
\overbrace{\qquad\qquad\qquad\qquad\qquad}^{m\text{ types of bulk–power suppliers } MG}\quad\overbrace{\qquad\qquad\qquad\qquad\qquad}^{r\text{ types of renewable technologies } RG} \\
\begin{array}{ccccccccccc}
G_1 & \cdots & G_j & \cdots & G_m & & G_{1+m} & \cdots & G_{j+m} & \cdots & G_{r+m}
\end{array} \\
Q = \left[\begin{array}{ccccc|ccccc}
q_{1,1} & \cdots & q_{1,j} & \cdots & q_{1,m} & q_{1,1+m} & \cdots & q_{1,j+m} & \cdots & q_{1,r+m} \\
\vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
q_{i,1} & \cdots & q_{i,j} & \cdots & q_{i,m} & q_{i,1+m} & \cdots & q_{i,j+m} & \cdots & q_{i,r+m} \\
\vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
q_{n,1} & \cdots & q_{n,j} & \cdots & q_{n,m} & q_{n,1+m} & \cdots & q_{n,j+m} & \cdots & q_{n,r+m}
\end{array}\right] \\
\underbrace{\qquad\qquad\qquad\qquad\qquad}_{Q^M}\quad\underbrace{\qquad\qquad\qquad\qquad\qquad}_{Q^R}
\end{array}
\tag{2.2}
$$

where, Q^M, Q^R – matrices of location and capacity size of bulk–power suppliers (MG) and renewable power generators (RG), respectively, j – power generator type index, m – total number of types of bulk–power suppliers, r – total number of types of renewable technologies and $[Q]_{i,j} = q_{i,j}$ – non–negative integer that specifies the number of units of the power generator type j allocated at node i:

$$
q_{i,j} = \begin{cases} q' \in \mathbb{Z}^* & \text{if } q' \text{ units of } G_j \text{ are allocated at node } i,\ \forall G_j \in \{G_1, \ldots, G_{m+r}\} \\ 0 & \text{otherwise} \end{cases}
\tag{2.3}
$$

Then, a proposed plan of DG integration is represented by the matrix Q^R and its concatenation to the fixed locations and sizes of the bulk–power generators, already present in the existing network, Q^M results in the complete static deployment of all power generator components $Q = [Q^M|Q^R]$. Therefore, the complete representation of a DG integrated is given by the pair $([Q], \{Y\})$. Any physical component, G or TD, is assumed to be affected by the stochastic occurrence of failures, conditioning dynamically the functionality of power generators and the lines through which the power flows. Furthermore, the magnitude of power available in each generation unit is subject to the intrinsic uncertain behavior of the corresponding primary energy source and, under the assumption that generators act as price–takers, the economic conditions depend on the variability of the power demands [4, 52].

The aforementioned uncertain conditions significantly affect the operation of a given DG–integrated network, therefore, modeling the diverse sources of variability becomes essential to

emulate the operational response of the system for a large representative combination of possible scenarios and, ultimately, be able to assess its probabilistic performance relative to preset target functions.

2.2 Uncertain operational inputs

In principle, analytical methods are preferable to Monte Carlo simulation (MCS) because of the possibility of achieving accurate solutions; however, their application usually requires simplifications in the modeling which may lead to unrealistic results. An example of this are analytical solutions for optimal DG planning that do not take into account uncertainty or intermittency in power generation and/or load profiles, and networks of low dimensionality [74]. On the other hand, MCS allows a more realistic modeling, because the operation of the network is not analytically solved but simulated, and the overall performance indicators are statistically estimated from virtual operational scenarios realizations [75]. MCS has been found quite suitable for the analysis of electric power networks with multiple sources of uncertainty, e.g., power generation, loads, component failures or degradation processes, etc. [37, 42, 43, 62, 63, 76], but at the expense of requiring more computational resources.

In the present thesis work, we adopt a non–sequential MCS, based on latin hypercube sampling (LHS) [77], to emulate the operation of the DG–integrated network, considering the realizations of uncertain operational variables as independent on previous realizations, so as to seize the advantages of MCS without overly increasing the computational efforts.

2.2.1 Power demands and energy price

Overall power demand profile in an electric power network, as well as single nodal load profiles, can be inferred from historical data as daily load curves, in which to each hour of the day corresponds one specific level of load [10, 52]. In addition, power demands can be considered uncertain following normal distributions [42, 76]. Here, both models are integrated, adopting normally distributed nodal load profiles for which their respective mean and standard deviation parameters vary depending on the hour of the day $t \in D = \{1, \ldots, 24\}$, as shown in Figure 2.2. Then, the power demand at node i is modeled as:

$$f_{i,t}(L_{i,t}|\mu_{i,t},\sigma_{i,t}) = \begin{cases} \dfrac{\phi(\xi(L_{i,t},\mu_{i,t},\sigma_{i,t}))}{\sigma_{i,t} Z(\mu_{i,t},\sigma_{i,t})} & \forall L_{i,t},\mu_{i,t},\sigma_{i,t} \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.4}$$

$$\xi(L_{i,t},\mu_{i,t},\sigma_{i,t}) = \frac{L_{i,t}-\mu_{i,t}}{\sigma_{i,t}}; \quad Z(\mu_{i,t},\sigma_{i,t}) = 1-\Phi\left(\frac{\mu_{i,t}}{\sigma_{i,t}}\right) \tag{2.4a}$$

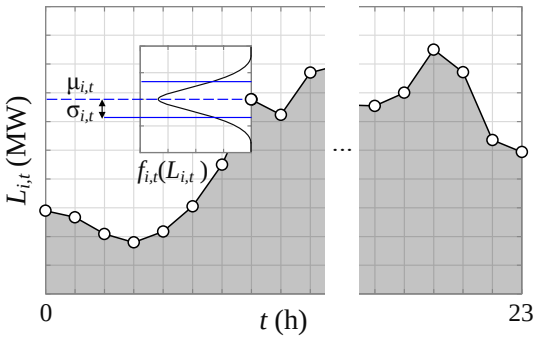

Figure 2.2 Example of a nodal daily load profile, hourly normally distributed

where, $f_{i,t}(L_{i,t}|\mu_{i,t},\sigma_{i,t})$ – truncated Normal probability density function, $L_{i,t}$ (MW) – power demand at node i at hour of the day t, $\mu_{i,t}$, $\sigma_{i,t}$ (MW) – Normal distribution mean and standard deviation, respectively, ϕ, Φ – standard Normal (*pdf*) and its cumulative distribution function (*cdf*), respectively.

The power generators in the network are assumed to be price–takers, for which the value of the energy price is correlated with the aggregated power demand. As an intermediate approximation of existing studies in [4, 52, 53] (Figure 2.3), the proportional correlation used in this study can be expressed as:

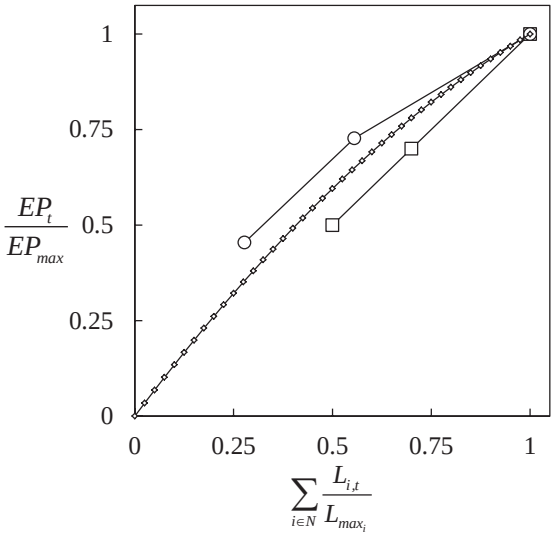

Figure 2.3 Proportional correlation energy price vs aggregated load

$$EP_t(L_{i,t}|EP_{max},L_{max_i}) = EP_{max}\sum_{i\in N}\frac{L_{i,t}}{L_{max_i}}\left(-0.38\left(\sum_{i\in N}\frac{L_{i,t}}{L_{max_i}}\right)+1.38\right) \tag{2.5}$$

where, EP_t (\$/MWh)– energy price at hour of the day t, EP_{max} (\$/MWh)– maximum value of energy price and L_{max_i} MW– maximum value of power load at node i.

2.2.2 Bulk–power supply

Bulk–power stands for the power supply coming from conventional power plants (MG) already existing in the network. These are rather stable and are connected to the network at sub–transmission or distribution transformers to provide the voltage level of the customers. The stochastic behavior of the available power in these sources is represented following normal distributions [20, 78], with small standard deviation and truncated by the maximum capacity of generation.

$$f_j(P_j|\mu_j,\sigma_j,P_{max_j}) = \begin{cases} \dfrac{\phi(\xi(P_j,\mu_j,\sigma_j))}{\sigma_j Z(\mu_j,\sigma_j)} & \forall P_j \in [0,P_{max_j}],\mu_j,\sigma_j \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.6}$$

$$\xi(P_j,\mu_j,\sigma_j) = \frac{P_j-\mu_j}{\sigma_j}; \quad Z(\mu_j,\sigma_j) = \Phi(\frac{P_{max_j}-\mu_j}{\sigma_j}) - \Phi(\frac{\mu_j}{\sigma_j}) \tag{2.6a}$$

where $\forall j \in \{j : G_j \in MG\}$, $f_j(P_j|\mu_j,\sigma_j,P_{max_j})$– truncated Normal *pdf*, P_j, P_{max_j} (MW) – available bulk power and maximum capacity of the MG generator type j, respectively, μ_j, σ_j (MW) – Normal distribution mean and standard deviation.

2.2.3 Solar photovoltaic generation

PV technologies (PV) converts solar irradiance into electric power through a set of solar cells configured. Commonly, solar irradiance uncertain behavior has been modeled using probabilistic distributions, obtained from long term weather historical data of a particular geographical area. The Beta distribution function has been found particularly suitable to model hourly solar irradiance [10, 79]. The intermittency in the solar irradiation is taken into account defining a daylight interval between 07.00 and 21.00 hours, setting a positive value of solar irradiation H if the value t of the hour of the day is in the subset of $D_L = \{7,\ldots,21\}$ of D, otherwise, the value of solar irradiance is assumed equal to 0 given that t is in the night interval. Thus, the Beta distribution function is adjusted considering the probability $p_L = P(t|t \in D_L)$ that t falls in the daylight interval:

$$f_i(H_i|\alpha_i,\beta_i,p_L,H_i^*) = \begin{cases} \dfrac{H_i^{(\alpha_i-1)}(1-H_i)^{(\beta_i-1)}}{(1-p_L)\mathrm{B}(\alpha_i,\beta_i)} & \forall H_i \in [H_i^*,1],\alpha_i,\beta_i > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.7}$$

$$\mathrm{B}(\alpha_i,\beta_i) = \frac{\Gamma(\alpha_i)\Gamma(\beta_i)}{\Gamma(\alpha_i+\beta_i)} \tag{2.7a}$$

where $f_i(H_i|\alpha_i,\beta_i,p_L,H_i^*)$ – adjusted Beta probability density function, H_i – solar irradiance at node i, α_i, β_i – shape parameters of the corresponding Beta distribution at node i, $H_i^* - p_L$ percentile of the non–adjusted Beta *pdf* $f_i(H_i|\alpha_i,\beta_i)$ and Γ – Gamma function.

Besides the dependence on solar irradiance, the available power output is determined by the technical characteristics of the PV cells and the ambient temperature on site. Then, the available power output provided by n_c solar cells can be obtained from the following equations:

$$P_{i,j}(H_i) = \begin{cases} P'_{i,j}(H_i) & \text{if } 0 \leq P'_{i,j}(H_i) \leq P_{max_j} \\ P_{max_j} & \text{if } P_{max_j} < P'_{i,j}(H_i) \end{cases} \tag{2.8}$$

$$P'_{i,j}(H_i) = n_c FF_j V(H_i) I(H_i) \times 10^{-6} \tag{2.8a}$$

$$T_C(H_i) = T_{A_i} + H_i(T_{No_j} + 20)/0.8 \tag{2.8b}$$

$$I(T_C) = H_i(I_{SC_j} + k_{I_j}(T_C - 25)) \tag{2.8c}$$

$$V(T_C) = V_{OC_j} + k_{V_j} T_C \tag{2.8d}$$

$$FF_j = \frac{V_{MPP_j} I_{MPP_j}}{V_{OC_j} I_{SC_j}} \tag{2.8e}$$

where $\forall j \in \{j : G_j \in PV\}$, $P_{i,j}$ (MW) – power output at node i, P_{max_j} (MW) – maximum power generation capacity, FF_j – fill factor, T_{A_i} (°C) – ambient temperature at node i, T_{No_j} (°C) – nominal cell operation temperature, I_{SC_j} (A) – short circuit current, k_{I_j} (mA/°C) – current temperature coefficient, V_{OC_j} (V) – open circuit voltage, k_{V_j} (mV/°C) – voltage temperature coefficient, and V_{MPP_j} (V), I_{MPP_j} (A) – voltage and current at maximum power, respectively.

2.2.4 Wind turbines generation

Wind power generation (W) is obtained from turbine–alternator devices that transform the kinetic energy of the wind into electric power. The stochastic behavior of the wind speed is commonly represented through probability distribution functions. The Weibull distribution has been widely used to model the randomness of the wind speed in various conditions [10, 20, 34, 73, 79, 80]:

$$f_i(U_i|\alpha_i^*, \beta_i^*) = \begin{cases} \dfrac{\beta_i^*}{\alpha_i^*} \left(\dfrac{U_i}{\alpha_i^*}\right)^{(\beta_i^* - 1)} \exp\left[-\left(\dfrac{U_i}{\alpha_i^*}\right)^{\beta_i^*}\right] & \forall U_i \geq 0, \alpha_i^*, \beta_i^* > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.9}$$

where, $f_i(U_i|\alpha_i^*, \beta_i^*)$ – Weibull probability density funtion, U_i (m/s) – wind speed at node i and α_i^*, β_i^* – scale and shape parameters of the Weibull distribution function at node i, respectively.

Similarly to PV type of technologies, the uncertainty associated to the wind speed and the technical features of a specific type of wind turbine characterize its power output function:

$$P_{i,j}(U_i) = \begin{cases} \dfrac{U_i - U_{CI_j}}{U_{A_j} - U_{CI_j}} P_{R_j} & \text{if } U_{CI_j} \leq U_i < U_{A_j} \\ P_{R_j} & \text{if } U_{A_j} \leq U_i \leq U_{CO_j} \\ 0 & \text{otherwise} \end{cases} \tag{2.10}$$

where $\forall j \in \{j : G_j \in W\}$, P_{R_j} (MW) – rated power and U_{CI_j}, U_{A_j}, U_{CO_j} (m/s) – cut–in, average and cut–out wind speeds, respectively.

2.2.5 Electrical vehicles

In this study, electric vehicles (EVs) are considered as battery electric vehicles with three possible operating states ρ: (−1) charging, (0) disconnected and (1) discharging [32]. When in charging state, an EV act as a power demand, whereas in discharging it injects power into the network. EVs operation is modeled considering them as 'block groups', i.e., EVs sharing similar operational patterns are aggregated into a single block. In fact, it has been observed that EVs present nearly stable daily usage schedules, in addition, modeling them as 'block groups' contributes to the need of avoiding the combinatorial explosion of the model [5].

The power output of one block of EVs is formulated by assigning residence time intervals t^{ρ} to each possible operating state and associating them with the percentage of trips that the vehicles perform by hour of a day [32]. This allows approximating the hourly probability distribution of the operating states per day, as shown Figure 2.4. Then, the random determination of the operating state $\rho_{j,t}$ of a block of EVs of type j, given a specific hour of the day t, is sampled from the corresponding probability $f_{j,t}(\rho_{j,t})$ associated to the occurrence of each state.

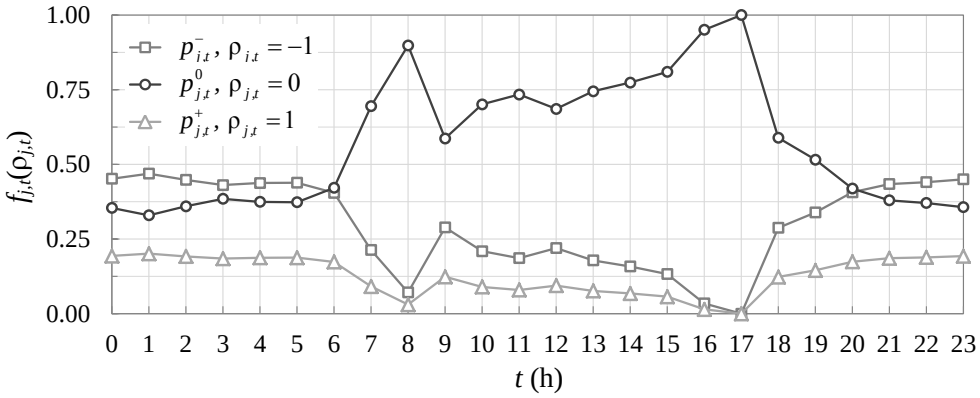

Figure 2.4 Example of hourly probability distribution of EV operating states per day

$$f_{j,t}(\rho_{j,t}|p^{-}_{j,t}, p^{0}_{j,t}, p^{+}_{j,t}) = \begin{cases} p^{-}_{j,t} & \text{if } \rho_{j,t} = -1 \\ p^{0}_{j,t} & \text{if } \rho_{j,t} = 0 \\ p^{+}_{j,t} & \text{if } \rho_{j,t} = 1 \end{cases} \tag{2.11}$$

where $\forall j \in \{j : G_j \in EV\}$, $p^{-}_{j,t}$, $p^{0}_{j,t}$, $p^{+}_{j,t}$ – hourly per day probability distribution of a block of EVs type j for the respective operating states, $\rho_{j,t} = -1$ (charging), $\rho_{j,t} = 0$ (disconnected), $\rho_{j,t} = 1$ (discharging).

Accordingly, the power output for a single EV is calculated using the expression (2.12) below:

$$P_{j,t}(\rho_{j,t}) = \rho_{j,t} P_{R_j} \quad \forall \Delta t \in [0, t^{\rho}] \tag{2.12}$$

where $\forall j \in \{j : G_j \in EV\}$, t^{ρ} – residence time interval for operating state ρ (h), P_{R_j} (MW) – rated power (MW).

2.2.6 Storage devices

Analogously to the EV case, storage devices (ST) are treated as batteries. In reality, these present two main operating states, charging and discharging [81]. However, in accordance to the non–sequential characteristic assumed for the MCS model, for this study the level of charge in the batteries is randomized and the state of discharging is the only one that takes place, making them independent on previous states of charge. The discharging time interval t^{J}_{j} is assigned according to the relation between the batteries rated power of type j, their energy density J_{S_j} and the random level of charge J they present. For this, the discharging action is carried out at a rate equal to the rated power. Then, the power output per unit of mass of active chemical in the battery M_{T_j} is estimated as follows:

$$f_j(J|, J_{S_j}, M_{T_j}) = \begin{cases} \dfrac{1}{J_{S_j} M_{T_j}} & \forall J \in \left[0, J_{S_j} M_{T_j}\right] \\ 0 & \text{otherwise} \end{cases} \tag{2.13}$$

$$P_j(J) = P_{R_j} \quad \forall \Delta t \in [0, t^{J}_{j}] \tag{2.14}$$

$$t^{J} = \frac{J}{P_{R_j}} \tag{2.14a}$$

where $\forall j \in \{j : G_j \in ST\}$, $f_j(J|, J_{S_j}, M_{T_j})$ – Uniform probability density function, J (MJ) – level of charge in the storage device, J_{S_j} (MJ/kg) – specific energy of the active chemical, M_{T_j} (kg) – mass of active chemical, P_{R_j} (MW) – rated power, t^{J}_{j} (h) – upper bound of discharging time interval.

2.2.7 Components availability state

Consistently with the non–sequential nature of the proposed MCS simulation, the availability states of the physical components in the network, generators (G) and T&D lines (TD), are direclty modeled by two–state stationary Markov chains [5, 37], defining two possible operating states: $\eta = 0$ if the corresponding component is non functional (failure) and $\eta = 1$ if the component is available to operate, i.e., be able to generate, transmit or distribute power accordingly to the class of component. Then, the discrete stationary distribution of operating states can be expressed as follows:

$$f_k(\eta_k|\lambda_{F_k}, \lambda_{R_k}) = \begin{cases} \lambda_{F_k}/(\lambda_{F_k} + \lambda_{R_k}) & \eta_k = 0 \\ \lambda_{R_k}/(\lambda_{F_k} + \lambda_{R_k}) & \eta_k = 1 \end{cases} \tag{2.15}$$

where $\forall k \in \{\{k : G_k \in G\} \cup \{k/k = (i, i') \in Y\}\}$, η_k – operating state of component k and λ_{F_k} (n/h) and λ_{R_k} (n/h) – failure and repair rates, respectively.

2.3 Stochastic operation modeling

2.3.1 Non–sequential Monte Carlo simulation

For a given DG–integrated network plan, denoted by the pair $([Q], \{Y\})$, recalling that $Q = [Q^M|Q^R]$, and that represents the locations and number of units of the different power generators and the T&D lines, each uncertain variable is randomly sampled several times by LHS [77] and the inverse transform method [75], for the realization of NS operational scenarios of duration t^{Δ}. For practicality, we define NS as multiple of 24, so each hour of the day t has the same number of realizations $NS/24$. The set ω contains all the sampled variables which constitute an operational scenario which, conjointly to the pair $([Q], \{Y\})$, set the stage for evaluating the response of the network in terms of available power usage, power demand satisfaction and the involved economics. Ω is defined as the set of all the NS realizations of ω, the respective notation is given by Equations (2.16) and (2.16).

Figure 2.5 shows schematically the sampling process of the uncertain variables from the models presented in the preeciding section.

$$\omega_s = \left\{t_s, L_{i,t_s}, EP_{t_s}, \left\{P_{i,j,s}, \forall G_j \in MG\right\}, \eta_{(i,i'),s}, \eta_{i,j,s}, H_{i,s}, U_{i,s}, \rho_{j,t,s}, J_s\right\} \tag{2.16}$$

$$\Omega = \{\omega_s : s \in \{1, 2, \ldots, NS\}\} \tag{2.17}$$

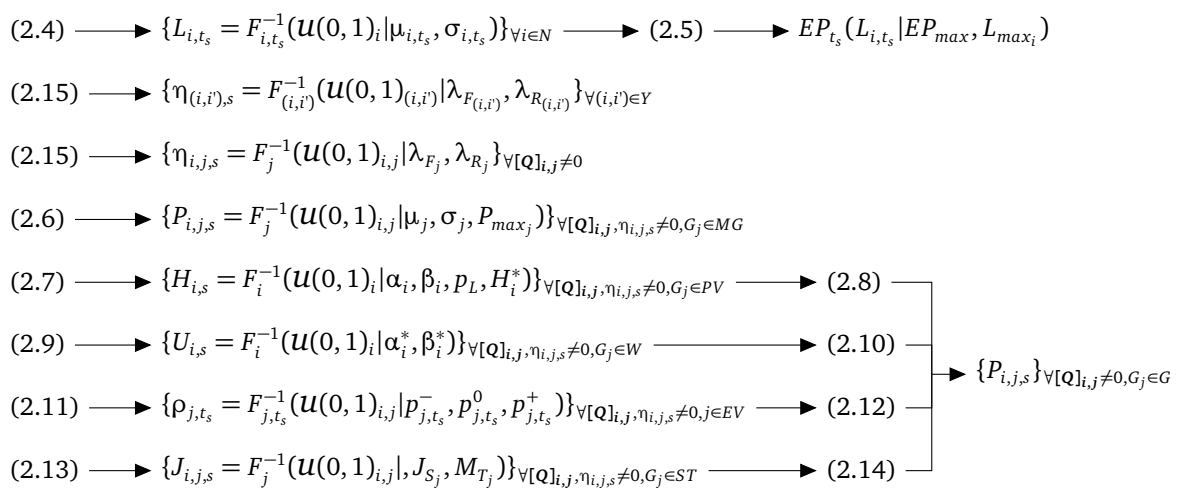

Figure 2.5 Sampling process

2.3.2 Optimal power flow

Power flow analysis is performed by the DC approximation [82] which takes into account only the active power, neglecting power losses, and assumes a flat voltage profile throughout the network. This allows transforming to linear the classic non–linear power flow equality constraints, gaining simplicity and computational tractability. DC power flow is often used in techno–economic analysis of electric power systems, more frequently in transmission [82, 83] but also in distribution networks [83].

The generic DC power flow equations are:

$$P_i = S_{ref}\sum_{i'\in N} B_{(i,i')}(\delta_i - \delta_{i'}) \quad \forall i\in N,(i,i')\in Y \tag{2.18}$$

$$\sum_{i\in N}\left(P_{G_i} - L_i - P_i\right) = 0 \quad \forall i\in N \tag{2.19}$$

where, S_{ref} (MV A) – reference apparent power in the network, P_i (MW) – active power leaving or entering node i, $B_{(i,i')}$ (p.u.) – susceptance of the T&D line (i,i'), δ_i – voltage angle at node i, P_{G_i} (MW) – active power injected or generated at node i and L_i (MW) – load at node i.

The assumptions are:

- the difference between voltage angles is small, i.e., $\sin(\Delta\delta)\approx\Delta\delta$, $\cos(\Delta\delta)\approx 1$.
- the resistance of the T&D lines are neglected, i.e., $R \ll X$, thus, that power losses are neglected as well.
- the voltage profile is flat, constant V, set to 1 (p.u.)

DC–OPF formulation

DC optimal power flow analysis (OPF) is run, taking in input the configuration $([Q],\{Y\})$ and each operating scenario $\omega_s \in \Omega$ and aiming at the minimization of the aggregated operating cost CO. The quantity CO has been defined in two different manners: (i) ${}^{G}CO$ considers solely the operating costs concerning generation of power, (ii) ${}^{GTD}CO$ accounts for the aggregation of the operating costs of generation, transmission and distribution and load shedding, including revenues per MW h sold. The present formulation of the power flow problem is:

MCS–OPF($[\mathbf{Q}],\{Y\},\omega_s$):

$$\min_{P_{U_s},\Delta\delta_s,LS_s} \quad CO_s^{\omega} \tag{2.20}$$

(i)

$${}^{\mathrm{G}}CO_s^{\omega} = t^{\Delta}\sum_{i\in N}\sum_{G_j\in G} COv_j P_{U_{i,j,s}} \tag{2.20i}$$

(ii)

$$\begin{aligned} {}^{\mathrm{GTD}}CO_s^{\omega} = \sum_{i\in N}\sum_{G_j\in G}(COv_j - EP_{t_s})P_{U_{i,j,s}} + (C_{LS}+EP_{t_s})\sum_{i\in N} LS_{i,s} + \\ S_{ref}\sum_{(i,i')\in Y} COv_{(i,i')}|B_{(i,i')}(\delta_{i,s}-\delta_{i',s})| \end{aligned} \tag{2.20ii}$$

s. t.

$$L_{i,t_s} - \sum_{G_j\in G} P_{U_{i,j,s}} - S_{ref}\sum_{i'\in N}\eta_{(i,i'),s}B_{(i,i')}(\delta_{i,s}-\delta_{i',s}) - LS_{i,s} = 0 \tag{2.21}$$

$$0 \le P_{U_{i,j,s}} \le \eta_{i,j,s}[\mathbf{Q}]_{i,j}P_{i,j,s} \tag{2.22}$$

$$S_{ref}|B_{(i,i')}(\delta_{i,s}-\delta_{i',s})| \le P_{max_{(i,i')}} \tag{2.23}$$

where $\forall t \in D, s \in \{1,2,\ldots,NS\}$, ${}^{\mathrm{G}}CO_s^{\omega}$ (\$) – aggregated operating cost of generation, ${}^{\mathrm{GTD}}CO_s^{\omega}$ (\$) – aggregated operating cost of generation, transmission and distribution, and load shedding, COv_j (\$/MWh) – variable operating cost of the power generator j, EP_{t_s} (\$/MWh) – energy price at hour t, $P_{U_{i,j,s}}$ (MW) – used power from the generator type j located at node i, $COv_{(i,i')}$ (\$/MWh) – variable operating cost of the T&D line (i,i'), C_{LS} (\$/MWh) – load shedding cost and $P_{max_{(i,i')}}$ (MW) – power rating of the T&D line (i,i'). The load shedding $LS_{i,s}$ (MW) at node i is defined as the amount of load disconnected to alleviate congestion in the T&D lines and/ or balance the demand of power with the available power supply.

The meaning of each constraint is:

- (2.21): power balance at node i.

- (2.22): bounds of the power generation units.
- (2.23): technical limits of the T&D lines (power rating).

The OPF problem is solved for each operational scenario ω_s, giving in output the corresponding values of minimum CO_s^{ω}. The set $CO^{\Omega} = \{CO_1^{\omega}, CO_2^{\omega}, \ldots, CO_{NS}^{\omega}\}$ is, then, considered as a sample of realizations of the probability density function of CO.

2.4 Performance evaluation

The proposed renewable DG–integrated network solutions $Q = [Q^M | Q^R]$ are evaluated with respect to performance indicators regarding economics and reliability of power supply. Specifically, the expected values of performance and the associated uncertainty–risk measures are considered as targets.

2.4.1 Global cost

The economic performance of the network, given a DG integration plan, is evaluated with respect to a global cost function (CG). The quantity CG is composed by two terms: the outcome operating cost of the MCS–OPF described in the previous section, CO^{Ω}, and the fixed investment and operating cost, $CI_j + COf_j$, associated to the renewable part of the proposed DG plan Q^R, i.e., $\forall j \in \{j : G_j \in RG\}$. The fixed investment and operating cost $CI_j + COf_j$ (\$) is prorated hourly over the lifetime of the project t^H. Thus, the global cost function for the set of operational scenarios Ω is given by the following equations differentiated according the two distinct definitions of CO^{Ω} (Equations (2.20i) and (2.20ii)):

(i)

$$ {}^{\mathrm{G}}CG^{\Omega} = {}^{\mathrm{G}}CO^{\Omega} + \frac{t^{\Delta}}{t^H} \sum_{i \in N} \sum_{G_j \in RG} \left[\left(CI_j + COf_j \right) [\boldsymbol{Q^R}]_{\boldsymbol{i,j}} - t^{\Delta} (inc + EP^{\Omega}) P^{\Omega}_{U_{i,j}} \right] \tag{2.24i} $$

(ii)

$$ {}^{\mathrm{GTD}}CG^{\Omega} = {}^{\mathrm{GTD}}CO^{\Omega} + \frac{1}{t^H} \sum_{i \in N} \sum_{G_j \in RG} \left(CI_j + COf_j \right) [\boldsymbol{Q^R}]_{\boldsymbol{i,j}} \tag{2.24ii} $$

where, EP^{Ω} – sample of NS realizations of EP_{t_s}, inc (\$/MWh) – incentive for power generation from DG sources and $P^{\Omega}_{U_{i,j}}$ – sample of NS realizations of $P_{U_{i,j,s}}$.

Analogously to CO^{Ω} definition, CG^{Ω} represents a sample of realizations of the probability density function of CG and performance indicators of interest can be obtained, relative to expected performance, uncertainty and risk.

2.4.2 Energy not supplied

The energy not supplied (ENS) is a common index for reliability of power supply evaluation [3, 10, 20, 34, 45, 84–87]. In the present work, its value is obtained directly from the OPF outputs in the form of the aggregation of all nodal load sheddings per scenario $\omega_s \in \Omega$:

$$ENS^{\Omega} = t^{\Delta} \sum_{i \in N} LS_i^{\Omega} \tag{2.25}$$

where, ENS^{Ω} – sample of NS realizations of ENS_s^{ω} (MW h) and LS_i^{Ω} – sample of NS realizations of $LS_{i,s}$ (MW h).

2.4.3 Uncertainty and risk measurement

The proposed framework introduces the CVaR and CVaR deviation (DCVaR) [66] to measure, respectively, the risk and uncertainty in the performance functions of interest: CG and ENS. The quantity DCVaR is a functional of the CVaR [88], which is a coherent risk measure broadly used in financial portfolio optimization and has been extended to engineering applications, including electric power systems analysis and, in particular, DG planning [61, 63, 64, 69–72].

The definitions and properties of CVaR and DCVaR for continuous and discrete general return (loss) functions are given in detail in [66, 88]. Here, only a graphical, but comprehensive view to understand the CVaR and DCVaR definitions is presented in Figure 2.6.

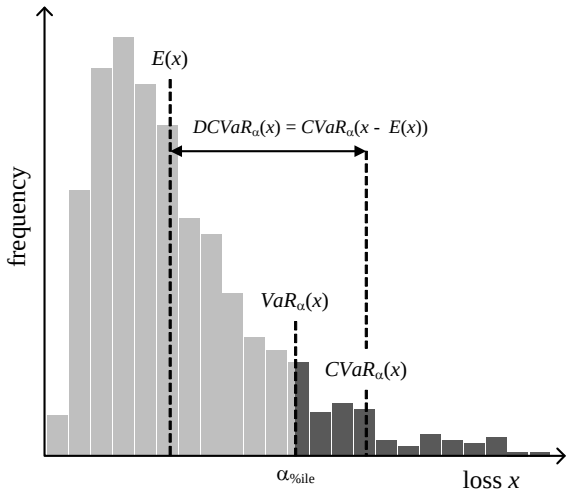

Figure 2.6 Graphic representation of $VaR_{\alpha}(x)$, $CVaR_{\alpha}(x)$ and $DCVaR_{\alpha}(x)$; $x =$ loss

For a discrete approximation of the probability function of the loss x, given a confidence level or α-percentile, the value–at–risk $VaR_{\alpha}(x)$ represents the smallest value of loss for which the probability that the loss does not exceed that threshold value is greater than or equal to α, whereas

$CVaR_\alpha(x)$ is the expected value of loss given that the loss is greater than or equal to the $VaR_\alpha(x)$. Thus, $CVaR_\alpha(x)$ provides a quantitative indication of the extent of the probability of occurrence of extreme non–desirable or risky scenarios of loss. The quantities $VaR_\alpha(x)$ and $CVaR_\alpha(x)$ can be expressed by the following equations:

$$VaR_\alpha(x) = \inf\{z : F_x(z) > \alpha\} \tag{2.26}$$

$$CVaR_\alpha(x) = E(x/x \geq VaR_\alpha(x)) \tag{2.27}$$

where, F_x – cumulative distribution function of the loss x.

With regards to $DCVaR_\alpha(x)$, this is a non symmetric deviation measure, as it accounts for the uncertainty associated to the loss exceeding its expected value. It is defined taking into account some important properties of the standard deviation [66], and is formulated as:

$$DCVaR_\alpha(x) = CVaR_\alpha(x - E(x)) \tag{2.28}$$

Furthermore, being a coherent risk measure, CVaR is a strictly expectation–bounded risk measure and it can be proved that a one–to–one relation exists with its corresponding deviation measure DCVaR [66]:

$$CVaR_\alpha(x) = E(x) + DCVaR_\alpha(x) \tag{2.29}$$

In the present framework, a specific configuration of the DG–integrated network ($[Q], \{Y\}$) can be considered as a generation portfolio, in which the renewable part $[Q^R]$ of $[Q]$ is the decision matrix. The corresponding assessed CG^Ω and ENS^Ω, obtained from the output MCS–OPF($[Q]$,$\{Y\}$,Ω), can be translated into the probability functions of loss; then, the quantities $CVaR_\alpha(CG^\Omega)$, $DCVaR_\alpha(CG^\Omega)$ and $CVaR_\alpha(ENS^\Omega)$, $DCVaR_\alpha(ENS^\Omega)$ represent the level of risk and uncertainty associated to the solution $[Q^R]$ with an expected global cost $ECG = E(CG^\Omega)$ and expected energy not supplied $EENS = E(ENS^\Omega)$, respectively.

3 Renewable DG–integrated electric power network planning

This Chapter presents the distinct optimization frameworks considered to address the optimal integration of DG in terms of selection of technology, sizing and allocation of renewable DG units. The corresponding decision matrix $[Q^R]$ is contained in the matrix $Q = [Q^M|Q^R]$ that stores the number and location of each type of power generator in the network.

Three distinct optimization strategies are formulated, defining correspondingly three different frameworks (FWs). Framework number 1 (FW1) focuses on controlling the risk associated to the expected performance of the DG plans, ECG and $EENS$, by measuring their respective CVaR values, $CVaR_\alpha(CG)$ and $CVaR_\alpha(ENS)$. FW2, aims at the control of uncertainty with respect to the ECG performance by targeting in conjunction the associated $DCVaR_\alpha(CG)$. Whereas FW3, presents a single objective strategy, aiming solely ECG to address the challenge of reducing the computational efforts required to implement heuristic optimization (HO) search engines with nested MCS–OPF and it is treated in details in the next Chapter 4.

3.1 Optimal DG technologies selection, sizing and allocation

3.1.1 Optimization strategies formulation

FW1: Multi–objective weighted *ECG*, *CVaR*(*CG*) & *EENS*, *CVaR*(*ENS*) minimization

The MOO problem consists in the concurrent minimization of the two objective functions measuring the CG and ENS. Specifically, their expected values and their CVaR values are combined, weighted by a factor $\beta \in [0, 1]$, which allows modulating the expected performance of the DG–integrated network and its associated risk.

Considering a set of randomly generated scenarios Ω, the MOO problem is formulated as follows:

$$\min_{[\mathbf{Q}^R]_{i,j}} \quad \beta E(CG^{\Omega}) + (1-\beta)CVaR_{\alpha}(CG^{\Omega})^{\dagger} \qquad (3.1)$$

$^{\dagger}CG^{\Omega} = {}^{\mathrm{G}}CG^{\Omega}$ by (2.24i)

$$\min_{[\mathbf{Q}^R]_{i,j}} \quad \beta E(ENS^{\Omega}) + (1-\beta)CVaR_{\alpha}(ENS^{\Omega}) \qquad (3.2)$$

s.t.

$$[\mathbf{Q}]_{i,j} \in \mathbb{Z}^* \qquad (2.3)$$

$$\sum_{i\in N}\sum_{G_j\in RG}\left(CI_j + COf_j\right)[\mathbf{Q}^R]_{i,j} \leq BGT \qquad (3.3)$$

$$\sum_{i\in N}[\mathbf{Q}^R]_{i,j} \leq \tau_j \quad \forall G_j \in RG \qquad (3.4)$$

$$\text{MCS–OPF}([\mathbf{Q}], \{Y\}, \Omega) \qquad (2.20)-(2.23)$$

The meaning of each constraint is:

- (2.3): the decision variable $[\mathbf{Q}^R]_{i,j}$ is a non–negative integer number.
- (3.3): the total fixed investment and operating cost of the DG units must be less than or equal to the available budget BGT (\$).
- (3.4): the total number of DG units to allocate of each technology j must be less than or equal to the maximum number of units available τ_j to be integrated.
- (2.20)−(2.23): the OPF equations of must be satisfied for all scenario $\omega_s \in \Omega$.

Constraint (3.4) can be translated into maximum allowed penetration factor $PF^{DG}_{max_j}$ of each DG technology j. Defining PF as 'the output active power of total capacity of DG divided by the aggregated maximum nodal loads' [23], constraint (3.4) can be rewritten as follows:

$$\frac{\sum_{i\in N}[\mathbf{Q}^R]_{i,j}E(P_j^{DG})}{\sum_{i\in N}L_{max_i}} = PF_j^{DG} \leq PF^{DG}_{max_j} = \frac{\tau_j E(P_j^{DG})}{\sum_{i\in N}L_{max_i}} \quad \forall G_j \in RG \qquad (3.5)$$

where, $\sum[\mathbf{Q}^R]_{i,j}$ – total number of units of DG technology j integrated in the network; $E(P_j^{DG})$ – expected power output of one unit of DG technology type j (MW); $\sum L_{max_i}$ – aggregated maximum nodal loads (MW).

FW2: Multi–objective *ECG* & *DCVaR*(*CG*) minimization

The practical aim of the MOO is the simultaneous minimization of representative indicators of the objective function CG, given by the expected value ECG and the associated uncertainty measure $DCVaR_\alpha(CG)$.

The general MOO problem for all set of randomly generated operational scenarios Ω is formulated as follows:

$$\min_{[\mathbf{Q}^R]_{i,j}} \quad E(CG^\Omega) \tag{3.6}$$

$$\min_{[\mathbf{Q}^R]_{i,j}} \quad DCVaR_\alpha(CG^\Omega)^\ddagger \qquad {}^\ddagger CG^\Omega =^{\text{GTD}} CG^\Omega \text{ by (2.24ii)} \tag{3.7}$$

(3.8)

s.t.

$$[\mathbf{Q}]_{i,j} \in \mathbb{Z}^* \tag{2.3}$$

$$\sum_{i \in N} \sum_{G_j \in RG} \frac{[\mathbf{Q}]_{i,j} P_{AV_j}}{L_{max_i}} \leq PF^{DG} \tag{3.9}$$

$$\text{MCS–OPF}([\mathbf{Q}], \{Y\}, \Omega) \tag{2.20)—(2.23}$$

The meaning of each constraint is the following:

- (2.3): the decision variables $[Q]_{i,j}$ are non–negative integer numbers.
- (3.9): the ratio of total amount of average renewable power integrated in the network must be less than or equal to the penetration factor PF^{DG}.
- (2.20)—(2.23): the OPF equations of must be satisfied for all scenario $\omega_s \in \Omega$.

FW3: Single–objective *ECG* minimization

This single objective optimization strategy ought to find the optimal plan of integration of renewable DG $[Q^R]$ by minimizing the expected value of global cost ECG. Considering a set of randomly generated scenarios Ω, the optimization problem is formulated as follows:

$$\min_{[\mathbf{Q}^R]_{i,j}} \quad E(CG^\Omega)^\ddagger \qquad {}^\ddagger CG^\Omega =^{\text{GTD}} CG^\Omega \text{ by (2.24ii)} \tag{3.10}$$

s.t.

$$[\mathbf{Q}]_{i,j} \in \mathbb{Z}^* \tag{2.3}$$

$$\sum_{i \in N} \sum_{G_j \in RG} \left(CI_j + COf_j\right)[\mathbf{Q^R}]_{i,j} \leq BGT \tag{3.11}$$

$$\sum_{i \in N} [\mathbf{Q^R}]_{i,j} \leq \tau_j \quad \forall G_j \in RG \tag{3.12}$$

$$\text{MCS–OPF}([\mathbf{Q}], \{Y\}, \Omega) \tag{2.20)—(2.23}$$

The meaning of each constraint is:

- (2.3): the decision variable $[Q^R]_{i,j}$ is a non–negative integer number.
- (3.11): the total investment and fixed operation and maintenance costs must be less than or equal to the available budget BGT.
- (3.12): the total number of renewable DG units of each technology j to be allocated must be less than or equal to the maximum number of units available for integration τ_j.
- (2.20)—(2.23): the OPF equations of must be satisfied for all scenario $\omega_s \in \Omega$.

3.2 Heuristic optimization & MCS–OPF simulation frameworks

The MOO optimization problems are non–linear and non–convex, i.e., a non–convex mixed–integer non–linear problem or non–convex MINLP. Non–linearity is due to the fact that all the objective functions involved cannot be written in the canonical form of a linear program, i.e., $C^T X$, where C is a vector of known coefficients and X the decision vector. In the present case, the decision matrix $[Q^R]$ enters the MCS–OPF flow simulation to obtain the probability density functions of ENS and/or CG, then, the objective functions are formed by performance indicators of interests: expected values, CVaR and DCVaR, in correspondence to the framework applied. Thus, the operations carried out on $[Q^R]$ through MCS–OPF, expected, CVaR and DCVaR values cannot not be represented as the product $C^T[Q^R]$. The problem is non–convex because the decision matrices $[Q^R]$ are integer–valued and, as it is known, the set of non–negative integers is non–convex.

Given the class of optimization problem in the proposed framework (non–convex MINLP), it is most likely to have multiple local minima. Moreover, the dimension of the distribution network can lead to a combinatorial explosion of the feasible space of the decision matrices $[Q^R]$ [16, 20].

Heuristic optimization algorithms (HO) have emerged as the most effective search engines for combinatorial optimization problems and they can deal with non–differentiable objective functions, discontinuous feasible spaces and non–convex conditions [20, 42]. Some of the best known techniques are: particle swarm optimization (PSO) [9, 42, 56, 76, 89], differential evolution (DE)

[51, 90, 91] and genetic algorithms (GA) [3, 16, 34, 37, 69]. In this view, the three non–convex MINLPs under uncertainties here proposed are solved by implementing HO search engines.

The MOO problem under uncertainties of FW1 is solved by the NSGA–II algorithm [92], in which the evaluation of the objective functions is performed by the developed MCS–OPF. In FW2, the MOO differential evolution (MOO–DE) algorithm is implemented, integrating a fast non–dominated sorting procedure and crowded–comparison operator [92] into the original single objective DE [93], and also evaluating the objective functions by MCS–OPF. FW3 performs the search for optimal solution by an original technique developed in this work, which integrates hierarchical clustering analysis (HCA) into the basic single objective differential evolution (DE) search, and is presented in Chapter 4.

For MOO frameworks, FW1 and FW2, the extension to MOO entails the integration of Pareto optimality concepts. In general terms, solving a MOO problem of the form:

$$\min_{X} \quad \{f_1(X), f_2(X), \ldots, f_m(X)\}$$
$$\text{s.t. } X \in \Lambda$$

with at least two conflicting objectives functions ($f_i : \Re^n \rightarrow \Re$) implies to find, within a set of acceptable solutions that belong to the non–empty feasible region $\Lambda \subseteq \Re^n$, the decision vectors $X \in \Lambda$ that satisfy the following [94]:

$$\neg X \in \Lambda / f_i(X) \leq f_i(X'), \forall i \in \{1, 2, \ldots, m\} \text{ and } f_i(X) < f_i(X') \text{ for at least one } i$$
$$\Downarrow$$
$$f_i(X) \prec f_i(X') \text{ i.e. } X \text{ dominates } X'$$

The vector X is called a Pareto optimal solution and the Pareto front is defined as the set $\{f(X) \in \Re^n\}$ such that X is Pareto optimal solution. The general NSGA–II and MOO–DE algorithms are described next.

3.2.1 Non–dominated sorting genetic algorithm II–based approach

NSGA–II is one of the most efficient MOO evolutionary algorithms in HO [95]. It uses an elitist approach by applying a fast non–dominated sorting and crowding–distance comparison operator, while the search for non–dominated solutions is performed based on two main genetic operators, namely mutation and crossover [92]. The general NSGA–II algorithm is summarized as follows:

Initialization

- Set the values of parameters:

- ▷ NP: population size.
- ▷ NG_{max}: maximum number of generations.
- ▷ p_C: crossover probability $\in [0,1]$.
- ▷ p_M: mutation probability $\in [0,1]$.

- ∘ Form the initial population POP^0, randomly generating NP decision matrices (individuals) X within the feasible space, $POP^0 = \{X_1^0, \ldots, X_k^0, \ldots, X_{NP}^0\}$
- ∘ Evaluate the objective functions $\left\{f_1(X_k^0), \ldots, f_{NF}(X_k^0)\right\}$ for each individual X_k^0, where NF – number of objective functions.
- ∘ Rank the individuals in POP^0, applying a fast non–dominated sorting procedure [92] with respect to the values of the objective functions and identify the non–dominated fronts $\{\mathscr{F}_1^0, \ldots, \mathscr{F}_{ND}^0\}$, where $\mathscr{F}_1^0$ is the best front and $\mathscr{F}_{ND}^0$ the less good front.
- ∘ Compute and assign the crowding–distance value (d_C) to each X_k^0 individual in POP^0 and sort, in ascending order with respect to d_C, the individuals belonging to the same non–domination–ranked group $\{\mathscr{F}_\ell^0\}_{\ell \in \{1,\ldots,ND\}}$.
- ∘ Apply binary tournament selection [92] to POP^0 based on the crowding distance to generate an intermediate population $POP^{0'}$ of size NP.

  Reproduction

  - ▷ Apply mutation and crossover operators to $POP^{0'}$, to create an offspring population $OPOP^0$ of size NP.
  - ▷ Evaluate the objective functions for each individual X_k^0 in $OPOP^0$.

Evolution loop

- ∘ Set generations count index $g = 1$.
- ∘ Set $POP^g = POP^0$ and $OPOP^g = OPOP^0$.
- ∘ While $g \leq NG_{max}$ (stopping criterion):
- ∘ Combine POP^g and $OPOP^g$ to obtain a population equals to their union $UPOP^g = POP^g \cup OPOP^g$.
- ∘ Apply a fast non–dominated sorting procedure on $UPOP^g$ and identify the new non–dominated fronts $\{\mathscr{F}_1^g, \ldots, \mathscr{F}_{ND}^g\}$.
- ∘ Compute and assign the crowding–distance value to each individual X_k^g in $UPOP^g$ and sort each non–domination–ranked group $\{\mathscr{F}_\ell^g\}_{\ell \in \{1,\ldots,ND\}}$.
- ∘ Select the best NP individuals X_k^g to create the next population of parents POP^{g+1}.
- ∘ Apply binary tournament selection to POP^{g+1} to generate an intermediate population $POP^{g+1'}$ of size NP.

  Reproduction

- ▷ Apply mutation and crossover operators to $POP^{g+1'}$, to create an offspring population $OPOP^{g+1}$ of size NP.
- ▷ Evaluate the objective functions for each individual X_k^{g+1} in $OPOP^{g+1}$.

- ◦ Set $g = g+1$ and verify the stopping criterion, if $g > NG_{max}$ then return POP^g, selecting the best front $\mathscr{F}_1^g$ as the optimal set of non–dominated solutions, otherwise continue the evolution loop.

In correspondence to the nomenclature used in this thesis, the process of searching the set of non–dominated solutions carried out by the NSGA–II MCS–OPF is presented schematically in Figure 3.1.

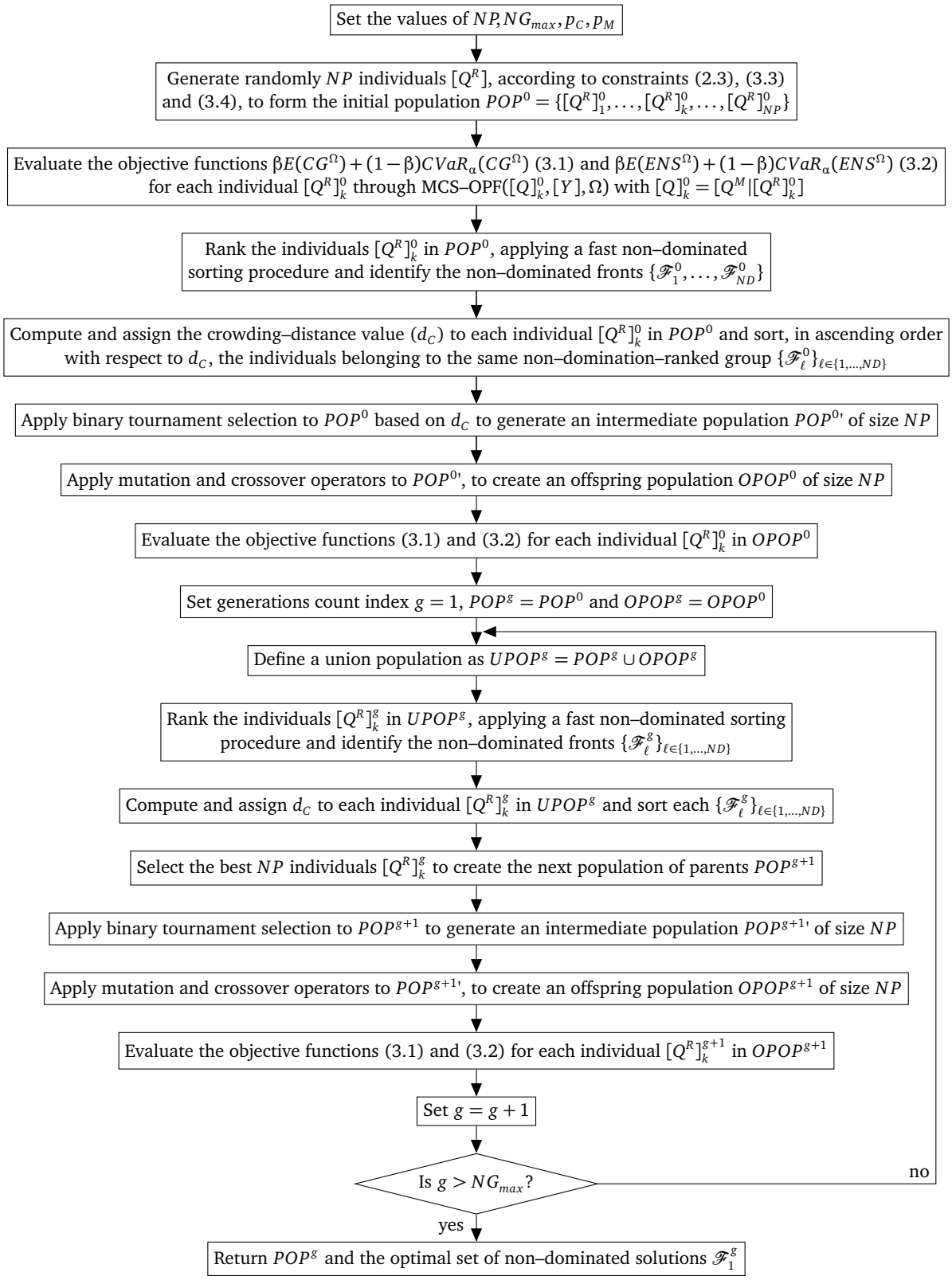

Figure 3.1 Flow chart of the proposed NSGA–II MCS–OPF framework

3.2.2 MOO differential evolution–based approach

DE is a population–based and parallel, direct search method, shown to be one of the most efficient evolutionary algorithms to solve complex optimization problems [93, 96, 97]. The implementation of the original version of DE involves two main phases: initialization and evolution. The extended MOO–DE, with fast non–dominated sorting and crowding–distance comparison operator [92] is summarized below [93].

Initialization

- Set the values of parameters:
  - NP: population size.
  - NG_{max}: maximum number of generations.
  - p_C: crossover coefficient $\in [0,1]$.
  - F: differential variation amplification factor $\in [0,2]$.
- Form the initial population POP^0, randomly generating NP decision matrices (individuals) X within the feasible space, $POP^0 = \{X_1^0, \ldots, X_k^0, \ldots, X_{NP}^0\}$
- Evaluate the objective functions $\{f_1(X_k^0), \ldots, f_{NF}(X_k^0)\}$ for each individual X_k^0, where NF – number of objective functions
- Rank the individuals in POP^0, applying a fast non–dominated sorting [92] procedure with respect to the values of the objective functions and identify the non–dominated fronts $\{\mathscr{F}_1^0, \ldots, \mathscr{F}_{ND}^0\}$, where $\mathscr{F}_1^0$ is the best front and $\mathscr{F}_{ND}^0$ the less good front.
- Compute and assign the crowding–distance value (d_C) [92] to each individual X_k^0 in POP^0 and sort, in ascending order with respect to d_C, the individuals belonging to the same non–domination–ranked group $\{\mathscr{F}_\ell^0\}_{\ell \in \{1,\ldots,ND\}}$.

Evolution loop

- Set generations count index $g = 1$.
- Set $POP^g = POP^0$.
- While $g \leq NG_{max}$ (stopping criterion):
- Set a repository population $RPOP$ as empty.

  Trial loop

  For each individual X_k^g in POP^g, $\forall k \in \{1, \ldots, NP\}$:
  - Sample from the uniform distribution three integer indexes in $\{1, \ldots, NP\}$ such that $k_1 \neq k_2 \neq k_3 \neq k$ and choose the corresponding three individuals $X_{k_1}^g, X_{k_2}^g, X_{k_3}^g$
  - Generate a mutant individual XM_k^g according to the following mutation operator:

$$XM_k^g = X_{k_1}^g + F(X_{k_2}^g - X_{k_3}^g) \tag{3.13}$$

- ▹ Apply crossover operator, initializing a randomly generated vector XC_k^g, whose dimensionality n is the same as that of X_k^g and each coordinate $xc_{k,i}^g$ follows a uniform distribution with outcome in $[0,1] \forall i \in \{1,\ldots,n\}$. In addition, generate randomly an integer index i^* in $\{1,\ldots,n\}$ from a uniform distribution to ensure that at least one coordinate from XM_k^g is exchanged to form trial individual XT_k^g, whose coordinates $xt_{k,i}^g$ are defined as follows:

$$xt_{k,i}^g = \begin{cases} xm_{k,i}^g & \text{if } xc_{k,i}^g \leq p_C \text{ or } i = i^* \\ x_{k,i}^g & \text{if } xc_{k,i}^g > p_C \text{ and } i \neq i^* \end{cases} \tag{3.14}$$

- ▹ Evaluate the objective functions for the trial individual $\{f_1(XT_k^g),\ldots,f_{NF}(XT_k^g)\}$; if XT_k^g dominates X_k^g, i.e., $\{f(XT_k^g) \prec f(X_k^g)\}$, XT_k^g replaces X_k^g in POP^g, otherwise retain X_k^g in POP^g and save XT_k^g in the repository population $RPOP$.

- ◦ Set a combined population $UPOP$ as $POP^g \cup RPOP$ and rank the individuals in $UPOP$, applying a fast non–dominated sorting procedure and identify the new non–dominated fronts $\{\mathscr{F}_1^g,\ldots,\mathscr{F}_{ND}^g\}$.
- ◦ Compute and assign the crowding–distance value d_C to each individual in $UPOP$ and sort each non–domination–ranked group $\{\mathscr{F}_\ell^g\}_{\ell\in\{1,\ldots,ND\}}$.
- ◦ Set POP^g as the first NP (best) individuals of the ranked and sorted population $UPOP$, $POP^g = \{X_k : X_k \in UPOP, k \in \{1,\ldots,NP\}\}$
- ◦ If the stopping criterion is reached return POP^g, otherwise set $g = g+1$.

Analogously to the NSGA–II based FW1 and, in correspondence to the nomenclature used in this thesis, search for the set of non–dominated solutions performed by MOO–DE MCS–OPF is presented schematically in Figure 3.2.

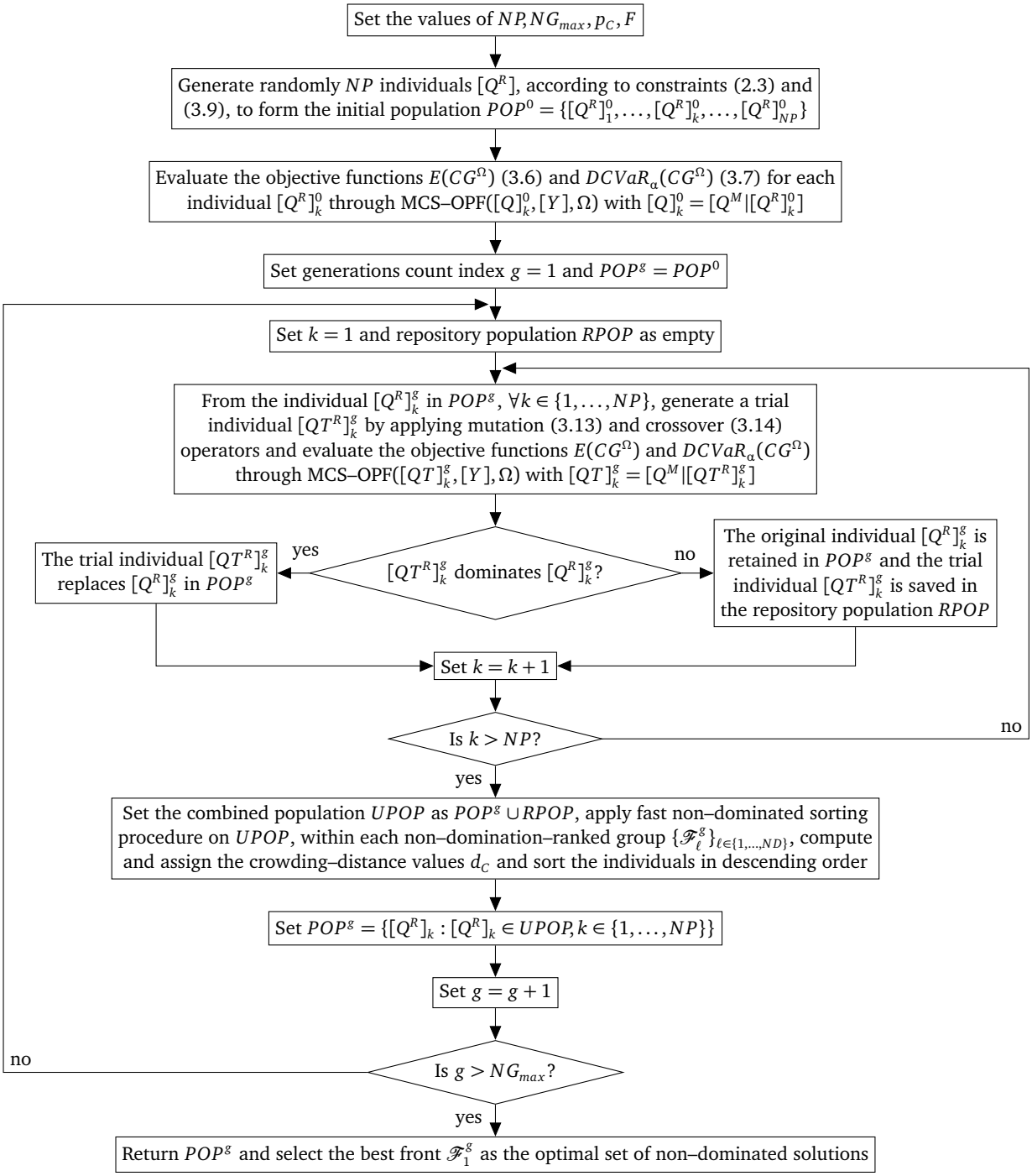

Figure 3.2 Flow chart of the proposed MOO–DE MCS–OPF framework

4 Computational Challenge

The computational challenge here presented consists in improving the performance of HO techniques used to solve the complex optimization problem of DG planning, when the objective function(s) is(are) evaluated by time consuming computational models, such as the developed MCS–OPF. For this, the integration of clustering into the HO search engine is considered.

4.1 Clustering in heuristic optimization

Clustering techniques like, k–means, fuzzy c–means and hierarchical clustering, among others, can be directed to the enhancement of the global and/or local searching ability of HO algorithms, and amounts to identifying groups of similar individuals and applying different evolution operators to those of a same cluster (group), e.g. for random generation of new individuals in the neighborhood of cluster centroids, or multi-parents crossover over new randomly generated individuals spread in the global feasible space [96–101]. In the literature, these approaches have been proved to be effective in improving convergence, but for 'light-weight' benchmark or not simulation–based objective function(s). When the computational time complexity associated to the evaluation of the objective function(s) is significant, even if convergence is improved by applying some of these methodologies, which imply a temporarily increment of the overall size of the population, the computational effort benefits may result counteracted. In addition, the accuracy of the clusters structures in representing the distribution of individuals must be controlled for performing clustering conveniently.

The main original contribution of the work here presented, lies in the development of the clustering strategy in a controlled manner. The implementation of such clustering strategy is done within a differential evolution (DE) optimization framework MCS–OPF model for the integration of renewable DG into an electric power system. The introduction of the clustering is hierarchically (i.e., hierarchical clustering analysis, HCA, [102]) by a controlled way of reducing the number of individuals to be evaluated during the DE search, therefore, improving the computational efficiency. Henceforth, the method is called hierarchical clustering differential evolution (HCDE).

HCA is introduced to build a hierarchical structure of grouping individuals (DG–integrated network plans) of the population that present closeness under the control of a specific linkage criterion based on defined distance metrics [102]. The HCA outcomes are the linkage distances at which the grouping actions take place, defining the different levels in the hierarchical structure. Two control parameters are introduced in the HCA, the cophenetic correlation coefficient (CCC) and a cutoff level coefficient of the linkage distances in the hierarchical structure of the groups (p_{co}). The CCC is a similarity coefficient that measures how representative is the proposed grouping structure by comparing their linkage distances with the original distances between all the individuals in the population. In the hierarchical structure, the linkage distance given by p_{co} sets the level at which the groups formed below it are considered to be 'close enough' to constitute independent clusters. The two parameters allow HCDE to adapt itself in each generation of the search, 'deciding' whether to perform clustering if the CCC is greater than or equal to a preset threshold (CCC_T) and cutting the hierarchical structure in independent clusters according to the linkage distance given by p_{co}. Then, the individual closest to the centroid of each cluster is taken as the feasible representative solution in the population that enters the evolution phase of the HCDE algorithm.

4.2 Hierarchical clustering & differential evolution

The original version of DE keeps the population size NP constant, making the computational performance dependent mainly on the number of objective function evaluations (NFE) carried out during the evolution phase of the algorithm. Then, the integration of HCA into DE is aimed at the reduction of the number of individuals that enter the evolution loop in each generation so as to decrease the number of objective function evaluations.

HCA links individuals or groups of individuals which are similar with respect to a specific property, translated into a metric of distance, obtaining a hierarchical structure. In practice, an agglomerative procedure is used, which in $NZ = NP - 1$ steps z fuses the closest pair or individuals or groups of individuals through a linkage function, e.g. single linkage (nearest neighbor distance), complete linkage (furthest neighbor), average linkage, among others, until the complete hierarchical structure is built. The base hierarchical clustering algorithm used in this study can be expressed as follows [102]:

> **Step 1:** Given a population POP $POP = \{X_1, \ldots, X_k, \ldots, X_{NP}\}$, form the set of singleton groups $\mathscr{C} = \{\mathscr{C}_\iota = \{X_k\}\}$, $\forall \iota = k \in \{1, \ldots, NP\}$ and calculate the linkage distances between all the NP groups using the average as linkage function and the Euclidean distance as metric: where, $d^z_{\iota,J}$ – average of the Euclidean distances between all the individuals X_k and $X_{k'}$ belonging to the groups $\mathscr{C}_\iota$ and $\mathscr{C}_J$, respectively, $N\mathscr{C}_z$ – number of groups at step z and $|\mathscr{C}_\iota|$, $|\mathscr{C}_J|$ – cardinalities of the groups $\mathscr{C}_\iota$ and $\mathscr{C}_J$, respectively.

$$D^z = \begin{bmatrix} 0 & d^z_{1,2} & \cdots & d^z_{1,J} & \cdots & d^z_{1,N\mathscr{C}_z} \\ & \ddots & \ddots & \vdots & \ddots & \vdots \\ & & 0 & d^z_{\iota,J} & \cdots & d^z_{\iota,N\mathscr{C}_z} \\ & & & 0 & \ddots & \vdots \\ & & & & \ddots & d^z_{N\mathscr{C}_z-1,N\mathscr{C}_z} \\ & & & & & 0 \end{bmatrix}$$

$$d^z_{\iota,J} = \frac{1}{|\mathscr{C}_\iota||\mathscr{C}_J|} \sum_{X_k \in \mathscr{C}_\iota} \sum_{X_{k'} \in \mathscr{C}_J} \sqrt{(X_k - X_{k'})^2} \tag{4.1}$$

$$N\mathscr{C}_z = NP - z + 1 \tag{4.1a}$$

$$\forall X_k, X_{k'} \in POP,\ z \in \{1,\ldots,NP-1\},\ \iota, J \in \{1,\ldots,N\mathscr{C}_z\}$$

Step 2: Fuse the first pair of groups $\mathscr{C}_{\iota'}$ and $\mathscr{C}_{J'}$, for which $d^1_{\iota',J'}$ is the minimum distance $\min(D^1)$ and form a new group $\mathscr{C}_{NP+1} = \{\mathscr{C}_{\iota'} \cup \mathscr{C}_{J'}\}$. Update the set of groups $\mathscr{C}$ replacing $\mathscr{C}_{\iota'}$ and $\mathscr{C}_{J'}$ by $\mathscr{C}_{NP+1}$, and calculate the linkage distances D^2 between all the $NP-1$ groups in $\mathscr{C}$ using (4.1).

Step 3: Fuse the second pair of groups $\mathscr{C}_{\iota'}$ and $\mathscr{C}_{J'}$ for which $d^2_{\iota',J'}$ is the minimum distance $\min(D^2)$, and form a new group $\mathscr{C}_{NP+2} = \{\mathscr{C}_{\iota'} \cup \mathscr{C}_{J'}\}$. As in the preceding step, update the set of groups $\mathscr{C}$ and calculate the linkage distances D^3 between all the $NP-2$ groups in $\mathscr{C}$ using (4.1).

$\vdots$

Step $NP-1$: Fuse the last pair of groups with linkage distance $d^{NP-1}_{\iota',J'}$, forming the last group $\mathscr{C}_{2NP-1} = \{\mathscr{C}_{\iota'} \cup \mathscr{C}_{J'}\}$ that contains all the individuals X.

The outcoming hierarchical (or tree) structure can be reported as a sorted table containing the $NP-1$ linkage distances relative to each pairing action of individuals/groups and be graphically illustrated as a dendrogram. Table 4.1 and Figure 4.1 present, respectively, the resultant linkage distances and dendrogram obtained from an example set of $NP = 8$ two-dimensional individuals X using the above introduced HCA algorithm.

Table 4.1 Example hierarchical structure outcome

Step z	Group	Groups linked	Linkage distance
1	$\mathscr{C}_9$	$\{\mathscr{C}_2 \cup \mathscr{C}_6\} = \{\{X_2\} \cup \{X_6\}\}$	$d^1_{2,6}$
2	$\mathscr{C}_{10}$	$\{\mathscr{C}_3 \cup \mathscr{C}_4\} = \{\{X_3\} \cup \{X_4\}\}$	$d^2_{3,4}$
3	$\mathscr{C}_{11}$	$\{\mathscr{C}_1 \cup \mathscr{C}_7\} = \{\{X_1\} \cup \{X_7\}\}$	$d^3_{1,7}$
4	$\mathscr{C}_{12}$	$\{\mathscr{C}_5 \cup \mathscr{C}_8\} = \{\{X_5\} \cup \{X_8\}\}$	$d^4_{5,8}$
5	$\mathscr{C}_{13}$	$\{\mathscr{C}_9 \cup \mathscr{C}_{11}\} = \{\{X_2, X_6\} \cup \{X_1, X_7\}\}$	$d^5_{9,11}$
6	$\mathscr{C}_{14}$	$\{\mathscr{C}_{10} \cup \mathscr{C}_{12}\} = \{\{X_3, X_4\} \cup \{X_5, X_8\}\}$	$d^6_{10,12}$
7	$\mathscr{C}_{15}$	$\{\mathscr{C}_{13} \cup \mathscr{C}_{14}\} = \{\{X_1, X_2, X_6, X_7\} \cup \{X_3, X_4, X_5, X_8\}\}$	$d^7_{13,14}$

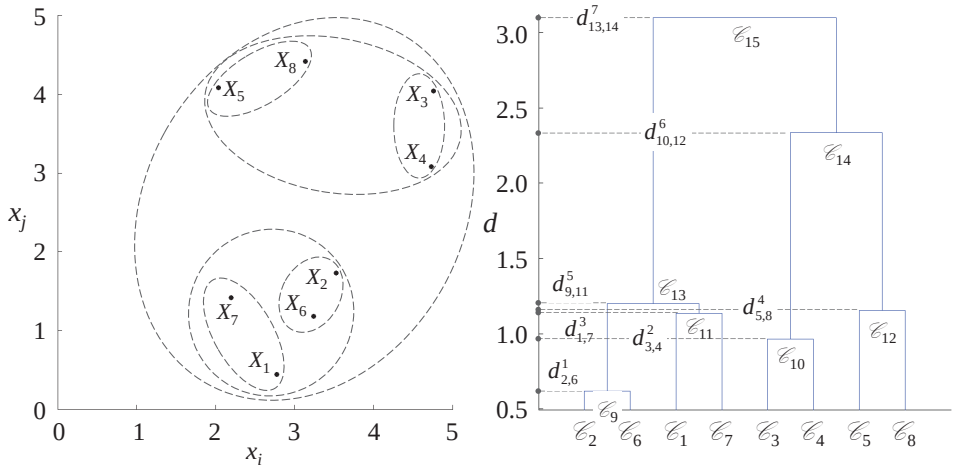

Figure 4.1 Example dendrogram for average linkage HCA

HCA builds the hierarchical structure through a linkage function introducing in each grouping action a larger or smaller degree of distortion with respect to the original distances between (ungrouped) individuals. The measurement of this distortion is important and the cophenetic correlation coefficient (CCC)is introduced to evaluate how representative is the hierarchical structure proposed by the HCA. The CCC can be obtained from Equations (4.2) and (4.3) below [102].

$$CCC = \frac{\sum_{k<k'}(d^1_{k,k'} - \bar{D^1})(l_{k,k'} - \bar{\mathscr{L}})}{\sqrt{\sum_{k<k'}(d^1_{k,k'} - \bar{D^1})^2 \sum_{k<k'}(l_{k,k'} - \bar{\mathscr{L}})^2}} \qquad \forall k,k' \in \{1,\ldots,NP\} \tag{4.2}$$

$$\mathscr{L} = \begin{bmatrix} 0 & l_{1,2} & \cdots & l_{1,k'} & \cdots & l_{1,NP} \\ & \ddots & \ddots & \vdots & \ddots & \vdots \\ & & 0 & l_{k,k'} & \cdots & l_{k,NP} \\ & & & 0 & \ddots & \vdots \\ & & & & \ddots & l_{NP-1,NP} \\ & & & & & 0 \end{bmatrix}$$

$$l_{k,k'} = d^{z^*}_{\iota,\jmath} \tag{4.3}$$

$$z^* = \{\min\ z : X_k, X_{k'} \in \mathscr{C}_{NP+z}\} \tag{4.3a}$$

$$\forall k,k' \in \{1,\ldots,NP\},\ \iota,\jmath \in 1,\ldots,2NP-1$$

where $\bar{D^1}$ – average of the original Euclidean distances $d^1_{\iota,\jmath}$ between all the individuals, $l_{k,k'}$ – linkage distance $d^{z^*}_{\iota,\jmath}$ where the pair of individuals X_k and $X_{k'}$ become members of the same group and $\bar{\mathscr{L}}$ – average of the resultant linkage distances $l_{\iota,\jmath}$ between all the individuals.

Recalling that the aim of nesting HCA into DE is to increase the computational performance by decreasing the NFE (times that the MCS–OPF is run) in each generation g, the presetting of a threshold CCC_T for the CCC value allows defining the level of representativeness required to the

hierarchical structure proposed. If by applying HCA over the population POP^g the corresponding CCC^g is such that $CCC^g \geq CCC_T$, the built hierarchical structure is considered an acceptable representation of the original distances amongst the individuals and the selection of a particular partition of the sets of groups can be performed, i.e., the determination of a specific number of clusters. On the contrary, if $CCC^g \leq CCC_T$, the hierarchical structure is considered not representative enough since it introduces unacceptable distortion that may affect the global searching process in the HCDE.

Whether the hierarchical structure is accepted, the clustering process itself takes place. As before stated, the HCA outcome linkage distances $d^z_{i,j}$ define each level (height) at which a pairing action is carried out. If the hierarchical structure is 'cut off' at a specific linkage distance d_{co}, all the groups that are formed below that level become independent clusters. In each generation g of HCDE, a d_{co} relative to the HCA outcome linkage distances for the corresponding POP^g, is determined from a preset cut–off level coefficient p_{co} of the linkage distances between the minimum $d^z_{i,j}$ that correspond to the first pairing action and the distance to form at least four clusters needed to perform the mutation process in the HCDE. Thus, d_{co} can be obtained from Equation (4.4). Figure 4.2 shows the cut–off distance representation for the example aforementioned, for which the formed clusters are $\{\mathscr{C}_2, \mathscr{C}_6\}$, $\{\mathscr{C}_1\}$, $\{\mathscr{C}_7\}$, $\{\mathscr{C}_3, \mathscr{C}_4\}$, $\{\mathscr{C}_5\}$ and $\{\mathscr{C}_8\}$.

$$d_{co} = d_{\min} + p_{co}(d_{NC=4} - d_{\min}) \tag{4.4}$$

$$d_{\min} = \min_z d^z_{i,j} \tag{4.4a}$$

$$d_{NC=4} = d_{\left(1-\frac{4}{NP}\right)\%\text{ile}} \tag{4.4b}$$

where, $d_{\min}$ – minimum linkage distance $d^z_{i,j}$ that correspond to the first pairing action and $d_{NC=4}$ – distance to form at least four clusters.

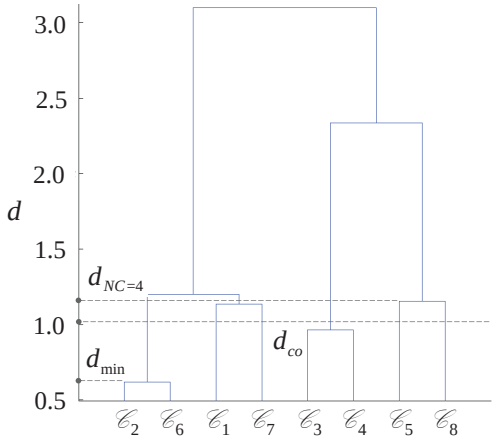

Figure 4.2 Example of cutoff distance calculation

The integration of HCA into DE and the definition of the parameters CCC_T and p_{co} allow HCDE adaptation at each generation, i.e., deciding whether to perform HCA and determining the clusters

to be taken. Then, the individuals closest to the centroids of the formed clusters are considered as the representatives of the group which they belong to and are taken in a reduced population that enters the evolution phase of the HCDE. The proposed HCDE algorithm is summarized schematically in the flowchart of Figure 4.3.

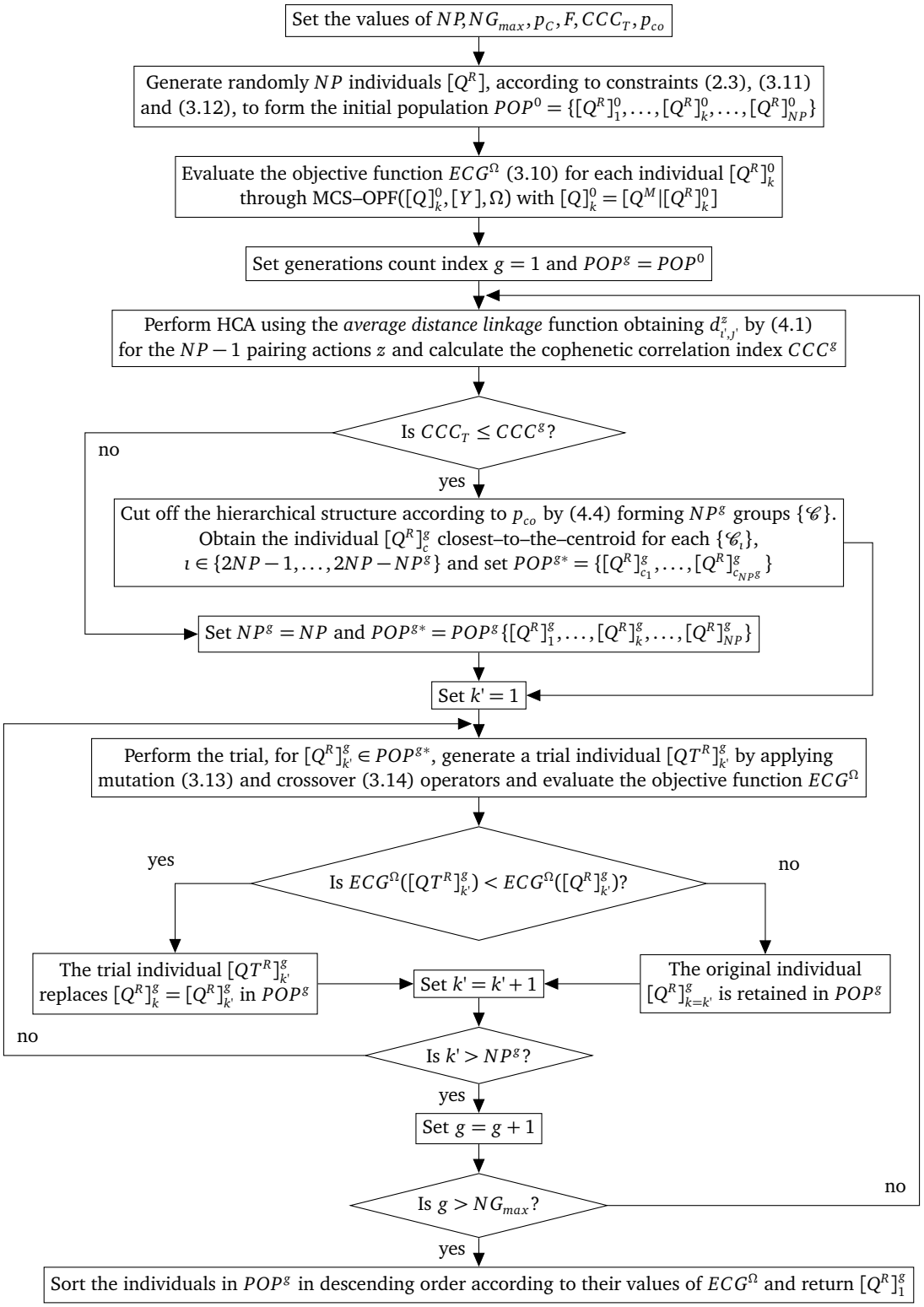

Figure 4.3 Flow chart of the proposed HCDE framework

5 Applications

In this Chapter, examples of application of the proposed frameworks are given. The main results obtained and the derived insights are presented, while encouraging the interested reader to consult the corresponding Papers (i)–(iii) reported in Part II for further details.

For practical ease of the presentation of the examples, tables and figures reporting the data associated to each application are provided in the corresponding Appendices A, B and C.

5.1 A risk–based MOO and MCS–OPF simulation framework for the integration of renewable DG and storage devices

This section introduces an example of application of the simulation and MOO framework for the integration of renewable DG and storage devices into an electric distribution network. The framework searches for the optimal size and location of the DG units, taking into account different sources of uncertainty: renewable resources availability, components failure and repair events, loads and grid power supply. The evaluation of the network performance is run by the developed MCS–OPF computational model and, as a response to the need of monitoring and controlling the risk associated to the performance of the optimal DG–integrated networks, the CVaR is integrated into the MOO optimization strategy which aims at the concurrent minimization of the expected values of CG and ENS, namely ECG and $EENS$, combined with their respective CVaR values, $CVaR(CG)$ and $CVaR(ENS)$, and weighed by a factor $beta$ which allows modulating expected performance and risk. The MOO is performed by the NSGA–II presented in Chapter 3, Section 3.2 and considering all the formulation relative to framework FW1 that can be found in the same Chapter 3, Subsection 3.1.1.

5.1.1 IEEE 13–bus test feeder case

The framework FW1 is applied to a distribution network derived from the IEEE 13 nodes [2, 103]. The spatial structure of the network has not been altered but the the regulator, capacitor, switch feeders of zero length are neglected. The network is chosen purposely small, but with all relevant characteristics for the analysis, e.g. comparatively low and high spot and distributed load values

and the presence of a power supply spot [103]. The original IEEE 13–bus test feeder is dimensioned such that the total power demand is satisfied without lines overloading. This is modified so that it becomes of interest to consider the integration of renewable DG units. Specifically, the location and values of some of the load spots and the power ratings values of some feeders have been modified in order to generate conditions of power congestion of the lines, leading to shortages of power supply to specific portions of the network.

The distribution network presents a radial structure of $n = 11$ nodes and, therefore, $(n-1) = 10$ feeders, as shown in Figure 5.1. The nominal voltage is $V = 4.16$ (kV), constant for the resolution of the DC–OPF problem. Four types of renewable DG technologies are considered to be integrated: solar photovoltaic (PV), wind turbines (W), electric vehicles (EV) and storage devices (ST).

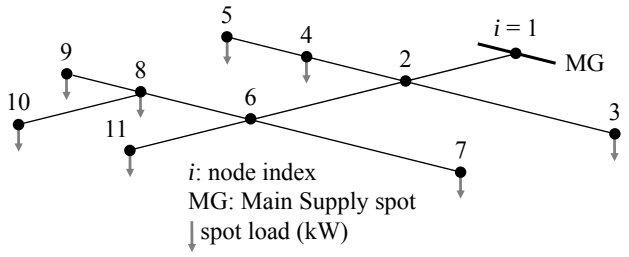

Figure 5.1 Radial 11–nodes distribution network

Five optimizations runs of the NSGA–II with the nested MCS–OPF algorithm are performed, each one with a different value of the weight parameter $\beta \in \{1, 0.75, 0.5, 0.25, 0\}$, to analyze different tradeoffs between optimal average performance and risk. From the equations that define the objective functions, (3.1) and (3.2), note that the value $\beta = 1$ corresponds to optimizing only the expected values ECG and $EENS$, whereas $\beta = 0$ corresponds to the opposite extreme case of optimizing only the CVaR values. Each NSGA–II run is set to perform $NG_{max} = 300$ generations over a population of $NP = 100$ chromosomes and, for the reproduction, the single–point crossover and mutation genetic operators are used. The crossover probability is $p_C = 1$, whereas the mutation probability is $p_M = 0.1$; the mutation can occur simultaneously in any bit of the chromosome.

Finally, $NS = 250$ random scenarios are simulated by the developed MCS–OPF with time step $\Delta t = 1$ (h), over an horizon of analysis of 10 years ($t^H = 87600$ (h)), in which the investment and fixed costs are prorated hourly.

5.1.2 Conditional value–at–risk: expected performance and risk trade–off

The Pareto fronts resulting from the NSGA–II MCS–OPF are presented in Figure 5.2 for the different values of β. The 'last generation' population is shown and the non–dominated solutions are marked in bold. Each non–dominated solution in the different Pareto fronts corresponds to an optimal

decision matrix $[Q^R]$ for the sizing and allocation of DG, i.e., an optimal DG–integrated network configuration $([Q], \{Y\})$ where $Q = [Q^M|Q^R]$.

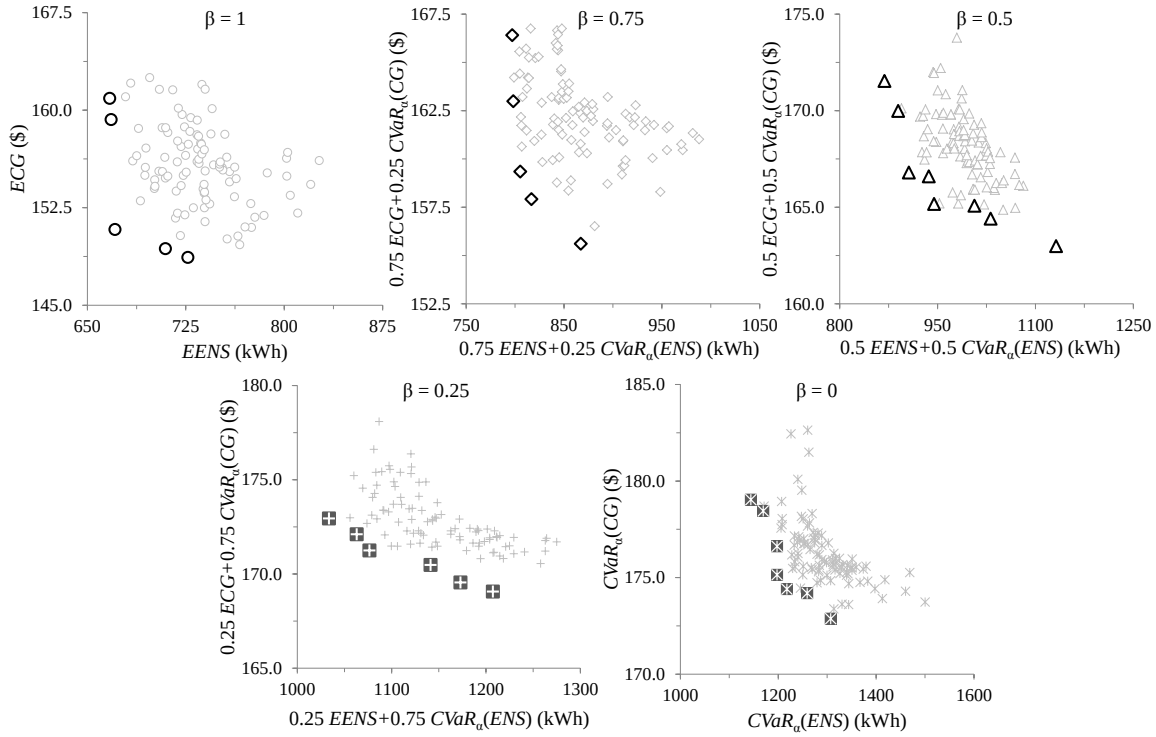

Figure 5.2 Pareto fronts for different values of β

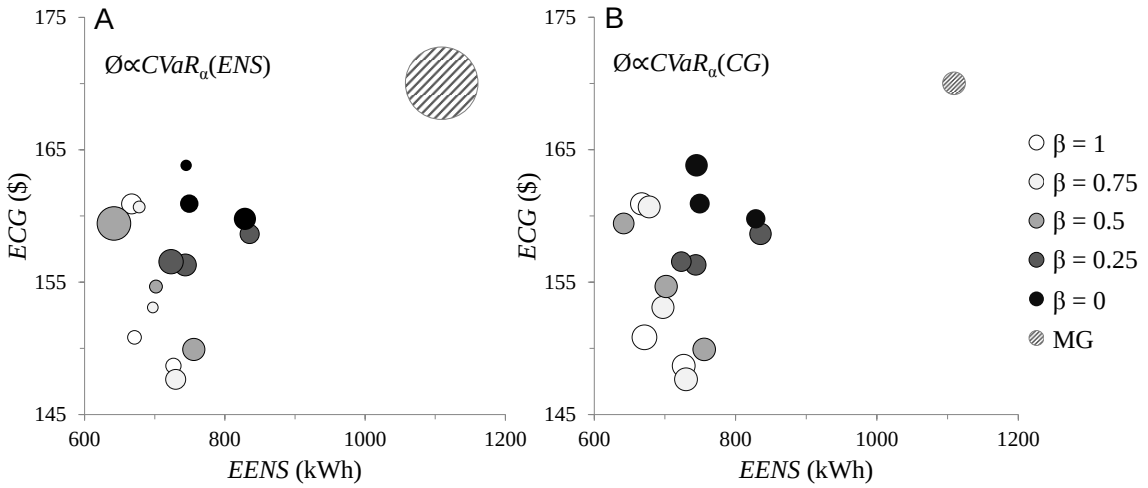

Figure 5.3 Bubble plots $EENS$ v/s ECG. Diameter of bubbles proportional to $CVaR(ENS)$ (A) and $CVaR(CG)$ (B)

In the Pareto fronts obtained, three representative non–dominated solutions are looked for the analysis: those with minimum values of the objective functions independently, denoted $[Q^R]_{\min f(CG)}$[†] and $[Q^R]_{\min f(ENS)}$[‡], respectively, and an intermediate compromised solution at the 'elbow' of the Pareto front, $[Q^R]_{f(CG)|f(ENS)}$. Table 5.1 presents the values of the objective functions, ECG, ENS and their respective CVaR values for the selected solutions. The ECG, $EENS$ and CVaR values of the case in which no DG is integrated in the network (MG case) is also reported.

Figure 5.3 shows a bubble plot representation of the selected optimal solutions. The axes report the $EENS$ and ECG values while the diameters of the bubbles are proportional to their respective CVaR values. The MG case is also plotted.

Table 5.1 Objective functions: expected and CVaR values of selected Pareto front solutions

	β	$f(CG)$ (\$)	$f(ENS)$ (kW h)	ECG (\$)	$CVaR(CG)$ (\$)	$EENS$ (kW h)	$CVaR(ENS)$ (kW h)
MG	–	–	–	170.27	179.24	1109.21	1656.53
$[Q^R]^{\min}_{f(CG)}$		160.91	666.95	160.91	185.11	666.95	1093.12
$[Q^R]^{f(CG)}_{f(ENS)}$	1	150.83	671.05	150.83	179.47	671.05	1185.53
$[Q^R]^{\min}_{f(ENS)}$		148.68	726.57	148.68	178.23	726.57	1279.37
$[Q^R]^{\min}_{f(CG)}$		166.41	797.07	160.68	183.62	677.74	1155.11
$[Q^R]^{f(CG)}_{f(ENS)}$	0.75	159.35	805.27	153.09	178.15	697.17	1129.62
$[Q^R]^{\min}_{f(ENS)}$		155.61	867.08	147.66	179.45	729.81	1278.94
$[Q^R]^{\min}_{f(CG)}$		171.54	868.61	159.43	183.64	641.68	1095.52
$[Q^R]^{f(CG)}_{f(ENS)}$	0.5	166.67	936.58	154.67	178.53	701.72	1171.47
$[Q^R]^{\min}_{f(ENS)}$		162.99	1131.64	150.45	175.58	843.53	1419.79
$[Q^R]^{\min}_{f(CG)}$		172.95	1033.65	156.55	178.42	723.19	1137.18
$[Q^R]^{f(CG)}_{f(ENS)}$	0.25	171.25	1076.53	156.32	176.24	743.61	1187.43
$[Q^R]^{\min}_{f(ENS)}$		169.07	1207.33	158.64	173.47	835.23	1331.34
$[Q^R]^{\min}_{f(CG)}$		179.03	1144.36	163.82	179.03	744.71	1144.31
$[Q^R]^{f(CG)}_{f(ENS)}$	0	176.62	1197.79	160.93	176.62	749.21	1197.74
$[Q^R]^{\min}_{f(ENS)}$		172.87	1307.33	159.78	172.87	828.55	1307.35

From Table 5.1 and Figure 5.3 it can be seen that, the MG case has an expected performance ($EENS = 1109.21$ (kW h) and $ECG = 170.27$ (\$)) inferior (high $EENS$ and ECG) to any case for which DG is optimally integrated. Furthermore, the $CVaR(ENS) = 1656.53$ (kW h) for the MG case is the highest, indicating the high risk of actually achieving the expected performance of ENS. This confirms that DG is capable of providing a gain of reliability of power supply and economic benefits, the risk of falling in scenarios of large amounts of ENS being reduced.

Comparing among the selected optimal DG–integrated networks, in general the expected performances of $EENS$ and ECG are progressively lower for increasing β. This to be expected: lowering the values of β, the MOO tends to search for optimal allocations and sizing $[Q^R]$ that sacrifice expected performance at the benefit of decreasing the level of risk (CVaR). These insights can serve

[†] $f(CG) = \beta ECG + (1-\beta)CVaR(CG)$ by (3.1)
[‡] $f(ENS) = \beta EENS + (1-\beta)CVaR(ENS)$ by (3.2)

the decision making process on the integration of renewable DG into the network, looking not only at the give–and–take between the values of $EENS$ and, but also at the level of risk of not achieving such expected performances due to the high variability.

5.1.3 Optimal DG–integrated network plans

Figure 5.4 shows the average total DG power allocated in the distribution network and its breakdown by type of DG technology for the optimal $[Q^R]$ as a function of β. It can be pointed out that the contribution of EV is practically negligible if compared with the other technologies. This is due to the fact that the probability that the EV is in a discharging state is much lower than that of being in the other two possible operating states, charging and disconnected (see Figure A.2 in Appendix A), combined with the fact that when EV is charging the effects are opposite to those desired.

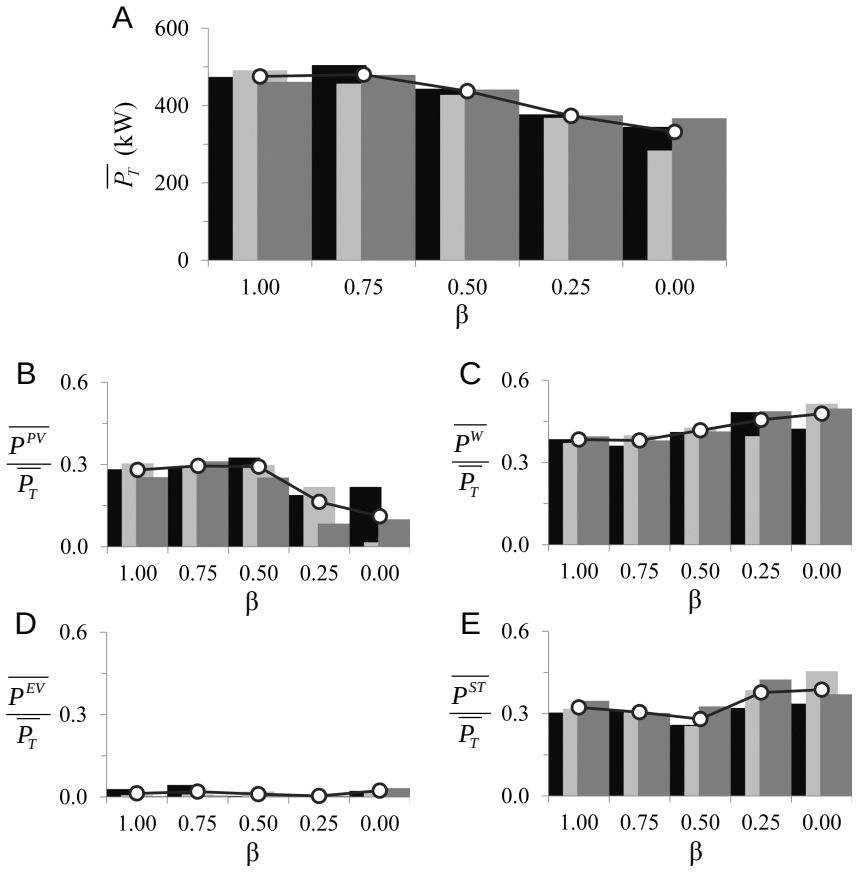

Figure 5.4 Average total DG power allocated (A) and its breakdown by type of DG: PV (B), W (C), EV (D) and ST (E)

The analysis of the results for different β values also allows highlighting the impact that each type of renewable DG technology has on the network performance. As can be noticed in Figure 5.4(A), the average total renewable DG power optimally allocated, increases progressively for increasing values

of β: this could mean that to obtain less 'risky' expected performances less renewable DG power needs to be installed. However, focusing on the individual fractions of average power allocated by PV, W and ST (Figure 5.4(B), (C) and (E), respectively), show that a reduction of the risk in the $EENS$ and ECG is achieved specifically diminishing the proportion of PV power (from $0.29_{\beta=1}$ to $0.11_{\beta=0}$) while increasing the W and ST (from $0.38_{\beta=1}$ to $0.48_{\beta=0}$ and from $0.31_{\beta=1}$ to $0.39_{\beta=0}$, respectively), but this increment of W and ST power is not enough to balance the loss of PV power due to the limits imposed by the constraints in the number of each DG technology to be installed given by τ_j. Thus, PV power supply is shown to most contribute to the achievement of optimal expected performances, but with higher levels of risk. On the other hand, privileging the integration of W and ST power supply provides more balanced optimal solutions in terms of expectations and of achieving these expectations.

Table 5.2 summarizes the minimum, average and maximum total renewable DG power allocated per node. The tendency is to install more localized sources (mainly nodes 4 and 8) of renewable DG power when the MOO searches only for the optimal expected performances ($\beta = 1$) and to have a more uniformly allocation of the power when searches for minimizing merely the CVaR ($\beta = 0$).

Table 5.2 Average, minimum and maximum total DG power allocated per node

$\bar{P_T}$ (kW)	β														
	1			0.75			0.5			0.25			0		
node	min	mean	max	min	mean	max	min	mean	max	min	mean	max	min	mean	max
1	12.08	34.44	54.77	1.15	22.40	38.56	0.00	19.23	40.98	0.00	39.03	121.00	3.00	17.33	34.71
2	2.30	40.72	69.73	0.00	49.95	77.70	36.50	58.40	123.36	3.00	63.61	132.93	0.00	42.54	84.09
3	0.00	24.83	46.45	14.80	41.79	85.03	0.00	37.94	105.11	4.00	36.87	98.53	1.00	32.84	77.78
4	76.00	110.00	133.41	1.15	67.40	133.63	0.58	38.04	80.13	6.15	20.73	61.85	0.00	39.85	85.86
5	22.60	52.39	77.08	28.90	60.66	98.59	12.63	89.39	143.50	3.30	23.49	54.25	1.00	24.97	79.64
6	12.33	55.56	85.46	10.45	21.22	38.95	2.00	27.68	106.26	12.15	53.78	84.43	0.00	50.64	116.85
7	8.00	16.52	35.38	39.38	64.07	104.05	0.00	52.03	159.73	0.00	34.09	92.81	5.00	18.51	39.23
8	79.03	111.20	146.63	30.00	74.57	114.41	0.00	40.60	146.06	4.00	37.94	102.60	1.00	39.49	119.38
9	0.00	20.03	68.73	4.00	74.07	107.88	0.00	46.72	85.61	0.00	44.06	94.08	0.00	32.86	74.53
10	0.00	9.07	25.35	0.00	1.58	7.88	0.00	11.88	58.69	0.00	8.58	43.40	0.00	30.12	83.45
11	0.00	9.98	17.68	0.00	3.04	13.20	0.00	4.74	23.45	0.00	8.99	45.95	0.00	7.31	51.17

5.1.4 Brief summary

The results obtained show the capability of the framework FW1 to identify Pareto optimal sets of renewable DG units allocations. The integration of the conditional value–at–risk into the framework and the performing of optimizations for different values of the weight parameter β has shown the possibility of optimizing expected economic and reliability of power supply performances while controlling the risk in its achievement. The contribution of each type of renewable DG technology

can also be analyzed, in terms of size and location, indicating which is more suitable for specific preferences of the decision makers.

5.2 Hierarchical clustering analysis and differential evolution (HCDE) for optimal integration of renewable DG

The current Section presents an example of application relative to the HCDE framework developed to address the computational challenge given by the expenses of large computational times required to solve DG planning optimization problems under uncertainty. Coherently to the presentation of HCDE framework in Chapter 4 and the single optimization strategy formulated in Chapter 3, Subsection 3.1.1 (FW3), the designed search engine embeds HCA within a DE search scheme to identify groups of similar individuals in the DE population to, then, calculate the objective function to be minimized, ECG, only for selected representative individuals of the groups. The HCDE evaluates the objective function by MCS–OPF and performs HCA in a controlled manner, by presetting two parameters: a threshold to the cophenetic correlation coefficient CCC_T and cutoff level coefficient p_{co}; which allows to decide whether or not is convenient to perform clustering and at what scale, respectively.

5.2.1 IEEE 13–bus test feeder case

This example regards a modification of the IEEE 13–bus test feeder distribution network [2] with the original spatial structure but neglecting the feeders of length zero, the regulator, capacitor and switch. The resulting network has $n = 11$ nodes and presents the relevant characteristics of interest for the analysis, e.g. the presence of a bulk–power supply spot and comparatively low and high spot, and distributed load values [103]. The nominal voltage V^{NET} is 4.16 (kV), kept constant for the resolution of the DC optimal power flow problem. An schematic view of the network is show in Figure 5.5. Four types of renewable DG technologies are considered to be integrated: solar photovoltaic (PV), wind turbines (W), electric vehicles (EV) and storage devices (ST).

A total of $NS = 500$ random scenarios are simulated by the MCS–OPF with time step $\Delta t = 1$ (h), over a horizon of analysis of 10 years ($t^H = 87600$ (h)), in which the investment and fixed costs are pro–rated hourly.

The DE iterations are set to perform $NG_{max} = 500$ generations over five different cases of population $NP \in \{10, 20, 30, 40, 50\}$. The differential variation amplification factor F is 1 to maintain the integer–valued definition of the individuals $[Q^R]$ after the mutation, whereas the crossover coefficient p_C is 0.1.

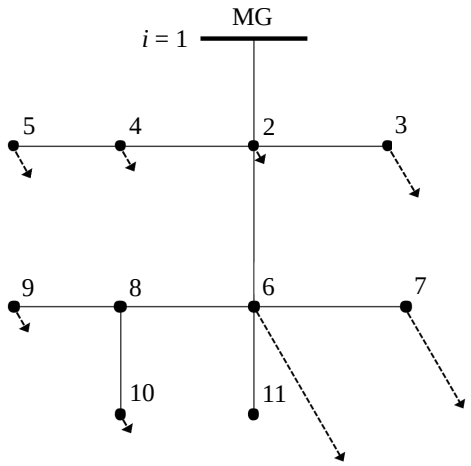

Figure 5.5 Radial 11-nodes distribution network

HCDE runs are performed under the same conditions set for DE (NG_{max}, F and p_C), but for the population size NP of 50 individuals. A sensitivity analysis is performed over the HCA control parameters, CCC_T and linkage distances cutoff level coefficient p_{co}, for all the nine possible pairs (CCC_T, p_{co}) with $CCC_T \in \{0.6, 0.7, 0.8\}$ and $p_{co} \in \{0.25, 0.50, 0.75\}$. Finally, for each of the five DE and nine HCDE settings, 20 realizations are carried out.

5.2.2 Quantification of the benefits of HCDE

The results of the DE MCS–OPF for the different population sizes $NP \in \{10, 20, 30, 40, 50\}$ are shown in Figure 5.6. The 50th percentile (%ile) or median values of the minimum global costs ECG_{min}, obtained from each experiment with fixed values of NP, are presented as functions of the respective numbers of objective function evaluations NFE; the error bars represent the 15th and 85th %iles.

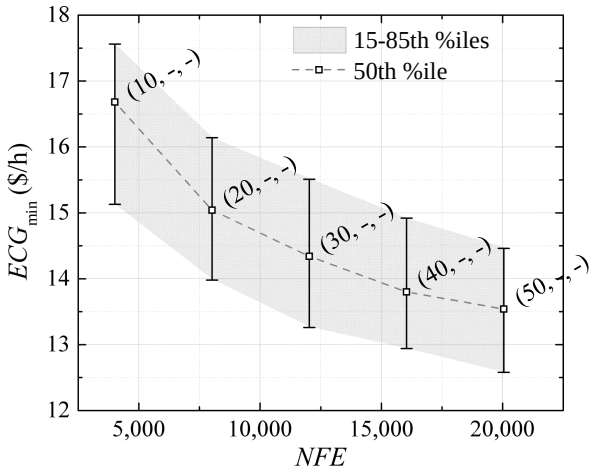

Figure 5.6 ECG_{min} vs NFE for $NP \in \{10, 20, 30, 40, 50\}$ set in DE

As expected, for the same number of generations set in the DE MCS–OPF, the larger the population size considered the lower the values of ECG_{min} obtained (better 'quality' of the minimum). Additionally, It can be observed marked tendencies in the reduction of both median and 15–85th %iles values of ECG_{min} for increasing NFE. Performing a curve fitting over these values, the following curves are gotten: $ECG_{\text{min;50th\%ile}} = 49.07NFE^{-0.13}$, $ECG_{\text{min;15th\%ile}} = 49.07NFE^{-0.115}$ and $ECG_{\text{min;85th\%ile}} = 49.07NFE^{-0.118}$, with the respective coefficients of determination $R^2_{\text{50th\%ile}} = 0.994$, $R^2_{\text{15th\%ile}} = 0.998$ and $R^2_{\text{85th\%ile}} = 0.998$. The fact that the difference between the values of the 15–85th %iles is constant indicates that the dispersion in the $ECG_{\text{min}}(NFE)$ does not depend on NP and can suggest that the global searching carried out by the DE is performed homogeneously in the feasible space that contains multiple local minima.

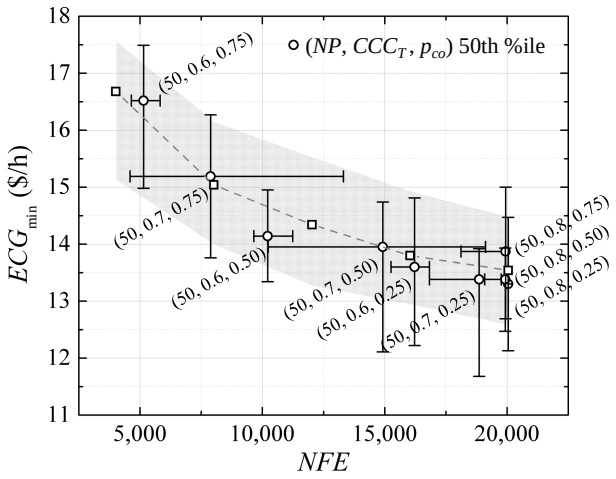

Figure 5.7 ECG_{min} vs NFE for each (NP, CCC_T, p_{co}) set in HCDE

Figure 5.7 reports the median ECG_{min} values corresponding to the HCDE MCS–OPF realizations superposed to the distribution of the median ECG_{min} and 15–85th %iles values of the base DE experiments represented by the square markers and shaded area, respectively. The vertical and horizontal error bars account for the 15–85th %iles of the outcome ECG_{min} and NFE values.

Focusing on CCC_T, it can be noticed that for the two extreme cases, $CCC_T = 0.6$ and 0.8, the dispersion of the number of objective function evaluations is relatively small. On the contrary, the cases with a $CCC_T = 0.7$ present high variability. This can be explained by the behavior of the CCC along each generation g in the evolution loop. Figure 5.8 shows the median, 15th and 85th %iles CCC values as a function of generation g derived from all HCDE MCS–OPF realizations. On the one hand, recalling that CCC_T is used to control whether it is convenient to perform HCA, the small NFE dispersion in the case with $CCC_T = 0.6$ is because clustering is practically been applied in all generations ($CCC_T \leq CCC^g$), thus disabling any effect generated by passing from populations with original size NP to reduced populations with $NP^g \leq NP$ and vice versa. On the other hand, the effect is also being avoided in the case $CCC_T = 0.8$ by not applying clustering. Indeed, in Figure 5.8

it can be observed that after the generation number 50 it is unlikely that by performing HCA the proposed hierarchical grouping structures represent well enough the population.

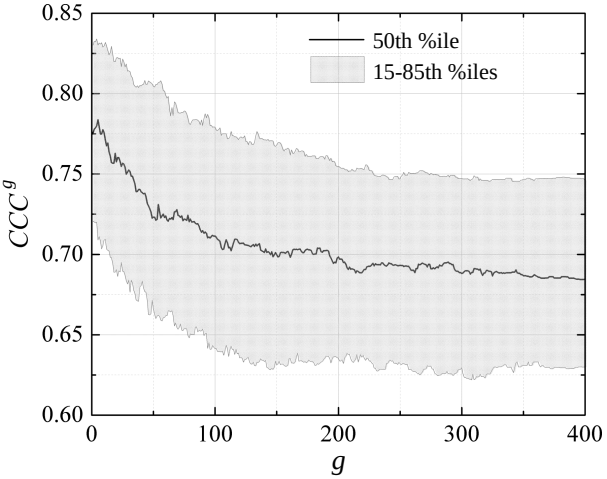

Figure 5.8 CCC behavior per generation g

Differently, the cases for which $CCC_T = 0.7$ present high dispersion in the NFE since the median values of CCC^g move in the neighborhood of the threshold throughout the major part of the evolution loop in the HCDE. Moreover, in general terms, the values of CCC^g 15–85th %iles maintain certain symmetry with respect to the median, i.e., performing or not HCA are equally likely events, producing high fluctuations in the number of individuals considered as population and, therefore, affecting in the same way the NFE.

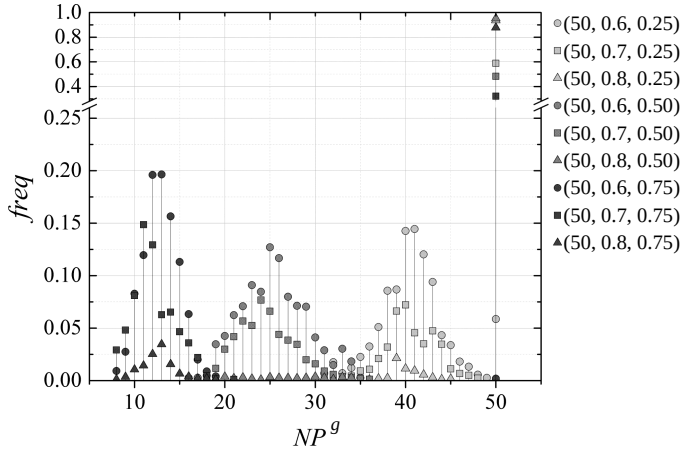

Figure 5.9 Empirical NP^g pdf for each (NP, CCC_T, p_{co}) set in HCDE

The above mentioned insights are noticeable also in Figure 5.9, which shows the empirical probability density functions (pdfs) of the population size NP^g per generation for each (NP, CCC_T, p_{co}) set in HCDE. Indeed, the average probabilities of performing HCA throughout the evolution cycle for the different values of $CCC_T = 0.6$, 0.7 and 0.8 are 0.980, 0.540 and 0.078, respectively.

Regarding the cutoff level coefficient of the linkage distances p_{co}, in Figure 5.9 it is possible to identify the three peaks of reduction in the population size, confirming the role of this control parameter in defining the level at which the hierarchical structures proposed are 'cut off' when the HCA takes place. In fact, lower values of p_{co} imply smaller reduction in the population size because of the higher demand of proximity between individuals or groups of individuals. In the opposite side, higher values of p_{co} allow forming clusters from individuals or groups which are relatively less similar.

From the results obtained for all the different DE and HCDE settings, six representative cases are the focus of the analysis (Figure 5.7). From the DE runs, the settings with extreme and middle population size $NP \in \{10, 30, 50\}$ are selected, whereas from HCDE, the cases (NP, CCC_T, p_{co}) set as (50, 0.6, 0.25), (50, 0.6, 0.50), (50, 0.7, 0.50) and (50, 0.7, 0.75) are chosen. The former (50, 0.6, 0.25) and (50, 0.6, 0.50) cases present significant reductions in the number of NFE, with small dispersion and loss of quality of the $ECG_{\min}$ obtained, compared to the results regarded by diminishing directly the fixed NP in DE from 50 to 10. Similarly, the cases (50, 0.7, 0.50) and (50, 0.7, 0.75) may lead to considerable reductions in NFE, with acceptable losses of $ECG_{\min}$, but subject to a high degree of variability that compromises the performance.

As for computational times, running on an Intel® Core™i7-3740QM (PC) 2.70 GHz without performing parallel computing, the average time to evaluate the objective function is 4.592 (s) for the $NS = 500$ scenarios in the MCS–OPF; for a fixed population of $NP = 50$ and its corresponding $NFE = 20050$, the total time for a single run is on average 25.574 (h). Taking into account this, under commonly used hardware configurations, the reductions in computational time that can be achieved by using HCDE with (50, 0.6, 0.25) and (50, 0.6, 0.50) settings are 19% and 49% for the median, 23% and 51% for the 15th %ile, and 16% and 43% for the 85th %ile, respectively.

5.2.3 Identification of time complexity conditions and limits

The integration of HCA into the DE algorithm introduces a significant time complexity, conditioning the reductions of computational efforts that can be obtained by applying the proposed HCDE MCS–OPF framework. Indeed, if performing HCA along all generations of DE and running the MCS–OPF on an eventually reduced population (depending on CCC_T and p_{co}) is computationally heavier than running the MCS–OPF over the complete population, the effects of the framework can be negligible or even negative. It is possible to formulate the condition to obtain reductions in the computational efforts by the proposed HCDE MCS–OPF framework, from the asymptotic time complexities of the main algorithms that compose it. Table 5.3 reports the independent asymptotic time complexities as functions of the generic size m of the input to each algorithm and of the parameters that define the dimensionality of the HCDE MCS–OPF framework [102, 104].

Table 5.3 Asymptotic time complexity of the algorithms

Algorithm			
PDIST	HC	MCS	OPF
Time complexity T			
$O(dm^2)$[a]	$O(m^2\log(m))$	$O(m)$	$O(\text{size}(A))$[b]
$O(n(m+r)\times NP^2)$	$O(NP^2\log(NP))$	$O(NS\times n(m+r))$	$O(NS\times(n(m+r))^2)$

where $n(m+r)$ represents the size of the DG-integrated network, i.e., the number of nodes n times the number of all the types or technologies of power generation available, m types bulk–power suppliers and r DG types, NP is the size of the complete population and NS is the number of scenarios in the MCS–OPF.

[a] Pairwise distance PDIST between all m vectors of size d.

[b] The matrix A comes from the canonical form $Ax \leq b$ of the linear programming of the DC–OPF problem approximation.

Comparing the asymptotic time complexities of the algorithms involved in the realization of the proposed framework with and without integrating HCA, the following inequalities must be fulfilled in order to obtain a reduction in the computational time by HCDE:

$$\overbrace{T(n(m+r),NP)}^{\text{PDIST}}+\overbrace{T(NP)}^{\text{HC}}+E(NP^g)\times\overbrace{T(NS,n(m+r))}^{\text{MCS–OPF}}<NP\times\overbrace{T(NS,n(m+r))}^{\text{MCS–OPF}}$$

$$\Downarrow$$

$$n(m+r)\times NP^2+NP^2\log(NP)+E(NP^g)\times NS\times(n(m+r))^2<NP\times NS\times(n(m+r))^2$$

$$\Downarrow$$

$$\kappa=\frac{NP}{NP\times n(m+r)}+\frac{NP\log(NP)}{NP\times(n(m+r))^2}+\epsilon<1 \tag{5.1}$$

$$\forall n,(m+r),NP,NS\in\mathbb{Z}^*,\epsilon=\frac{E(NP^g)}{NP}\in(0,1]$$

where ϵ is the expected ratio of the population NP^g evaluated along all generations g of DE to the total population NP and κ is the ratio of the asymptotic time complexities of HCDE to DE.

By Equation (5.1), the contribution of the terms related with the complexity of MCS–OPF, dependent on NS and $n(m+r)$, is considerably large for the fulfillment of the inequality conditions. In fact, when using DE, it is commonly accepted to set a size of the population NP not greater than ten times the size of the decision variables, in this case, $10n(m+r)$ [93], making the first two terms of κ strongly dependent on the number of scenarios NS. Moreover, given the complexity of the general problem, higher values of NS lead to a better approximation of the objective function via MCS–OPF, i.e., the more likely is to fulfill the condition and the greater can be the reduction of computation time. However, the value of ϵ depends on the probability of performing clustering in each generation and at what level, controlled by CCC_T and p_{co} respectively. In some cases, ϵ can be close to 1 (as it is inferred from Figure 5.9) implying negligible benefits. Table 5.4 shows the

values of the ratio κ for each $(NP,\ CCC_T,\ p_{co})$ set in HCDE considering the dimensionality of the present case study defined by the values of the parameters $n(m+r)=55$, $NS=500$, $NP=50$. The value of $1-\kappa$ can be interpreted as the expected asymptotic relative time reduction achieved by performing HCDE.

Table 5.4 Ratio κ for each $(NP,\ CCC_T,\ p_{co})$

$(NP,\ CCC_T,\ p_{co})$	$\frac{NP}{NP\times n(m+r)}$	$\frac{NP\log(NP)}{NP\times(n(m+r))^2}$	$\epsilon=\frac{E(NP^g)}{NP}$	κ	$1-\kappa$
(50, 0.6, 0.25)			0.817	0.819	0.181
(50, 0.7, 0.25)			0.921	0.923	0.077
(50, 0.8, 0.25)			0.987	0.989	0.011
(50, 0.6, 0.50)			0.510	0.512	0.488
(50, 0.7, 0.50)	1.818e−03	3.418e−05	0.738	0.740	0.260
(50, 0.8, 0.50)			0.978	0.979	0.021
(50, 0.6, 0.75)			0.259	0.261	0.739
(50, 0.7, 0.75)			0.487	0.488	0.512
(50, 0.8, 0.75)			0.909	0.911	0.089

5.2.4 Optimal DG–integrated network plans

Figure 5.10 shows the average total DG power allocated in the distribution network and the corresponding investment costs of the DE and HCDE MCS–OPF cases selected, choosing the corresponding optimal DG–integrated plans as the decision matrices $[Q^R]$ for which their $ECG_{\min}$ values are the closest to the median $ECG_{\min}$ value obtained for the twenty runs of each $(NP,\ CCC_T,\ p_{co})$ setting. It can be pointed out that in all the cases, the contribution of EV is practically negligible if compared with the other technologies. This is due to a combination of two facts: the probability that the EV is in a discharging state is much lower than that of being in the other two possible operating states, charging and disconnected (see Figure B.2 in Appendix B) and when EV is charging, the effects are opposite to those desired, i.e., it is acting as loads.

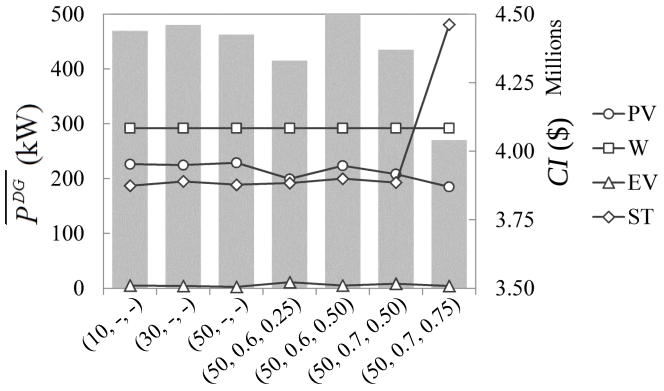

Figure 5.10 Average total DG power allocated and investment cost for representative $(NP,\ CCC_T,\ p_{co})$ settings

In all generality, both the investment cost CI and the average power installed by DG is comparable in all the cases, except for the setting (50, 0.7, 0.75) for which the level of clustering determined by $p_{co} = 0.75$, that translates into higher reductions of the population size, may lead to less similar local minima than the other settings.

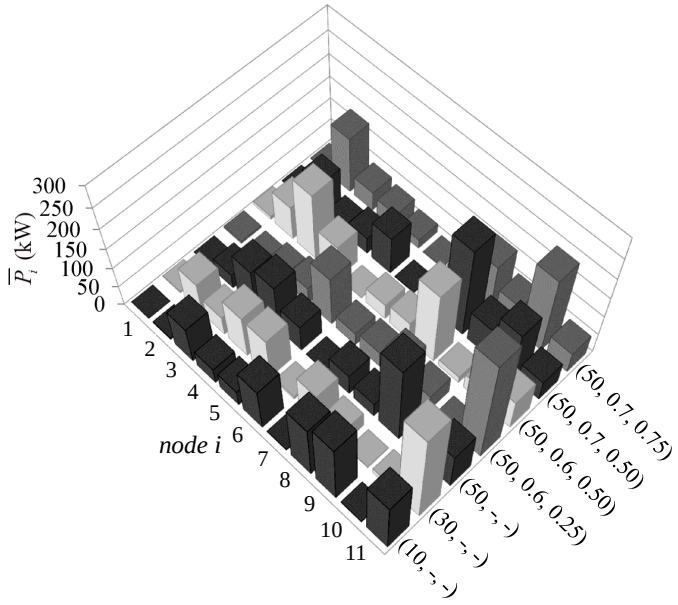

Figure 5.11 Nodal average total DG power for representative (NP, CCC_T, p_{co}) settings

The average total renewable DG power allocated per node is summarized in Figure 5.11. Even though all the ECG optimal decision matrices $[Q^R]$ show differences, the tendency is to install localized sources of renewable DG power between two identifiable portions of the distribution network, up and downstream the feeder (2,6) (Figure 5.5), giving preference to the second portion which presents higher and non-stream homogeneous nodal load profiles.

5.2.5 Brief summary

The results show that the the framework is effective in finding optimal DG–integrated network plans, with acceptable reductions of the computational efforts required while maintaining small dispersion and loss of quality in the minimum ECG obtained. The sensitivity analysis over the control parameters of the HCA suggest that the efficiency is improved with CCC_T that allows the clustering in almost all generations along the DE search, setting the level of clustering to no more than the 50% ($p_{co} = 0.5$) of the feasible linkage distances range in the hierarchical structure proposed.

5.3 A MOO and MCS–OPF simulation framework for risk–controlled integration of DG

In this Section, an example of application of the MOO framework for the integration of renewable DG into electric power networks is reported. The framework searches for the optimal size and location of different DG technologies, taking into account uncertain ties related to primary renewable resources availability, components failures, power demands and bulk-power supply. The network operation is emulated by the developed MCS–OPF computational model, assessing the system performance in terms of CG. To measure uncertainty, the DCVaR is introduced enabling the conjoint control of risk due to its axiomatic relation to the CVaR. A MOO strategy is adopted for the simultaneous minimization of the ECG and the associated deviation $DCVaR(CG)$. This is operatively implemented by the MOO–DE described in Chapter 3, Section 3.2 and respecting the formulation relative to framework FW2, to be found in the same cited Chapter.

5.3.1 IEEE 30 bus sub-transmission and distribution test system

The IEEE 30 bus sub-transmission and distribution test system is regarded for the analysis [11]. This constitutes a portion of the Midwestern U.S. electric power system which presents relevant characteristics of interest for the analysis, e.g., the presence of bulk-power supply spots different in type and with comparatively low and high nodal load profiles. An important consideration is that the synchronous condensers are neglected, given the DC assumptions made in the current thesis work for the resolution of the OPF problem.

The network consists of $n = 30$ nodes, a mesh deployment of 41 T&D lines and 2 transformers or bulk-power suppliers, as shown in Figure 5.12. The reference apparent power is $S_{ref} = 100$ (MV A). In this case, two renewable DG technologies are considered: solar photovoltaic (PV) and wind turbines (W).

Table 5.5 summarizes the main parameters set for the general MOO–DE and MCS–OPF framework.

Table 5.5 MOO-DE and MCS–OPF parameters

Parameter	Nomenclature	Value
Population size	NP	100
Maximum n° of generations	NG_{max}	600
Crossover coefficient	p_C	0.1
Differential variation amplification factor	F	1
N° of MCS-OPF scenarios	NS	24000
Scenario duration (h)	Δt	1

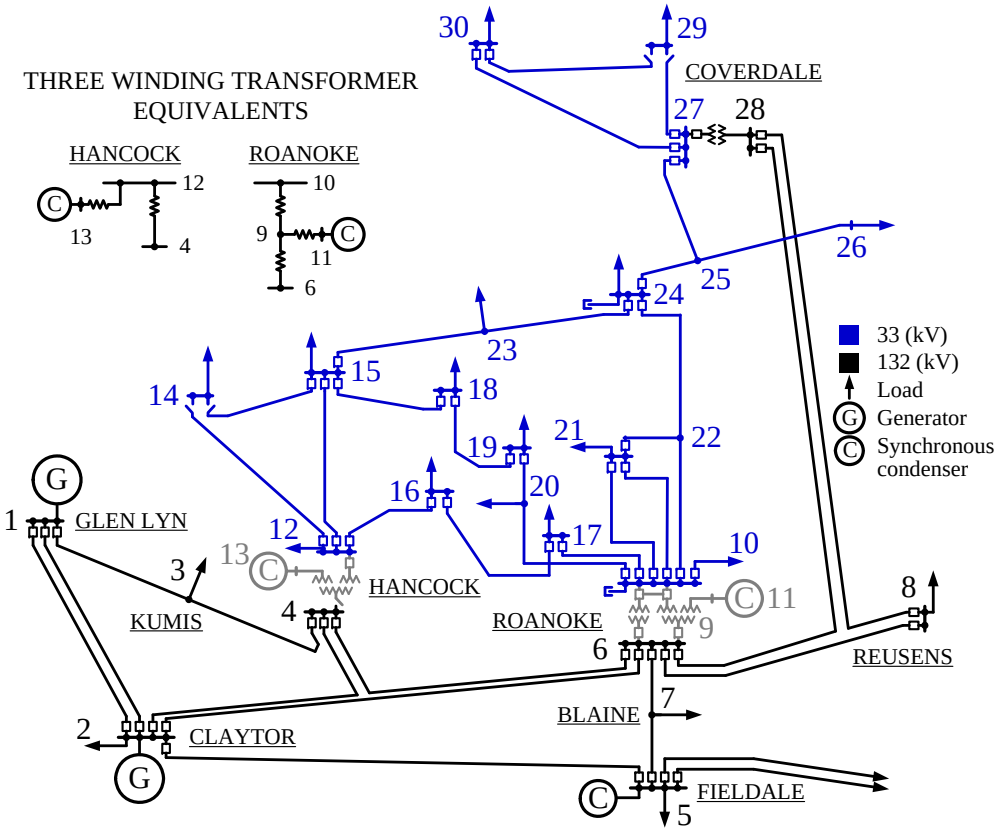

Figure 5.12 IEEE 30 bus sub-transmission and distribution test system diagram

5.3.2 Conditional value–at–risk deviation: expected performance, uncertainty and risk trade–off

The Pareto front resulting from the MOO-DE MCS–OPF is presented in Figure 5.13. The entire last generation population at convergence is shown by gray squares and the non–dominated solutions are the blue bullets. The base case (MG) in which no DG is integrated in the network is also shown as dark gray triangle. Each solution in the Pareto Front corresponds to a decision matrix $[Q^R]$ that indicates the number of units and locations of the different types of DG technologies integrated in the network.

Recalling the relation given by equation (2.29), $DCVaR_\alpha(x) = CVaR_\alpha(x - E(x))$, it is possible to draw a map of iso–CVaR curves in the plot of non–dominated solutions and so, to include risk into the trade–off between expected performance and uncertainty, represented by ECG and $DCVaR_\alpha(CG)$, respectively. Then, the compromised solution, namely $[Q^R]^{\min}_{CVaR(CG)}$ can be found, that minimizes risk as reported in Figure 5.14(A). The reciprocal case, i.e., a map of iso–DCVaR curves is drawn in the distribution of non–dominated solutions plotted as ECG vs $DCVaR\alpha(CG)$, and it is shown in Figure 5.14(B). Three representative non–dominated solutions are considered for the analysis: those with minimum values of ECG and $DCVaR_\alpha(CG)$, denoted $[Q^R]^{\min}_{ECG}$ and $[Q^R]^{\min}_{DCVaR(CG)}$ respectively, and $[Q^R]^{\min}_{CVaR(CG)}$ which, as mentioned earlier, minimizes risk.

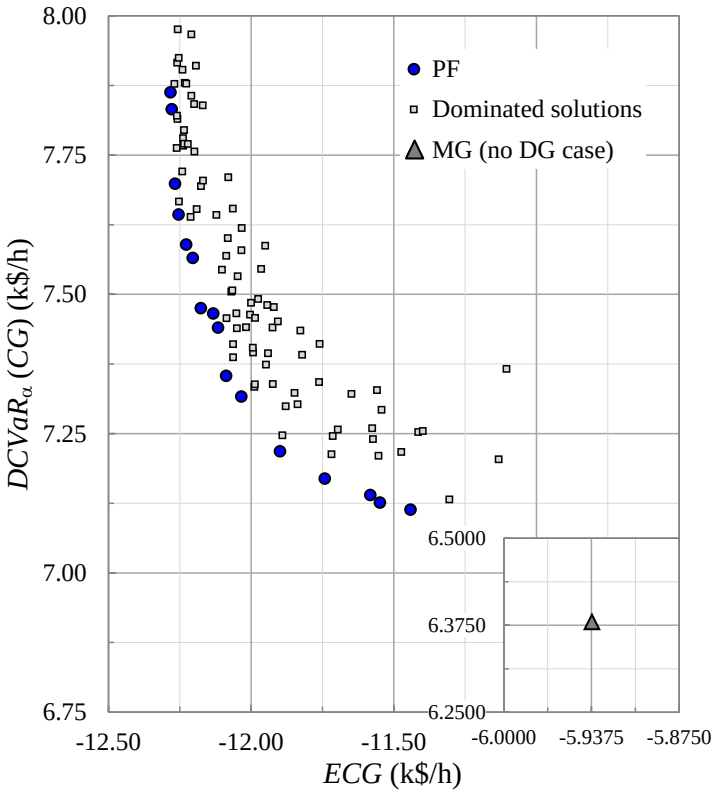

Figure 5.13 Set of non-dominated solutions: Pareto front

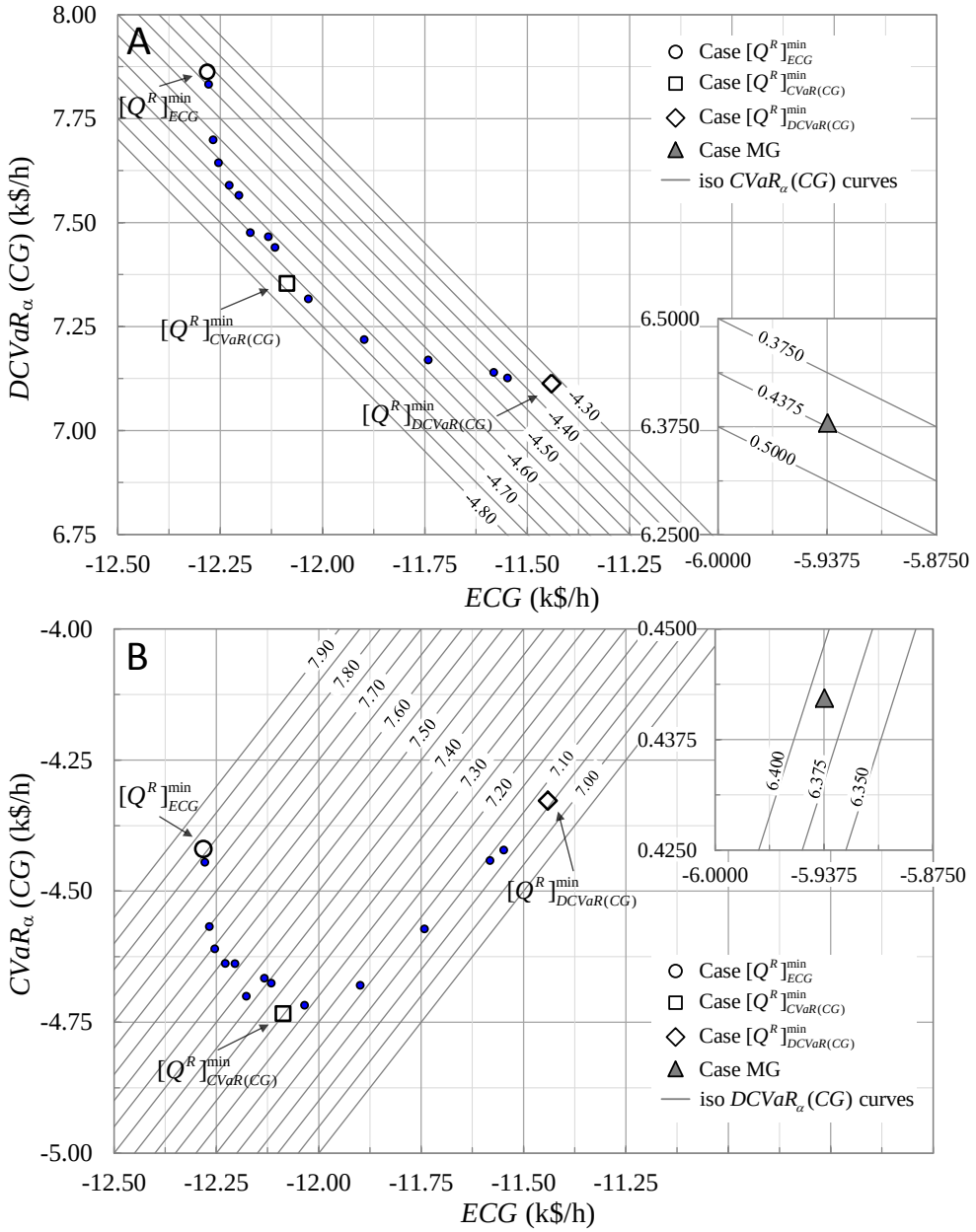

Figure 5.14 Set of non–dominated solutions with iso–CVaR (A) and iso–DCVaR (B) curves

In Figure 5.14, it can be observed that the three solutions of interest, $[Q^R]^{min}_{ECG}$, $[Q^R]^{min}_{CVaR}$ and $[Q^R]^{min}_{DCVaR}$, lead to considerable improvements in expected performance and risk with respect to the base case MG, in which no renewable generation is integrated into the network. However, the level of uncertainty in the ECG estimation is increased, in all DG–integrated solutions, because of their stochasticity. Even so, in comparison to the MG case, the increase in the level of uncertainty for all DG–integrated cases (on average 1.067 (k$/h)) is much less than the gain in both expected

performance and risk (on average −6.035 and −4.967 (k$/h)), respectively). This fact can be seen also in the empirical CG probability density functions (*pdf*) shown in Figure 5.15.

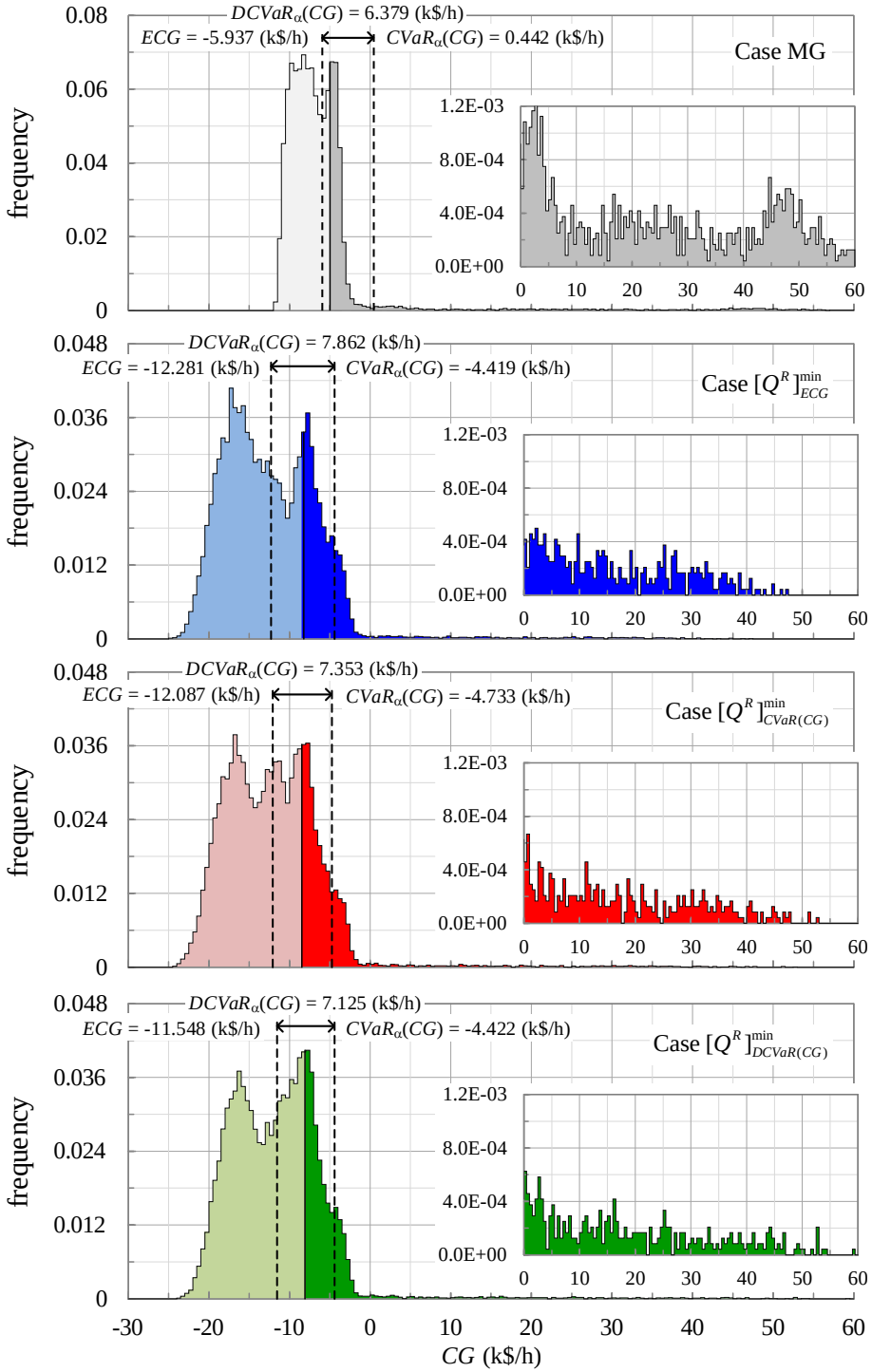

Figure 5.15 Empirical CG probability density functions

Furthermore, it can be inferred from Figure 5.15 that, in general, all CG empirical *pdf*s show three main peaks. This is due to the characteristics of the daily load profiles (Figure C.1 in Appendix C) which present three important ranges: a low power demand range during the night, between 23 and 6 hours and two high ranges of load taking place in the intervals 10 to 13 and 18 to 21 hours, respectively. Thus, the left peak of the distributions corresponds to the highest range of loads (18 to 21 hours), because higher levels of power demands imply more energy sold and, therefore, more profits or negative values of CG. Following the same logic, the central peak is due to the second highest range of loads (10 to 13 hours) and the right peak to the low range of power demand. As mentioned before, the three DG–integrated network cases improve the cost profiles because the usage of power is transferred from the bulk–power suppliers (MG) to the renewable generators (RG) as summarized in Table 5.6, presenting the ratio of power usage defined as the proportion of power used to satisfy the loads, determined by the MCS–OPF, over the power available. According to the IEEE 30 bus test systems information, the bulk–power supply arriving at node 1, G_1, comes from a conventional coal–fired power plant whereas G_2 is supplied by a hydro–power plant. The nature of the source of power supply conditions the operating costs, in particular, G_1 that presents the higher variable operating costs (Table C.3 in Appendix C). Nevertheless, the ratio of power usage associated to G_2 is in all cases 100%, even though its operating cost is higher than the renewable. This is because no investment is being paid for G_2, contrary to the DG technologies (PV and W). In this view, considering investment, PV and W are more convenient than the coal–fired power supply G_1 but not more than the hydro–power supply G_2.

Table 5.6 Ratio of power usage by type of generator

Case	G_1	G_2	PV	W
MG	77.93%	100.00%	–	–
$ECG^{\min}$	41.88%	100.00%	99.96%	99.94%
$CVaRCG^{\min}$	40.38%	100.00%	99.95%	99.42%
$DCVaRCG^{\min}$	40.71%	100.00%	100.00%	98.73%

The extreme CG scenarios encountered in the tail of the distributions are mainly produced by the occurrence of failures in the components of the system, power generators and T&D lines. In Figure 5.15, the stability of the CG to these non-desirable events can be pointed out. Since the cost objective function to be minimized in the OPF considers a load shedding cost, the occurrence of failures in the components, interrupting the power supply and/or the ability of distribute it, will impact the CG function depending on how much centralized is the power supply. Focusing on the MG base case, the power supply and its distribution depend on two generators and the T&D lines connected to them: the reliability of power supply is determined by few components and so, the eventual losses of functionality of these components can lead to high amounts of non–satisfied demands. This is precisely the effect observed in the tails of the CG distributions and, in particular, it is seen that distributing and diversifying generation helps to improve the risk impacts from multiple

failures in the network, even if the number of generation units in the system is increased and with it, the overall absolute likelihood of occurrence of failures.

5.3.3 Optimal DG–integrated network plans

Figure 5.16 shows the total average DG power integrated in the network for the three solutions under analysis and the corresponding obtained values of ECG and $CVaR_\alpha$. It can be pointed out that in all cases the DG power installed is almost equal to the limit value set by the penetration factor $PF^{DG} = 0.3$, approximately 135.5 (MW).

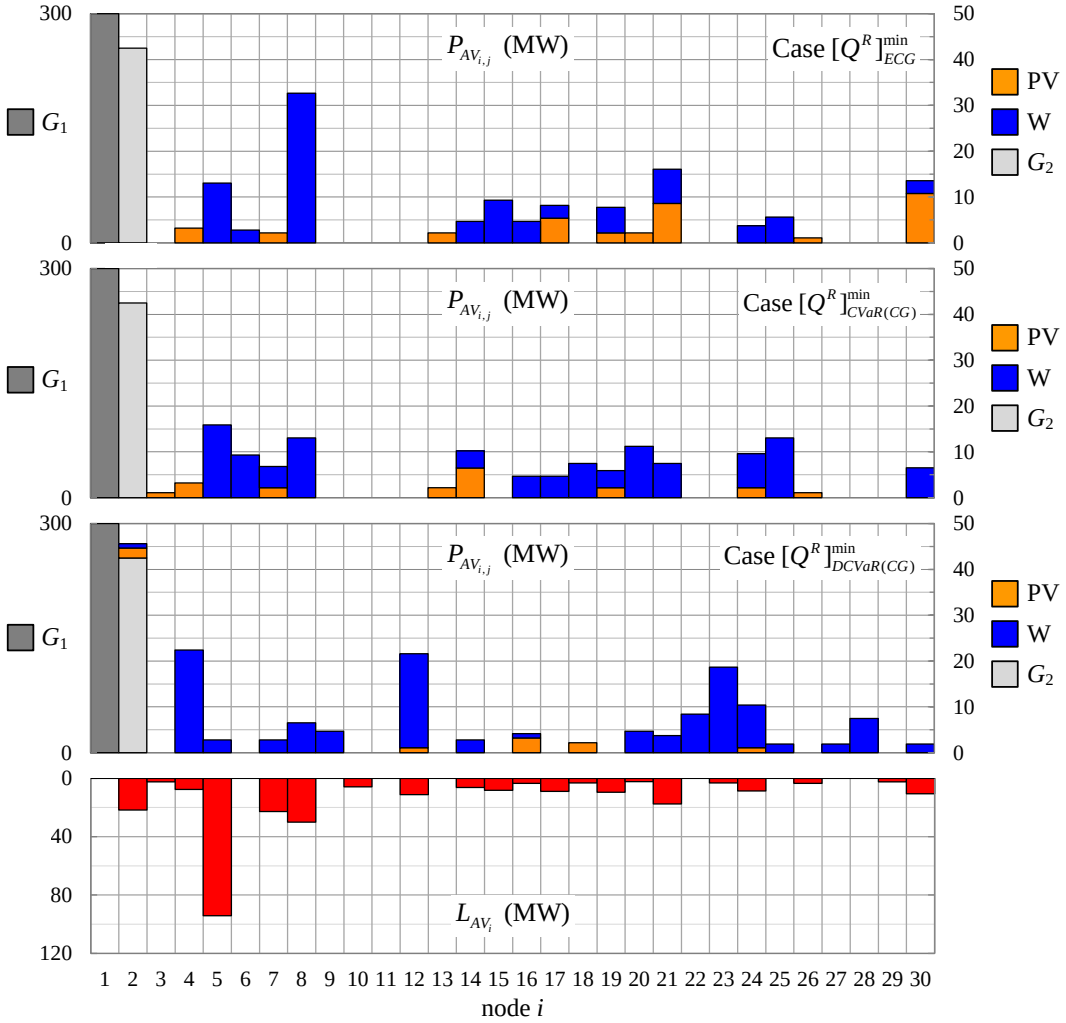

Figure 5.16 Total average renewable power installed. Cases $ECG^{\min}$ (A), $CVaRCG^{\min}$ (B) and $DCVaRCG^{\min}$ (C)

Furthermore, in general terms moving along the non–dominated solutions, starting from the one that minimizes ECG ($[Q^R]^{\min}_{ECG}$), passing to the compromising one that minimizes $CVaR_\alpha(CG)$ ($[Q^R]^{\min}_{CVaR(CG)}$) and ending with ($[Q^R]^{\min}_{DCVaR(CG)}$) that minimizes $DCVaR_\alpha(CG)$, the amount of PV

power integrated in the network decreases progressively, being replaced by W power. Then, it can be inferred that the more PV power generators are integrated, the better ECG performance is achieved. This seems physically coherent, taking into account that the average power outputs delivered by one generation unit of PV and W technologies are 1.07 and 0.93 (MW), respectively, and the cost performance benefits comparatively more from the PV MW h sold during the daylight interval within the two peaks ranges of power demand. Nevertheless, the amount of average integrated PV power is invariably smaller than W power, this is because PV generation units do not supply during the night interval, making convenient to integrate always a certain amount of W power and, thus, to avoid resorting to the more expensive coal–fired power supplier G_1. Moreover, it is precisely this lack of PV generation during the night interval that strongly conditions the trade–off between expected performance and uncertainty, ECG vs $DCVaR_{\alpha}(CG)$, i.e., the more PV power is integrated, the better ECG performance is achieved but the more uncertain is its estimation.

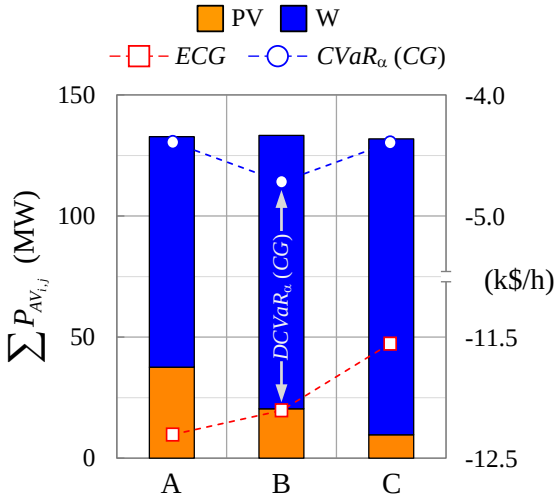

Figure 5.17 Nodal average power by type of generator

Concerning the compromising solution $[Q^R]^{\min}_{CVaR(CG)}$ that minimizes risk, as it was derived from the empirical CG distributions, the risk associated to a solution depends mainly to the occurrence of extreme non-desired events. Improvements in risk performance can, then, be achieved by solutions for which the allocation of DG power generation units decentralizes and diversifies to a large extent the supply. This insight is noticeable in Figure 5.17 that reports the nodal average power by type of generation for the three solutions of interest. For both extreme solutions, $[Q^R]^{\min}_{ECG}$ and $[Q^R]^{\min}_{DCVaR}$, the tendency is to integrate localized sources of renewable DG at two identifiable portions of the network, in the region close to nodes 2, 5, 7 and 8 of the sub–transmission portion of the network, and nodes 17, 19 and 21 in the distribution part, favoring the sub–transmission portion which presents higher and non homogeneous nodal load profiles. In a different manner, the solution $[Q^R]^{\min}_{CVaR(CG)}$ presents a more homogeneous deployment of DG power, allocating comparable generation capacities in both sub–transmission and distribution parts of the network.

5.3.4 Brief summary

We have presented a MOO framework for the integration of renewable distributed generation into an electric power network. Multiple uncertain operational inputs are taken into consideration: the inherent uncertain behavior of renewable energy sources and power demands, as well as the occurrence of failures of components. For managing the uncertainty and risk associated to the achievement of a certain level of expected global cost performance, we have introduced the conditional-value-at-risk deviation measure, which allows trading off the level of uncertainty and, given the axiomatic relation to the conditional-value-at-risk, enables conjointly the trade-off of risk by constructing an iso–risk map in the non–dominated set of solutions. The proposed framework integrates the multi-objective differential evolution as a search engine, Monte Carlo simulation to randomly generate realizations of the uncertain operational scenarios and optimal power flow to evaluate the network response. The optimization is done to simultaneously minimize the expected value of the global costs and the respective conditional–value–at–risk deviation.

A case study has been analyzed, based on the IEEE 30 bus sub–transmission and distribution test system. The results obtained show the capability of the framework to identify Pareto optimal solutions of renewable DG units allocations. Integrating the conditional value–at–risk deviation into the framework has shown effectively the possibility of optimizing expected performances while controlling the uncertainty and risk, analyzing, in addition, the contribution of each type of renewable DG technology on the level of uncertainty associated to the outcome performance of the optimal solutions and the importance of the deployment of the renewable generation capacity to lower the risk of incurring in non-desirable extreme scenarios. In this view, a complete and comprehensible spectrum of information can be supplied in support of specific preferences of the decision makers for their decision tasks.

6 Conclusions

In the present thesis, a modeling, simulation and optimization framework has been designed and developed, for the integration of renewable DG into electric power networks. The DG planning problem considered has been that of optimal selection of the technology, size and location of multiple DG units, subject to technical, operational and economic constraints. Key research questions addressed have been:

(i) Representation and treatment of the multiple uncertain operational inputs such as: the inherent variability in the availability of diverse primary renewable energy sources, bulk–power supply, power demands and components operating states.

(ii) Propagation of the uncertainties onto the system operational response, and control of the associated risk of incurring in extreme non–satisfactory performances.

(iii) Computational efforts resulting from the complex combinatorial optimization problem under uncertainty associated to renewable DG integration.

6.1 Original contributions

The original contributions of the work reside in:

(a) Design and implementation of a non–sequential Monte Carlo simulation (MCS) and optimal power flow (OPF) computational model, denoted MCS–OPF, that emulates the T&D network operation integrating a given renewable DG plan. Random realizations of operational scenarios are generated by latin hypercube (LHC) sampling from the different uncertain variables models, solving for each scenario a cost–based DC–OPF problem to assess the response of the DG–integrated network in terms of, available power usage, power demand satisfaction and operating cost (generation and T&D) to , then, evaluate the performance of the system with regards to the global economics, including DG investments and operation, and reliability of power supply, represented by the global cost (CG) and the energy not supplied (ENS), respectively. The simplifications involving the non–sequential character of the MCS and the use of DC–OPF approximation are done to not overly increasing the computational efforts and

gaining tractability, necessary to cope with the already complex DG planning optimization problem.

(b) With respect to the optimal technology selection, size and location of the renewable DG units, two distinct multi–objective optimization (MOO) strategies have been implemented by heuristic optimization (HO) search engines, in which the MCS–OPF model is nested to assess the performance of each DG–integrated network proposed along the evolutionary searching process and the values of the respective objective functions. Two HO search engines have been considered, the fast non–dominated sorting genetic algorithm II (NSGA–II) and a MOO differential evolution (MOO–DE), both capable of dealing with non-convex combinatorial problems, discontinuous search spaces and non-differentiable objective functions.

(c) Introduction of two indicators to measure and control the uncertainty and risk associated to the DG–integrated network solutions, namely conditional value–at–risk (CVaR) and conditional value–at–risk deviation (DCVaR), respectively. This is done, framing the DG planning problem as a portfolio optimization in which the different types of DG technologies are treated analogously to financial assets.

- By introducing CVaR, a direct risk-based MOO strategy is formulated and approached by the NSGA–II, allowing the possibility of concurrently minimizing expected performances, ECG and $EENS$, while controlling the risk in its achievement, $CVaR(CG)$ and $CVaR(ENS)$. The contribution of each type of renewable DG technology, on the expectation–risk trade–off, can also be analyzed, indicating which is more suitable for specific preferences of the decision makers.
- Alternatively, DCVaR allows the development of a distinct MOO optimization strategy, aiming at the simultaneous minimization of the considered objective functions: ECG and its corresponding $DCVaR(CG)$ value and implemented by the MOO–DE search. DCVaR acts as an enabler of trade-off between optimal expected performance and the associated uncertainty to achieve it and, given the axiomatic relation to the conditional–value–at–risk, allows the conjoint trade–off of risk by constructing an iso-risk map in the obtained non-dominated set of solutions. In addition, the contribution of each type of renewable DG technology on the level of uncertainty associated to the outcome performance of the optimal solutions can be analyzed, as well as the importance of the deployment of the DG capacity to lower the risk of incurring in non-desirable extreme scenarios. In this view, a complete and comprehensible spectrum of information can be supplied in support of the decision making tasks.

(d) To cope with the large computational efforts required by the developed MOO frameworks with nested MCS–OPF, an original technique is introduced which embeds hierarchical clustering analysis (HCA) within a DE search engine. The technique identifies, in a controlled manner, groups of similar individuals (DG plans) in the DE population and, then, evaluates ECG

performing MCS–OPF on selected representative individuals of the groups only, thus reducing the number of objective function evaluations in each iteration of the DE evolution loop. The introduction of two control parameters, namely the cophenetic correlation coefficient CCC and a cutoff level coefficient of the linkage distances p_{co}, allow the controlled adaptation during the search process and decision on whether or not to perform clustering and at which level of the hierarchical structure. The HCDE framework is capable of identifying optimal plans of renewable DG integration, leading to acceptable reductions in the number of objective function evaluations with small dispersion and loss of quality in the minimum ECG obtained.

6.2 Future work

Further developments can be thought of in all three main areas of focus of this thesis: modeling, simulation and optimization for the DG planning problem. Regarding the modeling of the distribution network and, in particular, the technical constraints imposed by power flow equations, more realistic assumptions can be considered, including the impact of DG in the reactive power balance of the network and, therefore, the voltage and power losses profiles, that are regulated and ought to be strictly controlled in electric power systems. Even if, at small scales no reactive power requirements are set to renewable power generators [25], high penetration levels of DG integration leads to the need of evaluating the potential contribution of renewable generation to power system voltage and reactive power regulation [105]. This eventually implies to use the AC power flow equations.

Another improvement in the modeling of a DG–integrated network is the representation of the evolution of the uncertain operational conditions, such as solar irradiation, wind speed, memory effect of batteries, load profiles, energy prices, etc. The forecast of these variables implies the prediction of future conditions given specific previous scenarios. Then, a sequential MCS model should be developed, in which the sampling is performed for each time step dependent on previous states conditions. As an example, load forecast uncertainty can be integrated properly building consecutive load scenarios and assigning corresponding probabilities of occurrence as presented by [16, 34]. Another interesting approach for load forecast uncertainty modeling is the geometric Brownian motion (GBM) stochastic process [42, 76].

The above–mentioned developments entail an increase in the computational efforts involved in the resolution of the DG planning. Hence, the motivation to explore and develop new methodologies to improve the computational performance still remain. In particular, a direct next step would be the extension of the HCDE search engine into MOO strategies by new controllable clustering techniques, measuring the variation in the quality of non–dominated solutions sets, e.g. by a hypervolume indicator [95].

A IEEE 13–bus test feeder data: Application 5.1

Table A.1 contains the technical characteristics of the different types of feeders considered: specifically, the indexes of the pairs of nodes that are connected by each feeder of the network, their length l, reactance X and their ampacity A.

Table A.1 Feeders characteristic and technical data [2]

Type	Node i	Node i'	l (km)	X (Ω/km)	A (A)
T1	1	2	0.61	0.37	365
T2	2	3	0.15	0.47	170
T3	2	4	0.15	0.56	115
T1	2	6	0.61	0.37	365
T3	4	5	0.09	0.56	115
T6	6	7	0.15	0.25	165
T4	6	8	0.09	0.56	115
T1	6	11	0.31	0.37	365
T5	8	9	0.09	0.56	115
T7	8	10	0.24	0.32	115

The nodal power demands are reported as daily profiles, normally distributed on each hour. The mean μ and variance σ values of the nodal daily profiles of the power demands are shown in Figure A.1(A) and (B), respectively.

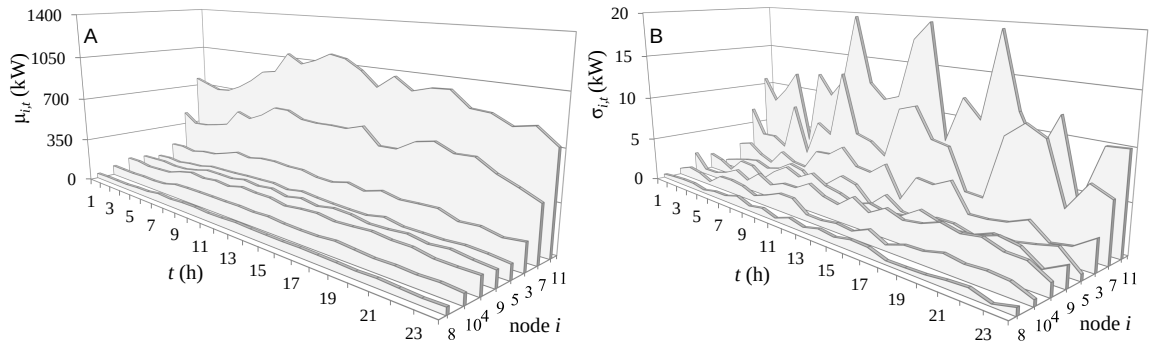

Figure A.1 Mean (A) and variance (B) values of nodal power demand daily profiles

The maximum active power capacity of the transformer (MG) and the parameters of the normal distribution that describe its variability are given in Table A.2.

Table A.2 Bulk–power supply parameters

Node i	P_{max} (kW)	Normal distribution parameters	
		μ (kW)	σ (kW)
1	1600	1200	27.5

The technical parameters of the four different types of DG technologies considered to be integrated into the distribution network (PV, W, EV and ST) are given in Table A.3. The values of the parameters of the Beta and Weibull distributions describing the variability of the solar irradiation and wind speed, are assumed constant in the whole network, i.e., the region of distribution is such that the weather conditions are the same for all nodes.

Table A.3 Parameters of PV, W, EV and ST technologies [3–5]

PV		W	
Beta distribution α	0.26	Weibull distribution α	11.25
Beta distribution β	0.73	Weibull distribution β	2
P_{max} (W)	50	P_R (kW)	50
T_A (°C)	30	U_{CI} (m/s)	3.8
T_{No} (°C)	43	U_A (m/s)	9.5
I_{SC} (A)	1.8	U_{CO} (m/s)	23.8
k_I (mA/°C)	1.4	EV	
V_{OC} (V)	55.5	P_R (kW)	6.3
k_V (mV/°C)	194	ST	
V_{MPP} (V)	38	P_R (kW)	0.275
I_{MPP} (A)	1.32	J_S (kJ/kg)	0.042

Failures and repair rates of the components of the distribution network are provided in Table A.4.

Table A.4 Failure rates of feeders, MG and DG units [3–6]

Type		λ_F (n/h)		λ_R (n/h)	
$MG \cup DG$	FD	$MG \cup DG$	TD	$MG \cup DG$	FD
MG	T1	3.33e−04	3.33e−04	0.021	0.198
PV	T2	4.05e−04	4.05e−04	0.013	0.162
W	T3	3.55e−04	3.55e−04	0.015	0.185
EV	T4	3.55e−04	3.55e−04	0.105	0.185
ST	T5	3.55e−04	3.55e−04	0.073	0.185
-	T6	-	4.00e−04	-	0.164
-	T7	-	3.55e−04	-	0.185

The hourly per day operating states probability profile of the EV is presented in Figure A.2.

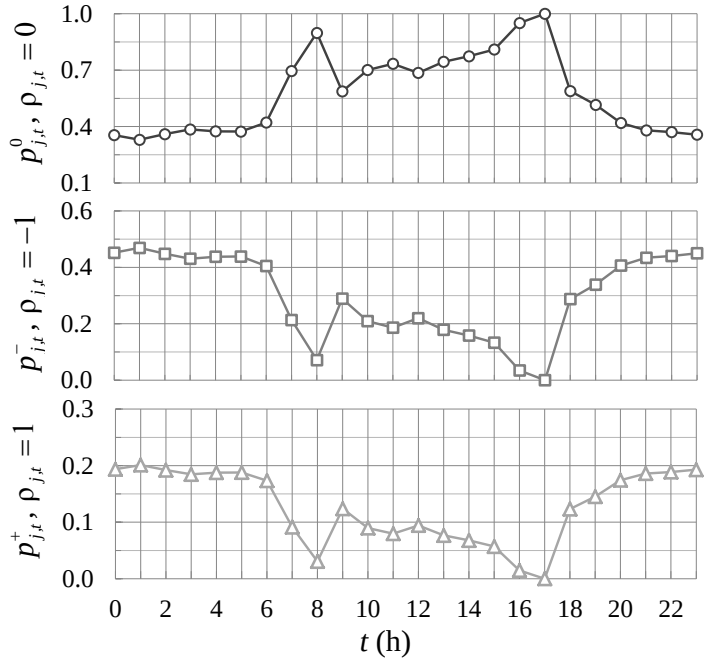

Figure A.2 Hourly per day probability data of EV operating states. $\rho = -1$: charging, $\rho = 0$: disconnected, $\rho = 1$: discharging

The values of the investment (CI) and fixed and variable Operational and Maintenance (COf and COv) costs of the MG and DG units are reported in Table A.5.

Table A.5 Investment, fixed O&M and variable O&M costs of MG and DG [6–8]

Type	$CI + COf$ (\$)	COv (\$/kWh)
MG	-	1.45e−01
PV	48	3.76e−05
W	113750	3.90e−02
EV	17000	2.20e−02
ST	135.15	4.62e−05

The total investment associated to DG-integrated network plan must be less than or equal to the limit budget; which is set to $BGT = 4500000$ (\$), and the total number of units of each type of DG (following the order [PV, W, EV, ST]) must be less than or equal to $\tau_j = [15000, 5, 200, 8000]$. The value of the incentive for renewable kW h supplied is taken as 0.024 (\$/kWh) [8]. The maximum value of the energy price EP_{max} is 0.11 (\$/kWh) [52, 53].

Concerning the calculation of the CVaR, the α–percentile is taken as $\alpha = 0.80$.

B IEEE 13–bus test feeder data: Application 5.2

Table B.1 contains the technical characteristics of the different types of feeders considered: specifically, the indexes of the pairs of nodes (i, i') that they connect, their length l, reactance X, ampacity A and failure and repair rates.

Table B.1 Feeders characteristic and technical data [2, 4, 9]

Type	Node i	Node i'	l (km)	X (Ω/km)	A (A)	λ_F (n/h)	λ_R (n/h)	COv ($/kWh)
T1	1	2	0.610	0.371	730	3.333e−04	0.198	1.970e−02
T2	2	3	0.152	0.472	340	4.050e−04	0.162	9.173e−03
T3	2	4	0.152	0.555	230	3.552e−04	0.185	6.205e−03
T1	2	6	0.610	0.371	730	3.333e−04	0.198	6.205e−03
T3	4	5	0.091	0.555	230	3.552e−04	0.185	6.205e−03
T6	6	7	0.152	0.252	329	4.048e−04	0.164	8.904e−03
T4	6	8	0.091	0.555	230	3.552e−04	0.185	1.970e−02
T1	6	11	0.305	0.371	730	3.333e−04	0.198	1.970e−02
T5	8	9	0.091	0.555	230	3.552e−04	0.185	9.173e−03
T7	8	10	0.244	0.318	175	3.552e−04	0.185	6.205e−03

The nodal power demands are built from the load data given in [2] and reported in Figure B.1 as daily profiles, normally distributed on each hour t with mean μ and standard deviation σ [10, 58].

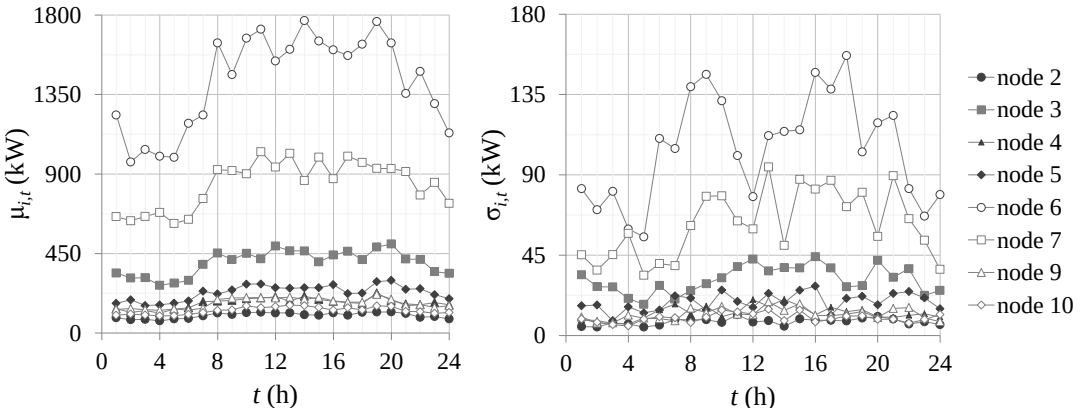

Figure B.1 Mean and standard deviation values of normally distributed nodal power demand daily profile

The technical parameters, failure and repair rates and costs of the MG and the four different types of DG technologies (PV, W, EV and ST) available to be integrated into the distribution network are given in Table B.2. For the corresponding application example, the distribution region is such that the solar irradiation and wind speed conditions are assumed uniform in the whole network, i.e., the values of the parameters of the corresponding Beta and Weibull distributions are assumed constant in the whole network.

Table B.2 Power sources parameters and technical data [3–8, 10]

Type j	Technical parameters	Distributions parameters, failure and repair rates	Costs
MG	$P_{max} = 4250$ (kW)	$\mu = 4000$ (kW) $\sigma = 125$ (kW) $\lambda_F = 4.00\text{e}{-}04$ (1/h) $\lambda_R = 1.30\text{e}{-}02$ (1/h)	$COv = 0.145$ (\$/kWh)
PV	$T_A = 30.00$ (°C) $T_{No} = 43.00$ (°C) $I_{SC} = 1.80$ (A) $V_{OC} = 55.50$ (V) $k_I = 1.40$ (mA/°C) $k_V = 194.00$ (mV/°C) $V_{MPP} = 38.00$ (V) $I_{MPP} = 1.32$ (A)	$\alpha = 0.26$ $\beta = 0.73$ $\lambda_F = 5.00\text{e}{-}04$ (1/h) $\lambda_R = 1.30\text{e}{-}02$ (1/h)	$CI + COf = 48$ (\$) $COv = 3.76\text{e}{-}05$ (\$/kWh)
W	$U_{CI+COf} = 3.80$ (m/s) $U_A = 9.50$ (m/s) $U_{CO} = 23.80$ (m/s) $P_R = 50.00$ (kW)	$\alpha = 11.25$ $\beta = 2$ $\lambda_F = 6.00\text{e}{-}04$ (1/h) $\lambda_R = 1.30\text{e}{-}02$ (1/h)	$CI + COf = 113,750$ (\$) $COv = 3.90\text{e}{-}02$ (\$/kWh)
EV	$P_R = 6.30$ (kW)	$\lambda_F = 2.00\text{e}{-}04$ (1/h) $\lambda_R = 9.70\text{e}{-}02$ (1/h)	$CI + COf = 17000$ (\$) $COv = 2.20\text{e}{-}02$ (\$/kWh)
ST	$P_R = 0.28$ (kW/kg) $J_S = 0.04$ (kJ/kg)	$\lambda_F = 3.00\text{e}{-}04$ (1/h) $\lambda_R = 7.30\text{e}{-}02$ (1/h)	$CI + COf = 135.15$ (\$) $COv = 4.62\text{e}{-}05$ (\$/kWh)

The hourly per day operating state probability profiles of the EV are presented in Figure B.2: p^0, p^- and p^+ correspond to the profiles of disconnected, charging and discharging states, respectively.

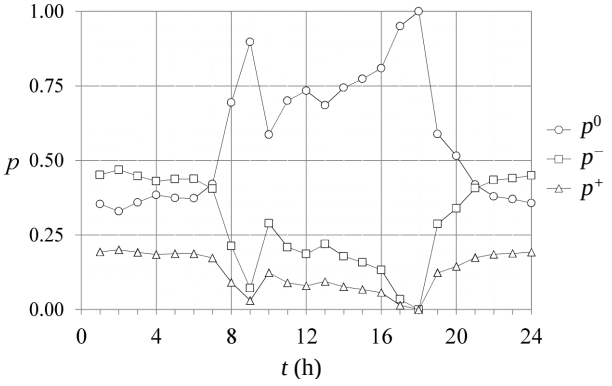

Figure B.2 Hourly per day probability data of EV operating states

The budget is set to $BGT = 4500000$ (\$) and the limit of units of the different DG technologies available to be purchased is $\tau = [20000, 8, 250, 10000]$. The maximum value of the energy price is $EP_{max} = 0.12$ (\$/kWh) [53] and the highest value of total demand ΣL_{max_i} is set to 4800 (kW). The opportunity cost for kW h not supplied C_{LS} is considered as twice of the maximum energy price.

C IEEE 30–bus sub–transmission & distribution system data: Application 5.3

Table C.1 summarizes the characteristics and technical data of the T&D lines, specifically: the indexes of the pair of nodes that they connect (i, i'), the susceptance values $B_{(i,i')}$, power rating $P_{max_{(i,i')}}$, failure $\lambda_{F_{(i,i')}}$ and repair $\lambda_{R_{(i,i')}}$ rates and operating cost $COv_{(i,i')}$.

The nodal power demands are built from the load data given in [11] and reported in Figure C.1 as daily profiles (accumulated according to the node indexes i), normally distributed on each hour t with mean $\mu_{i,t}$ and standard deviation $\sigma_{i,t}$.

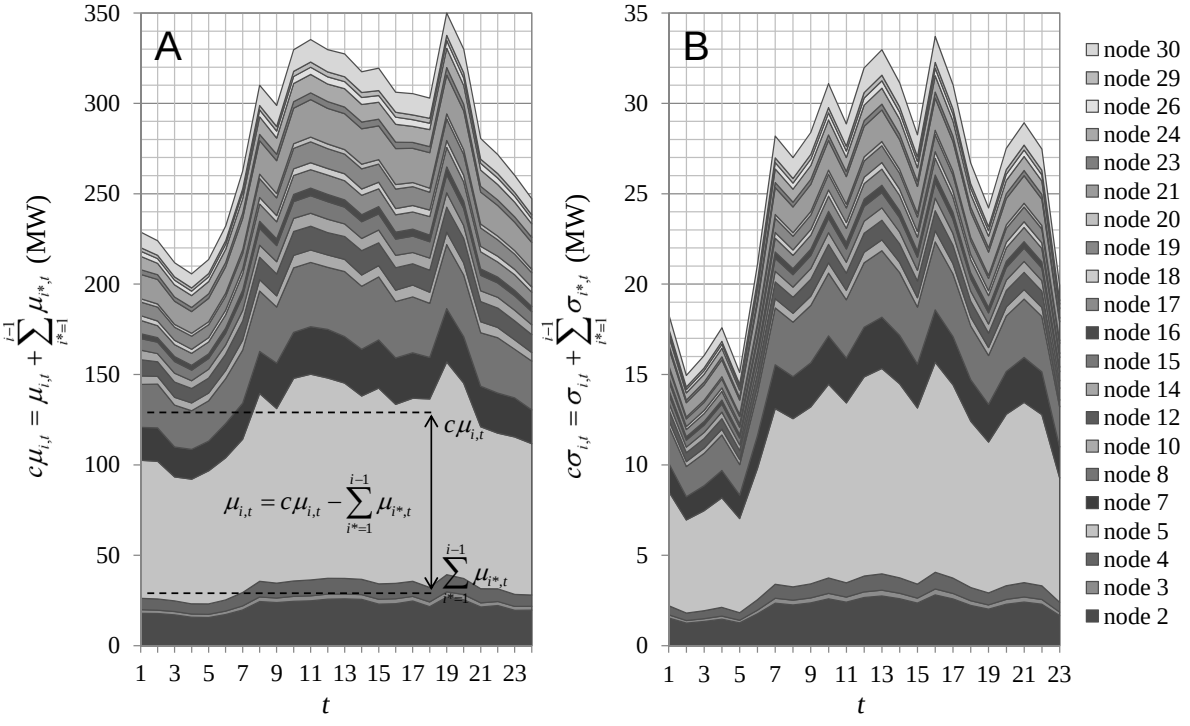

Figure C.1 Accumulated mean (A) and standard deviation (B) values of nodal load daily profiles

Table C.1 T&D lines characteristics and technical data [11–13]

i	i'	$B_{(i,i')}$ (p.u.)	$P_{max_{(i,i')}}$ (p.u.)	$\lambda_{F_{(i,i')}}$ (n/h)	$\lambda_{R_{(i,i')}}$ (n/h)	$COv_{(i,i')}$ ($/MWh)
1	2	17.24	1.30	2.85e−04	6.67e−02	4.95
1	3	5.41	1.30	2.85e−04	6.67e−02	1.61
2	4	5.75	0.65	4.28e−04	1.00e−01	2.61
3	4	26.32	1.30	2.85e−04	6.67e−02	5.97
2	5	5.05	1.30	2.85e−04	6.67e−02	2.53
2	6	5.68	0.65	4.28e−04	1.00e−01	3.68
4	6	24.39	0.90	3.73e−04	8.72e−02	6.06
5	7	8.62	0.70	4.17e−04	9.74e−02	0.70
6	7	12.20	1.30	2.85e−04	6.67e−02	0.93
6	8	23.81	0.32	5.01e−04	1.17e−01	0.19
6	9	4.81	0.65	4.28e−04	1.00e−01	5.21
6	10	1.80	0.32	5.01e−04	1.17e−01	4.14
9	10	9.09	0.65	4.28e−04	1.00e−01	8.78
9	11	4.81	0.65	4.28e−04	1.00e−01	4.81
4	12	3.91	0.65	4.28e−04	1.00e−01	2.93
12	13	7.14	0.65	4.28e−04	1.00e−01	1.41
12	14	3.91	0.32	5.01e−04	1.17e−01	11.10
12	15	7.69	0.32	5.01e−04	1.17e−01	6.66
14	15	5.00	0.16	5.36e−04	1.25e−01	13.70
12	16	5.03	0.32	5.01e−04	1.17e−01	13.70
10	17	11.76	0.32	5.01e−04	1.17e−01	13.20
16	17	5.18	0.16	5.36e−04	1.25e−01	8.71
15	18	4.57	0.16	5.36e−04	1.25e−01	3.46
18	19	7.75	0.16	5.36e−04	1.25e−01	12.20
10	20	4.78	0.32	5.01e−04	1.17e−01	5.40
19	20	14.71	0.32	5.01e−04	1.17e−01	3.52
10	21	13.33	0.32	5.01e−04	1.17e−01	11.50
10	22	6.67	0.32	5.01e−04	1.17e−01	9.76
21	22	41.67	0.32	5.01e−04	1.17e−01	8.08
15	23	4.95	0.16	5.36e−04	1.25e−01	1.68
22	24	5.59	0.16	5.36e−04	1.25e−01	7.82
23	24	3.70	0.16	5.36e−04	1.25e−01	16.40
24	25	3.04	0.16	5.36e−04	1.25e−01	7.98
25	26	2.63	0.16	5.36e−04	1.25e−01	9.13
25	27	4.78	0.16	5.36e−04	1.25e−01	5.56
6	28	16.67	0.32	5.01e−04	1.17e−01	1.15
8	28	5.00	0.32	5.01e−04	1.17e−01	0.36
27	28	2.53	0.65	4.28e−04	1.00e−01	6.17
27	29	2.41	0.16	5.36e−04	1.25e−01	14.40
27	30	1.66	0.16	5.36e−04	1.25e−01	14.00
29	30	2.21	0.16	5.36e−04	1.25e−01	11.40

The technical data and uncertain model parameters of the different types of bulk-power suppliers and DG technologies, available to be integrated into the network, are given in Table C.2. The number of photovoltaic cells per PV generation unit is $n_c = 20000$ and that the region covered by the system is such that the solar irradiation and wind speed conditions are uniform in the whole region, i.e., the values of the parameters of the corresponding Beta and Weibull distributions are taken equal for all nodes. The renewable power penetration factor is set to $PF^{DG} = 0.3$.

Table C.2 Power generators technical data and uncertain model parameters [3–5, 7, 10, 14]

Type		Technical parameters					Distribution parameters	
MG	G_1	P_{max_1} (MW)					Normal μ_1	Normal σ_1
		340					300	18.25
	G_2	P_{max_2} (MW)					Normal μ_2	Normal σ_2
		50					42.5	5
RG	PV	P_{AV} (MW)	P_{max} (W)	V_{OC}(V)	I_{SC} (A)	V_{MPP} (V)	Beta α	Beta β
		1.07	75	21.98	5.32	17.32	0.50	0.33
		I_{MPP} (A)	k_V (mV/°C)	k_I (mA/°C)	T_{No} (°C)	T_A (°C)		
		4.76	14.4	1.22	43	30		
	W	P_{AV} (MW)	P_R (MW)	U_{CI} (m/s)	U_A (m/s)	U_{CO} (m/s)	Weibull α	Weibull β
		0.93	1.5	5	15	25	15	2.2

Table C.3 reports the failure and repair rates, λ_F and λ_R, respectively, the investment and operating costs $CI + COf$ and the variable operating cost of the different types of power generators.

Table C.3 Power generators failure and repair rates and costs [3, 4, 7, 10, 14, 15]

Type		λ_F (n/h)	λ_R (n/h)	$CI + COf$ (M$/u)	COv ($/MWh)
MG	G_1	5.13e−04	2.77e−02	–	29.32
	G_2	6.84e−04	4.16e−02	–	8.92
RG	PV	6.27e−04	1.30e−02	2.20	9.69
	W	3.42e−04	9.00e−03	1.85	11.05

The maximum value of the energy price is $EP_{max} = 100$ ($/MWh) [4, 52, 53] and the corresponding highest value of total demand $\Sigma Lmax_i$ (MW) is set to 445 (MW). The load shedding cost C_{LS} ($/MWh) is considered as the maximum energy price. The horizon of analysis or lifetime of the project is 30 years, in which the investment and operating costs are hourly prorated. The confidence level or α-percentile considered to estimate the values $CVaR_\alpha$ and $DCVaR_\alpha$ is 75%, arbitrarily chosen.

References

[1] International Energy Agency, "Co_2 emissions from fuel combustion," 2014. Available at https://www.iea.org/publications.

[2] IEEE power and energy society. Distribution test feeders, Available at http:// ewh.ieee.org/ soc/pes/dsacom/testfeeders/index.html.

[3] M. Raoofat, "Simultaneous allocation of DGs and remote controllable switches in distribution networks considering multilevel load model," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 8, pp. 1429 – 1436, 2011.

[4] H. Falaghi, C. Singh, M.-R. Haghifam, and M. Ramezani, "DG integrated multistage distribution system expansion planning," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 8, pp. 1489 – 1497, 2011.

[5] Y.-F. Li and E. Zio, "A multi-state model for the reliability assessment of a distributed generation system via universal generating function," *Reliability Engineering & System Safety*, vol. 106, pp. 28 – 36, 2012.

[6] R. Webster, "Can the electricity distribution network cope with an influx of electric vehicles?," *Journal of Power Sources*, vol. 80, no. 1-2, pp. 217 – 225, 1999.

[7] K. Zou, A. Agalgaonkar, K. Muttaqi, and S. Perera, "Multi-objective optimisation for distribution system planning with renewable energy resources," in *Energy Conference and Exhibition (EnergyCon), 2010 IEEE International*, pp. 670–675, Dec 2010.

[8] F. Pilo, G. Celli, S. Mocci, and G. Soma, "Active distribution network evolution in different regulatory environments," in *Power Generation, Transmission, Distribution and Energy Conversion (MedPower 2010), 7th Mediterranean Conference and Exhibition on*, pp. 1–8, Nov 2010.

[9] S. Ganguly, N. Sahoo, and D. Das, "A novel multi-objective PSO for electrical distribution system planning incorporating distributed generation," *Energy Systems*, vol. 1, no. 3, pp. 291–337, 2010.

[10] Y. Atwa, E. El-Saadany, M. Salama, and R. Seethapathy, "Optimal renewable resources mix for distribution system energy loss minimization," *Power Systems, IEEE Transactions on*, vol. 25, pp. 360–370, Feb 2010.

[11] University of Washington, College of Engineering, Department of Electrical Engineering, Available at http://www.ee.washington.edu/research/pstca.

[12] P. Buijs, D. Bekaert, D. Van Hertem, and R. Belmans, "Needed investments in the power system to bring wind energy to shore in Belgium," in *PowerTech, 2009 IEEE Bucharest*, pp. 1–6, June 2009.

[13] O. Alsac and B. Stott, "Optimal load flow with steady-state security," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-93, pp. 745–751, May 1974.

[14] U.S. Energy Information Administration, Available at http://www.eia.gov/.

[15] Z. Liu, F. Wen, and G. Ledwich, "Optimal siting and sizing of distributed generators in distribution systems considering uncertainties," *Power Delivery, IEEE Transactions on*, vol. 26, pp. 2541–2551, Oct 2011.

[16] V. Martins and C. Borges, "Active distribution network integrated planning incorporating distributed generation and load response uncertainties," *Power Systems, IEEE Transactions on*, vol. 26, pp. 2164–2172, Nov 2011.

[17] C. Karger and W. Hennings, "Sustainability evaluation of decentralized electricity generation," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 3, pp. 583 – 593, 2009.

[18] S. Moosavian, N. Rahim, J. Selvaraj, and K. Solangi, "Energy policy to promote photovoltaic generation," *Renewable and Sustainable Energy Reviews*, vol. 25, no. 0, pp. 44 – 58, 2013.

[19] R. Viral and D. Khatod, "Optimal planning of distributed generation systems in distribution system: A review," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 7, pp. 5146–5165, 2012.

[20] A. Alarcon-Rodriguez, G. Ault, and S. Galloway, "Multi-objective planning of distributed energy resources: A review of the state-of-the-art," *Renewable and Sustainable Energy Reviews*, vol. 14, no. 5, pp. 1353 – 1366, 2010.

[21] G. Koutroumpezis and A. Safigianni, "Optimum allocation of the maximum possible distributed generation penetration in a distribution network," *Electric Power Systems Research*, vol. 80, no. 12, pp. 1421 – 1427, 2010.

[22] C. L. T. Borges and V. F. Martins, "Multistage expansion planning for active distribution networks under demand and distributed generation uncertainties," *International Journal of Electrical Power & Energy Systems*, vol. 36, no. 1, pp. 107 – 116, 2012.

[23] F. Ugranli and E. Karatepe, "Multiple-distributed generation planning under load uncertainty and different penetration levels," *International Journal of Electrical Power & Energy Systems*, vol. 46, pp. 132 – 144, 2013.

[24] J. Shin, J. Woo, S.-Y. Huh, J. Lee, and G. Jeong, "Analyzing public preferences and increasing acceptability for the renewable portfolio standard in Korea," *Energy Economics*, vol. 42, pp. 17 – 26, 2014.

[25] G. Allan, I. Eromenko, M. Gilmartin, I. Kockar, and P. McGregor, "The economics of distributed energy generation: A literature review," *Renewable and Sustainable Energy Reviews*, vol. 42, pp. 543 – 556, 2015.

[26] S. Javadian, M.-R. Haghifam, M. F. Firoozabad, and S. Bathaee, "Analysis of protection system's risk in distribution networks with DG," *International Journal of Electrical Power & Energy Systems*, vol. 44, no. 1, pp. 688 – 695, 2013.

[27] R. Poudineh and T. Jamasb, "Distributed generation, storage, demand response and energy efficiency as alternatives to grid capacity enhancement," *Energy Policy*, vol. 67, no. 0, pp. 222 – 231, 2014.

[28] International Energy Agency, "Key world energy statistics," 2014. Available at https://www.iea.org/publications.

[29] International Energy Agency, "Tracking clean energy progress," 2015. Available at https://www.iea.org/publications.

[30] A. M. Imran, M. Kowsalya, and D. Kothari, "A novel integration technique for optimal network reconfiguration and distributed generation placement in power distribution networks," *International Journal of Electrical Power & Energy Systems*, vol. 63, pp. 461 – 472, 2014.

[31] R. Sebastian, "Modelling and simulation of a high penetration wind diesel system with battery energy storage," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 3, pp. 767 – 774, 2011.

[32] K. Clement-Nyns, E. Haesen, and J. Driesen, "The impact of vehicle-to-grid on the distribution grid," *Electric Power Systems Research*, vol. 81, no. 1, pp. 185 – 192, 2011.

[33] S. Mirkhani and Y. Saboohi, "Stochastic modeling of the energy supply system with uncertain fuel price –A case of emerging technologies for distributed power generation," *Applied Energy*, vol. 93, pp. 668 – 674, 2012. (1) Green Energy; (2)Special Section from papers presented at the 2nd International Enery 2030 Conf.

[34] M. Shaaban, Y. Atwa, and E. El-Saadany, "DG allocation for benefit maximization in distribution networks," *Power Systems, IEEE Transactions on*, vol. 28, pp. 639–649, May 2013.

[35] J. J. Justo, F. Mwasilu, J. Lee, and J.-W. Jung, "AC-microgrids versus DC-microgrids with distributed energy resources: A review," *Renewable and Sustainable Energy Reviews*, vol. 24, pp. 387 – 405, 2013.

[36] A. S. Fini, M. P. Moghaddam, and M. Sheikh-El-Eslami, "A dynamic model for distributed energy resource expansion planning considering multi-resource support schemes," *International Journal of Electrical Power & Energy Systems*, vol. 60, pp. 357 – 366, 2014.

[37] C. L. T. Borges, "An overview of reliability models and methods for distribution systems with renewable energy distributed generation," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 4008 – 4015, 2012.

[38] G. P. Harrison, A. Piccolo, P. Siano, and A. R. Wallace, "Hybrid GA and OPF evaluation of network capacity for distributed generation connections," *Electric Power Systems Research*, vol. 78, no. 3, pp. 392 – 398, 2008.

[39] S.-H. Lee and J.-W. Park, "Selection of optimal location and size of multiple distributed generations by using Kalman filter algorithm," *Power Systems, IEEE Transactions on*, vol. 24, pp. 1393–1400, Aug 2009.

[40] J. Hernandez, D. Velasco, and C. Trujillo, "Analysis of the effect of the implementation of photovoltaic systems like option of distributed generation in Colombia," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 5, pp. 2290 – 2298, 2011.

[41] T. Kaya and C. Kahraman, "Multicriteria renewable energy planning using an integrated fuzzy VIKOR & AHP methodology: The case of Istanbul," *Energy*, vol. 35, no. 6, pp. 2517 – 2527, 2010. 7th International Conference on Sustainable Energy Technologies 7th International Conference on Sustainable Energy Technologies.

[42] M. Samper and A. Vargas, "Investment decisions in distribution networks under uncertainty with distributed generation - Part I: Model formulation," *Power Systems, IEEE Transactions on*, vol. 28, pp. 2331–2340, Aug 2013.

[43] E. Zio, M. Delfanti, L. Giorgi, V. Olivieri, and G. Sansavini, "Monte carlo simulation-based probabilistic assessment of dg penetration in medium voltage distribution networks," *International Journal of Electrical Power & Energy Systems*, vol. 64, pp. 852 – 860, 2015.

[44] L. Ochoa and G. Harrison, "Minimizing energy losses: Optimal accommodation and smart operation of renewable distributed generation," *Power Systems, IEEE Transactions on*, vol. 26, pp. 198–205, Feb 2011.

[45] M. Samper and A. Vargas, "Investment decisions in distribution networks under uncertainty with distributed generation - Part II: Implementation and results," *Power Systems, IEEE Transactions on*, vol. 28, pp. 2341–2351, Aug 2013.

[46] J.-H. Wu and Y.-H. Huang, "Electricity portfolio planning model incorporating renewable energy characteristics," *Applied Energy*, vol. 119, pp. 278 – 287, 2014.

[47] S.-Y. Lin and A.-C. Lin, "RLOPF (risk-limiting optimal power flow) for systems with high penetration of wind power," *Energy*, vol. 71, pp. 49 – 61, 2014.

[48] G. Celli, E. Ghiani, S. Mocci, and F. Pilo, "A multiobjective evolutionary algorithm for the sizing and siting of distributed generation," *Power Systems, IEEE Transactions on*, vol. 20, pp. 750–757, May 2005.

[49] G. Celli, S. Mocci, F. Pilo, and G. Soma, "A multi-objective approach for the optimal distributed generation allocation with environmental constraints," in *Probabilistic Methods Applied to Power Systems, 2008. PMAPS '08. Proceedings of the 10th International Conference on*, pp. 1–8, May 2008.

[50] G. Celli, F. Pilo, G. Soma, M. Gallanti, and R. Cicoria, "Active distribution network cost/benefit analysis with multi-objective programming," in *Electricity Distribution - Part 1, 2009. CIRED 2009. 20th International Conference and Exhibition on*, pp. 1–5, June 2009.

[51] H. Hejazi, M. Hejazi, G. Gharehpetian, and M. Abedi, "Distributed generation site and size allocation through a techno economical multi-objective differential evolution algorithm," in *Power and Energy (PECon), 2010 IEEE International Conference on*, pp. 874–879, Nov 2010.

[52] H. Ren and W. Gao, "A MILP model for integrated plan and evaluation of distributed energy systems," *Applied Energy*, vol. 87, no. 3, pp. 1001 – 1014, 2010.

[53] H. Ren, W. Zhou, K. Nakagami, W. Gao, and Q. Wu, "Multi-objective optimization for the operation of distributed energy systems considering economic and environmental aspects," *Applied Energy*, vol. 87, no. 12, pp. 3642 – 3651, 2010.

[54] W. El-Khattam, K. Bhattacharya, Y. Hegazy, and M. Salama, "Optimal investment planning for distributed generation in a competitive electricity market," *Power Systems, IEEE Transactions on*, vol. 19, pp. 1674–1684, Aug 2004.

[55] W. El-Khattam, Y. Hegazy, and M. Salama, "An integrated distributed generation optimization model for distribution system planning," *Power Systems, IEEE Transactions on*, vol. 20, pp. 1158–1165, May 2005.

[56] M. Gómez-González, A. López, and F. Jurado, "Optimization of distributed generation systems using a new discrete PSO and OPF," *Electric Power Systems Research*, vol. 84, no. 1, pp. 174 – 180, 2012.

[57] W. Ouyang, H. Cheng, X. Zhang, and L. Yao, "Distribution network planning method considering distributed generation for peak cutting," *Energy Conversion and Management*, vol. 51, no. 12, pp. 2394 – 2401, 2010.

[58] L. Wang and C. Singh, "Multicriteria design of hybrid power generation systems based on a modified particle swarm optimization algorithm," *Energy Conversion, IEEE Transactions on*, vol. 24, pp. 163–172, March 2009.

[59] A. Bhattacharya and S. Kojima, "Power sector investment risk and renewable energy: A japanese case study using portfolio risk optimization method," *Energy Policy*, vol. 40, pp. 69 – 80, 2012. Strategic Choices for Renewable Energy Investment.

[60] C. González-Pedraz, M. Moreno, and J. I. Pena, "Tail risk in energy portfolios," *Energy Economics*, vol. 46, pp. 422 – 434, 2014.

[61] D. Ramos, L. Camargo, E. Guarnier, and L. Witzler, "Minimizing market risk by trading hydro-wind portfolio: A complementarity approach," in *European Energy Market (EEM), 2013 10th International Conference on the*, pp. 1–8, May 2013.

[62] P. Vithayasrichareon, J. Riesz, and I. F. MacGill, "Using renewables to hedge against future electricity industry uncertainties - An Australian case study," *Energy Policy*, vol. 76, pp. 43 – 56, 2015.

[63] R. Fagiani, J. Barquín, and R. Hakvoort, "Risk-based assessment of the cost-efficiency and the effectivity of renewable energy support schemes: Certificate markets versus feed-in tariffs," *Energy Policy*, vol. 55, pp. 648 – 661, 2013. Special section: Long Run Transitions to Sustainable Economic Structures in the European Union and Beyond.

[64] S. Ahmed, M. Elsholkami, A. Elkamel, J. Du, E. B. Ydstie, and P. L. Douglas, "Financial risk management for new technology integration in energy planning under uncertainty," *Applied Energy*, vol. 128, pp. 75 – 81, 2014.

[65] L. Kitzing, "Risk implications of renewable support instruments: Comparative analysis of feed-in tariffs and premiums using a mean-variance approach," *Energy*, vol. 64, pp. 495 – 505, 2014.

[66] R. Rockafellar, S. Uryasev, and M. Zabarankin, "Generalized deviations in risk analysis," *Finance and Stochastics*, vol. 10, no. 1, pp. 51–74, 2006.

[67] M. Nazir and F. Bouffard, "Risk-sensitive investment in renewable distributed generation under uncertainty due to post-feed-in tariff policy," in *Developments in Renewable Energy Technology (ICDRET), 2012 2nd International Conference on the*, pp. 1–5, Jan 2012.

[68] J. Francis and D. Kim, *Modern Portfolio Theory: Foundations, Analysis, and New Developments*. Wiley Finance, Wiley, 2013.

[69] Y. Zhang and G. Giannakis, "Robust optimal power flow with wind integration using conditional value-at-risk," in *Smart Grid Communications (SmartGridComm), 2013 IEEE International Conference on*, pp. 654–659, Oct 2013.

[70] S. Martin, Y. Smeers, and J. Aguado, "A stochastic two settlement equilibrium model for electricity markets with wind generation," *Power Systems, IEEE Transactions on*, vol. 30, pp. 233–245, Jan 2015.

[71] S. Ghosh, J. Kalagnanam, D. Katz, M. Squillante, and X. Zhang, "Integration of demand response and renewable resources for power generation management," in *Innovative Smart Grid Technologies (ISGT), 2011 IEEE PES*, pp. –, Jan 2011.

[72] I. G. Moghaddam, M. Nick, F. Fallahi, M. Sanei, and S. Mortazavi, "Risk-averse profit-based optimal operation strategy of a combined wind farm-cascade hydro system in an electricity market," *Renewable Energy*, vol. 55, pp. 252 – 259, 2013.

[73] M. A. Tajeddini, A. Rahimi-Kian, and A. Soroudi, "Risk averse optimal operation of a virtual power plant using two stage stochastic programming," *Energy*, vol. 73, pp. 958 – 967, 2014.

[74] W.-S. Tan, M. Y. Hassan, M. S. Majid, and H. A. Rahman, "Optimal distributed renewable generation planning: A review of different approaches," *Renewable and Sustainable Energy Reviews*, vol. 18, pp. 626 – 645, 2013.

[75] E. Zio, *The Monte Carlo Simulation Method for System Reliability and Risk Analysis*. Springer Series in Reliability Engineering, Springer London, 2013.

[76] J. H. Zhao, J. Foster, Z.-Y. Dong, and K. P. Wong, "Flexible transmission network planning considering distributed generation impacts," *Power Systems, IEEE Transactions on*, vol. 26, pp. 1434–1443, Aug 2011.

[77] D. Kroese, T. Taimre, and Z. Botev, *Handbook of Monte Carlo Methods*. Wiley Series in Probability and Statistics, Wiley, 2013.

[78] A. Thornton and C. R. Monroy, "Distributed power generation in the United States," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 9, pp. 4809 – 4817, 2011.

[79] Y. Li and E. Zio, "Uncertainty analysis of the adequacy assessment model of a distributed generation system," *Renewable Energy*, vol. 41, pp. 235 – 244, 2012.

[80] Y.-C. Chang, T.-Y. Lee, C.-L. Chen, and R.-M. Jan, "Optimal power flow of a wind-thermal generation system," *International Journal of Electrical Power & Energy Systems*, vol. 55, pp. 312 – 320, 2014.

[81] F. Díaz-González, A. Sumper, O. Gomis-Bellmunt, and R. Villafáfila-Robles, "A review of energy storage technologies for wind power applications," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 4, pp. 2154 – 2171, 2012.

[82] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans, "Usefulness of DC power flow for active power flow analysis," in *Power Engineering Society General Meeting, 2005. IEEE*, pp. 454–459 Vol. 1, June 2005.

[83] D. Van Hertem, J. Verboomen, K. Purchala, R. Belmans, and W. Kling, "Usefulness of DC power flow for active power flow analysis with flow controlling devices," in *AC and DC Power Transmission, 2006. ACDC 2006. The 8th IEE International Conference on*, pp. 58–62, March 2006.

[84] R. Billinton and R. Allan, *Reliability evaluation of power systems*. Pitman Advanced Publishing Program, 1984.

[85] S. Haffner, L. Pereira, L. Pereira, and L. Barreto, "Multistage model for distribution expansion planning with distributed generation Part I: Problem formulation," *Power Delivery, IEEE Transactions on*, vol. 23, pp. 915–923, April 2008.

[86] S. Haffner, L. Pereira, L. Pereira, and L. Barreto, "Multistage model for distribution expansion planning with distributed generation Part II: Numerical results," *Power Delivery, IEEE Transactions on*, vol. 23, pp. 924–929, April 2008.

[87] L. Wang and C. Singh, "Multicriteria design of hybrid power generation systems based on a modified particle swarm optimization algorithm," *Energy Conversion, IEEE Transactions on*, vol. 24, pp. 163–172, March 2009.

[88] R. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1443 – 1471, 2002.

[89] J. Aghaei, K. M. Muttaqi, A. Azizivahed, and M. Gitizadeh, "Distribution expansion planning considering reliability and security of energy using modified PSO algorithm," *Energy*, vol. 65, pp. 398 – 411, 2014.

[90] H. Hejazi, A. Araghi, B. Vahidi, S. Hosseinian, M. Abedi, and H. Mohsenian-Rad, "Independent distributed generation planning to profit both utility and DG investors," *Power Systems, IEEE Transactions on*, vol. 28, pp. 1170–1178, May 2013.

[91] L. Arya, A. Koshti, and S. Choube, "Distributed generation planning using differential evolution accounting voltage stability consideration," *International Journal of Electrical Power & Energy Systems*, vol. 42, no. 1, pp. 196 – 207, 2012.

[92] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *Evolutionary Computation, IEEE Transactions on*, vol. 6, pp. 182–197, Apr 2002.

[93] R. Storn and K. Price, "Differential evolution - A simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.

[94] J. Branke, K. Deb, K. Miettinen, and R. Slowinski, *Multiobjective Optimization: Interactive and Evolutionary Approaches*. LNCS sublibrary: Theoretical computer science and general issues, Springer, 2008.

[95] R. Ak, Y. Li, V. Vitelli, E. Zio, E. L. Droguett, and C. M. C. Jacinto, "NSGA-II-trained neural network approach to the estimation of prediction intervals of scale deposition rate in oil & gas equipment," *Expert Systems with Applications*, vol. 40, no. 4, pp. 1205 – 1212, 2013.

[96] M.-Y. Cheng, D.-H. Tran, and Y.-W. Wu, "Using a fuzzy clustering chaotic-based differential evolution with serial method to solve resource-constrained project scheduling problems," *Automation in Construction*, vol. 37, pp. 88 – 97, 2014.

[97] R. Mukherjee, G. R. Patra, R. Kundu, and S. Das, "Cluster-based differential evolution with crowding archive for niching in dynamic environments," *Information Sciences*, vol. 267, pp. 58 – 82, 2014.

[98] Z. Cai, W. Gong, C. X. Ling, and H. Zhang, "A clustering-based differential evolution for global optimization," *Applied Soft Computing*, vol. 11, no. 1, pp. 1363 – 1379, 2011.

[99] G. Liu, Y. Li, X. Nie, and H. Zheng, "A novel clustering-based differential evolution with 2 multi-parent crossovers for global optimization," *Applied Soft Computing*, vol. 12, no. 2, pp. 663 – 681, 2012.

[100] S. Song and X. Yu, "Multi-peak function optimization using a hierarchical clustering based genetic algorithm," in *Intelligent Systems Design and Applications, 2006. ISDA '06. Sixth International Conference on*, vol. 1, pp. 425–428, Oct 2006.

[101] Y.-J. Wang, J.-S. Zhang, and G.-Y. Zhang, "A dynamic clustering based differential evolution algorithm for global optimization," *European Journal of Operational Research*, vol. 183, no. 1, pp. 56 – 73, 2007.

[102] B. S. Everitt, S. Landau, M. Leese, and D. Stahl, *Cluster Analysis*. John Wiley & Sons, Ltd, 2011.

[103] W. Kersting, "Radial distribution test feeders," *Power Systems, IEEE Transactions on*, vol. 6, pp. 975–985, Aug 1991.

[104] B. Korte and J. Vygen, *Combinatorial Optimization: Theory and Algorithms*. Algorithms and Combinatorics, Springer, 2007.

[105] A. Ellis, R. Nelson, E. Von Engeln, R. Walling, J. MacDowell, L. Casey, E. Seymour, W. Peter, C. Barker, B. Kirby, and J. Williams, "Reactive power performance requirements for wind and solar plants," in *Power and Energy Society General Meeting, 2012 IEEE*, pp. 1–8, July 2012.

II Papers

Paper (i)

R. Mena, M. Hennebel, Y.–F. Li, C. Ruiz, and E. Zio, "A risk-based simulation and multi–objective optimization framework for the integration of distributed renewable generation and storage," Renewable and Sustainable Energy Reviews, vol. 37, pp. 778-–793, 2014.

A Risk-Based Simulation and Multi-objective Optimization Framework for the Integration of Distributed Renewable Generation and Storage

Rodrigo Mena[a], Martin Hennebel[b], Yan-Fu Li[a], Carlos Ruiz[c], Enrico Zio[a,d,1]

[a]*Chair on Systems Science and the Energetic Challenge, Fondation Electricité de France at CentraleSupelec, Châtenay-Malabry Cedex, France*

[b]*CentraleSupelec, Department of Power & Energy Systems, Gif-Sur-Yvette, France*

[c]*Universidad Carlos III de Madrid, Department of Statistics, Madrid, Spain*

[d]*Politecnico di Milano, Energy Department, Milan, Italy*

Abstract

We present a simulation and multi–objective optimization framework for the integration of renewable generators and storage devices into an electrical distribution network. The framework searches for the optimal size and location of the distributed renewable generation units (DG). Uncertainties in renewable resources availability, components failure and repair events, loads and grid power supply are incorporated. A Monte Carlo simulation and optimal power flow (MCS–OPF) computational model is used to generate scenarios of the uncertain variables and evaluate the network electric performance. As a response to the need of monitoring and controlling the risk associated to the performance of the optimal DG–integrated network, we introduce the conditional value–at–risk (CVaR) measure into the framework. Multi–objective optimization (MOO) is done with respect to the minimization of the expectations of the global cost (C_g) and energy not supplied (ENS) combined with their respective $CVaR$ values. The multi–objective optimization is performed by the fast non–dominated sorting genetic algorithm NSGA–II. For exemplification, the framework is applied to a distribution network derived from the IEEE 13 nodes test feeder. The results show that the MOO MCS–OPF framework is effective in finding an optimal DG–integrated network considering multiple sources of uncertainties. In addition, from the perspective of decision making, introducing the CVaR as a measure of risk enables the evaluation of trade–offs between optimal expected performances and risks.

Keywords: Distribution network, renewable distributed generation, renewable energy, uncertainty, conditional value–at–risk, simulation, multi–objective optimization, genetic algorithm

[1] Corresponding author. Tel: +33 1 4113 1606; fax: +33 1 4113 1272.
E-mail addresses: enrico.zio@ecp.fr, enrico.zio@polimi.it (E. Zio).

1 Introduction

Over the last decade, the global energetic situation has been receiving a progressively greater attention. The adverse environmental effects of fossil fuels, the volatility of the energy market, the growing energy demand and the intensive reliance on centralized bulk–power generation have triggered a re/evolution towards cleaner, safer, diversified energy sources for reliable and sustainable electric power systems [1–6]. The challenges involved have stimulated both technological development of new equipment and devices, and efficiency improvements in design, planning, operation strategies and management across generation, transmission and distribution.

In this paper, we focus on distribution networks and the conceptual and operational transition they are facing. Indeed, the traditional passive operation with unidirectional flow supplied by a centralized generation/transmission system, is evolving towards an active operational setting with integration of distributed generation (DG) and possibly bidirectional power flows [7, 8].

DG is defined as 'an electric power source connected directly to the distribution network or on the customer site of the meter' [8–10] and in principle offers important technical and economical benefits. Under the assumption that the distribution network operators have control over the dispatching of the DG power, improvement of the reliability of power supply and reduction of the power losses and voltages drops can be achieved. Indeed, DG allocation on areas close to the customers allows the power flowing through shorter paths, and therefore, decreasing the amount of unsatisfied power demand and enhancing the power and voltage profiles. Thus, the eventual intermittence of the centralized power supply can be smoothed [11]. In addition, the modular structure of the DG technologies implies lower financial risks [12, 13] and thus the investments on the power system can be deferred [1, 3].

Most of the actual DG technologies make use of local renewable energy resources, such as wind power, solar irradiation, hydro–power, etc., which makes them even more attractive in view of the requested environmental sustainability (e.g. the Kyoto Protocol [7, 14, 15]). Given the intermittent character of these energy sources, their implementation needs to be accompanied by efficient energy storage technologies. Attentive DG planning is needed to seize the potential advantages associated to DG integration, taking into account specific technical, operational and economic constraints, sources and loads forecasts and regulations. If the practice of selection, sizing and allocation of the different available technologies is not performed attentively, the installation of multiple renewable DG units could produce serious operational complications, in fact, counteracting the potential benefits. Degradation of control and protection devices, reduction of power quality and reliability on the supply, increment in the voltage instability and all related negative impacts on the costs, could become impediments for integration of DG [1–3, 8, 10, 14, 16–20].

Viewing DG planning as a fundamental baseline of advancement, many efforts have been made to solve the associated problem of DG allocation and sizing. Objective functions considered for

the optimization are of economic, operational and technical type. Among the first type, cost–based objective functions have been used considering the costs of energy and fuel for generation, investments, operation and maintenance, energy purchase from the transmission system, energy losses, emissions, taxes, incentives, incomes, etc. [1–3, 7, 8, 11, 13, 14, 16–27]. The second type of operational objective functions mainly revolves around indexes such as the contingency load loss index (CLLI) [23], expected value of non–distributed energy cost (ECOST), system average interruption duration index (SAIDI), system average interruption frequency index (SAIFI) [7, 16, 28], expected energy not supplied (EENS) [28, 29], among others. Regarding the third type of objective functions, technical performance indicators include energy losses [1, 30] and total voltage deviation (TVD) [18].

Power Flow (PF) equations are typically solved within the optimization problem to evaluate the objective functions, while respecting constraints and incorporating non–convex and non–linear conditions. Given the complexity of the optimization problem, heuristic optimization techniques belonging to the class of Evolutionary Algorithms (EAs) have been proposed as a most effective way of solution [10], including particle swarm optimization (PSO) [23, 24, 27, 31, 32], differential evolution (DEA) [18] and genetic algorithms (GA) [3, 7, 11, 13, 14, 16, 26, 33, 34].

An additional difficulty associated to the problem is the proper modeling of the uncertainties inherent to the behavior of primary renewable energy sources and the unexpected operating events (failures or stoppages) that can affect the generation units. These uncertainties come on top of those already present in the network, such as intermittence and fluctuation in the main power supply due to unavailability of the transmission system, overloads and interruptions of the power flow in the feeders, failures in the control and protection devices, variability in the power loads and energy prices, etc. These uncertainties are incorporated into the modeling by generating a random set of scenarios by Monte Carlo simulation (MCS); the optimization is, then, executed to obtain the optimal expected or cumulative value(s) of the objective function(s) under the set of scenarios considered [2, 3, 7, 16, 28, 32, 34, 35].

In the search for the optimal DG–integrated network, the use of only mean or cumulative values as objective function(s) of the optimization hinders the possibility of controlling the risk of the optimal solution(s): the optimal DG–integrated network may on average satisfy the performance objectives but be exposed to high–risk scenarios with non–negligible probabilities [1, 7, 16, 24, 28, 36].

The original contributions of this work reside in: addressing the optimal renewable DG technology selection, sizing and allocation problem within a simulation and multi–objective optimization (MOO) framework that allows for assessing and controlling risk; introducing the conditional value–at–risk (CVaR) as a measure of the risk associated to each objective function of the optimization [37, 38]. The main sources of uncertainty are taken into account through the implementation of a MCS and OPF (MCS–OPF) resolution engine nested in a MOO based on NSGA–II [39]. The aim of the MOO is, specifically, the simultaneous minimization of the expected global cost (EC_g) and expected

energy not supplied ($EENS$), and corresponding $CVaR$ values. A weighting factor β is introduced to leverage the impact of the CVaR in the search of the final Pareto optimal renewable DG integration solutions. The proposed framework provides a new spectrum of information for well–supported decision making enabling the trade–off between optimal expected performance and the associated risk to achieve it.

2 Distributed generation network simulation model

This section introduces the MCS–OPF model, including the definition of the DG structure and configuration, the presentation of the uncertainty sources and their treatment, the MCS for scenarios generation and the OPF formulation for evaluating the performance of the distribution network, in terms of the objective functions of the MOO problem. The outputs of the MCS–OPF model are the probability density functions of the energy not supplied (ENS) and the global cost (C_g) of the network, and their respective $CVaR$ values.

2.1 Distributed generation network structure and configuration

Four main classes of components are considered in the distribution network: nodes, feeders, renewable DG units and main power supply spots (MS). The nodes can be understood as fixed spatial locations at which generation units and loads can be allocated. Feeders connect different nodes and through them the power is distributed. Renewable DG units and main power supply spots are power sources; in the case of electric vehicles and storage devices they can also act as loads when they are in charging state. The locations of the main supply spots are fixed. The MOO aims at optimally allocating renewable DG units at the different nodes.

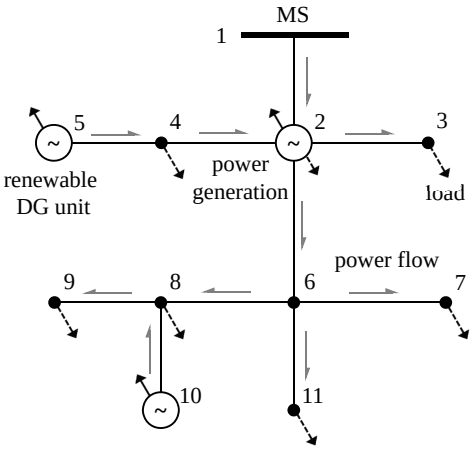

Figure 1: Example of distribution network configuration

Figure 1 shows an example of configuration of a distribution network adapted from the IEEE 13 nodes test feeder [40], for which the regulator, capacitor, switch and the feeders with length equals to zero are neglected.

Each component in the distribution network has its own features and operating states that determine its performance. Assuming stationary conditions of the operating variables, the network operation is characterized by the location and magnitude of power available, the loads and the mechanical states of the components, because degradation or failures can have a direct impact on the power availability (in the DG units, feeders and/or main supply).

The renewable DG technologies considered in this work include solar photovoltaic (PV), wind turbines (W), electric vehicles (EV) and storage devices (batteries) (ST). The power output of each of these technologies is inherently uncertain. PV and W generation are subject to variability through their dependence on environmental conditions, i.e., solar irradiance and wind speed. Dis/connection and dis/charging patterns in EV and ST, respectively, further influence the uncertainty in the power outputs from the DG units. Also generation and distribution interruptions caused by failures are regarded as significant.

The following notation is used for sets and subsets of components in the distribution network: N – set of all nodes; MS – set of all types of main supply power sources; DG – set of all DG technologies; PV – set of all photovoltaic technologies; W – set of all wind technologies; EV – set of all electric vehicle technologies; ST – set of all storage technologies; FD – set of all feeders.

The configurations of power sources allocated in the network, indicating the size of power capacity and the location, is given in matrix form:

$$\Xi = \begin{array}{c} \begin{array}{cccccccccccc} \multicolumn{5}{c}{\underbrace{}_{\Xi^{MS}}} & \multicolumn{7}{c}{\underbrace{}_{\Xi^{DG}}} \end{array} \\ \begin{array}{cccccccccccc} \mathrm{MS}_1 & \cdots & \mathrm{MS}_j & \cdots & \mathrm{MS}_m & \mathrm{DG}_1 & \cdots & \mathrm{DG}_j & \cdots & \mathrm{DG}_d & & \\ \xi_{1,1} & \cdots & \xi_{1,j} & \cdots & \xi_{1,m} & \xi_{1,m+1} & \cdots & \xi_{1,m+j} & \cdots & \xi_{1,m+d} & & \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & & \\ \xi_{i,1} & \cdots & \xi_{i,j} & \cdots & \xi_{i,m} & \xi_{i,m+1} & \cdots & \xi_{i,m+j} & \cdots & \xi_{i,m+d} & & \text{node } i \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & & \\ \xi_{n,1} & \cdots & \xi_{n,j} & \cdots & \xi_{n,m} & \xi_{n,m+1} & \cdots & \xi_{n,m+j} & \cdots & \xi_{n,m+d} & & \end{array} \\ \begin{array}{cc} \text{fixed size and} & \text{decision matrix of type, size} \\ \text{location of MS} & \text{and location of DG units} \end{array} \end{array} \tag{1}$$

where Ξ – configuration matrix of type, size and location of the power sources allocated in the distribution network; Ξ^{MS} – size and location of main supply, fixed part of the configuration matrix; Ξ^{DG} – type size and location of DG units, decision variable part of the configuration matrix; n – number of nodes in the network, $|N|$; m – number of main supply type (transformers), $|MS|$; d –

number of DG technologies, $|DG|$.

$$\xi_{i,j} = \begin{cases} \zeta & \text{number of units of MS type or DG technology } j \text{ are allocated at node } i \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$$\forall i \in N, j \in MS \cup DG, \zeta \in \mathbb{Z}^*$$

Feeders deployment is described by the set of pairs of nodes connected:

$$FD = \{(1,2), \ldots, (i,i')\} \quad \forall (i,i') \in N \times N, (i,i') \text{ is a feeder} \tag{3}$$

Any configuration $\{\Xi, FD\}$ of power sources $\Xi = [\Xi^{MS}|\Xi^{DG}]$ and feeders FD of the distribution network are affected by uncertainty, so that the operation and performance of the distribution network is strongly dependent on the network configuration and scenarios. Furthermore, if the distribution network acts as a 'price taker', the variability of the economic conditions, particularly the price of the energy, is also an influencing factor [13, 19, 20]. For these reasons, it is imperative to represent and account for the uncertainties in the optimal allocation results for informed and conscious decision–making.

2.2 Uncertainty Modeling

2.2.1 Photovoltaic generation

PV technology converts the solar irradiance into electrical power through a set of solar cells configured as panels. Commonly, solar irradiance has been modeled using probabilistic distributions, derived from the weather historical data of a particular geographical area. The Beta distribution function [41, 42] is used in this paper:

$$f_{pv}(s) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} s^{(\alpha-1)}(1-s)^{(\beta-1)} & \forall s \in [0,1], \alpha_i, \beta_i > 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where s – solar irradiance; f_{pv} – beta probability density function; α, β – parameters of the beta probability density function. The parameters of the Beta probability density function can be inferred from the estimated mean μ and standard deviation σ of the random variable s as follows [1]:

$$\beta = (1-\mu)\left(\frac{\mu(1+\mu)}{\sigma^2} - 1\right) \tag{5}$$

$$\alpha = \frac{\mu\beta}{1-\mu} \tag{6}$$

Besides dependence on solar irradiation, PV depends also on the features of the solar cells that

constitute the panels and on ambient temperature on site. The power outputs from a single solar cell is obtained from the following equations [41, 42]:

$$T_c = T_a + s\left(\frac{N_{oT} - 20}{0.8}\right) \tag{7}$$

$$I = s(I_{sc} + k_i(T_c - 25)) \tag{8}$$

$$V = V_{oc} + k_v T_c \tag{9}$$

$$FF = \frac{V_{MPP} I_{MPP}}{V_{oc} I_{sc}} \tag{10}$$

$$P^{pv}(s) = n_{cells} FFVI \tag{11}$$

where T_a – ambient temperature (°C); N_{oT} – nominal cell operating temperature (°C); T_c – cell temperature (°C); I_{sc} – short circuit current (A); k_i – current temperature coefficient (mA/°C); V_{oc} – open circuit voltage (V); k_v – voltage temperature coefficient (mV/°C); V_{MPP} – voltage at maximum power (V); I_{MPP} – current at maximum power (A); FF – fill factor; n_{cells} – number of photovoltaic cells; $P^{pv}(s)$ – PV power output (W).

2.2.2 Wind generation

Wind generation is obtained from turbine–alternator devices that transform the kinetic energy of the wind into electrical power. The stochastic behavior of the wind speed is commonly represented through probability distribution functions. In particular, the Rayleigh distribution has been found suitable to model the randomness of the wind speed in various conditions [1, 42]:

$$f_w(ws) = \frac{2ws}{\sigma} \exp\left[-\left(\frac{ws}{\sigma}\right)^2\right] \tag{12}$$

where ws – wind speed (m/s); f_w – Rayleigh probability density function; σ – scale parameter of the Rayleigh distribution function.

Then, for a given wind speed value, the power output of one wind turbine can be determined as [1, 41, 42]:

$$P^w(ws) = \begin{cases} P^w_{RTD} \dfrac{ws - ws_{ci}}{ws_a - ws_{ci}} & \text{if } ws_{ci} \le ws < ws_a \\ P^w_{RTD} & \text{if } ws_a \le ws \le ws_{co} \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

where ws_{ci} – cut–in wind speed (m/s); ws_a – rated wind speed (m/s); ws_{co} – cut–out wind speed (m/s); P^w_{RTD} – rated power (kW); $P^w(ws)$ – wind power output (kW).

2.2.3 Electric vehicles

In this work, EV are considered as battery electric vehicles with three possible operating states: charging, discharging (i.e., injecting power into the distribution network) and disconnected [43]. To model their pattern of operation, they are considered as a 'block group', aggregating their single operating states into an overall performance. The main reasons for this aggregation are the observed nearly stable daily usage schedule of EV and the need of avoiding the combinatorial explosion of the model [42].

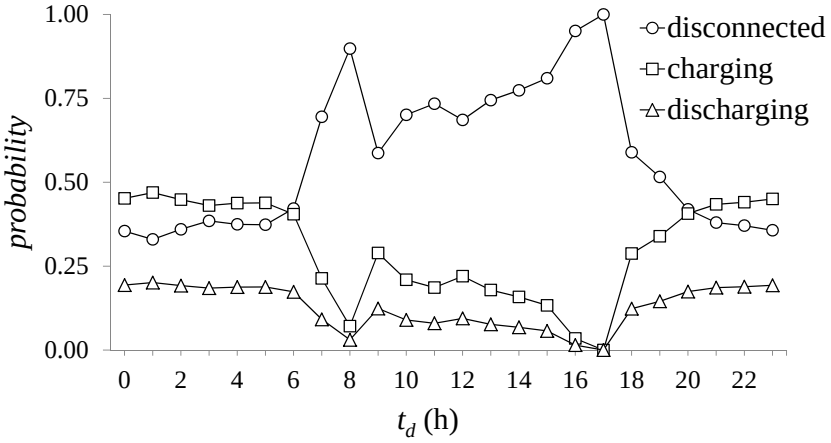

Figure 2: Hourly probability distribution of EV operating states per day

The power output of one block of EV is formulated by assigning residence time intervals to each possible operating state and associating them with the percentage of trips that the vehicles perform by hour of a day [43]. This allows approximating the hourly probability distribution of the operating states per day, as shown Figure 2. In a given (random) scenario of operational conditions, the determination of the operating state of a block of EV, of a specific hour of the day, is sampled randomly from the corresponding probability distribution. Accordingly, the power output for a unit or block group of EV is calculated using the expressions (14) and (15) below:

$$f_{ev}(t_d, op) = \begin{cases} p_{dch}(t_d) & \text{if } op = \text{discharging} \\ p_{ch}(t_d) & \text{if } op = \text{charging} \\ p_{dtd}(t_d) & \text{if } op = \text{disconnected} \end{cases} \tag{14}$$

$$\forall op \in OPs = \{\text{charging, discharging, disconnected}\}$$

$$P^{ev}(op) = \begin{cases} P^{ev}_{RTD} & \text{if } op = \text{discharging} \\ -P^{ev}_{RTD} & \text{if } op = \text{charging} \\ 0 & \text{if } op = \text{disconnected} \end{cases} \tag{15}$$

$$\forall t \in [0, t_{Rop}], op \in OPs = \{\text{charging, discharging, disconnected}\}$$

where t_d – hour of the day (h); t_{Rop} – residence time interval for operating state op (h); f_{ev} – operating state probability density function; P^{ev}_{RTD} – rated power (kW).

2.2.4 Storage devices

Analogously to the EV case, storage devices are treated as batteries. In reality, these present two main operating states, charging and discharging [44]. However, for this study the level of charge in the batteries is randomized and the state of discharging is the only one that is allowed. This is done to simplify the behavior of the batteries, making it independent on the previous state of charge. The discharging time interval is assigned according to the relation between the batteries rated power, their energy density and the random level of charge they present. For this, the discharging action is carried out at a rate equal to the rated power. Then, the power output per unit of mass of active chemical in the battery M_T is estimated as follows:

$$f_{st}(Q^{st}) = \begin{cases} \dfrac{1}{SE \times M_T} & \forall Q^{st} \in [0, SE \times M_T] \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

$$t'_R(Q^{st}) = \frac{Q^{st}}{P^s t_{RTD}} \tag{17}$$

$$P^{st}(t_R) = P^{st}_{RTD} \quad \forall t_R \in [0, t'_R] \tag{18}$$

where Q^{st} – level of charge in the battery (kJ); SE – specific energy of the active chemical (J/kg); M_T – total mass of the active chemical in the battery (kg); f_{st} – uniform probability density function; P^{st}_{RTD} – rated power (kW); t'_R – discharging time interval (h).

2.2.5 Main power supply

The MS spots in the distribution network are the power stations connected to the transmission system. The distribution transformers are located on these spots and provide the voltage level of the customers. The stochasticity of the available main supplies of power is represented following

normal distributions [10,45], truncated by the maximum capacity of the transformers.

$$f_{ms}(P^{ms}) = \begin{cases} \dfrac{\frac{1}{\sigma^{ms}}\phi\left(\frac{P^{ms}-\mu^{ms}}{\sigma^{ms}}\right)}{\Phi\left(\frac{P^{ms}_{cap}-\mu^{ms}}{\sigma^{ms}}\right)-\Phi\left(-\frac{\mu^{ms}}{\sigma^{ms}}\right)} & \forall P^{ms} \in [0, P^{ms}_{cap}] \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

where P^{ms} – available main power supply (kW); μ^{ms} – Normal distribution mean; σ^{ms} – Normal distribution standard deviation; f_{ms} – Normal probability density function; P^{ms}_{cap} – maximum capacity of the transformer (kW); ϕ – standard Normal probability density function; Φ – cumulative distribution function of ϕ.

2.2.6 Mechanical states of the components

Renewable DG units, MS spots and feeders are subject to wearing and degradation processes. These processes can trigger unexpected events, even failures, interrupting or reducing the specific functionality of each component. Frequently, the stochastic behavior of failures, repairs and maintenance actions is modeled using Markov models [28,42]. In this work, a two–state model is implemented in which the components can be in the mutually exclusive states: *available to operate* and *under repair* (failure state). Assuming the duration of each state as exponentially distributed, the mechanical state of a component can be randomly generated as follows:

$$mc = \begin{cases} 1 & \text{if the component is available to operate} \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

$$\forall \text{ component } \in \{\Xi, FD\}$$

$$f_{mc}(mc) = \frac{(1-mc)\lambda^F + mc\lambda^R}{\lambda^F + \lambda^R} \quad \forall mc \in \{0,1\} \tag{21}$$

where mc – binary mechanical state variable, λ^F – failure rate (failures/h), λ^R – repair rate (repairs/h), f_{mc} – mechanical state probability mass function.

2.2.7 Demand of power

Overall demands of power, as well as single load profiles in the nodes of the distribution network, can be obtained as daily load curves in which to each hour corresponds one specific level of load, inferred from historical data [1,14,19]. In addition, power demands profiles can be considered uncertain following normal distributions [34].

Within the proposed modeling framework, the nodal demands of power are defined by integrating

the two models mentioned above, i.e. adopting the general daily load profile and considering the hourly levels of load as normally distributed. Figure 3 schematizes the previous assumption for a generic node i.

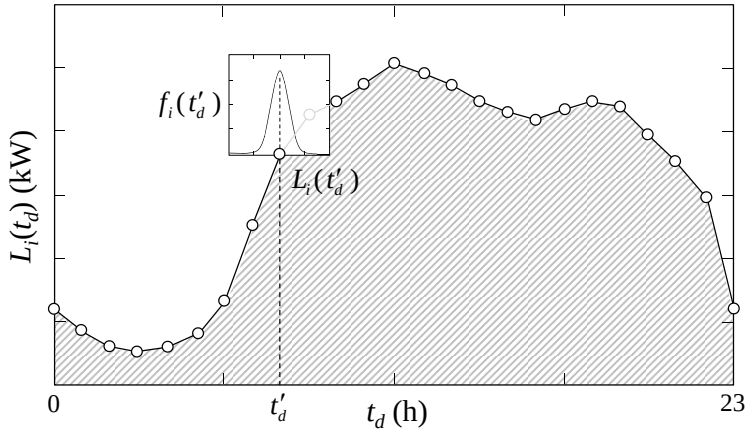

Figure 3: Daily load profile. Hourly normally distributed load

In this manner, the nodal demand of power is deducted from the overall demand in the network, and modeled as:

$$f_{L_i}(L_i, t_d) = \begin{cases} \dfrac{\dfrac{1}{\sigma_i(t_d)}\phi\left(\dfrac{L_i - \mu_i(t_d)}{\sigma_i(t_d)}\right)}{1 - \Phi\left(-\dfrac{\mu_i(t_d)}{\sigma_i(t_d)}\right)} & \forall i \in N, L_i \in [0, \infty] \\ 0 & \text{otherwise} \end{cases} \tag{22}$$

where t_d – hour of the day (h), L_i – power demand in node i (kW), μ_i – normal distribution mean of power demand in node i, σ_i – normal distribution standard deviation of power demand in node i, f_{L_i} – normal probability density function of power demand in node i.

2.3 Monte Carlo simulation

Most of the techniques used for evaluating the performance of renewable DG–integrated distribution networks are of two classes: analytical methods and MCS [28]. The implementation of analytical methods is always preferable, in theory, because of the possibility of achieving closed exact solutions, but in practice; it often requires strongly simplifying assumptions that may lead to unrealistic results: power network applications exist but for non–fluctuating or non–intermittent generation and/or load profiles, and low dimensionality of the network, gaining traceability with reduced computational efforts [32]. Different, MCS techniques allow considering more realistic models that analytical methods do, because simplifying assumptions are not necessary to solve the model, since *de facto* the model is not solved but simulated and the quantities of interest are estimated

from the statistics of the virtual simulation runs [46]. For this reason MCS is quite adequate for application on the analysis of distribution networks with significant randomness or variability in the sources of power supply and loads, failure occurrence and strong dependence on the power flows as a consequence of congestion conditions in the feeders, etc. [3, 31, 33, 41, 42, 47]; the price to pay for this is the possibly considerable increment in the use of computational resources, and various methods exist to tackle this problem [46].

Given the multiple sources of uncertainties considered in the proposed framework and the proven advantages of MCS for adequacy assessment of power distribution networks with uncertainties [3, 31, 33, 41, 42, 47], we adopt a non–sequential MCS to emulate the operation of a distribution network, sampling the uncertain variables without considering their time dependence, so as to reduce the computational problem.

For a given structure and configuration of the distribution network $\{\Xi, FD\}$, i.e., for the fixed Ξ^{MS} and FD deployments and the proposed renewable DG integration plan denoted by Ξ^{DG}, each uncertain variable is randomly sampled. The set ϑ of sampled variables constitutes an operational scenario, in correspondence of which the distribution network operation is modeled by OPF and its performance evaluated. The two inputs to the OPF model are the network configuration $\{\Xi, FD\}$ and the operational conditions scenario ϑ.

$$\vartheta = [t_d, P^{ms}_{i,j}, L_i, s_i, ws_i, op^{ev}_{i,j}, Q^{st}_{i,j}, mc_{i,j}, mc_{(i,i')}] \quad \forall i, i' \in N, j \in MS \cup DG, (i,i') \in FD \tag{23}$$

where t_d – hour of the day (h), randomly sampled from a discrete uniform distribution $U(1, 24)$.

Figure 4 shows an example of the matrix form construction of the DG–integrated distribution network, considering a simple case of $n = 3$ nodes. The network contains one MS spot at node $i = 1$, defining the fixed part Ξ^{MS} of the configuration matrix, whereas, the decision variable Ξ^{DG} proposes a renewable DG integration plan Ξ^{DG} that built from the number of units ξ of each DG technology allocated. In this way, the network configuration $\{\Xi, FD\}$ is composed by the matrix $\Xi = [\Xi^{MS}|\Xi^{DG}]$ and the deployment of feeders. Then, given the spatial representation $\{\Xi, FD\}$, the sampling of the scenario ϑ determines the operational conditions to perform power flow analysis, i.e., distribute the power available P^{ϑ}_{Ga} to supply appropriately the demands L_i. The available power in the power source type j at node i, $P^{\vartheta}_{Ga_{i,j}}$, is function of the number of units allocated $\xi_{i,j}$, the mechanical state $mc_{i,j}$ and the specific unitary power output function associated to the generation unit j, formulated in equations (24) and (25).

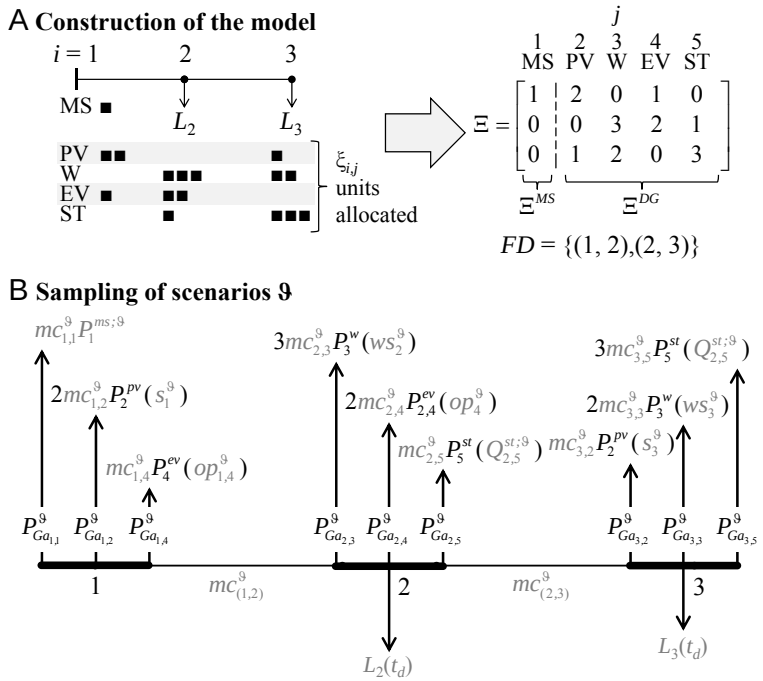

Figure 4: Example of the matrix form construction of a DG–integrated network (A) and schema of the operating state definition from the sampled variables (B)

$$P^{\vartheta}_{Ga_{i,j}} = \xi_{i,j} mc^{\vartheta}_{i,j} G_j(\vartheta) \tag{24}$$

$$G_j(\vartheta) = \begin{cases} P^{ms;\vartheta}_j & \text{if } j \in MS \\ P^{pv}_j(s^{\vartheta}_i) & \text{if } j \in PV \\ P^{w}_j(ws^{\vartheta}_i) & \text{if } j \in W \quad \forall i \in N \\ P^{ev}_j(op^{\vartheta}_{i,j}) & \text{if } j \in EV \\ P^{st}_j(Q^{st;\vartheta}_{i,j}) & \text{if } j \in ST \end{cases} \tag{25}$$

In the proposed non–sequential MCS procedure, the intermittency in the solar irradiation is taken into account defining a night interval between 22.00 and 06.00 hours, i.e., if the value of the hour of the day t_d (h), sampled from a discrete uniform distribution $U(1,24)$, falls in the night interval, there is no solar irradiation. Regarding the wind speed, its variability is considered by sampling positive values from a Rayleigh probability density function fitted on historical data and whose parameters as such that the probability of absence of wind is zero. Since it is not reasonable to force the historical profile of the wind speed to follow a distribution that admits intermittency, a common alternative technique is to model the wind by a Markov Chain. Indeed, it is possible to accurately represent the wind speed by a stationary Markov process if the historical profile of wind speed data is sufficiently large e.g. years [28]. The intermittency is, then, represented by the first state of the chain with wind speed equals to zero, and the sampling of the wind speed states in the non–sequential MCS of the proposed framework, can be performed using the steady–state

probabilities of the Markov Chain.

An important issue in modeling the operation of power systems is how to represent the evolution of uncertain operating conditions, such as solar irradiation, wind speed, load profiles, energy prices, among others. As an example, the load forecast implies the prediction of future power demands given specific previous conditions. Therefore, to consider load forecast uncertainty within the proposed MCS framework, it would be necessary to change to a sequential simulation model, in which the uncertain renewable energy resources, main power supply and loads must be sampled at each time step. In particular, load forecast uncertainty can be integrated properly building consecutive load scenarios and assigning corresponding probabilities of occurrence as presented by [7] and [48]. Another interesting approach for load forecast uncertainty modelling is the geometric Brownian motion (GBM) stochastic process [31, 49].

2.4 Optimal power flow

Power flow analysis is performed by DC OPF [50] which takes into account the active power flows, neglecting power losses, and assumes a constant value of the voltage throughout the network. This allows transforming to linear the classic non–linear power flow formulation, gaining simplicity and computational tractability. For this reason, DC power flow is often used in techno–economic analysis of power systems, more frequently in transmission [50, 51] but also in distribution networks [51].

The DC power flow generic formulation is:

$$P_i = \sum_{i' \in N} B_{(i,i')}(\delta_i - \delta_{i'}) \quad \forall i \in N, (i,i') \in FD \tag{26}$$

$$\sum_{i \in N} \left(P_{G_i} - L_i - P_i\right) = 0 \quad \forall i \in N \tag{27}$$

where, P_i – active power leaving node i (kW); $B_{(i,i')}$ – susceptance of the feeder (i,i') (Ω^{-1}); δ_i – voltage angle at node i; P_{G_i} – active power injected or generated at node i (kW); L_i – load at node i (kW).

The assumptions are:

- the difference between voltage angles is small, i.e., $\sin(\Delta\delta) \approx \Delta\delta$, $\cos(\Delta\delta) \approx 1$
- the feeders resistance are neglected, i.e., $R << X$, which implies that power losses in the feeder are also neglected
- the voltage profile is flat (constant V, set to 1 p.u.)

Then, for a given configuration $\{\Xi, FD\}$ and operational scenario ϑ the formulation of the OPF

problem is:

$$\min \quad C_{O\&M}^{net;\vartheta}(P_{Gu}^{\vartheta}) = \sum_{i \in N} \sum_{j \in MS \cup DG} C_{O\&M_j^v} P_{Gu_{i,j}}^{\vartheta} t^S \tag{28}$$

s.t.

$$L_i^{\vartheta} - \sum_{j \in MS \cup DG} P_{Gu_{i,j}}^{\vartheta} - \sum_{i' \in N} mc_{(i,i')}^{\vartheta} B_{(i,i')} (\delta_i^{\vartheta} - \delta_{i'}^{\vartheta}) - LS_i^{\vartheta} = 0 \quad \forall i, i' \in N, (i,i') \in FD \tag{29}$$

$$P_{Gu_{i,j}}^{\vartheta} \leq P_{Ga_{i,j}}^{\vartheta} \quad \forall i \in N, j \in MS \cup DG \tag{30}$$

$$0 \leq P_{Gu_{i,j}}^{\vartheta} \quad \forall i \in N, j \in MS \cup DG \tag{31}$$

$$mc_{(i,i')}^{\vartheta} B_{(i,i')} (\delta_i^{\vartheta} - \delta_{i'}^{\vartheta}) \leq V \times Amp_{(i,i')} \quad \forall i, i' \in N, (i,i') \in FD \tag{32}$$

$$-mc_{(i,i')}^{\vartheta} B_{(i,i')} (\delta_i^{\vartheta} - \delta_{i'}^{\vartheta}) \leq V \times Amp_{(i,i')} \quad \forall i, i' \in N, (i,i') \in FD \tag{33}$$

where, t^S – duration of the scenario (h); $C_{O\&M}^{net;\vartheta}$ – operating and maintenance costs of the total power supply and generation (\$); $C_{O\&M_j^v}$ – operating and maintenance variable costs of the power source j (\$/kWh); $mc_{(i,i')}^{\vartheta}$ – mechanical state of the feeder (i,i'); $B_{(i,i')}$ – susceptance of the feeder (i,i') (Ω^{-1}); $mc_{i,j}^{\vartheta}$ – mechanical state of the power source j at node i; $P_{Ga_{i,j}}^{\vartheta}$ – available power in the source j at node i (kW); $P_{Gu_{i,j}}^{\vartheta}$ power produced by source j at node i (kW); LS_i^{ϑ} – load shedding at node i (kW); V – nominal voltage of the network (kV); $Amp_{(i,i')}$ – ampacity of the feeder (i,i') (A).

The load shedding in the node i, LS_i, is defined as the amount of load(s) disconnected in node i to alleviate overloaded feeders and/or balance the demand of power with the available power supply [52]. The OPF objective is the minimization of the operating and maintenance costs associated to the generation of power for a given scenario ϑ of duration t^S. Equation (29) corresponds to the power balance equation at node i, while equations (30) and (31) are the bounds of the power generation and equations (32) and (33) account for the technical limits of the feeders.

2.5 Performance indicators

Given a set Υ of ns sampled operational scenarios ϑ_ℓ, $\ell \in \{1, \ldots, ns\}$, the OPF is solved for each scenario $\vartheta_\ell \in \Upsilon$, giving in output the values of ENS and global cost.

2.5.1 Energy not supplied

ENS is a common index for reliability evaluation in power systems [1, 10, 11, 48, 49, 52–55]. In the present work, its value is obtained directly from the OPF output in the form of the aggregation of

all nodal load sheddings per scenario ϑ_ℓ:

$$ENS^{\vartheta_\ell} = \sum_{i \in N} LS_i^{\vartheta_\ell} \times t^S \quad \forall \vartheta_\ell \in \Upsilon \tag{34}$$

$$ENS^{\Upsilon} = \{ENS^{\vartheta_1}, \ldots, ENS^{\vartheta_\ell}, \ldots, ENS^{\vartheta_{ns}}\} \tag{35}$$

2.5.2 Global cost

The C_g of the distribution network is formed by two terms, fixed and variable costs. The former term includes those costs paid at the beginning of the operation after the installation of the DG (conception of Ξ^{DG}). They are the investment–installation cost and the operation–maintenance fixed cost. The variable term refers to the operating and maintenance costs. Note that these costs are dependent on the power generation and supply, which are a direct output of the OPF (eq. (28)). In addition, this term considers revenues associated to the renewable sources incentives. Considering the distribution network as a 'price taker' entity, the profits depend on the value of the energy price that is correlated with the total load in the network. Three different ranges of load are considered for the daily profile. For each range, a correlation value of energy price is considered as shown in Figure 5(A).

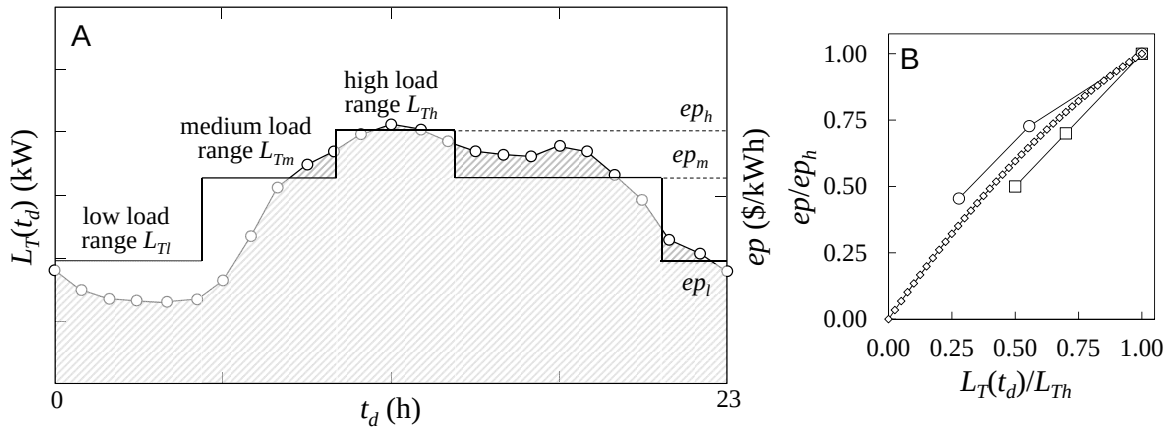

Figure 5: Example of load ranges definition for a generic daily load profile (A) and correlation energy price–total load (B) [13, 19, 20]

In Figure5(B) the correlation between energy price and total load is presented as the proportion of their maximum values. As an intermediate approximation of existing studies (e.g. [13, 19, 20]), the line with square–markers represents the proportional correlation used in this study, which can be expressed as:

$$ep = ep_h \left(-0.38 \left(\frac{L_T(t_d)}{L_{Th}} \right)^2 + 1.38 \frac{L_T(t_d)}{L_{Th}} \right) \tag{36}$$

Thereby, the global cost function for a scenario ϑ_ℓ is given by:

$$C_g^{\vartheta_\ell} = \sum_{i\in N}\sum_{j\in DG}(C_{inv_j} + C_{O\&M_j^f})\left(\frac{t^S}{t^h}\right) + C_{O\&M}^{net;\vartheta_\ell} - (inc + ep(L^{\vartheta_\ell}))\sum_{i\in N}\sum_{j\in DG} P_{Gu_{i,j}}^{\vartheta_\ell} t^S \tag{37}$$

$$C_g^{\Upsilon} = \{C_g^{\vartheta_1}, \ldots, C_g^{\vartheta_\ell}, \ldots, C_g^{\vartheta_{ns}}\} \tag{38}$$

where C_{inv_j} – investment cost of the DG technology j (\$); $C_{O\&M_j^f}$ – operating and maintenance fixed costs of the DG technology j (\$); t^h – horizon of analysis (h); inc – incentive for generation from renewable sources (\$/kWh); ep – energy price (\$/kWh); $C_g^{\vartheta_\ell}$ – global cost (\$).

2.5.3 Risk

In [38], the importance of measuring risk when optimizing under uncertainty and including it as part of the objective function(s) or constraints is emphasized. The proposed MOO framework introduces the CVaR as a coherent measure of the risk associated to the objective functions of interest. The CVaR has been broadly used in financial portfolio optimization either to reduce or minimize the probability of incurring in large losses [37, 38]. This risk measurement allows evaluating how 'risky' is the selection of a solution leading to a determined value of expected losses.

We can consider a fixed configuration of the distribution network $\{\Xi, FD\}$ including the integration of DG units as a portfolio. The assessed expectations of ENS^{Υ} and C_g^{Υ}, found from the MCS–OPF applied to the set of scenarios Υ, are estimations of the losses; then, $CVaR(ENS^{\Upsilon})$ and $CVaR(C_g^{\Upsilon})$ represent the risk associated to the solutions with these expectations.

The definition of CVaR for continuous and discrete general loss functions is given in detail in [38]. Here a simplified and intuitive manner to understand the CVaR definition and its derivation according to [56] is presented.

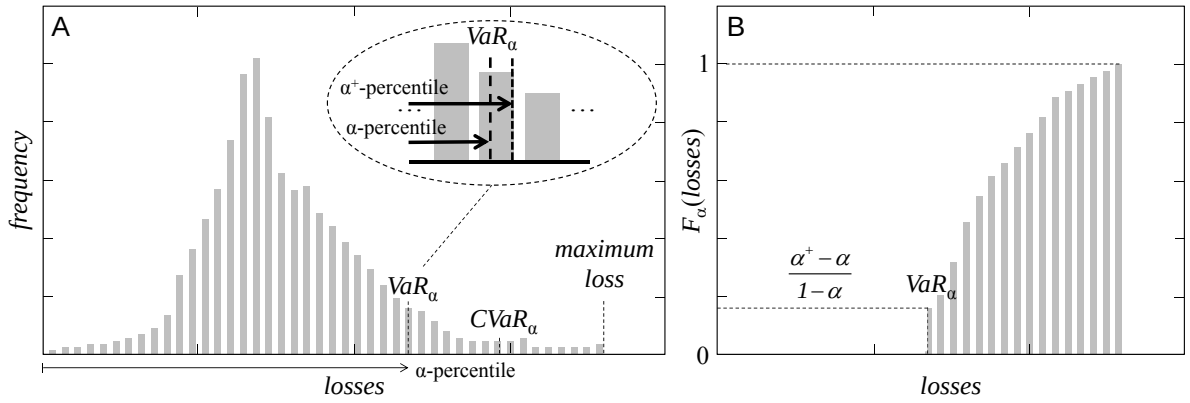

Figure 6: Graphic representation of the CVaR

As shown in Figure 6(A), for a discrete approximation of the probability of the losses, given

a confidence level or α-percentile, the value–at–risk VaR_α represents the smallest value of losses for which the probability that the losses do not exceed the value of VaR_α is greater than or equal to α. Thus, from the cumulative distribution function F(losses) is possible to construct the α-tail cumulative distribution function F_α(losses) for the losses, such that (Figure 6(B)):

$$F_\alpha(losses) = \begin{cases} \frac{F(losses)-\alpha}{1-\alpha} & \text{if } VaR_\alpha \leq losses \\ 0 & \text{otherwise} \end{cases} \tag{39}$$

The α-tail cumulative distribution function represents the risk 'beyond the VaR' and its mean value corresponds to the $CVaR_\alpha$.

Among other risk measures, the CVaR has been commonly used to assess the financial impact associated to different sources of uncertainty on electricity markets behavior. Some interesting approaches in the use of diverse risk measures for electricity markets modelling can be found in [49, 57, 58].

3 DG units selection, sizing and allocation

This section presents the general formulation of the MOO problem considered previously. As introduced, the practical aim of the MOO is to find the optimal integration of DG in terms of selection, sizing and allocation of the different renewable generation units (including EV and ST). The corresponding decision variables are contained in Ξ^{DG} of the configuration matrix Ξ.

The MOO problem consists in the concurrent minimization of the two objective functions measuring the C_g and ENS, and their associated risk. Specifically, their expected values and their CVaR values are combined, weighted by a factor $\beta \in [0, 1]$, which allows modulating the expected performance of the distribution network and its associated risk.

3.1 MOO problem formulation

Considering a set of randomly generated scenarios ϒ, the optimization problem is formulated as follows:

$$\min \quad f_1 = \beta EENS^{\Upsilon} + (1-\beta) CVaR_\alpha(ENS^{\Upsilon}) \tag{40}$$

$$\min \quad f_2 = \beta EC_g^{\Upsilon} + (1-\beta) CVaR_\alpha(C_g^{\Upsilon}) \tag{41}$$

s.t.

$$\xi_{i,j} = \begin{cases} \zeta & \text{number of units of MS type or DG technology } j \text{ are allocated at node } i \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$$\forall i \in N, j \in MS \cup DG, \zeta \in \mathbb{Z}^*$$

$$\sum_{i \in N} \sum_{j \in DG} \xi_{i,j} (C_{inv_j} + C_{O\&M_j^f}) \leq BGT \tag{42}$$

$$\sum_{i \in N} \xi_{i,j} \leq \tau_j \quad \forall j \in DG \tag{43}$$

$$\text{OPF}(\{\Xi, FD\}, \Upsilon) \tag{28 -33}$$

where EC_g and $EENS$ denote the expected values of C_g and ENS, respectively.

The meaning of each constraint is, (2) – the decision variable $\xi_{i,j}$ is a non–negative integer number; (42) – the total costs of investment and fixed operation and maintenance of the DG units must be less or equal to the available budget BGT; (43) – the total number of DG units to allocate of each technology j must be less or equal to the maximum number of units available τ_j to be integrated; (28)-(33) – all the equations of OPF must be satisfied for all scenarios in Υ.

Constraint (43) can be translated into maximum allowed penetration factor $PF_{max_j}^{DG}$ of each DG technology j. Defining PF as 'the output active power of total capacity of DG divided by the total network load' [59], constraint (43) can be rewritten as follows:

$$\underbrace{\frac{\sum_{i \in N} \xi_{i,j} EP_j^{DG}}{EL_T}}_{PF_j^{DG}} \leq \underbrace{\frac{\tau_j EP_j^{DG}}{EL_T}}_{PF_{max_j}^{DG}} \quad \forall j \in DG \tag{44}$$

where $\sum_{i \in N} \xi_{i,j}$ – total number of units of DG technology j integrated in the network; EP_j^{DG} – expected power output of one unit of DG technology j (kW); EL_T – expected total load (kW).

The MOO optimization problem is non–linear and non–convex, i.e., a non–convex mixed–integer non–linear problem or non–convex MINLP. It is non–linear because the objective functions given by equations (40) and (41) cannot be written in the canonical form of a linear program, i.e., $C^T X$, where C is a vector of known coefficients and X the decision vector. In the present case, the decision matrix Ξ^{DG} enters the MCS–OPF flow simulation to obtain the probability mass functions of C_g and ENS and, then, the objective functions are formed from the corresponding expected and CVaR values. Thus, the operations applied on Ξ^{DG} through MCS–OPF, expectation and CVaR cannot not be represented as the product $C^T \Xi^{DG}$. The problem is non–convex because the decision matrices Ξ^{DG} are integer–valued (constraint (2)) and, as it is known, the set of non–negative integers is non–convex.

Given the class of optimization problem in the proposed framework (non–convex MINLP), it is most likely to have multiple local minima. Moreover, the dimension of the distribution network can lead to a combinatorial explosion of the feasible space of the decision matrices Ξ^{DG} [7, 10], incrementing the number of possible local minima and hindering the possibility of benchmarking the optimal solutions obtained. However, an approximated but straightforward alternative is to perform several realizations of the framework obtaining different optimal solutions under the same optimization and simulation conditions (parameters) and, thus, compare them regarding the optimal decision matrices and their associated value of the objective functions.

This process was performed for the proposed case study. Indeed, the optimal decision matrices Ξ^{DG} are different in all the cases, when the optimization and simulation framework is performed under the same conditions but, nonetheless, practically the same Pareto optimal values of EC_g and $EENS$ are eventually obtained. This reflects that equally expected performances $(EC_g, EENS)$ can be obtained for different Ξ^{DG} considering the large amount of feasible combinations, which is what is of interest for practical applications.

3.1.1 NSGA–II with nested MCS–OPF

The combinatorial MOO problem under uncertainties is solved by the NSGA–II algorithm [39], in which the evaluation of the objective functions is performed by the developed MCS–OPF. The NSGA–II is one of the most efficient evolutionary algorithms to solve MOO problems [60]. The extension to MOO entails the integration of Pareto optimality concepts. In general terms, solving a MOO problem of the form:

$$\begin{aligned} \min_{X} \quad & \{f_1(X), f_2(X), \ldots, f_k(X)\} \\ & \text{s.t. } X \in \Lambda \end{aligned} \tag{45}$$

with at least two conflicting objectives functions ($f_i : \Re^n \rightarrow \Re$) implies to find, within a set of acceptable solutions that belong to the non–empty feasible region $\Lambda \subseteq \Re^n$, the decision vectors $X \in \Lambda$ that satisfy the following [61]:

$$\begin{gathered} \neg X \in \Lambda : f_i(X) \leq f_i(X'),\ \forall i \in \{1, \ldots, k\} \text{ and } f_i(X) < f_i(X') \text{ for at least one } i \\ \Downarrow \\ f_i(X) \prec f_i(X') \text{ i.e. } X \text{ dominates } X' \end{gathered} \tag{46}$$

The vector X is called a Pareto optimal solution and the Pareto front is defined as the set $\{f(X) \in \Re^n : X \text{ is Pareto optimal solution}\}$.

The process of searching the non–dominated solutions set, carried out by the NSGA–II MCS–OPF, can be summarized as shown in Figure 7.

The interested reader can consult [62–64] to compare the proposed framework to alternative MOO analytical approaches in energy applications.

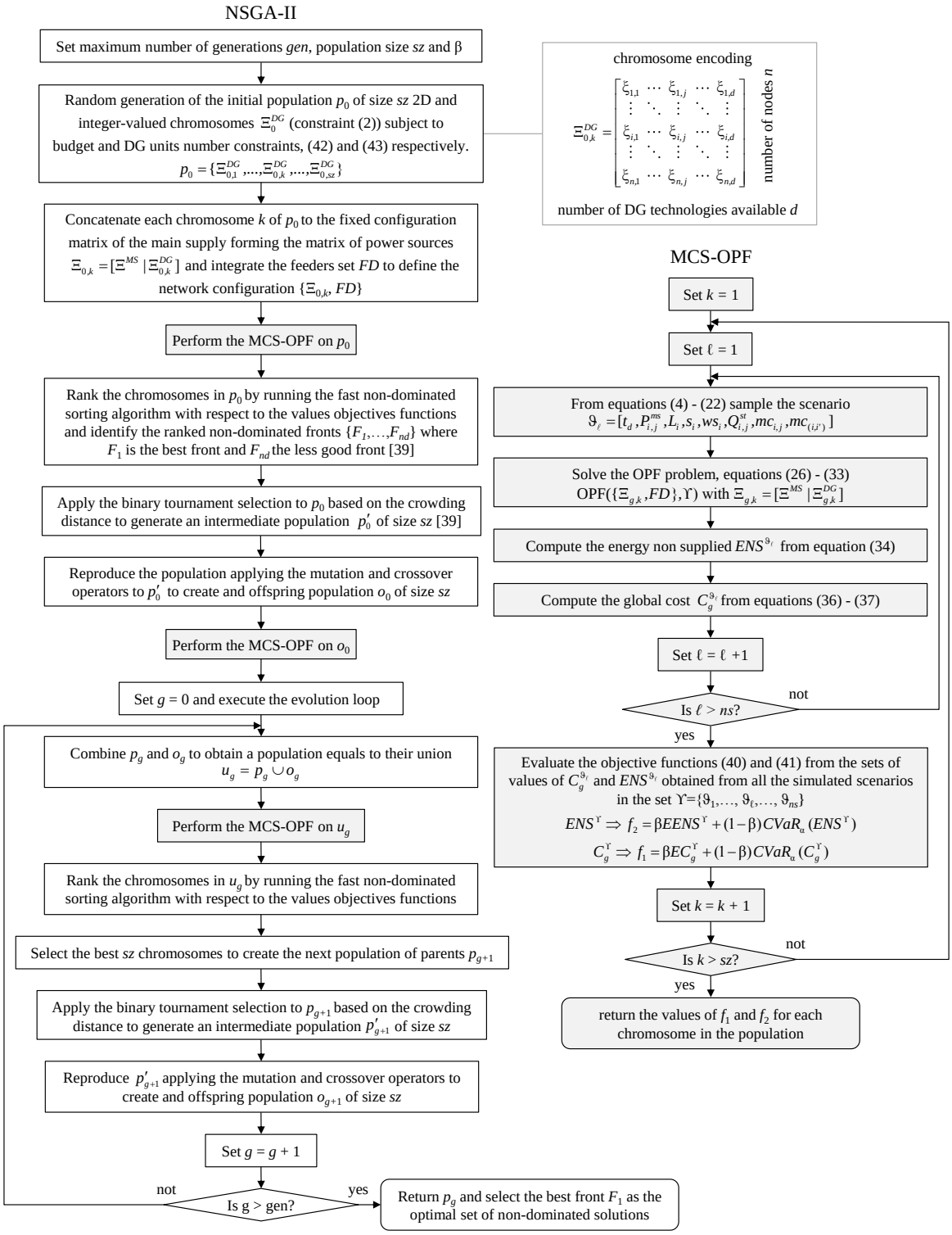

Figure 7: Flow chart of NSGA–II MCS–OPF MOO framework

4 Case study

We consider a distribution network adapted from the IEEE 13 nodes test feeder [40, 65]. The spatial structure of the network has not been altered but we neglect the regulator, capacitor and switch, and remove the feeders of zero length. The network is chosen purposely small, but with all relevant characteristics for the analysis, e.g. comparatively low and high spot and distributed load values and the presence of a power supply spot [65]. The original IEEE 13 nodes test feeder is dimensioned such that the total power demand is satisfied without lines overloading. We modify it so that it becomes of interest to consider the integration of renewable DG units. Specifically, the location and values of some of the load spots and the ampacity values of some feeders have been modified in order to generate conditions of power congestion of the lines, leading to shortages of power supply to specific portions of the network.

4.1 Distribution network description

The distribution network presents a radial structure of $n = 11$ nodes and $fd = (n-1) = 10$ feeders, as shown in Figure 8. The nominal voltage is $V = 4.16$ (kV), constant for the resolution of the DC optimal power flow problem.

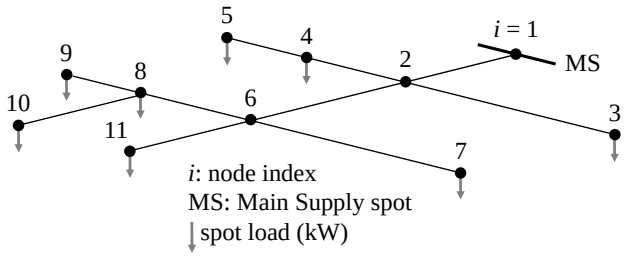

Figure 8: Radial 11–nodes distribution network

Table 1: Feeders characteristic and technical data [40]

type	node i	node i'	length (km)	X (Ω/km)	Amp (A)
T1	1	2	0.61	0.37	365
T2	2	3	0.15	0.47	170
T3	2	4	0.15	0.56	115
T1	2	6	0.61	0.37	365
T3	4	5	0.09	0.56	115
T6	6	7	0.15	0.25	165
T4	6	8	0.09	0.56	115
T1	6	11	0.31	0.37	365
T5	8	9	0.09	0.56	115
T7	8	10	0.24	0.32	115

Table 1 contains the technical characteristics of the different types of feeders considered: specifically, the indexes of the pairs of nodes that are connected by each feeder of the network, their length, reactance X and their ampacity Amp.

Concerning the main power supply spot, the maximum active power capacity of the transformer and the parameters of the normal distribution that describe its variability are given in Table 2.

Table 2: Main power supply parameters

node i	P_{cap}^{ms} (kW)	Normal distribution parameters	
		μ^{ms} (kW)	σ^{ms} (kW)
1	1600	1200	27.5

The nodal power demands are reported as daily profiles, normally distributed on each hour. The mean μ and variance σ values of the nodal daily profiles of the power demands are shown in Figure 9(A) and (B), respectively.

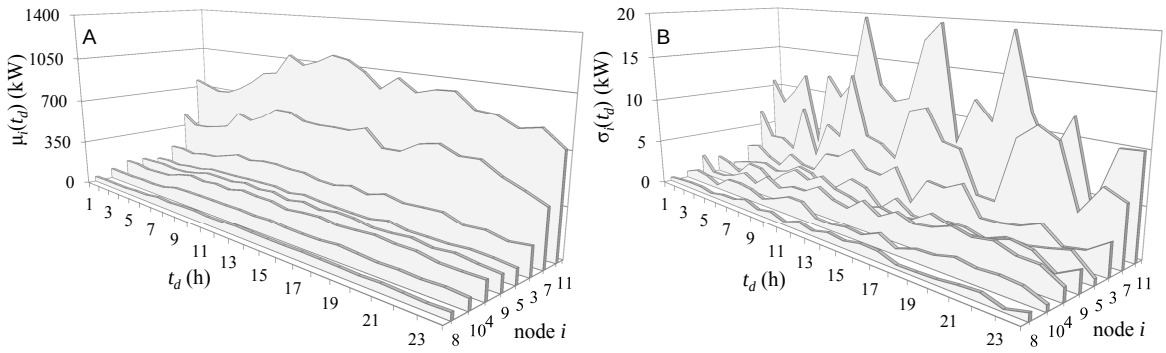

Figure 9: Mean (A) and variance (B) values of nodal power demand daily profiles

Table 3: Parameters of PV, W, EV and ST technologies [11, 13, 42]

PV		W	
Beta distribution α	0.26	Rayleigh distribution σ	7.96
Beta distribution β	0.73	P_{RTD}^{w} (kW)	50
T_a (°C)	30	ws_{ci} (m/s)	3.8
N_{oT} (°C)	43	ws_a (m/s)	9.5
I_{sc} (A)	1.8	ws_{co} (m/s)	23.8
k_i (mA/°C)	1.4	EV	
V_{oc} (V)	55.5	P_{RTD}^{ev} (kW)	6.3
k_v (mV/°C)	194	ST	
V_{MPP} (V)	38	P_{RTD}^{st} (kW)	0.275
I_{MPP} (A)	1.32	SE (kJ/kg)	0.042

The technical parameters of the four different types of DG technologies available to be integrated into the distribution network (PV, W, EV and ST) are given in Table 3. The values of the parameters

of the Beta and Rayleigh distributions describing the variability of the solar irradiation and wind speed, are assumed constant in the whole network, i.e., the region of distribution is such that the weather conditions are the same for all nodes.

The hourly per day operating states probability profile of the EV is presented in Figure 10 and failures and repair rates of the components of the distribution network are provided in Table 4.

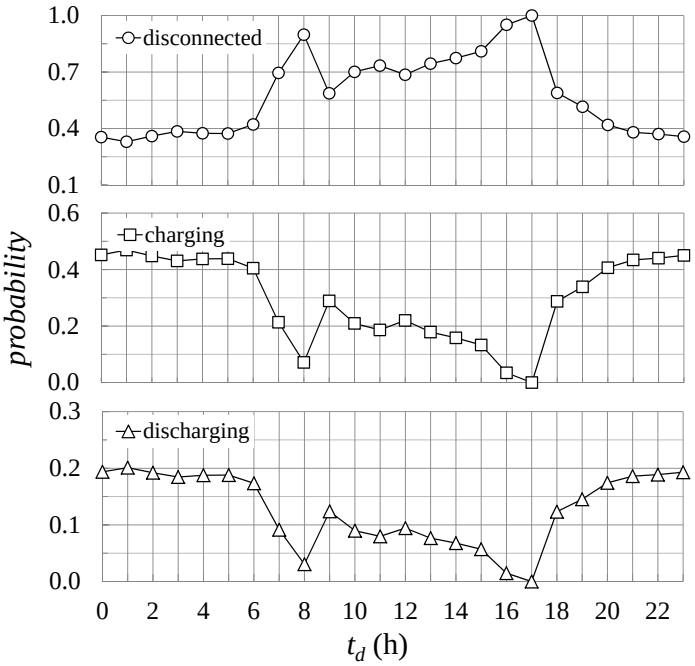

Figure 10: Hourly per day probability data of EV operating states

Table 4: Failure rates of feeders, MS and DG units [11, 13, 42, 66]

type		λ^F (failures/h)		λ^R (repairs/h)	
$MS \cup DG$	FD	$MS \cup DG$	FD	$MS \cup DG$	FD
MS	T1	3.33e−04	3.33e−04	0.021	0.198
PV	T2	4.05e−04	4.05e−04	0.013	0.162
W	T3	3.55e−04	3.55e−04	0.015	0.185
EV	T4	3.55e−04	3.55e−04	0.105	0.185
ST	T5	3.55e−04	3.55e−04	0.073	0.185
-	T6	-	4.00e−04	-	0.164
-	T7	-	3.55e−04	-	0.185

The values of the investment (C_{inv}) and fixed and variable Operational and Maintenance ($C_{O\&M^f}$ and $C_{O\&M^v}$) costs of the MS and DG units are reported in Table 5. Consistently with the constraints (42) and (43) of the MOO problem, the total investment associated to a decision variable Ξ^{DG} (proposed by the NSGA–II) must be less than or equal to the limit budget; which is set to $BGT = 4500000$ (\$), and the total number of units of each type of DG (following the order [PV, W,

EV, ST]) must be less than or equal to $\tau_j = [15000, 5, 200, 8000]$. The value of the incentive for renewable kWh supplied is taken as 0.024 ($/kWh) [34]. The maximum value of the energy price ep_h is 0.11 ($/kWh) [19,20]. Concerning the calculation of the CVaR, the α-percentile is taken as $\alpha = 0.80$.

Table 5: Investment, fixed O&M and variable O&M costs of MS and DG [27,34,66]

type	$C_{inv} + C_{O\&M^f}$ ($)	$C_{O\&M^v}$ ($/kWh)
MS	-	1.45e−01
PV	48	3.76e−05
W	113750	3.90e−02
EV	17000	2.20e−02
ST	135.15	4.62e−05

Five optimizations runs of the NSGA–II with the nested MCS–OPF algorithm have been performed, each one with a different value of the weight parameter β ∈ {1, 0.75, 0.5, 0.25, 0}, to analyze different tradeoffs between optimal average performance and risk. From equations (40) and (41), note that the value $\beta = 1$ corresponds to optimizing only the expected values of ENS and C_g, whereas $\beta = 0$ corresponds to the opposite extreme case of optimizing only the CVaR values. Each NSGA–II run is set to perform $g = 300$ generations over a population of $sz = 100$ chromosomes and, for the reproduction, the single–point crossover and mutation genetic operators are used. The crossover probability is $pco = 1$, whereas the mutation probability is $pmu = 0.1$; the mutation can occur simultaneously in any bit of the chromosome.

Finally, $sn = 250$ random scenarios are simulated by the MCS–OPF with time step $t^S = 1$ (h). Over an horizon of analysis of 10 years ($t^h = 87600$ (h)), in which the investment and fixed costs are prorated hourly.

4.2 Results and discussion

The Pareto fronts resulting from the NSGA–II MCS–OPF are presented in Figure 11 for the different values of β. The 'last generation' population is shown and the non–dominated solutions are marked in bold. Each non–dominated solution in the different Pareto fronts corresponds to an optimal decision matrix Ξ^{DG} for the sizing and allocation of DG, i.e., an optimal DG–integrated network configuration $\{\Xi, FD\}$ where $\Xi = [\Xi^{MS}|\Xi^{DG}]$.

In the Pareto fronts obtained, we look of three representative non–dominated solutions for the analysis: those with minimum values of the objective functions f_1 and f_2 independently ($\Xi^{DG}_{\min f_1}$ and $\Xi^{DG}_{\min f_2}$, respectively) and an intermediate solution at the 'elbow' of the Pareto front. Table 6 presents the values of the objective functions, $EENS$, EC_g and their respective CVaR values for the selected solutions. The $EENS$, EC_g and CVaR values of the case in which no DG is integrated in the network (MS case) is also reported.

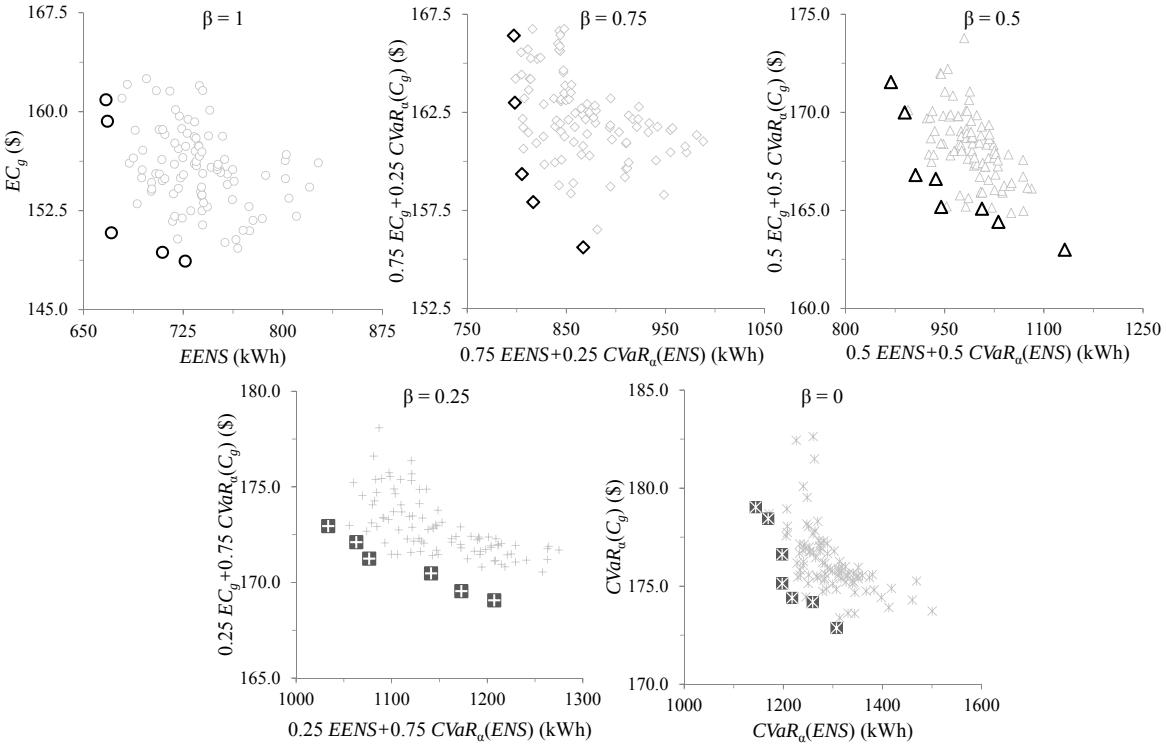

Figure 11: Pareto fronts for different values of β

Table 6: Objective functions: expected and CVaR values of selected Pareto front solutions

	β	f_1 (kW h)	f_2 ($)	$EENS$ (kW h)	$CVaR(ENS)$ (kW h)	EC_g ($)	$CVaR(C_g)$ ($)
MS	-	-	-	1109.21	1656.53	170.27	179.24
$\Xi^{DG}_{\min f_1}$		666.95	160.91	666.95	1093.12	160.91	185.11
Ξ^{DG}_{elbow}	1	671.05	150.83	671.05	1185.53	150.83	179.47
$\Xi^{DG}_{\min f_2}$		726.57	148.68	726.57	1279.37	148.68	178.23
$\Xi^{DG}_{\min f_1}$		797.07	166.41	677.74	1155.11	160.68	183.62
Ξ^{DG}_{elbow}	0.75	805.27	159.35	697.17	1129.62	153.09	178.15
$\Xi^{DG}_{\min f_2}$		867.08	155.61	729.81	1278.94	147.66	179.45
$\Xi^{DG}_{\min f_1}$		868.61	171.54	641.68	1095.52	159.43	183.64
Ξ^{DG}_{elbow}	0.5	936.58	166.67	701.72	1171.47	154.67	178.53
$\Xi^{DG}_{\min f_2}$		1131.64	162.99	843.53	1419.79	150.45	175.58
$\Xi^{DG}_{\min f_1}$		1033.65	172.95	723.19	1137.18	156.55	178.42
Ξ^{DG}_{elbow}	0.25	1076.53	171.25	743.61	1187.43	156.32	176.24
$\Xi^{DG}_{\min f_2}$		1207.33	169.07	835.23	1331.34	158.64	173.47
$\Xi^{DG}_{\min f_1}$		1144.36	179.03	744.71	1144.31	163.82	179.03
Ξ^{DG}_{elbow}	0	1197.79	176.62	749.21	1197.74	160.93	176.62
$\Xi^{DG}_{\min f_2}$		1307.33	172.87	828.55	1307.35	159.78	172.87

Figure 12 shows a bubble plot representation of the selected optimal solutions. The axes report the $EENS$ and EC_g values while the diameters of the bubbles are proportional to their respective CVaR values. The MS case is also plotted.

From Table 6 and Figure 12 it can be seen that, the MS case has an expected performance ($EENS = 1109.21$ (kW h) and $EC_g = 170.27$ (\$)) inferior (high $EENS$ and EC_g) to any case for which DG is optimally integrated. Furthermore, the $CVaR(ENS) = 1656.53$ (kW h) for the MS case is the highest, indicating the high risk of actually achieving the expected performance of energy not supplied. This confirms that DG is capable of providing a gain of reliability of power supply and economic benefits, the risk of falling in scenarios of large amounts of energy not supplied being reduced.

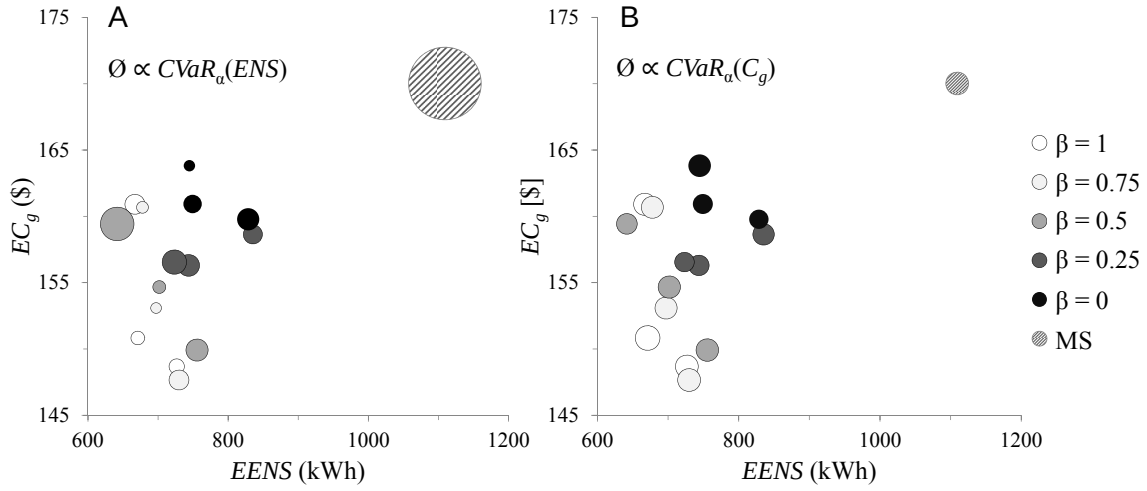

Figure 12: Bubble plots $EENS$ v/s EC_g. Diameter of bubbles proportional to $CVaR(ENS)$ (A) and $CVaR(Cg)$ (B)

Comparing among the selected optimal DG–integrated networks, in general the expected performances of $EENS$ and EC_g are progressively lower for increasing β. This to be expected: lowering the values of β, the MOO tends to search for optimal allocations and sizing Ξ^{DG} that sacrifice expected performance at the benefit of decreasing the level of risk (CVaR). These insights can serve the decision making process on the integration of renewable DG into the network, looking not only at the give–and–take between the values of $EENS$ and, but also at the level of risk of not achieving such expected performances due to the high variability.

Figure 13 shows the average total DG power allocated in the distribution network and its breakdown by type of DG technology for the optimal Ξ^{DG} as a function of β. It can be pointed out that the contribution of EV is practically negligible if compared with the other technologies. This is due to the fact that the probability that the EV is in a discharging state is much lower than that of being in the other two possible operating states, charging and disconnected (see Figure 10), combined with the fact that when EV is charging the effects are opposite to those desired.

The analysis of the results for different β values also allows highlighting the impact that each type of renewable DG technology has on the network performance. As can be noticed in Figure 13(A), the average total renewable DG power optimally allocated, increases progressively for increasing values of β: this could mean that to obtain less 'risky' expected performances less renewable DG power

needs to be installed. However, focusing on the individual fractions of average power allocated by PV, W and ST (Figure 13(B), (C) and (E), respectively), show that a reduction of the risk in the $EENS$ and EC_g is achieved specifically diminishing the proportion of PV power (from $0.29_{\beta=1}$ to $0.11_{\beta=0}$) while increasing the W and ST (from $0.38_{\beta=1}$ to $0.48_{\beta=0}$ and from $0.31_{\beta=1}$ to $0.39_{\beta=0}$, respectively), but this increment of W and ST power is not enough to balance the loss of PV power due to the limits imposed by the constraints in the number of each DG technology to be installed given by τ_j. Thus, PV power supply is shown to most contribute to the achievement of optimal expected performances, but with higher levels of risk. On the other hand, privileging the integration of W and ST power supply provides more balanced optimal solutions in terms of expectations and of achieving these expectations.

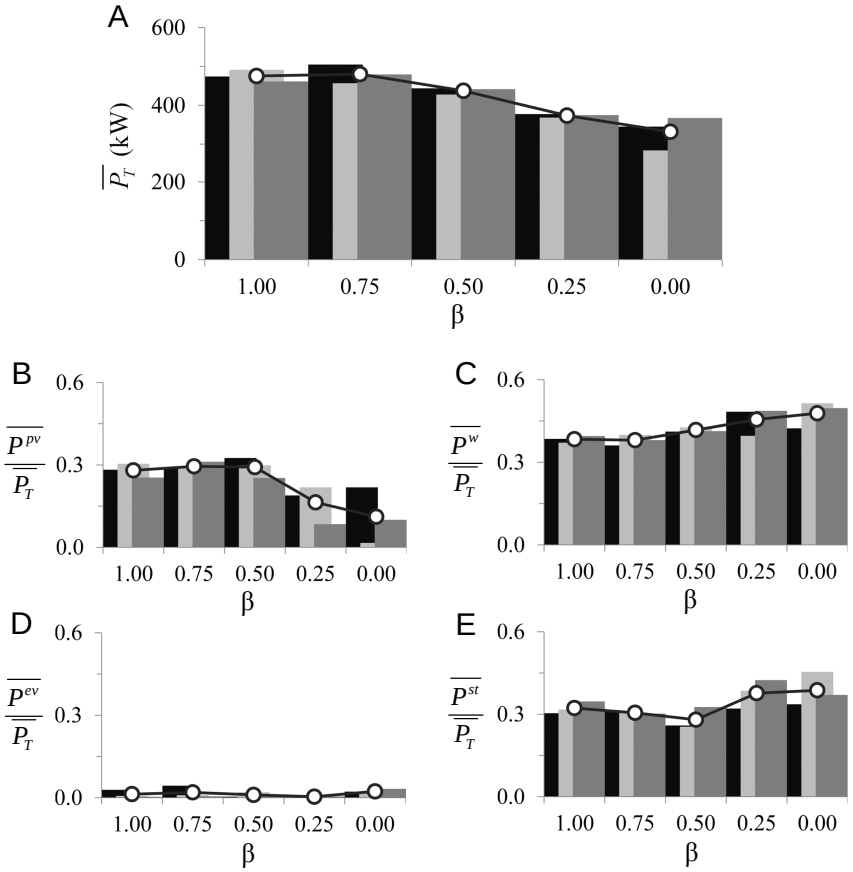

Figure 13: Average total DG power allocated (A) and its breakdown by type of DG: PV (B), W (C), EV (D) and ST (E)

Table 7 summarizes the minimum, average and maximum total renewable DG power allocated per node. The tendency is to install more localized sources (mainly nodes 4 and 8) of renewable DG power when the MOO searches only for the optimal expected performances ($\beta = 1$) and to have a more uniformly allocation of the power when searches for minimizing merely the CVaR ($\beta = 0$).

Table 7: Average, minimum and maximum total DG power allocated per node

P_T (kW)	β														
	1			0.75			0.5			0.25			0		
node	min	mean	max	min	mean	max	min	mean	max	min	mean	max	min	mean	max
1	12.08	34.44	54.77	1.15	22.40	38.56	0.00	19.23	40.98	0.00	39.03	121.00	3.00	17.33	34.71
2	2.30	40.72	69.73	0.00	49.95	77.70	36.50	58.40	123.36	3.00	63.61	132.93	0.00	42.54	84.09
3	0.00	24.83	46.45	14.80	41.79	85.03	0.00	37.94	105.11	4.00	36.87	98.53	1.00	32.84	77.78
4	76.00	110.00	133.41	1.15	67.40	133.63	0.58	38.04	80.13	6.15	20.73	61.85	0.00	39.85	85.86
5	22.60	52.39	77.08	28.90	60.66	98.59	12.63	89.39	143.50	3.30	23.49	54.25	1.00	24.97	79.64
6	12.33	55.56	85.46	10.45	21.22	38.95	2.00	27.68	106.26	12.15	53.78	84.43	0.00	50.64	116.85
7	8.00	16.52	35.38	39.38	64.07	104.05	0.00	52.03	159.73	0.00	34.09	92.81	5.00	18.51	39.23
8	79.03	111.20	146.63	30.00	74.57	114.41	0.00	40.60	146.06	4.00	37.94	102.60	1.00	39.49	119.38
9	0.00	20.03	68.73	4.00	74.07	107.88	0.00	46.72	85.61	0.00	44.06	94.08	0.00	32.86	74.53
10	0.00	9.07	25.35	0.00	1.58	7.88	0.00	11.88	58.69	0.00	8.58	43.40	0.00	30.12	83.45
11	0.00	9.98	17.68	0.00	3.04	13.20	0.00	4.74	23.45	0.00	8.99	45.95	0.00	7.31	51.17

5 Conclusions

We have presented a risk-based simulation and multi-objective optimization framework for the integration of renewable generation into a distribution network. The inherent uncertain behavior of renewable energy sources and variability in the loads are taken into account, as well as the possibility of failures of network components. For managing the risk of not achieving expected performances due to the multiple sources of uncertainty, the conditional value-at-risk is introduced in the objective functions, weighed by a β parameter which allows trading off the level of risk. The proposed framework integrates the Non-dominated Sorting Genetic Algorithm II as a search engine, Monte Carlo simulation to randomly generate realizations of the uncertain operational scenarios and Optimal Power Flow to model the electrical distribution network flows. The optimization is done to simultaneously minimize the energy not supplied and global cost, combined with their respective conditional value-at-risk values in an amount controlled by β.

To exemplify the proposed framework, a case study has been analyzed derived from the IEEE 13 nodes test feeder. The results obtained show the capability of the framework to identify Pareto optimal sets of renewable DG units allocations. Integrating the conditional value-at-risk into the framework and performing optimizations for different values of β has shown the possibility of optimizing expected performances while controlling the uncertainty in its achievement. The contribution of each type of renewable DG technology can also be analyzed, indicating which is more suitable for specific preferences of the decision makers.

References

[1] Y. Atwa, E. El-Saadany, M. Salama, and R. Seethapathy, "Optimal renewable resources mix for distribution system energy loss minimization," *Power Systems, IEEE Transactions on*, vol. 25, pp. 360–370, Feb 2010.

[2] G. Celli, E. Ghiani, S. Mocci, and F. Pilo, "A multiobjective evolutionary algorithm for the sizing and siting of distributed generation," *Power Systems, IEEE Transactions on*, vol. 20, pp. 750–757, May 2005.

[3] Z. Liu, F. Wen, and G. Ledwich, "Optimal siting and sizing of distributed generators in distribution systems considering uncertainties," *Power Delivery, IEEE Transactions on*, vol. 26, pp. 2541–2551, Oct 2011.

[4] M. F. Akorede, H. Hizam, and E. Pouresmaeil, "Distributed energy resources and benefits to the environment," *Renewable and Sustainable Energy Reviews*, vol. 14, no. 2, pp. 724 – 734, 2010.

[5] K. Alanne and A. Saari, "Distributed energy generation and sustainable development," *Renewable and Sustainable Energy Reviews*, vol. 10, no. 6, pp. 539 – 558, 2006.

[6] C. R. Karger and W. Hennings, "Sustainability evaluation of decentralized electricity generation," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 3, pp. 583 – 593, 2009.

[7] V. Martins and C. Borges, "Active distribution network integrated planning incorporating distributed generation and load response uncertainties," *Power Systems, IEEE Transactions on*, vol. 26, pp. 2164–2172, Nov 2011.

[8] R. Viral and D. Khatod, "Optimal planning of distributed generation systems in distribution system: A review," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 7, pp. 5146–5165, 2012.

[9] G. Koutroumpezis and A. Safigianni, "Optimum allocation of the maximum possible distributed generation penetration in a distribution network," *Electric Power Systems Research*, vol. 80, no. 12, pp. 1421 – 1427, 2010.

[10] A. Alarcon-Rodriguez, G. Ault, and S. Galloway, "Multi-objective planning of distributed energy resources: A review of the state-of-the-art," *Renewable and Sustainable Energy Reviews*, vol. 14, no. 5, pp. 1353 – 1366, 2010.

[11] M. Raoofat, "Simultaneous allocation of DGs and remote controllable switches in distribution networks considering multilevel load model," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 8, pp. 1429 – 1436, 2011.

[12] S.-H. Lee and J.-W. Park, "Selection of optimal location and size of multiple distributed generations by using Kalman filter algorithm," *Power Systems, IEEE Transactions on*, vol. 24, pp. 1393–1400, Aug 2009.

[13] H. Falaghi, C. Singh, M.-R. Haghifam, and M. Ramezani, "DG integrated multistage distribution system expansion planning," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 8, pp. 1489 – 1497, 2011.

[14] G. Celli, S. Mocci, F. Pilo, and G. Soma, "A multi-objective approach for the optimal distributed generation allocation with environmental constraints," in *Probabilistic Methods Applied to Power Systems, 2008. PMAPS '08. Proceedings of the 10th International Conference on*, pp. 1–8, May 2008.

[15] Y. Mohammed, M. Mustafa, N. Bashir, and A. Mokhtar, "Renewable energy resources for distributed power generation in nigeria: A review of the potential," *Renewable and Sustainable Energy Reviews*, vol. 22, pp. 257 – 268, 2013.

[16] C. L. T. Borges and V. F. Martins, "Multistage expansion planning for active distribution networks under demand and distributed generation uncertainties," *International Journal of Electrical Power & Energy Systems*, vol. 36, no. 1, pp. 107 – 116, 2012.

[17] G. Celli, F. Pilo, G. Soma, M. Gallanti, and R. Cicoria, "Active distribution network cost/benefit analysis with multi-objective programming," in *Electricity Distribution - Part 1, 2009. CIRED 2009. 20th International Conference and Exhibition on*, pp. 1–5, June 2009.

[18] H. Hejazi, M. Hejazi, G. Gharehpetian, and M. Abedi, "Distributed generation site and size allocation through a techno economical multi-objective differential evolution algorithm," in *Power and Energy (PECon), 2010 IEEE International Conference on*, pp. 874–879, Nov 2010.

[19] H. Ren and W. Gao, "A MILP model for integrated plan and evaluation of distributed energy systems," *Applied Energy*, vol. 87, no. 3, pp. 1001 – 1014, 2010.

[20] H. Ren, W. Zhou, K. Nakagami, W. Gao, and Q. Wu, "Multi-objective optimization for the operation of distributed energy systems considering economic and environmental aspects," *Applied Energy*, vol. 87, no. 12, pp. 3642 – 3651, 2010.

[21] W. El-Khattam, K. Bhattacharya, Y. Hegazy, and M. Salama, "Optimal investment planning for distributed generation in a competitive electricity market," *Power Systems, IEEE Transactions on*, vol. 19, pp. 1674–1684, Aug 2004.

[22] W. El-Khattam, Y. Hegazy, and M. Salama, "An integrated distributed generation optimization model for distribution system planning," *Power Systems, IEEE Transactions on*, vol. 20, pp. 1158–1165, May 2005.

[23] S. Ganguly, N. Sahoo, and D. Das, "A novel multi-objective PSO for electrical distribution system planning incorporating distributed generation," *Energy Systems*, vol. 1, no. 3, pp. 291–337, 2010.

[24] M. Gómez-González, A. López, and F. Jurado, "Optimization of distributed generation systems using a new discrete PSO and OPF," *Electric Power Systems Research*, vol. 84, no. 1, pp. 174 – 180, 2012.

[25] G. P. Harrison, A. Piccolo, P. Siano, and A. R. Wallace, "Hybrid GA and OPF evaluation of network capacity for distributed generation connections," *Electric Power Systems Research*, vol. 78, no. 3, pp. 392 – 398, 2008.

[26] W. Ouyang, H. Cheng, X. Zhang, and L. Yao, "Distribution network planning method considering distributed generation for peak cutting," *Energy Conversion and Management*, vol. 51, no. 12, pp. 2394 – 2401, 2010.

[27] K. Zou, A. Agalgaonkar, K. Muttaqi, and S. Perera, "Multi-objective optimisation for distribution system planning with renewable energy resources," in *Energy Conference and Exhibition (EnergyCon), 2010 IEEE International*, pp. 670–675, Dec 2010.

[28] C. L. T. Borges, "An overview of reliability models and methods for distribution systems with renewable energy distributed generation," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 4008 – 4015, 2012.

[29] L. Wang and C. Singh, "Multicriteria design of hybrid power generation systems based on a modified particle swarm optimization algorithm," *Energy Conversion, IEEE Transactions on*, vol. 24, pp. 163–172, March 2009.

[30] L. Ochoa and G. Harrison, "Minimizing energy losses: Optimal accommodation and smart operation of renewable distributed generation," *Power Systems, IEEE Transactions on*, vol. 26, pp. 198–205, Feb 2011.

[31] J. H. Zhao, J. Foster, Z.-Y. Dong, and K. P. Wong, "Flexible transmission network planning considering distributed generation impacts," *Power Systems, IEEE Transactions on*, vol. 26, pp. 1434–1443, Aug 2011.

[32] W.-S. Tan, M. Y. Hassan, M. S. Majid, and H. A. Rahman, "Optimal distributed renewable generation planning: A review of different approaches," *Renewable and Sustainable Energy Reviews*, vol. 18, pp. 626 – 645, 2013.

[33] A. Alarcon-Rodriguez, E. Haesen, G. Ault, J. Driesen, and R. Belmans, "Multi-objective planning framework for stochastic and controllable distributed energy resources," *Renewable Power Generation, IET*, vol. 3, pp. 227–238, June 2009.

[34] F. Pilo, G. Celli, S. Mocci, and G. Soma, "Active distribution network evolution in different regulatory environments," in *Power Generation, Transmission, Distribution and Energy Conversion (MedPower 2010), 7th Mediterranean Conference and Exhibition on*, pp. 1–8, Nov 2010.

[35] A. Soroudi and M. Ehsan, "A possibilistic - probabilistic tool for evaluating the impact of stochastic renewable and controllable power generation on energy losses in distribution networks - A case study," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 1, pp. 794 – 800, 2011.

[36] H. Hejazi, A. Araghi, B. Vahidi, S. Hosseinian, M. Abedi, and H. Mohsenian-Rad, "Independent distributed generation planning to profit both utility and DG investors," *Power Systems, IEEE Transactions on*, vol. 28, pp. 1170–1178, May 2013.

[37] A. Melnikov and I. Smirnov, "Dynamic hedging of conditional value-at-risk," *Insurance: Mathematics and Economics*, vol. 51, no. 1, pp. 182 – 190, 2012.

[38] R. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1443 – 1471, 2002.

[39] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *Evolutionary Computation, IEEE Transactions on*, vol. 6, pp. 182–197, Apr 2002.

[40] IEEE power and energy society. Distribution test feeders, Available at http:// ewh.ieee.org/ soc/pes/dsacom/testfeeders/index.html.

[41] Y. Li and E. Zio, "Uncertainty analysis of the adequacy assessment model of a distributed generation system," *Renewable Energy*, vol. 41, pp. 235 – 244, 2012.

[42] Y.-F. Li and E. Zio, "A multi-state model for the reliability assessment of a distributed generation system via universal generating function," *Reliability Engineering & System Safety*, vol. 106, pp. 28 – 36, 2012.

[43] K. Clement-Nyns, E. Haesen, and J. Driesen, "The impact of vehicle-to-grid on the distribution grid," *Electric Power Systems Research*, vol. 81, no. 1, pp. 185 – 192, 2011.

[44] F. Díaz-González, A. Sumper, O. Gomis-Bellmunt, and R. Villafáfila-Robles, "A review of energy storage technologies for wind power applications," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 4, pp. 2154 – 2171, 2012.

[45] A. Thornton and C. R. Monroy, "Distributed power generation in the United States," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 9, pp. 4809 – 4817, 2011.

[46] E. Zio, *The Monte Carlo Simulation Method for System Reliability and Risk Analysis*. Springer Series in Reliability Engineering, Springer London, 2013.

[47] Y. Hegazy, M. Salama, and A. Chikhani, "Adequacy assessment of distributed generation systems using monte carlo simulation," *Power Systems, IEEE Transactions on*, vol. 18, pp. 48–52, Feb 2003.

[48] M. Shaaban, Y. Atwa, and E. El-Saadany, "DG allocation for benefit maximization in distribution networks," *Power Systems, IEEE Transactions on*, vol. 28, pp. 639–649, May 2013.

[49] M. Samper and A. Vargas, "Investment decisions in distribution networks under uncertainty with distributed generation - Part II: Implementation and results," *Power Systems, IEEE Transactions on*, vol. 28, pp. 2341–2351, Aug 2013.

[50] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans, "Usefulness of DC power flow for active power flow analysis," in *Power Engineering Society General Meeting, 2005. IEEE*, pp. 454–459 Vol. 1, June 2005.

[51] D. Van Hertem, J. Verboomen, K. Purchala, R. Belmans, and W. Kling, "Usefulness of DC power flow for active power flow analysis with flow controlling devices," in *AC and DC Power Transmission, 2006. ACDC 2006. The 8th IEE International Conference on*, pp. 58–62, March 2006.

[52] R. Billinton and R. Allan, *Reliability evaluation of power systems*. Pitman Advanced Publishing Program, 1984.

[53] S. Haffner, L. Pereira, L. Pereira, and L. Barreto, "Multistage model for distribution expansion planning with distributed generation Part I: Problem formulation," *Power Delivery, IEEE Transactions on*, vol. 23, pp. 915–923, April 2008.

[54] S. Haffner, L. Pereira, L. Pereira, and L. Barreto, "Multistage model for distribution expansion planning with distributed generation Part II: Numerical results," *Power Delivery, IEEE Transactions on*, vol. 23, pp. 924–929, April 2008.

[55] L. Wang and C. Singh, "Multicriteria design of hybrid power generation systems based on a modified particle swarm optimization algorithm," *Energy Conversion, IEEE Transactions on*, vol. 24, pp. 163–172, March 2009.

[56] U. S. VaR vs CVaR in risk management and optimization, CARISMA conference; 2010 (Presentation), Available at: http://www.ise.ufl.edu/uryasev/publications/.

[57] A. Ahmadi, M. Charwand, and J. Aghaei, "Risk-constrained optimal strategy for retailer forward contract portfolio," *International Journal of Electrical Power & Energy Systems*, vol. 53, pp. 704 – 713, 2013.

[58] M. Gitizadeh, M. Kaji, and J. Aghaei, "Risk based multiobjective generation expansion planning considering renewable energy sources," *Energy*, vol. 50, pp. 74 – 82, 2013.

[59] F. Ugranli and E. Karatepe, "Multiple-distributed generation planning under load uncertainty and different penetration levels," *International Journal of Electrical Power & Energy Systems*, vol. 46, pp. 132 – 144, 2013.

[60] R. Ak, Y. Li, V. Vitelli, E. Zio, E. L. Droguett, and C. M. C. Jacinto, "NSGA-II-trained neural network approach to the estimation of prediction intervals of scale deposition rate in oil & gas equipment," *Expert Systems with Applications*, vol. 40, no. 4, pp. 1205 – 1212, 2013.

[61] J. Branke, K. Deb, K. Miettinen, and R. Slowinski, *Multiobjective Optimization: Interactive and Evolutionary Approaches*. LNCS sublibrary: Theoretical computer science and general issues, Springer, 2008.

[62] J. Aghaei and N. Amjady, "A scenario-based multiobjective operation of electricity markets enhancing transient stability," *International Journal of Electrical Power & Energy Systems*, vol. 35, no. 1, pp. 112 – 122, 2012.

[63] J. Aghaei, N. Amjady, and H. Shayanfar, "Multi-objective electricity market clearing considering dynamic security by lexicographic optimization and augmented epsilon constraint method," *Applied Soft Computing*, vol. 11, no. 4, pp. 3846 – 3858, 2011.

[64] J. Aghaei, M. A. Akbari, A. Roosta, and A. Baharvandi, "Multiobjective generation expansion planning considering power system adequacy," *Electric Power Systems Research*, vol. 102, pp. 8 – 19, 2013.

[65] W. Kersting, "Radial distribution test feeders," *Power Systems, IEEE Transactions on*, vol. 6, pp. 975–985, Aug 1991.

[66] R. Webster, "Can the electricity distribution network cope with an influx of electric vehicles?," *Journal of Power Sources*, vol. 80, no. 1-2, pp. 217 – 225, 1999.

Paper (ii)

R. Mena, M. Hennebel, Y.–F. Li and E. Zio, "Self–adaptable hierarchical clustering analysis and differential evolution for optimal integration of renewable distributed generation," Applied Energy, vol. 133, pp. 388—402, 2014.

Self-adaptable hierarchical clustering analysis and differential evolution for optimal integration of renewable distributed generation

Rodrigo Mena[a], Martin Hennebel[b], Yan-Fu Li[a], Enrico Zio[a,c,1]

[a]*Chair on Systems Science and the Energetic Challenge, Fondation Electricité de France at CentraleSupeléc, Châtenay-Malabry Cedex, France*

[b]*CentraleSupeléc, Department of Power & Energy Systems, Gif-Sur-Yvette, France*

[c]*Politecnico di Milano, Energy Department, Milan, Italy*

Abstract

In a previous paper, we have introduced a simulation and optimization framework for the integration of renewable generators into an electrical distribution network. The framework searches for the optimal size and location of the distributed renewable generation units (DG). Uncertainties in renewable resources availability, components failure and repair events, loads and grid power supply are incorporated. A Monte Carlo simulation and optimal power flow (MCS–OPF) computational model is used to generate scenarios of the uncertain variables and evaluate the network electric performance with respect to the expected value of the global cost (ECG). The framework is quite general and complete, but at the expenses of large computational times for the analysis of real systems. In this respect, the work of the present paper addresses the issue and introduces a purposely tailored, original technique for reducing the computational efforts of the analysis. The originality of the proposed approach lies in the development of a new search engine for performing the minimization of the ECG, which embeds hierarchical clustering analysis (HCA) within a differential evolution (DE) search scheme to identify groups of similar individuals in the DE population and, then, ECG is calculated for selected representative individuals of the groups only, thus reducing the number of objective function evaluations. For exemplification, the framework is applied to a distribution network derived from the IEEE 13 nodes test feeder. The results show that the newly proposed hierarchical clustering differential evolution (HCDE) MCS–OPF framework is effective in finding optimal DG–integrated network configurations with reduced computational efforts.

Keywords: Renewable distributed generation, uncertainty, simulation, optimization, differential evolution, hierarchical clustering analysis

[1] Corresponding author. Tel: +33 1 4113 1606; fax: +33 1 4113 1272.
E-mail addresses: enrico.zio@ecp.fr, enrico.zio@polimi.it (E. Zio).

1 Introduction

Renewable distributed generation (DG) requires the selection of the different available technologies, and their sizing and allocation onto the power distribution network, considering the specific economic, operational and technical constraints [1–5]. This can become a complex optimization problem, depending on the size of the distribution network and the number of renewable DG technologies available, that can lead to combinatorial explosion [1, 3, 6–9]. Furthermore, for each renewable DG plan considered, the power flow problem needs to be solved to assess the response of the distribution network in terms of power and voltage profiles, available power usage, power demand satisfaction, economic performances, etc., with possibly significant computation times.

Heuristic optimization techniques belonging to the class of Evolutionary Algorithms (EAs), like honey bee mating [10], particle swarm optimization (PSO) [9, 11–13], differential evolution (DE) [14, 15] and genetic algorithms (GA) [2, 3, 16, 17], have been considered for the solution to this problem, since they can deal straightforwardly with non–convex combinatorial problems, discontinuous search spaces and non–differentiable objective functions [1, 9].

To improve the performance of EAs for the complex optimization problem of DG planning, we consider the integration of clustering [18–23]. This can be directed to the enhancement of the global and/or local searching ability of the algorithm, and amounts to identifying groups of similar individuals and applying different evolution operators to those of a same cluster (group) [18, 20–22], e.g. for random generation of new individuals in the neighborhood of cluster centroids [23], or multi–parents crossover over new randomly generated individuals spread in the global feasible space [19]. Even if convergence is improved, some of these methodologies increase temporarily the overall size of the population and, thus, the computational effort. In addition, the accuracy of the clusters structures in representing the distribution of individuals must be controlled for performing clustering conveniently.

The main original contribution of the work here presented, lies in the development of the clustering strategy in a controlled manner. The implementation of such clustering strategy is done within a Monte Carlo simulation and optimal power flow (MCS–OPF) model and differential evolution (DE) optimization framework [24] previously developed by the authors for the integration of renewable generators into an electrical distribution network: the framework searches for the optimal size and location of the distributed renewable generation units (DG) [25]. Optimality of the DG plan is sought with respect to the expected global cost (ECG). The introduction of the clustering is hierarchically (i.e., hierarchical clustering analysis, HCA, [26]) by a controlled way of reducing the number of individuals to be evaluated during the DE search, therefore, improving the computational efficiency. Henceforth, we call our method hierarchical clustering differential evolution (HCDE).

HCA is introduced to build a hierarchical structure of grouping individuals of the population

that present closeness under the control of a specific linkage criterion based on defined distance metrics [26]. The HCA outcomes are the linkage distances at which the grouping actions take place, defining the different levels in the hierarchical structure. Two control parameters are introduced in the HCA, the cophenetic correlation coefficient (CCC) and a cutoff level coefficient of the linkage distances in the hierarchical structure of the groups (p_{co}). The CCC is a similarity coefficient that measures how representative is the proposed grouping structure by comparing their linkage distances with the original distances between all the individuals in the population. In the hierarchical structure, the linkage distance given by p_{co} sets the level at which the groups formed below it are considered to be 'close enough' to constitute independent clusters. The two parameters allow HCDE to adapt itself in each generation of the search, 'deciding' whether to perform clustering if the CCC is greater than or equal to a preset threshold (CCC_{th}) and cutting the hierarchical structure in independent clusters according to the linkage distance given by p_{co}. Then, the individual closest to the centroid of each cluster is taken as the feasible representative solution in the population that enters the evolution phase of the HCDE algorithm. Figure 1 summarizes schematically the structure of the proposed framework.

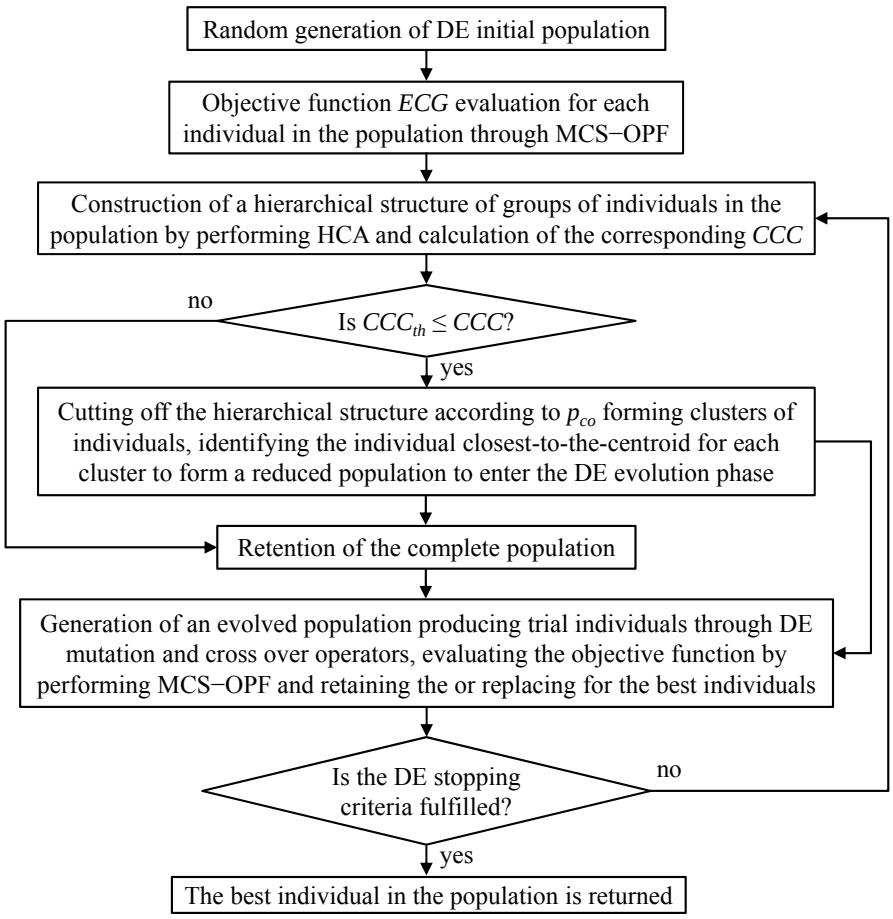

Figure 1: HCDE framework schema

We test the approach on a case study based on the IEEE 13 nodes test feeder distribution network [27], completing the study with a sensitivity analysis to investigate the effects of the parameters controlling the clustering, namely CCC and p_{co}.

For practical ease of the presentation of the approach, in the next section we provide the basic elements of the model of the distribution network considered as case study and we briefly summarize the MCS–OPF model taken from [25]. In Section 3, we embed this in the HCDE for renewable DG selection, sizing and allocation. Finally, in Section 4 we present the numerical results of the case study and in Section 5 we draw some conclusions on the work performed.

2 Renewable DG–integrated network model

The operation of the renewable DG–integrated network is considered to be dictated by the location and magnitude of the power available in the different sources, the loads and the operating states of the components. Uncertainty is present in the states of operation of the components, due to stochasticity of degradation and failures, and in the behavior of the renewable energy sources. These uncertainties have a direct impact on the power available (from the DG units, main supply spots and/or feeders) to satisfy power demands, which are, in turn, also subject to fluctuations. Furthermore, if the distribution network is considered as a 'price taker' entity, the uncertain behavior of the power demand impacts directly over the energy price [4, 5, 28]. Consequently, an attentive modeling of the uncertainties in renewable DG planning is imperative for well–supported decision–making.

Monte Carlo simulation (MCS) has already been used to emulate the stochastic operating conditions and evaluate the performance of power distribution networks [19, 28–30]. In the present paper, non–sequential MCS is used to randomly sample the modeled uncertain variables for a specific renewable DG plan, without dependence on previous operating conditions, characterizing the network operation in terms of location and magnitudes of power available and loads. Then, the performance of the DG–integrated network is evaluated through the optimal power flow model.

2.1 Monte Carlo and optimal power flow simulation

In the proposed framework, the renewable DG technologies considered are of four types: solar photovoltaic (PV), wind turbines (W), electric vehicles (EV) and storage devices (ST); these are represented by the set DG that contains all the dg types of technologies. As for main power supply spots or transformers (MS), the set MS indicates the ms different types of MS considered in the network.

The DG–integrated network deployment is represented by the location and capacity size of the power sources, as indicated in matrix form in Equation (1) below, where $\xi_{i,j}$ indicates the number

of units of main supply spots or DG technology j that are allocated at a node i:

$$\Xi = \left[\begin{array}{ccccc|ccccc} \xi_{1,1} & \cdots & \xi_{1,j} & \cdots & \xi_{1,ms} & \xi_{1,1+ms} & \cdots & \xi_{1,j+ms} & \cdots & \xi_{1,dg+ms} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \xi_{i,1} & \cdots & \xi_{i,j} & \cdots & \xi_{i,ms} & \xi_{i,1+ms} & \cdots & \xi_{i,j+ms} & \cdots & \xi_{i,dg+ms} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \xi_{n,1} & \cdots & \xi_{n,j} & \cdots & \xi_{n,ms} & \xi_{n,1+ms} & \cdots & \xi_{n,j+ms} & \cdots & \xi_{n,dg+ms} \end{array}\right] = [\Xi^{MS}|\Xi^{DG}] \quad (1)$$

$$\forall \xi i, j \in \mathbb{Z}^*, i \in N, j \in PS$$

where N and $PS = MS \cup DG$ are the set of nodes in the network and the set of all power sources, whose cardinalities are n and $ps = ms + dg$, respectively.

The set of feeders FD is defined by all the pairs of nodes (i, i') connected by a distribution line $\forall (i, i') \in N \times N$.

The considered uncertain conditions that determine the operation of the DG–integrated network are accounted for using different stochastic models, as summarized in Table 1. The interested reader can consult [25] for further details.

Table 1: Uncertain conditions models in the DG–integrated network operation

Variable	Nomenclature	States and units	Model	Parameters
Hour of the day	t_d	(h)	Discrete uniform distribution	$[1, 24]$
Mechanical state	$mc_{i,j}$	(0): under repair (1): operating	Two–state Markov	λ_j^F, λ_j^R $\lambda_{i,i'}^F, \lambda_{i,i'}^R$
Main power supply	$P_{i,j}^{MS}$	(kW)	Truncated normal distribution $0 \leq P_{i,j}^{MS} \leq P_{cap_j}^{MS}$	$\mu_j^{MS}, \sigma_j^{MS}$ $P_{cap_j}^{MS}$
Solar irradiance	s_i	$[0, 1]$	Beta distribution	$\alpha_i^{PV}, \beta_i^{PV}$
Wind speed	ws_i	(m/s)	Rayleigh distribution	σ_i^W
EV operating state	$op_{i,j}^{EV}$	(-1) : charging (0): disconnected (1): discharging	'Block groups' Hourly probability distribution of EV operating states per day: p^- (charging) p^0 (disconnected) p^+ (discharging)	t_d
ST level of charge	$Q_{i,j}^{ST}$	(kJ)	Uniform distribution	$[SE_j^{ST} M_{T_{i,j}}^{ST}]$
Nodal power demand	L_i	(kW)	Daily nodal load profiles, hourly normally distributed load. Truncated normal distribution $0 \leq L_i \leq \infty$	$\mu_i^L(t_d), \sigma_i^L(t_d)$

where $\forall i, i' \in N$, $j \in PS$, $(i, i') \in FD$, λ_j^F and λ_j^R (1/h) are the failure and repair rates of the power source j, respectively, $\lambda_{i,i'}^F$ and $\lambda_{i,i'}^R$ (1/h) are the failure and repair rates of the feeder (i, i'), respectively, μ_j^{MS} and σ_j^{MS} (kW) are the normal distribution mean and standard deviation associated to the main supply j, $P_{cap_j}^{MS}$ (kW) is the maximum capacity of the transformer j, α_i^{PV} and β_i^{PV} are the parameters of the Beta probability density function of the solar irradiance at node i, σ_i^W is the scale parameter of the Rayleigh distribution function of the wind speed at node i, SE_j^{ST} (kJ/kg) is the specific

energy of the active chemical in the battery type j, $M^{ST}_{T_{i,j}}$ (kg) is the mass of active chemical in the battery type j at node i, $\mu^L_i(t_d)$ and $\sigma^L_i(t_d)$ (kW) are the hourly mean and standard deviation of the normal distribution of the power load at node i.

Concerning the hour of the day t_d (h), sampled from a discrete uniform distribution $U(1,24)$, the night interval is defined between 22.00 and 06.00 h. If the value of t_d falls in the night interval, there is no solar irradiation.

The resulting realization of one operational scenario of duration ts (h), for the given DG plan denoted by $\{FD,\Xi\}$, consists in the random sampling of each uncertain variable (Table 1), here indicated by the vector ϑ below:

$$\vartheta = [\,t_d, mc_{i,j}, mc_{i,i'}, L_i, P^{MS}_{i,j}, s_i, ws_i, op^{EV}_{i,j}, Q^{ST}_{i,j}\,] \tag{2}$$

To evaluate the performance of the distribution network the OPF model receives as input the location and magnitude of the available power in the power sources and demanded at the loads, which are set by the operating conditions defined by $\{FD,\Xi\}$ and ϑ. The nodal power loads L_i are directly sampled, whereas the available power in the power sources (MS and DG) depends on the uncertain variables that represent the behavior of the energy sources, the specific technical characteristics of each type of technology and the mechanical states. The available power in each type of power source considered is modeled by the functions summarized in Table 2, for a given configuration $\{FD,\Xi\}$, operating scenario ϑ and a generic node i.

Table 2: Available power functions of the power sources (PS) [25, 29, 30]

PS type j	Parameters	Available power function (kW)	
MS	-	$Pa^{MS;\vartheta}_{i,j} = \xi_{i,j} mc^{\vartheta}_{i,j} P^{MS;\vartheta}_{i,j}$	(3)
PV	T_{a_i} N_{oT_j} I_{sc_j} V_{oc_j} k_{I_j}, k_{V_j} V_{MPP_j}, I_{MPP_j}	$Pa^{PV;\vartheta}_{i,j}(s^{\vartheta}_i) = \xi_{i,j} mc^{\vartheta}_{i,j} FF_j V^{\vartheta}_{i,j} I^{\vartheta}_{i,j} \times 10^{-3}$ $T^{\vartheta}_{c_{i,j}} = Ta_i + s^{\vartheta}_i (N_{oT_j} - 20)/0.8$ $I^{\vartheta}_{i,j} = s^{\vartheta}_i (I_{sc_j} + k_{I_j}(T^{\vartheta}_{c_{i,j}} - 25))$ $V^{\vartheta}_{i,j} = V_{oc_j} + k_{V_j} T^{\vartheta}_{c_{i,j}}$ $FF_j = (V_{MPP_j} I_{MPP_j})/(V_{oc_j} I_{sc_j})$	(4)
W	ws_{ci_j} ws_{a_j} ws_{co_j} $P^W_{R_j}$	$Pa^{W;\vartheta}_{i,j}(ws^{\vartheta}_i) = \xi_{i,j} mc^{\vartheta}_{i,j} \times \begin{cases} P^W_{R_j} \frac{ws^{\vartheta}_i - ws_{ci_j}}{ws_{a_j} - ws_{ci_j}} & \text{if } ws_{ci_j} \le ws^{\vartheta}_i < ws_{a_j} \\ P^W_{R_j} & \text{if } ws_{a_j} \le ws^{\vartheta}_i \le ws_{co_j} \\ 0 & \text{otherwise} \end{cases}$	(5)
EV	$t_{op^{EV;\vartheta}_{i,j}}, P^{EV}_{R_j}$	$Pa^{EV;\vartheta}_{i,j}(op^{EV;\vartheta}_{i,j}, t) = \xi_{i,j} mc^{\vartheta}_{i,j} op^{EV;\vartheta}_{i,j} P^{EV}_{R_j} \quad \forall t \in [0, t_{op^{EV;\vartheta}_{i,j}}]$	(6)
ST	$P^{ST}_{R_j}$	$Pa^{ST;\vartheta}_{i,j}(t) = \xi_{i,j} mc^{\vartheta}_{i,j} P^{ST}_{R_j} \quad \forall t \in [0, t^{ST;\vartheta}_{R_{i,j}}]$ $t^{ST;\vartheta}_{R_{i,j}}(Q^{ST;\vartheta}_{i,j}) = Q^{ST;\vartheta}_{i,j} / P^{ST}_{R_j}$	(7)

In Table 2, $Pa^{MS;\vartheta}_{i,j}$ (kW), $\xi_{i,j}$ and $mc^{\vartheta}_{i,j}$ denote the available power, the units and the mechanical

state of the power source of type j allocated at node i. For solar photovoltaic technologies $j \in PV$, the parameter T_{a_i} (°C) is the ambient temperature at node i, N_{oT_j} (°C) is the nominal cell operation temperature, I_{sc_j} (A) is the short circuit current, V_{oc_j} (V) is the open circuit voltage, k_{V_j} (mV/°C) is the voltage temperature coefficient, k_{I_j} (mA/°C) is the current temperature coefficients and V_{MPP_j} (V) and I_{MPP_j} (A) are the voltage and current at maximum power point, respectively. For wind turbines of types $j \in W$, ws_{ci_j}, ws_{a_j} and ws_{co_j} (m/s) are the cut–in, rated and cut–out wind speeds, respectively, and $P^W_{R_j}$ (kW) is the rated power of the turbine. For electric vehicles $j \in EV$, $t_{op^{EV;\vartheta}_{i,j}}$ (h) is the time of residence in the operating state $op^{EV;\vartheta}_{i,j}$ and $P^{EV}_{R_j}$ (kW) is the rated power. For storage devices $j \in ST$, $t^{ST;\vartheta}_{R_{i,j}}$ (h) is the upper bound of the discharging time interval and $P^{ST}_{R_j}$ (kW) is the rated power.

Under the operating conditions set forth, the given configuration of the renewable DG–integrated network $\{FD, \Xi\}$ and the scenario ϑ, the OPF objective is the minimization of the operating cost associated to the generation and distribution of power, considering the revenues per kW h sold. Power flow analysis is performed by DC modeling, neglecting power losses and assuming the voltage throughout the network as constant, linearizing the classic non–linear power flow formulation by accounting solely for active power flows [31, 32]. The present formulation of the DC optimal power flow problem is:

$$\begin{aligned} \min \quad & Co^{\vartheta}(Pu, \Delta\delta) = \sum_{i \in N} \sum_{j \in PS} (Cov^{PS}_j - ep^{\vartheta}) Pu_{i,j} \\ & + \sum_{(i,i') \in FD} Cov^{FD}_{i,i'} |B_{i,i'}(\delta_i - \delta_{i'})| + (Cop + ep^{\vartheta}) \sum_{i \in N} LS_i \end{aligned} \tag{8}$$

s.t.

$$L^{\vartheta}_i - LS_i - \sum_{j \in PS} Pu_{i,j} - \sum_{i \in N} mc^{\vartheta}_{i,i'} B_{i,i'}(\delta_i - \delta_{i'}) = 0 \tag{9}$$

$$0 \leq Pu_{i,j} \leq Pa^{PS;\vartheta}_{i,j} \tag{10}$$

$$|B_{i,i'}(\delta_i - \delta_{i'})| \leq V^{NET} A^{FD}_{i,i'} \tag{11}$$

where $\forall i, i' \in N$, $j \in PS$, $(i, i') \in FD$ and the operating scenario ϑ, Co^{ϑ} (\$/h) is the operating cost of the total power supply and distribution, Cov^{PS}_j (\$/kWh) is the variable operating cost of the power source j, ep^{ϑ} (\$/kWh) is the energy price, $Pu_{i,j}$ (kW) is the used power from the source of type j at node i, $Cov^{FD}_{i,i'}$ (\$/kWh) and $B_{i,i'}$ (Ω^{-1}) are the variable operating cost and the susceptance of the feeder (i, i'), respectively, δ_i is the voltage angle at node i, Cop (\$/kWh) is the opportunity cost for kW h not supplied, V^{NET} (kV) is the nominal voltage of the network and $A_{i,i'}$ (A) is the ampacity of the feeder (i, i'). The load shedding LS_i (kW) is defined as the amount of load disconnected at node i to alleviate congestions in the feeders and/or balance the demand of power with the available power supply.

The distribution network is considered as a 'price taker' entity, assuming a correlation between the total demand of power and the energy price ep (\$/kWh). Then, the energy price is calculated from an intermediate correlation proposed by [4, 5, 28]:

$$ep(TL) = ep_h \left(-0.38 \left(\frac{TL(t_d)}{TL_h} \right)^2 + 1.38 \frac{TL(t_d)}{TL_h} \right) \tag{12}$$

where ep_h is the energy price corresponding to the highest value of total demand considered TL_h. The total demand of power $TL(t_d)$ at the hour of the day t_d is the summation of all the nodal loads $L_i(t_d)$ (Table 1).

The constraint given by Equation (9) corresponds to the power balance equation at node i, whereas Equations (10) and (11) represent the bounds of the power generation and technical limits of the feeders, respectively.

One realization of the MCS–OPF consists of the sampling of NS operating scenarios ϑ regarded as the set $\Upsilon = \{\vartheta_1, \ldots, \vartheta_h, \ldots, \vartheta_{NS}\}$ for each of which the optimal power flow problem is solved, giving in output the values of the minimum operating cost of the total power supply and distribution $Co^{\Upsilon} = \{Co^{\vartheta_1}, \ldots, Co^{\vartheta_h}, \ldots, Co^{\vartheta_{NS}}\}$.

2.2 Expected global cost ECG

The proposed renewable DG–integrated network solutions are evaluated with respect to the expected global cost ECG. The global cost CG is composed by two terms: the fixed investment and operation (maintenance) costs Ci (\$), which are prorated hourly over the life of the project th (h), and the operating costs Co^{Υ} (\$/h) that is the outcome of the MCS–OPF (Equation (8)) described in the precedent Section 2.1. Thus, the global cost function for a scenario ϑ is given by:

$$CG^{\vartheta} = Ci + Co^{\vartheta} \quad \forall \vartheta \in \Upsilon \tag{13}$$

$$Ci = \frac{1}{th} \sum_{i \in N} \sum_{j \in DG} \xi_{i,j} ci_j \tag{14}$$

where ci_j (\$) is the investment cost of the DG technology type j.

Then, the global cost $CG^{\Upsilon} = \{CG^{\vartheta_1}, \ldots, CG^{\vartheta_h}, \ldots, CG^{\vartheta_{NS}}\}$ is considered as realizations of the probability mass function of CG, and from multiple realizations the expected value ECG^{Υ} can be obtained.

3 Renewable DG selection, sizing and allocation

The aim of the proposed simulation and optimization framework is to find the optimal plan of integration of renewable DG in terms of selection, sizing and allocation of generation units from different technologies available (PV, W, EV and ST). The corresponding decision variables are contained in Ξ^{DG} of the configuration matrix Ξ defined in Equation (1).

3.1 Optimization problem formulation

Considering a network configuration (FD, Ξ) and a set of randomly generated scenarios Υ, the optimization problem is formulated as follows:

$$\min \quad ECG^{\Upsilon} \tag{15}$$

s.t.

$$\xi_{i,j} \in \mathbb{Z}^* \tag{1}$$

$$\sum_{i \in N} \sum_{j \in DG} \xi_{i,j} ci_j \leq BGT \tag{16}$$

$$\sum_{i \in N} \xi_{i,j} \leq \tau_j \tag{17}$$

$$\text{MCS} - \text{OPF}((FD, \Xi), \Upsilon) \tag{18}$$

The meaning of each constraint $\forall i, i' \in N,\ j \in PS,\ (i, i') \in FD,\ \tau_j \in \mathbb{Z}^*$ is:

- (1): the decision variable $\xi_{i,j}$ is a non–negative integer number.
- (16): the total investment and fixed operation and maintenance costs must be less than or equal to the available budget BGT.
- (17): the total number of renewable DG units of each technology j to be allocated must be less than or equal to the maximum number of units available for integration τ_j.
- (18): all Equations (8)–(11) of MCS–OPF must be satisfied.

3.2 Hierarchical clustering differential evolution (HCDE)

The complex combinatorial optimization problem of DG planning under uncertainties described above is solved by integrating DE with HCA to reduce computational efforts, whereby the evaluation of the objective function is performed by the MCS–OPF presented in Section 2.1.

DE is a population–based and parallel, direct search method, shown to be one of the most efficient

evolutionary algorithms to solve complex optimization problems [19, 21, 24]. The implementation of the original version of DE involves two main phases: initialization and evolution, summarized below for completeness of the paper [24]:

Initialization

- Set the values of parameters:
  - NP: population size
  - G_{max}: maximum number of generations
  - Coc: crossover coefficient $\in [0,1]$
  - F: differential variation amplification factor $\in [0,2]$
- Generate randomly NP individuals X (decision vectors) within the feasible space, to form the initial population $POP^0 = \{X_1^0, \ldots, X_k^0, \ldots, X_{NP}^0\}$
- Evaluate the objective function $f(X) = y$ for each individual

Evolution loop

- Set generations count index $G = 1$
- Set $POP^G = POP^0$
- While $G \leq G_{max}$ (stopping criterion)

  Trial loop

  For each individual X_k^G in POP^G, $\forall k \in \{1, \ldots, NP\}$:
  - Sample from the uniform distribution three integer indexes r_1, r_2, r_3, with $k \neq r_1 \neq r_2 \neq r_3$ and choose the corresponding three individuals $X_{r_1}^G, X_{r_2}^G, X_{r_3}^G$
  - *Mutation*: Generate a mutant individual V_k^G according to:

$$V_k^G = X_{r_1}^G + F(X_{r_2}^G - X_{r_3}^G) \tag{19}$$

  - *Crossover*: initialize a randomly generated vector U_k^G, whose dimensionality dim is the same as that of X_k^G and each coordinate $u_{k,i}^G$ follows a uniform distribution with outcome in $[0,1]$ $\forall i \in \{1, \ldots, dim\}$. In addition, generate randomly an integer index $ri \in \{1, \ldots, dim\}$ from a uniform distribution to ensure that at least one coordinate from V_k^G is exchanged to form a trial individual XT_k^G, whose coordinates are defined as follows:

$$xt_{k,i}^G = \begin{cases} v_{k,i}^G & \text{if } u_{k,i}^G \leq Cco \text{ or } i = ri \\ x_{k,i}^G & \text{if } u_{k,i}^G > Cco \text{ and } i \neq ri \end{cases} \tag{20}$$

  - *Selection*: evaluate the objective function for the trial individual $f(XT_k^G)$; if $f(XT_k^G) < f(X_k^G)$ (minimization), then XT_k^G replaces X_k^G in the population POP^G, otherwise X_k^G is retained

- Set $G = G + 1$
- Once the stopping criterion is reached, sort the individuals in $POP^{G_{max}}$ in descending order according to their values of the objective function and return $X^{G_{max}}$

The original version of DE keeps the population size NP constant, making the computational performance dependent mainly on the number of objective function evaluations carried out during the evolution phase of the algorithm. Then, the integration of HCA into DE is aimed at the reduction of the number of individuals that enter the evolution loop in each generation so as to decrease the number of objective function evaluations.

HCA links individuals or groups of individuals which are similar with respect to a specific property, translated into a metric of distance, obtaining a hierarchical structure. In practice, we use an agglomerative procedure which in $sp = NP - 1$ steps fuses the closest pair or individuals or groups of individuals through a linkage function, e.g. single linkage (nearest neighbor distance), complete linkage (furthest neighbor), average linkage, among others, until the complete hierarchical structure is built. The base hierarchical clustering algorithm used in this study can be expressed as follows [26]:

Step 0: Given a population POP $POP = \{X_1, \ldots, X_k, \ldots, X_{NP}\}$, form the set of singleton groups $O = \{O_p = \{X_k\}\}$, $\forall p = k \in \{1, \ldots, NP\}$ and calculate the linkage distances between all the NP groups using the average as linkage function and the Euclidean distance as metric:

$$D^1 = \begin{bmatrix} 0 & d^1_{1,2} & \cdots & d^1_{1,q} & \cdots & d^1_{1,NP} \\ & \ddots & \ddots & \vdots & \ddots & \vdots \\ & & 0 & d^1_{p,q} & \cdots & d^1_{p,NP} \\ & & & 0 & \ddots & \vdots \\ & & & & \ddots & d^1_{NP-1,NP} \\ & & & & & 0 \end{bmatrix} \quad d^1_{p,q} = \frac{\sum_{X_{kp} \in O_p} \sum_{X_{kq} \in O_q} \sqrt{(X_{kp} - X_{kq})^2}}{|O_p||O_q|} \tag{21}$$

$$\forall p, q \in \{1, \ldots, NP\}, kp, kq \in \{1, \ldots, NP\}$$

where $d^1_{p,q}$ is the average of the Euclidean distances between all the individuals X_k belonging to the groups O_p and O_q, respectively.

Step 1: Fuse the first pair of groups $O_{p'}$ and $O_{1'}$, for which $d^1_{p',q'}$ is the minimum distance $\min(D^1)$ and form a new group $O_{NP+1} = \{O_{p'} \cup O_{q'}\}$. Update the set of groups O replacing $O_{p'}$ and $O_{q'}$ by O_{NP+1}, and calculate the linkage distances D^2 between all the $NP - 1$ groups in O using (21).

Step 2: Fuse the second pair of groups $O_{p'}$ and $O_{q'}$ for which $d^2_{p',q'}$ is the minimum distance $\min(D^2)$, and form a new group $O_{NP+2} = \{O_{p'} \cup O_{q'}\}$. As in the preceding step, update the set of groups O and calculate the linkage distances D^3 between all the $NP - 2$ groups in O using (21).

$\vdots$

Step $NP-1$: Fuse the last pair of groups with linkage distance $d^{NP-1}_{p',q'}$, forming the last group $O_{2NP-1} = \{O_{p'} \cup O_{q'}\}$ that contains all the individuals X.

The outcoming hierarchical (or tree) structure can be reported as a sorted table containing the $NP-1$ linkage distances relative to each pairing action of individuals/groups and be graphically illustrated as a dendrogram. Table 3 and Figure 2 present, respectively, the resultant linkage distances and dendrogram obtained from an example set of $NP = 8$ two–dimensional individuals X using the above introduced HCA algorithm.

Table 3: Example hierarchical structure outcome

Step sp	Group	Groups linked	Linkage distance
1	O_9	$\{O_2 \cup O_6\} = \{\{X_2\} \cup \{X_6\}\}$	$d^1_{2,6}$
2	O_{10}	$\{O_3 \cup O_4\} = \{\{X_3\} \cup \{X_4\}\}$	$d^2_{3,4}$
3	O_{11}	$\{O_1 \cup O_7\} = \{\{X_1\} \cup \{X_7\}\}$	$d^3_{1,7}$
4	O_{12}	$\{O_5 \cup O_8\} = \{\{X_5\} \cup \{X_8\}\}$	$d^4_{5,8}$
5	O_{13}	$\{O_9 \cup O_{11}\} = \{\{X_2, X_6\} \cup \{X_1, X_7\}\}$	$d^5_{9,11}$
6	O_{14}	$\{O_{10} \cup O_{12}\} = \{\{X_3, X_4\} \cup \{X_5, X_8\}\}$	$d^6_{10,12}$
7	O_{15}	$\{O_{13} \cup O_{14}\} = \{\{X_1, X_2, X_6, X_7\} \cup \{X_3, X_4, X_5, X_8\}\}$	$d^7_{13,14}$

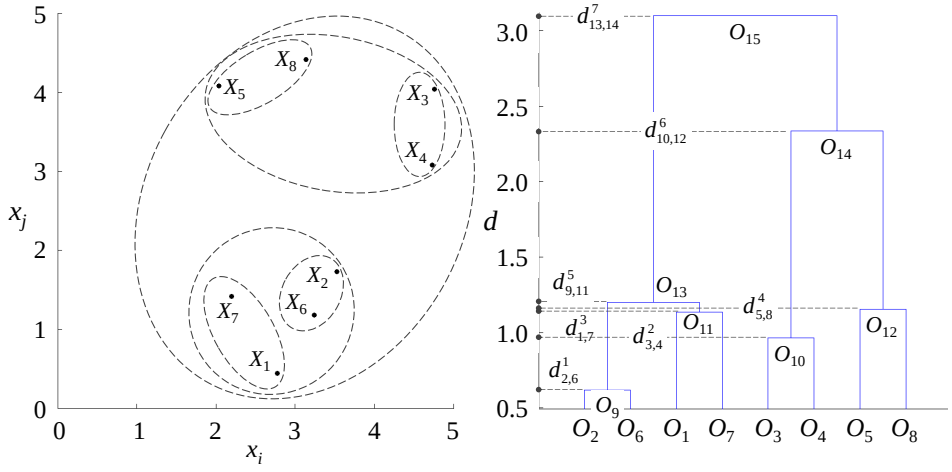

Figure 2: Example dendrogram for average linkage HCA

As stated above, HCA builds the hierarchical structure through a linkage function introducing in each grouping action a larger or smaller degree of distortion with respect to the original distances between (ungrouped) individuals. The measurement of this distortion is important and the cophenetic correlation coefficient (CCC)is introduced to evaluate how representative is the

hierarchical structure proposed by the HCA. The CCC can be obtained from Equations (22) and (23) below [26].

$$CCC = \frac{\sum_{p<q}(d^1_{p,q} - \bar{D^1})(h_{p,q} - \bar{H})}{\sqrt{\sum_{p<q}(d^1_{p,q} - \bar{D^1})^2 \sum_{p<q}(h_{p,q} - \bar{H})^2}} \quad \forall p,q \in \{1,\ldots,NP\} \tag{22}$$

$$H = \begin{bmatrix} 0 & h_{1,2} & \cdots & h_{1,q} & \cdots & h_{1,NP} \\ & \ddots & \ddots & \vdots & \ddots & \vdots \\ & & 0 & h_{p,q} & \cdots & h_{p,NP} \\ & & & 0 & \ddots & \vdots \\ & & & & \ddots & h_{NP-1,NP} \\ & & & & & 0 \end{bmatrix} \quad h_{p,q} = d^{sp^*}_{p',q'}$$

$$sp^* = \{\min(sp) : X_p \wedge X_q \in O_{NP+sp}\}, \tag{23}$$

$$\forall p,q \in \{1,\ldots,NP\}, p',q' \in \{1,\ldots,2NP-1\}, sp \in \{1,\ldots,NP-1\}$$

where $\bar{D^1}$ is the average of the original Euclidean distances $d^1_{p,q}$ between all the individuals, $h_{p,q}$ is the linkage distance $d^{sp^*}_{p',q'}$ where the pair of individuals X_p and X_q become members of the same group and $\bar{H}$ is the average of the resultant linkage distances $h_{p,q}$ between all the individuals.

Recalling that the aim of nesting HCA into DE is to increase the computational performance by decreasing the number of individuals to be evaluated in each generation G, the presetting of a threshold CCC_{th} for the CCC value allows defining the level of representativeness required to the hierarchical structure proposed. If the CCC^G obtained from applying HCA over the corresponding population POP^G is higher than or equal to the threshold CCC_{th}, the built hierarchical structure is considered an acceptable representation of the original distances amongst the individuals and the selection of a particular partition of the sets of groups can be performed, i.e., the determination of a specific number of clusters. Conversely, if CCC^G is less than CCC_{th}, the hierarchical structure is considered not representative enough since it introduces unacceptable distortion that may affect the global searching process in the HCDE.

Whether the hierarchical structure is accepted, the clustering process itself takes place. As before stated, the HCA outcome linkage distances $d^{sp}_{p',q'}$ define each level (height) at which a pairing action takes place. If the hierarchical structure is 'cut off' at a specific linkage distance d_{CO}, all the groups that are formed below the level d_{CO} become independent clusters. In each generation G of HCDE, a d_{CO} relative to the HCA outcome linkage distances for the corresponding POP^G, is determined from a preset cutoff level coefficient p_{co} of the linkage distances between the minimum $d^{sp}_{p',q'}$ that correspond to the first pairing action and the distance to form at least four clusters needed to perform the mutation process in the HCDE. Thus, d_{CO} can be obtained from Equations (24) and (25). Figure 3 shows the cutoff distance representation for the example aforementioned, for which

the formed clusters are $\{O_2, O_6\}$, $\{O_1\}$, $\{O_7\}$, $\{O_3, O_4\}$, $\{O_5\}$ and $\{O_8\}$.

$$d_{CO} = d_{\min} + p_{co}(d_{NC=4} - d_{\min}) \tag{24}$$

$$d_{NC=4} = d_{(1-\frac{4}{NP})\%\text{ile}} \tag{25}$$

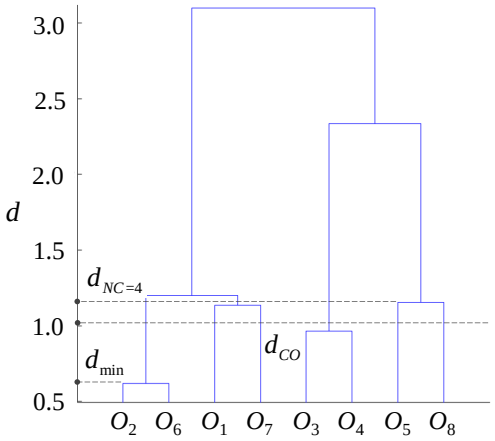

Figure 3: Example of cutoff distance calculation

The integration of HCA into DE and the definition of the parameters CCC_{th} and p_{co} allow HCDE adaptation at each generation, i.e., deciding whether to perform HCA and determining the clusters to be taken. Then, the individuals closest to the centroids of the formed clusters are considered as the representatives of the group which they belong to and are taken in a reduced population that enters the evolution phase of the HCDE. The proposed HCDE algorithm is summarized schematically in the flowchart of Figure 4.

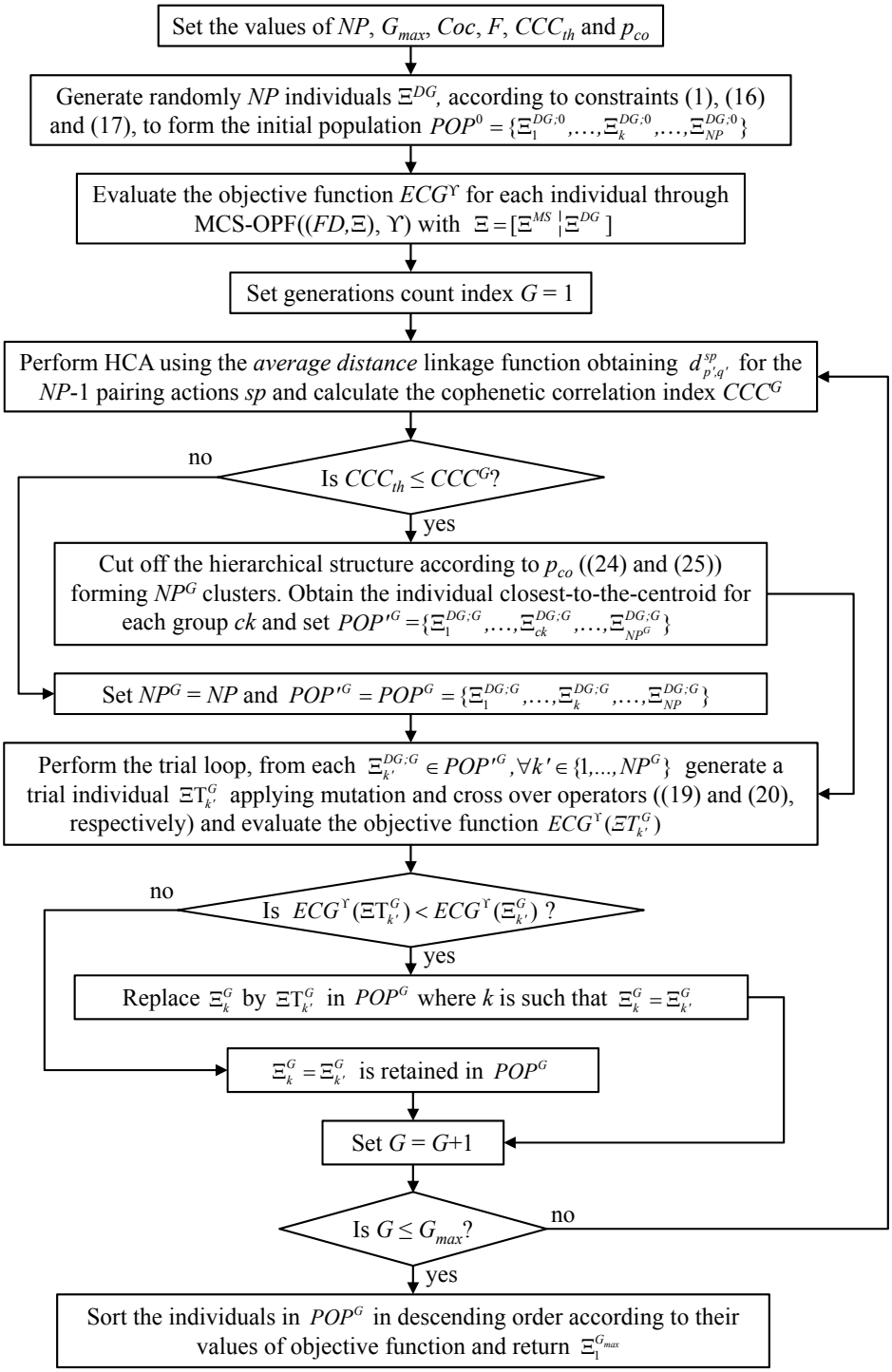

Figure 4: Flowchart of the framework

4 Case study

We consider a modification of the IEEE 13 nodes test feeder distribution network [27] with the original spatial structure but neglecting the feeders of length zero, the regulator, capacitor and switch. The resulting network has 11 nodes and presents the relevant characteristics of interest for the analysis, e.g. the presence of a main power supply spot and comparatively low and high spot, and distributed load values [33].

4.1 Distribution network description

The distribution network presents a radial structure of $n = 11$ nodes as shown in Figure 5. The nominal voltage V^{NET} is 4.16 (kV), kept constant for the resolution of the DC optimal power flow problem.

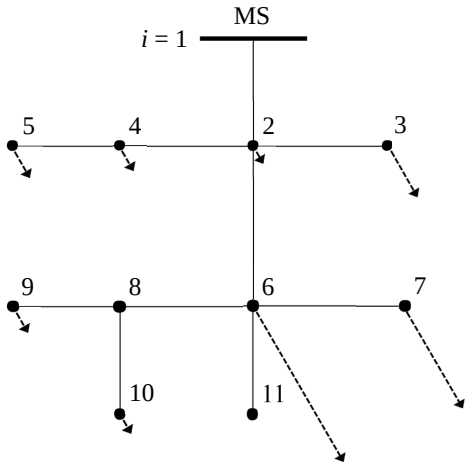

Figure 5: Radial 11–nodes distribution network

Table 4 contains the technical characteristics of the different types of feeders considered: specifically, the indexes of the pairs of nodes (i, i') that they connect, their length l, reactance X^{FD}, ampacity A^{FD} and failure and repair rates.

The nodal power demands are built from the load data given in [27] and reported in Figure 6 as daily profiles, normally distributed on each hour t_d with mean μ^L and standard deviation σ^L [29,34].

The technical parameters, failure and repair rates and costs of the MS and the four different types of DG technologies (PV, W, EV and ST) available to be integrated into the distribution network are given in Table 5. For the present case study, the distribution region is such that the solar irradiation and wind speed conditions are assumed uniform in the whole network, i.e., the values of the parameters of the corresponding Beta and Rayleigh distributions are assumed constant in the whole network.

Table 4: Feeders characteristic and technical data [11, 27, 28]

Type	Node i	Node i'	l (km)	X^{FD} (Ω/km)	A^{FD} (A)	λ^F (1/h)	λ^R (1/h)	Cov^{FD} ($/kWh)
T1	1	2	0.610	0.371	730	3.333e−04	0.198	1.970e−02
T2	2	3	0.152	0.472	340	4.050e−04	0.162	9.173e−03
T3	2	4	0.152	0.555	230	3.552e−04	0.185	6.205e−03
T1	2	6	0.610	0.371	730	3.333e−04	0.198	6.205e−03
T3	4	5	0.091	0.555	230	3.552e−04	0.185	6.205e−03
T6	6	7	0.152	0.252	329	4.048e−04	0.164	8.904e−03
T4	6	8	0.091	0.555	230	3.552e−04	0.185	1.970e−02
T1	6	11	0.305	0.371	730	3.333e−04	0.198	1.970e−02
T5	8	9	0.091	0.555	230	3.552e−04	0.185	9.173e−03
T7	8	10	0.244	0.318	175	3.552e−04	0.185	6.205e−03

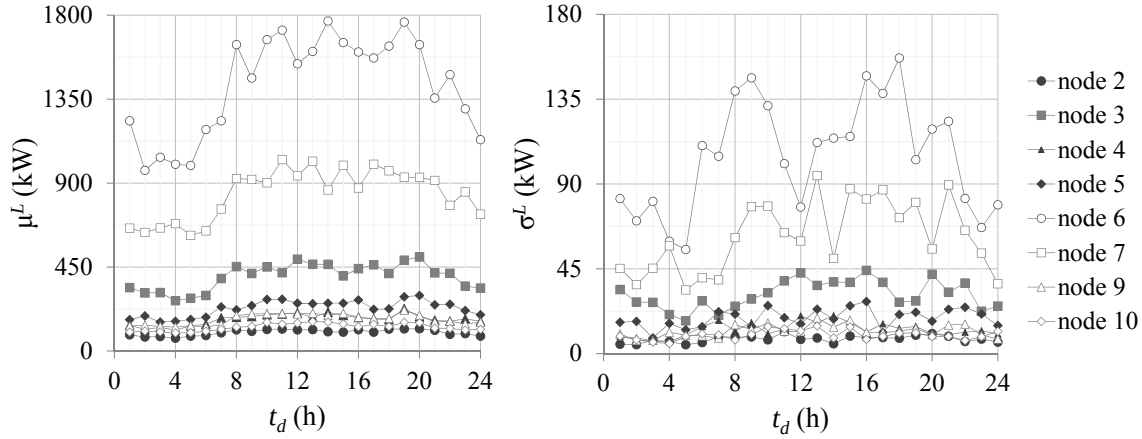

Figure 6: Mean and standard deviation values of normally distributed nodal power demand daily profile

The hourly per day operating state probability profiles of the EV are presented in Figure 7: p^0, p^- and p^+ correspond to the profiles of disconnected, charging and discharging states, respectively.

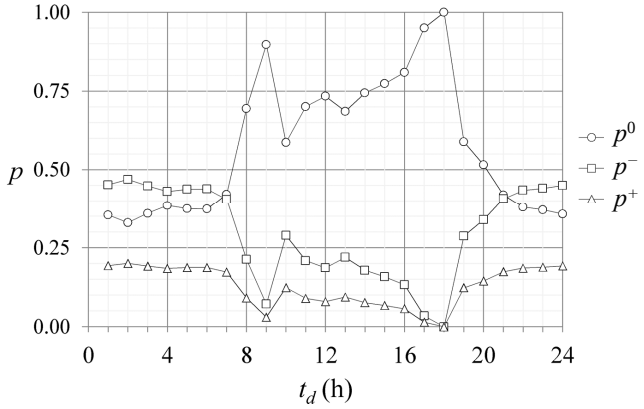

Figure 7: Hourly per day probability data of EV operating states

Table 5: Power sources parameters and technical data [13, 17, 28, 29, 35–37]

Type j	Technical parameters	Distributions parameters, failure and repair rates	Costs
MS	$P_{cap}^{MS} = 4250$ (kW)	$\mu^{MS} = 4000$ (kW) $\sigma^{MS} = 125$ (kW) $\lambda^{F} = 4.00\text{e}{-}04$ (1/h) $\lambda^{R} = 1.30\text{e}{-}02$ (1/h)	$Cov = 0.145$ (\$/kWh)
PV	$T_a = 30.00$ (°C) $N_{oT} = 43.00$ (°C) $I_{sc} = 1.80$ (A) $V_{oc} = 55.50$ (V) $k_I = 1.40$ (mA/°C) $k_V = 194.00$ (mV/°C) $V_{MPP} = 38.00$ (V) $I_{MPP_j} = 1.32$ (A)	$\alpha^{PV} = 0.26$ $\beta^{PV} = 0.73$ $\lambda^{F} = 5.00\text{e}{-}04$ (1/h) $\lambda^{R} = 1.30\text{e}{-}02$ (1/h)	$Ci = 48$ (\$) $Cov = 3.76\text{e}{-}05$ (\$/kWh)
W	$ws_{ci} = 3.80$ (m/s) $ws_a = 9.50$ (m/s) $ws_{co} = 23.80$ (m/s) $P_R^W = 50.00$ (kW)	$\sigma^{W} = 7.96$ $\lambda^{F} = 6.00\text{e}{-}04$ (1/h) $\lambda^{R} = 1.30\text{e}{-}02$ (1/h)	$Ci = 113,750$ (\$) $Cov = 3.90\text{e}{-}02$ (\$/kWh)
EV	$P_R^{EV} = 6.30$ (kW)	$\lambda^{F} = 2.00\text{e}{-}04$ (1/h) $\lambda^{R} = 9.70\text{e}{-}02$ (1/h)	$Ci = 17,000$ (\$) $Cov = 2.20\text{e}{-}02$ (\$/kWh)
ST	$P_R^{ST} = 0.28$ (kW/kg) $SE^{ST} = 0.04$ (kJ/kg)	$\lambda^{F} = 3.00\text{e}{-}04$ (1/h) $\lambda^{R} = 7.30\text{e}{-}02$ (1/h)	$Ci = 135.15$ (\$) $Cov = 4.62\text{e}{-}05$ (\$/kWh)

Coherently with constraints (16) and (17), the budget is set to $BGT = 4,500,000$ (\$) and the limit of units of the different DG technologies available to be purchased is $\tau = [20000, 8, 250, 10000]$. The maximum value of the energy price is $ep_h = 0.12$ (\$/kWh) [5][5] and the highest value of total demand TL_h is set to 4800 (kW). The opportunity cost for kW h not supplied Cop is considered as twice of the maximum energy price.

A total of $NS = 500$ random scenarios are simulated by the MCS–OPF with time step $ts = 1$ (h), over a horizon of analysis of 10 years ($th = 87,600$ (h)), in which the investment and fixed costs are pro–rated hourly.

The DE iterations are set to perform $G_{max} = 500$ generations over five different cases of population $NP \in \{10, 20, 30, 40, 50\}$. The differential variation amplification factor F is 1 to maintain the integer–valued definition of the individuals after the mutation, whereas the crossover coefficient Coc is 0.1.

HCDE runs are performed under the same conditions set for DE (G_{max}, F and Coc), but for the population size NP of 50 individuals. A sensitivity analysis is performed over the HCA control parameters, namely the cophenetic correlation coefficient CCC_{th} and linkage distances cutoff level coefficient p_{co}, for all the nine possible pairs (CCC_{th}, p_{co}) with $CCC_{th} \in \{0.6, 0.7, 0.8\}$ and $p_{co} \in \{0.25, 0.50, 0.75\}$. Finally, for each of the five DE and nine HCDE settings, twenty realizations are carried out.

4.2 Results and discussion

The results of the DE MCS–OPF for the different population sizes $NP \in \{10, 20, 30, 40, 50\}$ are shown in Figure 8. The 50th percentile (%ile) or median values of the minimum global costs ECG_{min}, obtained from each experiment with fixed values of NP, are presented as functions of the respective numbers of objective function evaluations NFE; the error bars represent the 15th and 85th %iles.

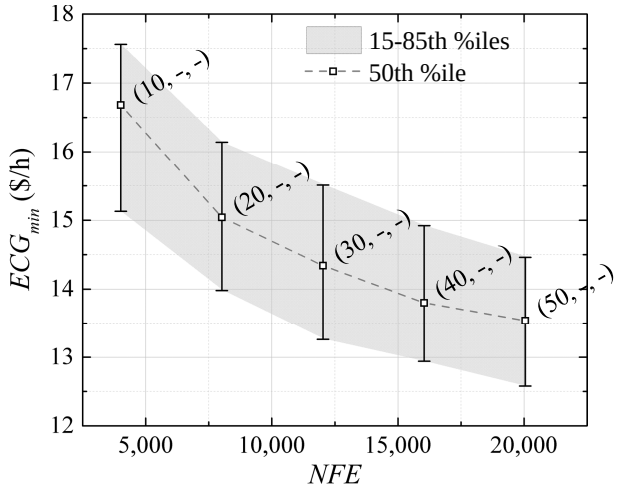

Figure 8: ECG_{min} vs NFE for $NP \in \{10, 20, 30, 40, 50\}$ set in DE

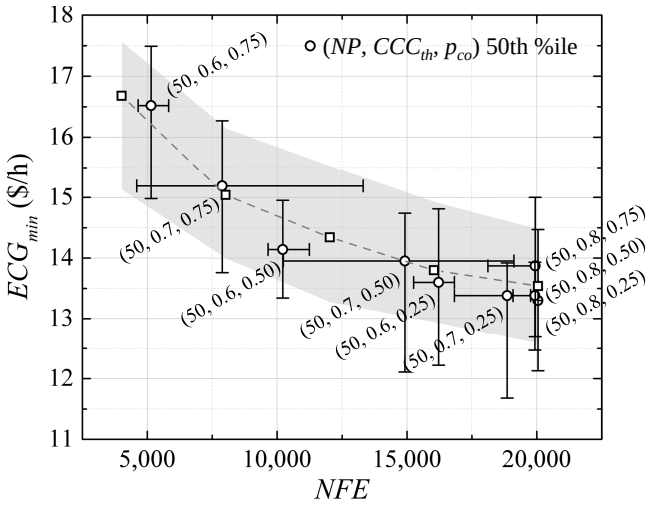

Figure 9: ECG_{min} vs NFE for each (NP, CCC_{th}, p_{co}) set in HCDE

As expected, for the same number of generations set in the DE MCS–OPF, the larger the population size considered the lower the values of ECG_{min} obtained (better 'quality' of the minimum). Additionally, we can observe marked tendencies in the reduction of both median and 15–85th %iles values of ECG_{min} for increasing NFE. Performing a curve fitting over these values, we get: $ECG_{\text{min;50th\%ile}} = 49.07NFE^{-0.13}$, $ECG_{\text{min;15th\%ile}} = 49.07NFE^{-0.115}$ and $ECG_{\text{min;85th\%ile}} = 49.07NFE^{-0.118}$, with the respective coefficients of determination $R^2_{\text{50th\%ile}} = 0.994$, $R^2_{\text{15th\%ile}} = 0.998$ and $R^2_{\text{85th\%ile}} = 0.998$.

The fact that the difference between the values of the 15–85th %iles is constant indicates that the dispersion in the $ECG_{\min}(NFE)$ does not depend on NP and can suggest that the global searching performed by the DE is performed homogeneously in the feasible space that contains multiple local minima.

Figure 9 reports the median $ECG_{\min}$ values corresponding to the HCDE MCS–OPF realizations superposed to the distribution of the median $ECG_{\min}$ and 15–85th %iles values of the base DE experiments represented by the square markers and shaded area, respectively. The vertical and horizontal error bars account for the 15–85th %iles of the outcome $ECG_{\min}$ and NFE values.

Focusing on CCC_{th}, it can be noticed that for the two extreme cases, $CCC_{th} = 0.6$ and 0.8, the dispersion of the number of objective function evaluations is relatively small. On the contrary, the cases with a $CCC_{th} = 0.7$ present high variability. This can be explained by the behavior of the CCC along each generation G in the evolution loop. Figure 10 shows the median, 15th and 85th %iles CCC values as a function of generation G derived from all HCDE MCS–OPF realizations. On the one hand, recalling that CCC_{th} is used to control whether it is convenient to perform HCA, the small NFE dispersion in the case with $CCC_{th} = 0.6$ is because clustering is practically been applied in all generations ($CCC_{th} \leq CCC^G$), thus disabling any effect generated by passing from populations with original size NP to reduced populations with $NP^G \leq NP$ and vice versa. On the other hand, the effect is also being avoided in the case $CCC_{th} = 0.8$ by not applying clustering. Indeed, in Figure 10 it can be observed that after the generation 50 it is unlikely that by performing HCA the proposed hierarchical grouping structures represent well enough the population.

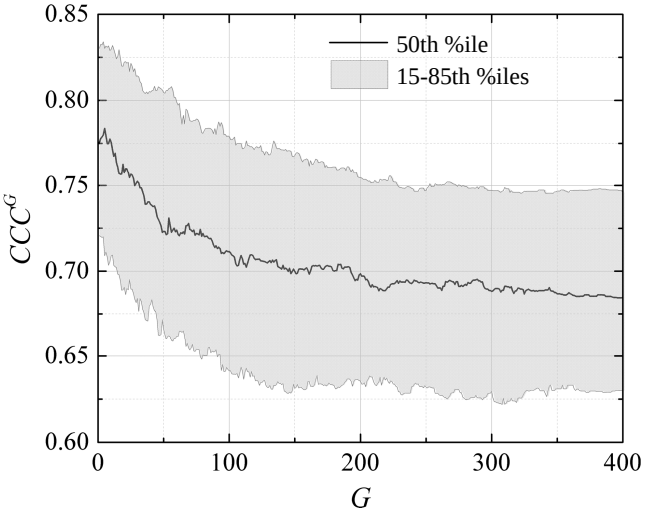

Figure 10: CCC behavior per generation G

Differently, the cases for which $CCC_{th} = 0.7$ present high dispersion in the NFE since the median values of CCC^G move in the neighborhood of the threshold throughout the major part of the evolution loop in the HCDE. Moreover, in general terms, the values of CCC^G 15–85th %iles

maintain certain symmetry with respect to the median, i.e., performing or not HCA are equally likely events, producing high fluctuations in the number of individuals considered as population and, therefore, affecting in the same way the NFE.

The above mentioned insights are noticeable also in Figure 11, which shows the empirical probability density functions (pdfs) of the population size NP^G per generation for each (NP, CCC_{th}, p_{co}) set in HCDE. Indeed, the average probabilities of performing HCA throughout the evolution cycle for the different values of $CCC_{th} = 0.6$, 0.7 and 0.8 are 0.980, 0.540 and 0.078, respectively.

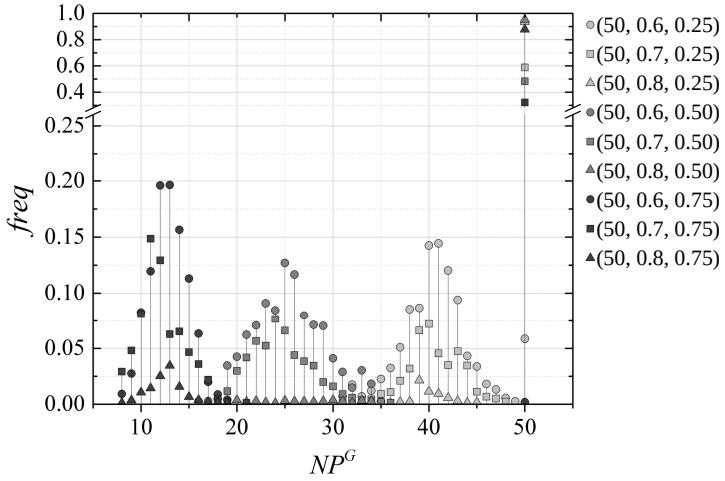

Figure 11: Empirical NP^G pdf for each (NP, CCC_{th}, p_{co}) set in HCDE

Regarding the cutoff level coefficient of the linkage distances p_{co}, in Figure 11 it is possible to identify the three peaks of reduction in the population size, confirming the role of this control parameter in defining the level at which the hierarchical structures proposed are 'cut off' when the HCA takes place. In fact, lower values of p_{co} imply smaller reduction in the population size because of the higher demand of proximity between individuals or groups of individuals. In the opposite side, higher values of p_{co} allow forming clusters from individuals or groups which are relatively less similar.

From the results obtained for all the different DE and HCDE settings, we look for six representative cases for the analysis (Figure 9). From the DE runs, we select the settings with extreme and middle population size $NP \in \{10, 30, 50\}$, whereas from HCDE we choose the cases (NP, CCC_{th}, p_{co}) set as (50, 0.6, 0.25), (50, 0.6, 0.50), (50, 0.7, 0.50) and (50, 0.7, 0.75). The former (50, 0.6, 0.25) and (50, 0.6, 0.50) cases present significant reductions in the number of NFE, with small dispersion and loss of quality of the minimum ECG obtained, compared to the results obtained by diminishing directly the fixed NP in DE from 50 to 10. Similarly, the cases (50, 0.7, 0.50) and (50, 0.7, 0.75) may lead to considerable reductions in NFE, with acceptable losses of ECG_{min}, but subject to a high degree of variability that compromises the performance.

As for computational times, running on an Intel® Core™i7–3740QM (PC) 2.70 GHz without performing parallel computing, the average time to evaluate the objective function is 4.592 (s) for the $NS = 500$ scenarios in the MCS–OPF; for a fixed population of $NP = 50$ and its corresponding $NFE = 20,050$, the total time for a single run is on average 25.574 (h). Taking into account this, under commonly used hardware configurations, the reductions in computational time that can be achieved by using HCDE with (50, 0.6, 0.25) and (50, 0.6, 0.50) settings are 19% and 49% for the median, 23% and 51% for the 15th %ile, and 16% and 43% for the 85th %ile, respectively.

The integration of HCA into the DE algorithm introduces a significant time complexity, conditioning the reductions of computational efforts that can be obtained by applying the proposed HCDE MCS–OPF framework. Indeed, if performing HCA along all generations of DE and running the MCS–OPF on an eventually reduced population (depending on CCC_{th} and p_{co}) is computationally heavier than running the MCS–OPF over the complete population, the effects of the framework can be negligible or even negative. It is possible to formulate the condition to obtain reductions in the computational efforts by the proposed HCDE MCS–OPF framework, from the asymptotic time complexities of the main algorithms that compose it. Table 6 reports the independent asymptotic time complexities as functions of the generic size m of the input to each algorithm and of the parameters that define the dimensionality of the HCDE MCS–OPF framework [26, 38].

Table 6: Asymptotic time complexity of the algorithms

Algorithm			
PDIST	HC	MCS	OPF
Time complexity T			
$O(dm^2)$[a]	$O(m^2 \log(m))$	$O(m)$	$O(\text{size}(A))$[b]
$O(nps \times NP^2)$	$O(NP^2 \log(NP))$	$O(NS \times nps)$	$O(NS \times nps^2)$

[a] Pairwise distance PDIST between all m vectors of size d.
[b] The matrix A comes from the canonical form $Ax \leq b$ of the linear programming of the DC OPF problem approximation.

where nps represents the size of the DG–integrated network, i.e., the number of nodes n times the number of all the technologies of power generation available ps, NP is the size of the complete population and NS is the number of scenarios in the MCS–OPF.

Comparing the asymptotic time complexities of the algorithms involved in the realization of the proposed framework with and without integrating HCA, the following inequalities must be fulfilled in order to obtain a reduction in the computational time by HCDE:

$$\begin{gathered} T^{\text{PDIST}}(nps, NP) + T^{\text{HC}}(NP) + E[NP^G] \times T^{\text{MCS-OPF}}(NS, nps) < NP \times T^{\text{MCS-OPF}}(NS, nps) \\ \Downarrow \\ nps \times NP^2 + NP^2 \log(NP) + E[NP^G] \times NS \times nps^2 < NP \times NS \times nps^2 \\ \Downarrow \\ \kappa = \frac{NP}{NP \times nps} + \frac{NP \log(NP)}{NP \times nps^2} + \epsilon < 1 \quad \forall n, ps, NP, NS \in \mathbb{Z}^*, \epsilon = \frac{E[NP^G]}{NP} \in (0, 1] \end{gathered} \tag{26}$$

where ϵ is the expected ratio of the population NP^G evaluated along all generations G of DE to the total population NP and κ is the ratio of the asymptotic time complexities of HCDE to DE.

From Equation (26), we can observe that the contribution of the terms related with the complexity of MCS–OPF, dependent on NS and nps, is considerably large for the fulfillment of the inequality conditions. In fact, when using DE, it is commonly accepted to set a size of the population NP not greater than ten times the size of the decision variables, in this case, $10nps$ [24], making the first two terms of κ strongly dependent on the number of scenarios NS. Moreover, given the complexity of the general problem, higher values of NS lead to a better approximation of the objective function via MCS–OPF, i.e., the more likely is to fulfill the condition and the greater can be the reduction of computation time. However, the value of ϵ depends on the probability of performing clustering in each generation and at what level, controlled by CCC_{th} and p_{co} respectively. In some cases, ϵ can be close to 1 (as we inferred from Figure 11) implying negligible benefits. Table 7 shows the values of the ratio κ for each (NP, CCC_{th}, p_{co}) set in HCDE considering the dimensionality of the present case study defined by the values of the parameters $nps = 55$, $NS = 500$, $NP = 50$. The value of $1-\kappa$ can be interpreted as the expected asymptotic relative time reduction achieved by performing HCDE.

Table 7: Ratio κ for each (NP, CCC_{th}, p_{co})

(NP, CCC_{th}, p_{co})	$\frac{NP}{NP \times nps}$	$\frac{NP\log(NP)}{NP \times nps^2}$	$\epsilon = \frac{E[NP^G]}{NP}$	κ	$1-\kappa$
(50, 0.6, 0.25)			0.817	0.819	0.181
(50, 0.7, 0.25)			0.921	0.923	0.077
(50, 0.8, 0.25)			0.987	0.989	0.011
(50, 0.6, 0.50)			0.510	0.512	0.488
(50, 0.7, 0.50)	1.818e−03	3.418e−05	0.738	0.740	0.260
(50, 0.8, 0.50)			0.978	0.979	0.021
(50, 0.6, 0.75)			0.259	0.261	0.739
(50, 0.7, 0.75)			0.487	0.488	0.512
(50, 0.8, 0.75)			0.909	0.911	0.089

Figure 12 shows the convergence curves for the DE and HCDE cases selected, for the twenty runs performed for each (NP, CCC_{th}, p_{co}) setting: no significant differences can be found among the convergence curves except for the expected behavior of converging to lower values of $ECG_{\min}$ for settings which imply a larger population size.

Figure 13 shows the average total DG power allocated in the distribution network and the corresponding investment costs of the DE and HCDE MCS–OPF cases selected, choosing the corresponding optimal DG–integrated plans as the decision matrices Ξ^{DG} for which their $ECG_{\min}$ values are the closest to the median $ECG_{\min}$ value obtained for the twenty runs of each (NP, CCC_{th}, p_{co}) setting. It can be pointed out that in all the cases, the contribution of EV is practically negligible if compared with the other technologies. This is due to a combination of two facts: the probability

that the EV is in a discharging state is much lower than that of being in the other two possible operating states, charging and disconnected (see Figure 7) and when EV is charging, the effects are opposite to those desired, i.e., it is acting as loads.

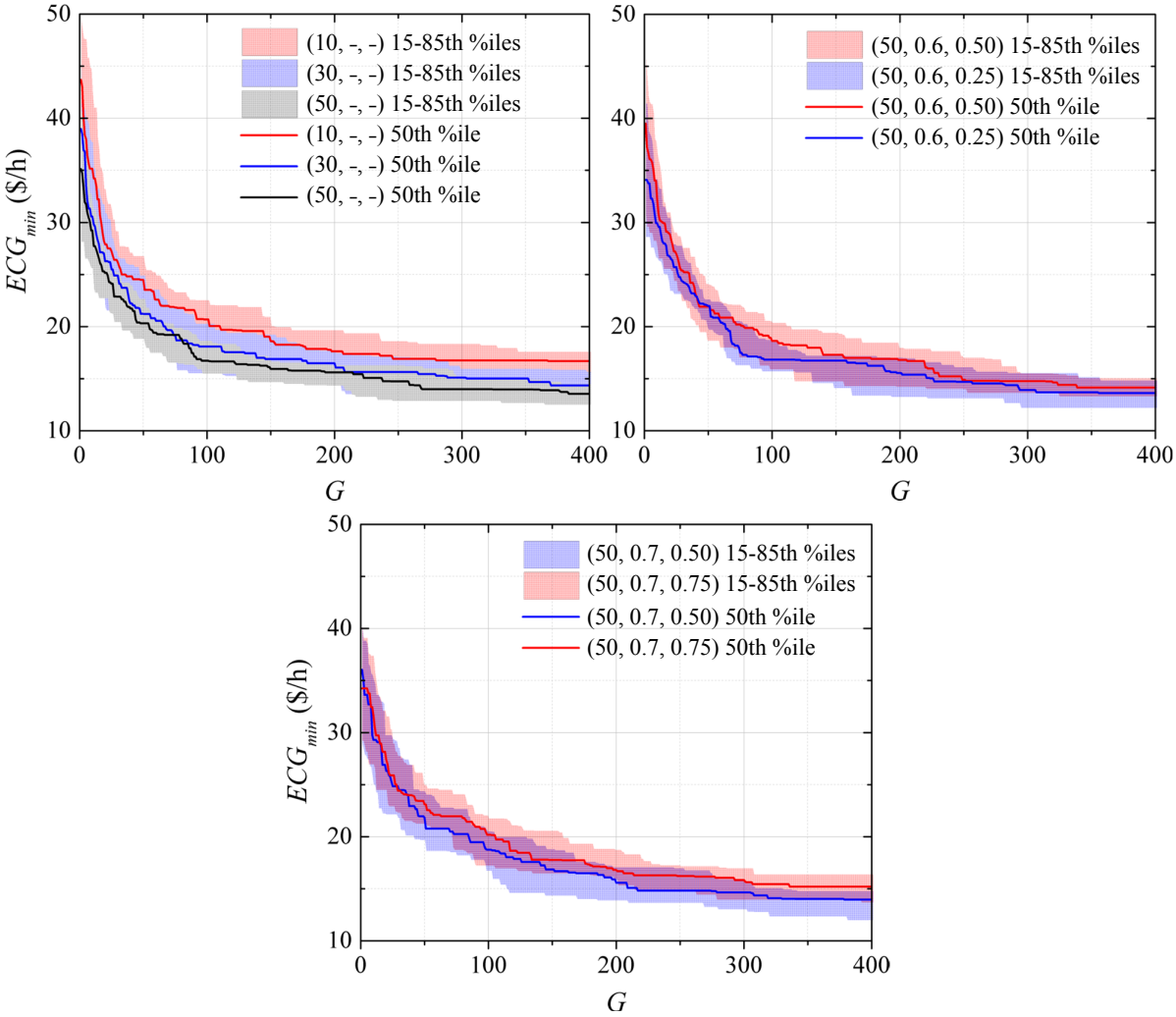

Figure 12: Convergence curves for representative (NP, CCC_{th}, p_{co}) settings

In all generality, both the investment cost Ci and the average power installed by DG is comparable in all the cases, except for the setting (50, 0.7, 0.75) for which the level of clustering determined by $p_{co} = 0.75$, that translates into higher reductions of the population size, may lead to less similar local minima than the other settings.

The average total renewable DG power allocated per node is summarized in Figure 14. Even though all the ECG optimal decision matrices Ξ^{DG} show differences, the tendency is to install localized sources of renewable DG power between two identifiable portions of the distribution network, up and downstream the feeder (2, 6) (Figure 5), giving preference to the second portion which presents higher and non–stream homogeneous nodal load profiles.

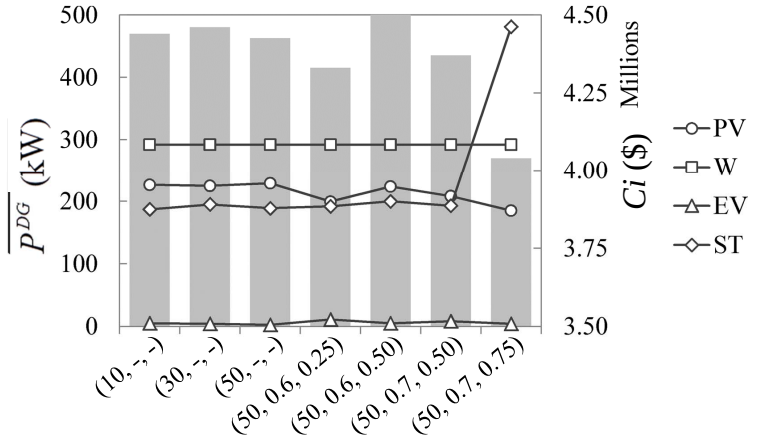

Figure 13: Average total DG power allocated and investment cost for representative (NP, CCC_{th}, p_{co}) settings

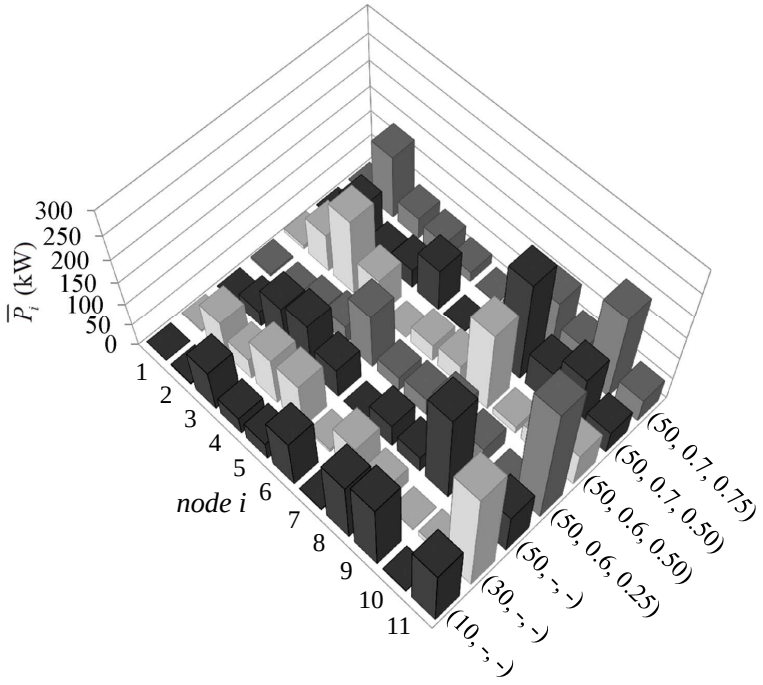

Figure 14: Nodal average total DG power for representative (NP, CCC_{th}, p_{co}) settings

5 Conclusions

In a previous paper, we have presented a simulation and optimization framework for the planning of integration of renewable generation into a distribution network. The optimization is considered with respect the objective of minimizing the expected global cost of the system. The inherent uncertain behavior of renewable energy sources, variability in the main power supply and loads, as well as the possibility of failures of network components are included in a Monte Carlo simulation, which samples realizations of the uncertain operational scenarios for the optimal power flow.

The framework is quite general and complete in the characteristics of the realistic system scenarios considered. However, this is at the expenses of the computational time required for the overall optimization.

In this respect, in the present paper we have addressed the problem of computational efficiency in the resolution of the renewable DG planning optimization problem. We have done so by an original introduction of a controlled clustering strategy, with, the main original contributions being:

- The integration of differential evolution and hierarchical clustering analysis for grouping similar individuals from a given population and selecting representatives to be evaluated for each group, thus reducing the number of objective function evaluations during the optimization.
- The introduction of two control parameters, namely the cophenetic correlation coefficient and a cutoff level coefficient of the linkage distances, for allowing controlled adaptation during the search process and decision on whether or not to perform clustering and at which level of the hierarchical structure built.

A case study has been analyzed derived from the IEEE 13 nodes test feeder. The results obtained show the capability of the framework to identify optimal plans of renewable DG integration. The sensitivity analysis over the control parameters of the hierarchical clustering shows that the efficiency is improved with cophenetic correlation thresholds that allow the clustering in almost all generations along the differential evolution, setting the level of clustering to no more than the fifty percent of the feasible linkage distances range in the hierarchical structure proposed. Indeed, this is shown to lead to acceptable reductions in the number of objective function evaluations, with small dispersion and loss of quality in the minimum global cost obtained.

References

[1] A. Alarcon-Rodriguez, G. Ault, and S. Galloway, "Multi-objective planning of distributed energy resources: A review of the state-of-the-art," *Renewable and Sustainable Energy Reviews*, vol. 14, no. 5, pp. 1353 – 1366, 2010.

[2] C. L. T. Borges, "An overview of reliability models and methods for distribution systems with renewable energy distributed generation," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 4008 – 4015, 2012.

[3] V. Martins and C. Borges, "Active distribution network integrated planning incorporating distributed generation and load response uncertainties," *Power Systems, IEEE Transactions on*, vol. 26, pp. 2164–2172, Nov 2011.

[4] H. Ren and W. Gao, "A MILP model for integrated plan and evaluation of distributed energy systems," *Applied Energy*, vol. 87, no. 3, pp. 1001 – 1014, 2010.

[5] H. Ren, W. Zhou, K. Nakagami, W. Gao, and Q. Wu, "Multi-objective optimization for the operation of distributed energy systems considering economic and environmental aspects," *Applied Energy*, vol. 87, no. 12, pp. 3642 – 3651, 2010.

[6] C. Chen, S. Duan, T. Cai, B. Liu, and G. Hu, "Optimal allocation and economic analysis of energy storage system in microgrids," *Power Electronics, IEEE Transactions on*, vol. 26, pp. 2762–2773, Oct 2011.

[7] S. Haffner, L. Pereira, L. Pereira, and L. Barreto, "Multistage model for distribution expansion planning with distributed generation Part I: Problem formulation," *Power Delivery, IEEE Transactions on*, vol. 23, pp. 915–923, April 2008.

[8] S. Haffner, L. Pereira, L. Pereira, and L. Barreto, "Multistage model for distribution expansion planning with distributed generation Part II: Numerical results," *Power Delivery, IEEE Transactions on*, vol. 23, pp. 924–929, April 2008.

[9] M. Samper and A. Vargas, "Investment decisions in distribution networks under uncertainty with distributed generation - Part II: Implementation and results," *Power Systems, IEEE Transactions on*, vol. 28, pp. 2341–2351, Aug 2013.

[10] T. Niknam, S. I. Taheri, J. Aghaei, S. Tabatabaei, and M. Nayeripour, "A modified honey bee mating optimization algorithm for multiobjective placement of renewable energy resources," *Applied Energy*, vol. 88, no. 12, pp. 4817 – 4830, 2011.

[11] S. Ganguly, N. Sahoo, and D. Das, "A novel multi-objective PSO for electrical distribution system planning incorporating distributed generation," *Energy Systems*, vol. 1, no. 3, pp. 291–337, 2010.

[12] M. Gómez-González, A. López, and F. Jurado, "Optimization of distributed generation systems using a new discrete PSO and OPF," *Electric Power Systems Research*, vol. 84, no. 1, pp. 174 – 180, 2012.

[13] K. Zou, A. Agalgaonkar, K. Muttaqi, and S. Perera, "Multi-objective optimisation for distribution system planning with renewable energy resources," in *Energy Conference and Exhibition (EnergyCon), 2010 IEEE International*, pp. 670–675, Dec 2010.

[14] H. Hejazi, A. Araghi, B. Vahidi, S. Hosseinian, M. Abedi, and H. Mohsenian-Rad, "Independent distributed generation planning to profit both utility and DG investors," *Power Systems, IEEE Transactions on*, vol. 28, pp. 1170–1178, May 2013.

[15] H. Hejazi, M. Hejazi, G. Gharehpetian, and M. Abedi, "Distributed generation site and size allocation through a techno economical multi-objective differential evolution algorithm," in *Power and Energy (PECon), 2010 IEEE International Conference on*, pp. 874–879, Nov 2010.

[16] M. Shaaban, Y. Atwa, and E. El-Saadany, "DG allocation for benefit maximization in distribution networks," *Power Systems, IEEE Transactions on*, vol. 28, pp. 639–649, May 2013.

[17] M. Raoofat, "Simultaneous allocation of DGs and remote controllable switches in distribution networks considering multilevel load model," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 8, pp. 1429 – 1436, 2011.

[18] Z. Cai, W. Gong, C. X. Ling, and H. Zhang, "A clustering-based differential evolution for global optimization," *Applied Soft Computing*, vol. 11, no. 1, pp. 1363 – 1379, 2011.

[19] M.-Y. Cheng, D.-H. Tran, and Y.-W. Wu, "Using a fuzzy clustering chaotic-based differential evolution with serial method to solve resource-constrained project scheduling problems," *Automation in Construction*, vol. 37, pp. 88 – 97, 2014.

[20] G. Liu, Y. Li, X. Nie, and H. Zheng, "A novel clustering-based differential evolution with 2 multi-parent crossovers for global optimization," *Applied Soft Computing*, vol. 12, no. 2, pp. 663 – 681, 2012.

[21] R. Mukherjee, G. R. Patra, R. Kundu, and S. Das, "Cluster-based differential evolution with crowding archive for niching in dynamic environments," *Information Sciences*, vol. 267, pp. 58 – 82, 2014.

[22] S. Song and X. Yu, "Multi-peak function optimization using a hierarchical clustering based genetic algorithm," in *Intelligent Systems Design and Applications, 2006. ISDA '06. Sixth International Conference on*, vol. 1, pp. 425–428, Oct 2006.

[23] Y.-J. Wang, J.-S. Zhang, and G.-Y. Zhang, "A dynamic clustering based differential evolution algorithm for global optimization," *European Journal of Operational Research*, vol. 183, no. 1, pp. 56 – 73, 2007.

[24] R. Storn and K. Price, "Differential evolution - A simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.

[25] R. Mena, M. Hennebel, Y.-F. Li, C. Ruiz, and E. Zio, "A risk-based simulation and multi-objective optimization framework for the integration of distributed renewable generation and storage," *Renewable and Sustainable Energy Reviews*, vol. 37, pp. 778 – 793, 2014. (I.F. 5.51).

[26] B. S. Everitt, S. Landau, M. Leese, and D. Stahl, *Cluster Analysis*. John Wiley & Sons, Ltd, 2011.

[27] IEEE power and energy society. Distribution test feeders, Available at http:// ewh.ieee.org/ soc/pes/dsacom/testfeeders/index.html.

[28] H. Falaghi, C. Singh, M.-R. Haghifam, and M. Ramezani, "DG integrated multistage distribution system expansion planning," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 8, pp. 1489 – 1497, 2011.

[29] Y. Atwa, E. El-Saadany, M. Salama, and R. Seethapathy, "Optimal renewable resources mix for distribution system energy loss minimization," *Power Systems, IEEE Transactions on*, vol. 25, pp. 360–370, Feb 2010.

[30] Y. Li and E. Zio, "Uncertainty analysis of the adequacy assessment model of a distributed generation system," *Renewable Energy*, vol. 41, pp. 235 – 244, 2012.

[31] D. Van Hertem, J. Verboomen, K. Purchala, R. Belmans, and W. Kling, "Usefulness of DC power flow for active power flow analysis with flow controlling devices," in *AC and DC Power Transmission, 2006. ACDC 2006. The 8th IEE International Conference on*, pp. 58–62, March 2006.

[32] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans, "Usefulness of DC power flow for active power flow analysis," in *Power Engineering Society General Meeting, 2005. IEEE*, pp. 454–459 Vol. 1, June 2005.

[33] W. Kersting, "Radial distribution test feeders," *Power Systems, IEEE Transactions on*, vol. 6, pp. 975–985, Aug 1991.

[34] L. Wang and C. Singh, "Multicriteria design of hybrid power generation systems based on a modified particle swarm optimization algorithm," *Energy Conversion, IEEE Transactions on*, vol. 24, pp. 163–172, March 2009.

[35] Y.-F. Li and E. Zio, "A multi-state model for the reliability assessment of a distributed generation system via universal generating function," *Reliability Engineering & System Safety*, vol. 106, pp. 28 – 36, 2012.

[36] F. Pilo, G. Celli, S. Mocci, and G. Soma, "Active distribution network evolution in different regulatory environments," in *Power Generation, Transmission, Distribution and Energy Conversion (MedPower 2010), 7th Mediterranean Conference and Exhibition on*, pp. 1–8, Nov 2010.

[37] R. Webster, "Can the electricity distribution network cope with an influx of electric vehicles?," *Journal of Power Sources*, vol. 80, no. 1-2, pp. 217 – 225, 1999.

[38] B. Korte and J. Vygen, *Combinatorial Optimization: Theory and Algorithms*. Algorithms and Combinatorics, Springer, 2007.

Paper (iii)

R. Mena, M. Hennebel, Y.–F. Li and E. Zio, "A multi-objective optimization framework for risk–controlled integration of renewable generation into electric power systems," Energy, (*under review*)

A Multi-objective Optimization Framework for Risk-controlled Integration of Renewable Generation into Electric Power Systems

Rodrigo Mena[a], Martin Hennebel[b], Yan-Fu Li[a], Enrico Zio[a,c,1]

[a]*Chair on Systems Science and the Energetic Challenge, Fondation Electricité de France at CentraleSupeléc, Châtenay-Malabry Cedex, France*

[b]*CentraleSupeléc, Department of Power & Energy Systems, Gif-Sur-Yvette, France*

[c]*Politecnico di Milano, Energy Department, Milan, Italy*

Abstract

We introduce a multi–objective optimization (MOO) framework for the integration of renewable distributed generation (DG) into electric power networks. The framework searches for the optimal size and location of different DG technologies, taking into account uncertain ties related to primary renewable resources availability, components failures, power demands and bulk–power supply. A non–sequential Monte Carlo simulation and optimal power flow (MCS–OPF) computational model is developed to emulate the network operation by generating random scenarios from the diverse sources of uncertainty, and assess the system performance in terms of global cost (CG). To measure uncertainty in the system performance, we introduce the conditional value–at–risk deviation (DCVaR) which, due to its axiomatic relation to the CVaR, allows the conjoint control of risk. A MOO strategy can, then, be adopted for the concurrent minimization of the expected global cost (ECG) and the associated deviation $DCVaR(CG)$. In our work this is operatively implemented by a heuristic search engine based on differential evolution (MOO–DE). An example of application of the proposed framework is given with regards to the IEEE 30 bus test system, where in DCVaR is shown capable of enabling and controlling tradeoffs between optimal expected economic performance, uncertainty and risk.

Keywords: Renewable distributed generation, uncertainty, risk, optimization, differential evolution, conditional value–at–risk, conditional value–at–risk deviation.

[1] Corresponding author. Tel: +33 1 4113 1606; fax: +33 1 4113 1272.
E-mail addresses: enrico.zio@ecp.fr, enrico.zio@polimi.it (E. Zio).

1 Introduction

The integration of renewable distributed generation (DG) into electric power systems is playing a relevant role in the strategies to decentralize and diversify the overall power generation and, in the efforts to contribute to the pursuit of environmental sustainability. In principle, the allocation of renewable DG units directly on 'the customer site of the meter' and/or on strategically–defined points of the sub–transmission grids can improve reliability of power supply, alleviating the restrictive dependence on the bulk power generated by large–scale conventional power plants [1, 2]. Furthermore, improvements of power losses and voltage stability profiles can be achieved, as well as reductions of investment risks and transmission costs [2–4].

DG planning is subject to complex economic, regulatory, technical and operational constraints, which must be attentively considered in order not to generate complications in the operative power system that may end up counteracting DG's potential benefits [5–7]. One of the main difficulties associated to the selection of technologies, dimensioning of power capacities and definition of the location of the different renewable generation units, is the modeling of the intrinsic uncertain behavior of the primary renewable energy sources (e.g. solar irradiance, wind speed and water inflow, in the case of solar photovoltaic, wind turbines and hydro–power technologies, respectively) and of the stochastic occurrence of unexpected events on the power generation units, such as failures and stoppages. Indeed, these sources of uncertainty are adjoined to those already present in the operative power system: failures and stoppages of transmission and distribution (T&D) lines and conventional generators, variability in the power demands and energy prices, fluctuations in the bulk power supply, etc. Consequently, for any proposed plan of DG–integrated network, the stochastic operational conditions need to be emulated coherently to the various sources of uncertainty to assess its expected performance by solving power flow equations [1, 8–10].

The selection of DG–integrated network plans can be framed as an optimization problem, whose complexity is driven by the size of the network and number of available DG technologies, that can lead to combinatorial explosion [9, 11, 12] and, by the aforementioned stochasticity and uncertainty. Optimality of the plan is customarily sought with regards to economic targets, such as minimization of costs of carbon dioxide emissions, fuel and transportation, energy losses, operation and maintenance, investment, etc.. The solutions identified must, of course, comply with technical constraints like generation capacities, T&D lines rating and voltage drops [1, 5].

Evolutionary algorithms (EAs) have emerged as the most effective search engines for combinatorial optimization problems and they can deal with non–differential objective functions, discontinuous feasible spaces and non–convex conditions [9, 11]. Some of the best known techniques are: particle swarm optimization (PSO) [9, 13–16], differential evolution (DE) [17–19] and genetic algorithms (GA) [2, 12, 20–22].

The uncertainty in the physical parameters and variables of a DG–integrated network propagates

onto the uncertainty in its operational response and the risk of incurring in non–desirable outcomes or non–satisfactory performance. This demands a framework of evaluation and control, for allowing properly informed and confident decision making. For evaluation purposes, various deviation and risk measures have been introduced into the DG planning optimization frameworks, framing the problem as a portfolio optimization in which the different types of DG technologies are treated analogously to financial assets [9, 23–30].

In portfolio optimization theory, variance is the most commonly used deviation measure for estimating uncertainty, although other indicators, like the mean absolute deviation (MAD), have also been employed [31]. Considering a probabilistic performance function associated to a certain portfolio, which usually is defined as a return (or loss) function, its variance is considered as an indicator of the likelihood of achieving the mean value of performance of the portfolio. This mean–variance approach can be applied to various portfolio optimization strategies, like the (single–objective) minimization of mean loss (maximization of mean return) with constrained variance, the (single–objective) minimization of variance with constrained mean loss or the simultaneous (multi–objective) minimization of mean loss and variance which enables the investors to select (an) optimal portfolio(s) by trading–off or conditioning mean loss and uncertainty (variance) on the Pareto front of dominant solutions. Mean–variance approaches have been applied to DG planning [9, 23, 24, 27, 30, 32] even though, one important warning must be considered from a risk perspective: the variance measure includes symmetrically the values of performance that fall short of or exceed the mean value; in the search for optimal DG technologies portfolios, lower levels of uncertainty (variance) in the performance function can be obtained with portfolios that lead to rarer occurrences of both beneficial and/or non–desired (risky) scenarios. Then, for controlling the risk side, it is necessary to introduce additional indicators that provide information on the extent of asymmetry of the performance function, weighting accordingly the risky part of it. For instance, skewness and kurtosis indicators [33] have been traditionally integrated into mean–variance approaches, for both financial and engineering applications, to estimate the asymmetry and peakedness of a probabilistic performance function, respectively, and allow controlling conjointly mean performance, uncertainty of it and risk.

Similarly to mean–variance approaches, and likewise derived from portfolio optimization theory, direct risk–based frameworks have been formulated and applied to tackle DG planning problems under uncertain conditions. As for mean–variance approaches, similar optimization strategies can be implemented to target expected performance and risk separately, as objectives or constraints, or to search for an efficient Pareto front of them. For this, the most widely used risk measures are the value–at–risk (VaR) and conditional value–at–risk (CVaR) [22, 26, 28, 29, 34–37], both with focus on the non–desirable performance outcomes given by a portfolio relative to a specific confidence level or percentile: VaR is defined as the threshold value of performance at the confidence level, whereas CVaR is the expected value of the performances that fall beyond the VaR [38]. For optimization under uncertainty, CVaR has gained more interest than VaR because of its preferable properties: among

others, CVaR maintains consistency with VaR and is a coherent risk measure; it can be expressed by a minimization formula and aims towards conservatism, which prevails in risk management [38].

Even though integrated portfolio optimization approaches present robustness advantages by conjointly controlling the level of uncertainty and risk associated to the mean value of performance of different portfolios, they can considerably increase the complexity of the concurrent optimization problem that, in our case of interest, is already complicated given its combinatorial nature due to the size of a DG–integrated network, and the presence of uncertainty in the operational conditions. Considering the mean value and the necessary deviation and risk measures as representative objectives of a portfolio increases significantly the number of objective functions to be simultaneously optimized, further constraining the feasible space and, eventually, hindering in understanding the information delivered to the decision–makers.

In this work, we take the challenge of developing an optimization framework for the integration of DG into a electric power network accounting for uncertain operational inputs, and assessing and controlling uncertainty and risk. For this, we exploit some relevant concepts and tools from portfolio optimization, but avoiding increasing the complexity of the problem by adopting suitable performance indicators and, eventually, providing a spectrum of comprehensible information for well–supported decision making. The main original contribution of this work lies in the introduction of the CVaR deviation (DCVaR) [31] as a measure of uncertainty, used along with the expected value as performance indicators of the global cost function (CG) associated to each portfolio of DG technologies. A multi–objective optimization (MOO) strategy is, then, developed, aiming at the simultaneous minimization of both objective functions: the expected global cost (ECG) and its corresponding $DCVaR$. The numerical implementation of the approach is based on a Monte Carlo–optimal power flow (MCS–OPF) simulation model nested in a MOO search engine based on differential evolution (DE) [39]. The MCS–OPF model emulates the operation of each DG–integrated network proposed by MOO–DE, generating random scenarios from the diverse uncertain operational inputs and assessing the global cost performance CG of each, and, then, evaluating the two objective functions. The proposed framework searches for optimal technologies, size and location of DG units, and enables the direct trade–off between optimal expected performance and the associated uncertainty to achieve it. Furthermore, thanks to the axiomatic relation between CVaR and DCVaR, it integrates in the optimal Pareto front the level of risk, given by the values of $CVaR$.

We apply the framework on a case study based on the IEEE 30 nodes test feeder sub–transmission and distribution network [40], considering solar photovoltaic and wind turbines DG technologies, discussing the effectiveness and implications of of its practical application.

2 DG–Integrated power network simulation model

In this section, we present the Monte Carlo and optimal power flow (MCS–OPF) simulation model. We introduce the representation of the DG–integrated network, the modeling of the different uncertainty sources considered, the process of generating the random operational scenarios and the formulation of the OPF problem.

2.1 Network configuration

For modeling a DG–integrated network, one needs the definition of the type (or classes) of components involved and their topological location. In this work, we consider three main classes of components: power generators G, transmission and distribution lines TD and power demands (loads). The class of power generators considers all the different types of conventional bulk power suppliers MG and renewable technologies RG. The topology of the network is described as a graph, i.e., a set of nodes and the various connections that link them. The nodes represent spatial points at which generation components (MG and RG) and loads are located or can be allocated, whereas the connections between nodes are the transmission and distribution lines. We indicate a node by the index i and the set of all nodes by $N = \{i : i \in \{1,2,\ldots,n\}\}$, where n is the total number of nodes in the network. Consequently, we define the set of transmission and distribution lines as $Y = \{(i,i') : \text{nodes } i \text{ and } i' \text{ are connected}, \forall i,i' \in N\}$.

Assuming stationary operational conditions, the network performance is considered dependent on the fixed location and magnitude of the power available in each generation unit and loads and, the technical limits of the T&D lines. To indicate the location and capacity size of the different types of generation units present in the network, we define the following matrix Q, $\forall i \in N$ as follows:

$$
Q = [Q^M | Q^R] = \begin{array}{c}
\begin{array}{cccccccccc}
\multicolumn{5}{c}{\overbrace{\qquad\qquad\qquad\qquad\qquad}^{m \text{ types of bulk power suppliers } MG}} & \multicolumn{5}{c}{\overbrace{\qquad\qquad\qquad\qquad\qquad}^{r \text{ types of renewable technologies } RG}}
\end{array} \\
\begin{array}{cccccccccc}
G_1 & \cdots & G_j & \cdots & G_m & G_{1+m} & \cdots & G_{j+m} & \cdots & G_{r+m}
\end{array} \\
\left[\begin{array}{ccccc|ccccc}
q_{1,1} & \cdots & q_{1,j} & \cdots & q_{1,m} & q_{1,1+m} & \cdots & q_{1,j+m} & \cdots & q_{1,r+m} \\
\vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
q_{i,1} & \cdots & q_{i,j} & \cdots & q_{i,m} & q_{i,1+m} & \cdots & q_{i,j+m} & \cdots & q_{i,r+m} \\
\vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
q_{n,1} & \cdots & q_{n,j} & \cdots & q_{n,m} & q_{n,1+m} & \cdots & q_{n,j+m} & \cdots & q_{n,r+m}
\end{array}\right]
\end{array}
\overbrace{\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad}^{m+r \text{ types of power generators } G}
\tag{1}
$$

where $[Q]_{i,j} = q_{i,j}$ is a non–negative integer that specifies the number of units of the power generator type j allocated at node i:

$$
q_{i,j} = \begin{cases} q' \in \mathbb{Z}^* & \text{if } q' \text{ units of } G_j \text{ are allocated at node } i, \forall G_j \in G = \{G_1,\ldots,G_{m+r}\} \\ 0 & \text{otherwise} \end{cases} \tag{2}
$$

Then, the static configuration of the physical components G and TD of a DG–integrated network is represented by the pair $([Q], \{Y\})$. Any physical component is assumed to be affected by the stochastic occurrence of failures, conditioning dynamically the functionality of power generators and the paths through which the power flows. Moreover, the magnitude of power available in each generator is subject to the inherent uncertain behavior of the corresponding primary energy source and, considering the DG–integrated network as a 'price taker' entity, the economic conditions depend on the variability of the power demands [41, 42]. Hence, to evaluate the operating performance of a given DG–integrated network, affected by significant uncertain conditions, it is essential to model the different sources of variability and emulate the response of the network for a large representative combination of possible scenarios.

2.2 Uncertainty modeling

Analytical methods and Monte Carlo simulation (MCS) are among the most common techniques for evaluating the performance of DG–integrated networks [2]. In theory, analytical methods are preferable because of the possibility of achieving closed form solutions; however, their application often requires simplifications in the modeling which may lead to unrealistic results. For instance, analytical solutions for optimal DG planning consider non–uncertain or non–intermittent power generation and/or load profiles, and networks of low dimensionality [43]. On the contrary, MCS allows a more realistic modeling, because the performance of the network is not analytically solved but simulated, and the overall performance indicators are statistically estimated from virtual operational scenarios realizations [44]. MCS has been found quite adequate for the analysis of distribution networks with a significant number of sources of randomness or variability, e.g., power generation, loads, component failures or degradation processes, etc. [2, 8, 9, 16, 27, 28], but at the expense of incrementing the use of computational resources.

In the present framework, we adopt a non–sequential MCS, based on latin hypercube sampling (LHS) [45], to emulate the operation of the DG–integrated network, considering the operation variables as independent on previous uncertain conditions, so as to seize the advantages of MCS without overly increasing the computational efforts.

The considered uncertain conditions that determine the operation of the DG–integrated network are accounted for using different stochastic models, as presented below.

2.2.1 Power demands

The aggregated profile of power demand in an electric power network, as well as the single nodal profiles, can be represented as daily load curves inferred from historical data [42, 46], and can be considered uncertain following normal distributions [9, 16]. We model the nodal power demand profiles by integrating both ideas, i.e., for a specific hour of the day $t \in D = \{1, \ldots, 24\}$ the

corresponding power demand at node i, denoted $L_{i,t}$ (MW), is normally distributed with mean $\mu_{i,t}$ (MW), standard deviation $\sigma_{i,t}$ (MW) and truncated at 0:

$$f_{i,t}(L_{i,t}|\mu_{i,t},\sigma_{i,t}) = \begin{cases} \dfrac{\phi(\xi(L_{i,t},\mu_{i,t},\sigma_{i,t}))}{\sigma_{i,t}Z(\mu_{i,t},\sigma_{i,t})} & \forall L_{i,t},\mu_{i,t},\sigma_{i,t} \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

$$\xi(L_{i,t},\mu_{i,t},\sigma_{i,t}) = \frac{L_{i,t}-\mu_{i,t}}{\sigma_{i,t}}; \quad Z(\mu_{i,t},\sigma_{i,t}) = 1-\Phi\left(\frac{\mu_{i,t}}{\sigma_{i,t}}\right)$$

where, ϕ and Φ are the standard Normal probability density function and its cumulative distribution function, respectively.

As aforementioned, the network is assumed as a 'price taker' entity, for which the value of the energy price is correlated with the aggregated power demand in the network. As an intermediate approximation of existing studies (e.g. [41, 42, 47]), the proportional correlation used in this study can be expressed as:

$$EP_t(L_{i,t}|EP_{max},L_{max_i}) = EP_{max}\sum_{i\in N}\frac{L_{i,t}}{L_{max_i}}\left(-0.38\left(\sum_{i\in N}\frac{L_{i,t}}{L_{max_i}}\right)+1.38\right) \tag{4}$$

where, EP_t (\$/MWh) is the energy price at the hour of the day t and EP_{max} (\$/MWh) is the maximum value of energy price correspondent to the aggregated value of maximum nodal power loads L_{max_i} (MW).

2.2.2 Bulk power generation

Bulk power generation stands for the power supply coming from conventional power plants (MG) already existing in the network. These sources of power supply are rather stable and are connected to the network at sub–transmission or distribution transformers. Their stochastic behavior is represented following normal distributions [11, 48], with small standard deviation and truncated by the maximum capacity of generation.

$$f_j(P_j|\mu_j,\sigma_j,P_{max_j}) = \begin{cases} \dfrac{\phi(\xi(P_j,\mu_j,\sigma_j))}{\sigma_j Z(\mu_j,\sigma_j)} & \forall P_j \in [0,P_{max_j}],\mu_j,\sigma_j \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

$$\xi(P_j,\mu_j,\sigma_j) = \frac{P_j-\mu_j}{\sigma_j}; \quad Z(\mu_j,\sigma_j) = \Phi(\frac{P_{max_j}-\mu_j}{\sigma_j})-\Phi(\frac{\mu_j}{\sigma_j})$$

where $\forall j \in \{j : G_j \in MG\}$, P_j (MW) and P_{max_j} (MW) are the available bulk power and maximum capacity of the MG generator type j, respectively, and μ_j (MW) and σ_j (MW) the corresponding Normal distribution mean and standard deviation.

2.2.3 Solar photovoltaic generation

Solar photovoltaic technologies (PV) transforms solar irradiance into electric power through panels of solar cells. Commonly for long term periods of analysis, solar irradiance uncertain behavior has been modeled using probabilistic distributions, obtained from weather historical data of a particular geographical area. In particular, the Beta distribution function has been found suitable to model hourly solar irradiance [46, 49] and, therefore, is adopted in this paper. Moreover, the intermittency in the solar irradiation is taken into account defining a daylight interval between 07.00 and 21.00 hours, i.e., if the value t of the hour of the day is in the subset of $D_L = \{7, \ldots, 21\}$ of D, the solar irradiation H is a positive value, otherwise t is in the night interval $D_N = \{22, \ldots, 24, 1, \ldots, 6\}$ and the value of solar irradiation is assumed equal to 0. Then, we adjust the Beta distribution function considering the probability that t is in the daylight interval $p_L = P(t|t \in D_L)$, as follows:

$$f_i(H_i|\alpha_i, \beta_i, p_L, H_i^*) = \begin{cases} \dfrac{\Gamma(\alpha_i+\beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i)} \dfrac{H_i^{(\alpha_i-1)}(1-H_i)^{(\beta_i-1)}}{(1-p_L)\mathrm{B}(\alpha_i,\beta_i)} & \forall H_i \in [H_i^*, 1], \alpha_i, \beta_i > 0 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where $\forall j \in \{j : G_j \in PV\}$, H_i is the solar irradiance at node i, α_i and β_i are the shape parameters of the corresponding Beta probability density function at node i and H_i^* is the p_L percentile of the non–adjusted Beta distribution $f_{i,j}(H_i|\alpha_i, \beta_i)$.

Given the technical characteristics of PV cells and the model of solar irradiance, the probabilistic power output of a single cell is obtained from the following equation:

$$P_{i,j}(H_i) = \begin{cases} P'_{i,j}(H_i) & \text{if } 0 \le P'_{i,j}(H_i) \le P_{max_j} \\ P_{max_j} & \text{if } P_{max_j} < P'_{i,j}(H_i) \end{cases} \tag{7}$$

$$P'_{i,j}(H_i) = n_c FF_j V(H_i) I(H_i) \times 10^{-6}$$

$$T_C(H_i) = T_{A_i} + H_i(T_{No_j} + 20)/0.8$$

$$I(T_C) = H_i(I_{SC_j} + k_{I_j}(T_C - 25))$$

$$V(T_C) = V_{OC_j} + k_{V_j} T_C$$

$$FF_j = (V_{MPP_j} I_{MPP_j})/(V_{OC_j} I_{SC_j})$$

where referring to PV type j, $P_{i,j}$ (MW) is the power output at node i, P_{max_j} (MW) is the maximum power generation capacity, n_c is the number of photovoltaic cells, FF_j is the fill factor, T_{A_i} (°C) is the ambient temperature at node i, T_{No_j} (°C) is the nominal cell operation temperature, I_{SC_j} (A) is the short circuit current, k_{I_j} (mA/°C) is the current temperature coefficient, V_{OC_j} (V) is the open circuit voltage, k_{V_j} (mV/°C) is the voltage temperature coefficient, and V_{MPP_j} (V) and I_{MPP_j} (A) are the voltage and current at maximum power, respectively.

2.2.4 Wind turbines generation

Wind power generation (W) is obtained from turbine–alternator devices that transform the kinetic energy of the wind into electric power. The stochastic behavior of the wind speed is commonly represented through probability distribution functions. The Weibull distribution has been widely used to model the randomness of the wind speed in various conditions [3, 11, 21, 37, 46, 49]:

$$f_i(U_i|\alpha_i,\beta_i) = \begin{cases} \frac{\beta_i}{\alpha_i}\left(\frac{U_i}{\alpha_i}\right)^{(\beta_i-1)} \exp\left[-\left(\frac{U_i}{\alpha_i}\right)^{\beta_i}\right] & \forall U_i \geq 0, \alpha_i, \beta_i > 0 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where $\forall j \in \{j : G_j \in W\}$, U_i (m/s) is the wind speed at node i and α_i and β_i are the scale and shape parameters of the Weibull distribution function at node i, respectively.

Similarly to PV type of technologies, the uncertainty associated to the wind speed and the technical features of a specific type of wind turbine characterize its power output function that can be determined as follows:

$$P_{i,j}(U_i) = \begin{cases} \frac{U_i - U_{CI_j}}{U_{A_j} - U_{CI_j}} P_{R_j} & \text{if } U_{CI_j} \leq U_i < U_{A_j} \\ P_{R_j} & \text{if } U_{A_j} \leq U_i \leq U_{CO_j} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

where referring to W type j, P_{R_j} (MW) is the rated power and U_{CI_j}, U_{A_j} and U_{CO_j} (m/s) are the cut–in, average and cut–out wind speeds, respectively.

2.2.5 Components availability state

Consistently with the non–sequential nature of the MCS simulation proposed in the present framework, the availability states of the physical components in the network, generators and T&D lines, are straightforwardly modeled by two–state stationary Markov chains [2, 50], defining two possible operating states: $\eta = 0$ if the corresponding component is non functional (failure) and $\eta = 1$ if the component is available to operate, i.e., generate, transmit or distribute power accordingly to the class of component. Then, the discrete stationary distribution of operating states can be expressed as follows:

$$f_k(\eta_k|\lambda_{F_k},\lambda_{R_k}) = \begin{cases} \lambda_{F_k}/(\lambda_{F_k} + \lambda_{R_k}) & \eta_k = 0 \\ \lambda_{R_k}/(\lambda_{F_k} + \lambda_{R_k}) & \eta_k = 1 \end{cases} \tag{10}$$

where $\forall k \in \{\{k : G_k \in G\} \cup \{k/k = (i,i') \in Y\}\}$, η_k is the operating state of component k and λ_{F_k} and λ_{R_k} are the corresponding failure and repair rates, respectively.

2.3 Monte Carlo and optimal power flow simulation

For a given DG–integrated network, represented by the pair $([Q],\{Y\})$ that contains the locations and number of units of the different power generators and the T&D lines, each uncertain variable is randomly sampled several times by LHS [45] and the inverse transform method [44], for the realization of NS operational scenarios of duration t^{Δ}. For practicality, we define NS as multiple of 24, so each hour of the day has the same number of realizations $NS/24$. We denote by ω the set of sampled variables, which constitutes an operational scenario and by Ω the sets of all the NS realizations of ω.

$$\omega_s = \{t_s, L_{i,t_s}, EP_{t_s}, \{P_{i,j,s}/G_j \in MG\}, H_{i,s}, U_{i,s}, \eta_{i,j,s}, \eta_{(i,i'),s}\} \tag{11}$$

$$\Omega = \{w_s : s \in \{1,2,\ldots,NS\}\} \tag{12}$$

For each scenario ω_s, the configuration $([Q],\{Y\})$ and the sampled variables set the stage for evaluating the response of the network in terms of available power usage, power demand satisfaction and the involved economics. For this, DC optimal power flow analysis (OPF) is run, neglecting power losses, assuming constant the voltage throughout the network and accounting solely for active power flows [51, 52]. The OPF receives $([Q],\{Y\})$ and ω_s as inputs and aims the minimization of the aggregated operating cost of generation, transmission and distribution and load shedding, including revenues per MW h sold. The present formulation of the power flow problem is:

MCS–OPF$([Q],\{Y\},\omega)$:

$$\begin{aligned}\min_{P_{U_s},\Delta\delta_s,LS_s} \quad & CO_s^{\omega} = \sum_{i\in N}\sum_{G_j\in G}(COv_j - EP_{t_s})P_{U_{i,j,s}} + \\ & (C_{LS} + EP_{t_s})\sum_{i\in N} LS_{i,s} + S_{ref}\sum_{(i,i')\in Y} COv_{(i,i')}|B_{(i,i')}(\delta_{i,s} - \delta_{i',s})|\end{aligned} \tag{13}$$

s.t.

$$L_{i,t_s} - \sum_{G_j\in G} P_{U_{i,j,s}} - S_{ref}\sum_{i'\in N}\eta_{(i,i'),s}B_{(i,i')}(\delta_{i,s} - \delta_{i',s}) - LS_{i,s} = 0 \tag{14}$$

$$0 \le P_{U_{i,j,s}} \le \eta_{i,j,s}[\boldsymbol{Q}]_{i,j}P_{i,j,s} \tag{15}$$

$$S_{ref}\,|B_{(i,i')}(\delta_{i,s} - \delta_{i',s})| \le P_{max_{(i,i')}} \tag{16}$$

where $\forall t \in D, s \in \{1,2,\ldots,NS\}$, CO_s^{ω} (\$/h) is the aggregated operating cost of generation, transmission and distribution, and load shedding, COv_j (\$/MWh) is the variable operating cost of the power generator j, EP_{t_s} (\$/MWh) is the energy price at hour t, $P_{U_{i,j,s}}$ (MW) is the used power from the generator type j located at node i, S_{ref} (MV A) is the reference apparent power in the network, $COv_{(i,i')}$ (\$/MWh) and $B_{(i,i')}$ (p.u.) are the variable operating cost and susceptance of the T&D line (i,i'), respectively, $\delta_{i,s}$ is the voltage angle at node i, C_{LS} (\$/MWh) is the load shedding cost and

$P_{max_{(i,i')}}$ (MW) is the power rating of the T&D line (i, i'). The load shedding $LS_{i,s}$ (MW) at node i is defined as the amount of load disconnected to alleviate congestion in the feeders and/ or balance the demand of power with the available power supply.

Equation (14) corresponds to the power balance at node i, whereas equation (15) represents the bounds of the power generation units and equation (16) considers the technical limits of the feeders.

As above–mentioned, OPF is solved for each operational scenario ω_s, giving in output the respective values of minimum CO_s^{ω}. The set $CO^{\Omega} = \{CO_1^{\omega}, CO_2^{\omega}, \ldots, CO_{NS}^{\omega}\}$ is, then, considered as a sample of realizations of the probability function of CO.

2.3.1 DG–integrated network performance evaluation

The proposed renewable DG–integrated network solutions $Q = [Q^M|Q^R]$ are evaluated with respect to performance indicators of the global cost CG function. The quantity CG is composed of two terms: the outcome operating cost of the MCS–OPF described in the previous section, CO^{Ω}, and the fixed investment and operating cost, $CI_j + COf_j$, associated to the renewable part of the proposed DG plan Q^R, i.e., $\forall j \in \{j : G_j \in RG\}$. The quantity $CI_j + COf_j$ (\$) is prorated hourly over the lifetime of the project t^h. Thus, the global cost function for the set of operational scenarios Ω is given by:

$$CG^{\Omega} = CO^{\Omega} + \frac{1}{t^H} \sum_{i \in N} \sum_{G_j \in RG} (CI_j + COf_j)[\mathbf{Q^R}]_{i,j} \tag{17}$$

Analogously to CO^{Ω} definition, CG^{Ω} represents a sample of realizations of the probability function of CG and performance indicators of interest can be obtained, relative to expected performance, uncertainty and risk.

2.4 Uncertainty and Risk Assessment

The proposed MOO framework introduces the CVaR deviation (DCVaR) [31] to measure the uncertainty in the performance function (CG) of interest. The quantity DCVaR is a functional of the CVaR [38], which is a coherent risk measure broadly used in financial portfolio optimization and has been extended to engineering applications, including electric power systems analysis and, in particular, DG planning [22, 26, 28, 29, 34–36]. The axiomatic relation between DCVaR and CVaR allows optimizing simultaneously uncertainty and risk, by targeting just one of these two indicators.

The definitions and properties of CVaR and DCVaR for continuous and discrete general return (loss) functions are given in detail in [31,38]. Here, we limit ourselves to presenting only a graphical, but comprehensive view to understand the CVaR and DCVaR definitions, as shown in Figure 1.

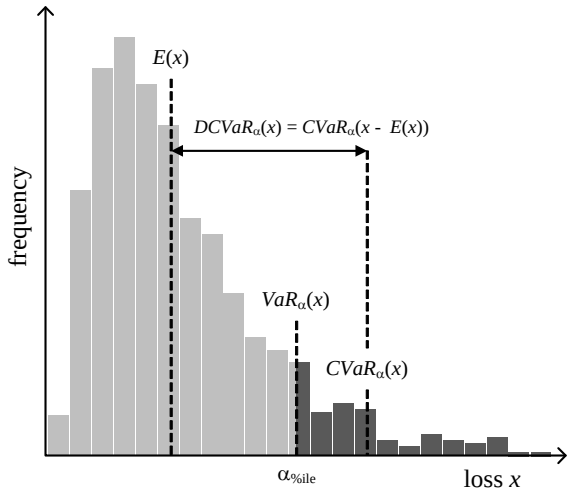

Figure 1: Graphic representation of $VaR_\alpha(x)$, $CVaR_\alpha(x)$ and $DCVaR_\alpha(x)$; x = loss

For a discrete approximation of the probability function of the loss x, given a confidence level or α-percentile, the value–at–risk $VaR_\alpha(x)$ represents the smallest value of loss for which the probability that the loss does not exceed that threshold value is greater than or equal to α, whereas $CVaR_\alpha(x)$ is the expected value of loss given that the loss is greater than or equal to the $VaR_\alpha(x)$. Thus, $CVaR_\alpha(x)$ provides a quantitative indication of the extent of the probability of occurrence of extreme non–desirable or risky scenarios of loss. The quantities $VaR_\alpha(x)$ and $CVaR_\alpha(x)$ can be expressed by the following equations:

$$VaR_\alpha(x) = \inf\{z : F_x(z) > \alpha\} \tag{18}$$

$$CVaR_\alpha(x) = E(x/x \geq VaR_\alpha(x)) \tag{19}$$

With regards to $DCVaR_\alpha(x)$, this is a non symmetric deviation measure, as it accounts for the uncertainty associated to the loss exceeding its expected value. It is defined taking into account some important properties of the standard deviation [31], and is formulated as:

$$DCVaR_\alpha(x) = CVaR_\alpha(x - E(x)) \tag{20}$$

Furthermore, being a coherent risk measure, CVaR is a strictly expectation–bounded risk measure and it can be proved that a one–to–one relation exists with its corresponding deviation measure DCVaR [31]:

$$CVaR_\alpha(x) = E(x) + DCVaR_\alpha(x) \tag{21}$$

In the present framework, we can consider a specific configuration of the DG–integrated network $([Q], \{Y\})$ as a generation portfolio, in which the renewable part $[Q^R]$ of $[Q]$ is the decision matrix. The corresponding assessed CG^Ω, obtained from the output MCS–OPF($[Q]$,$\{Y\}$,Ω), can be translated

into the probability function of loss (CG); then, the quantities $CVaR_\alpha(CG^\Omega)$ and $DCVaR_\alpha(CG^\Omega)$ represent the level of risk and uncertainty associated to the solution $[Q^R]$ with an expected global cost $ECG = E(CG^\Omega)$, respectively.

3 Renewable generation selection, sizing and allocation

The practical aim of the MOO is to find the optimal integration of DG in terms of selection, sizing and allocation of the different renewable generation units (PV and W). The corresponding decision matrix $[Q^R]$ is contained in the matrix $Q = [Q^M|Q^R]$ that stores the number and location of each type of power generator in the network. The MOO problem consists in the simultaneous minimization of two objective functions, given by the indicators ECG and $DCVaR_\alpha(CG)$.

3.1 Multi–objective optimization problem formulation

The general multi–objective optimization problem for all set of randomly generated operational scenarios Ω is formulated as follows:

$$\min_{[\mathbf{Q^R}]_{i,j}} \quad E(CG^\Omega) \tag{22}$$

$$\min_{[\mathbf{Q^R}]_{i,j}} \quad DCVaR_\alpha(CG^\Omega) \tag{23}$$

s.t.

$$[\mathbf{Q}]_{i,j} \in \mathbb{Z}^* \tag{2}$$

$$\sum_{i \in N} \sum_{G_j \in RG} \frac{[\mathbf{Q}]_{i,j} P_{AV_j}}{L_{max_i}} \leq PF \tag{24}$$

$$\text{MCS—OPF}([\mathbf{Q}], \{Y\}, \Omega) \tag{13 -16}$$

The meaning of each constraint is the following: (2) the decision variables $[Q]_{i,j}$ are non–negative integer numbers; (15) the ratio of total amount of average renewable power integrated in the network must be less or equal to the penetration factor PF; (13)–(16) all the power flows equations must be satisfied.

3.1.1 Multi–objective differential evolution

The non–convex mixed–integer non–linear MOO problem under uncertainties is solved by the multi–objective differential evolution (MOO–DE) algorithm, integrating a fast non–dominated sorting procedure and crowded–comparison operator [53] into the original single objective DE [39], and

evaluating the objective functions by the developed MCS–OPF. The extension to MOO entails the integration of Pareto optimality concepts. In general terms, solving a MOO problem of the form:

$$\min_{X} \quad \{f_1(X), f_2(X), \ldots, f_m(X)\}$$
$$\text{s.t. } X \in \Lambda$$

with at least two conflicting objectives functions ($f_i : \Re^n \rightarrow \Re$) implies to find, within a set of acceptable solutions that belong to the non–empty feasible region $\Lambda \subseteq \Re^n$, the decision vectors $X \in \Lambda$ that satisfy the following [54]:

$$\neg X \in \Lambda / f_i(X) \leq f_i(X'), \forall i \in \{1, 2, \ldots, m\} \text{ and } f_i(X) < f_i(X') \text{ for at least one } i$$
$$\Downarrow$$
$$f_i(X) \prec f_i(X') \text{ i.e. } X \text{ dominates } X'$$

The vector X is called a Pareto optimal solution and the Pareto front is defined as the set $\{f(X) \in \Re^n\}$ such that X is Pareto optimal solution. The general MOO–DE algorithm is summarized as follows:

Initialization

- Set the values of parameters:
  - NP: population size
  - Gen_{max}: maximum number of generations
  - COc: crossover coefficient $\in [0, 1]$
  - F: differential variation amplification factor $\in [0, 2]$
- Form the initial population POP^0, randomly generating NP decision matrices (individuals) X within the feasible space, $POP^0 = \{X_1^0, \ldots, X_k^0, \ldots, X_{NP}^0\}$
- Evaluate the objective functions $\{f_1(X_k^0), \ldots, f_{NF}(X_k^0)\}$ for each individual X_k^0, where NF is the number of objective functions
- Rank the individuals in POP^0, applying a fast non–dominated sorting procedure with respect to the values of the objective functions
- Compute and assign the crowding–distance value (d_C) to each individual in POP^0 and sort, in ascending order, with respect to d_C, the individuals belonging to the same non–domination–ranked group

Evolution loop

- Set generations count index $g = 1$
- Set $POP^g = POP^0$

- While $g \leq Gen_{max}$ (stopping criterion)
- Set a repository population $RPOP$ as empty

  Trial loop

  For each individual X_k^g in POP^g, $\forall k \in \{1,\ldots,NP\}$

  - Sample from the uniform distribution three integer indexes in $\{1,\ldots,NP\}$ such that $k_1 \neq k_2 \neq k_3 \neq k$ and choose the corresponding three individuals $X_{k_1}^g, X_{k_2}^g, X_{k_3}^g$
  - Generate a mutant individual XM_k^g according to the following mutation operator:

$$XM_k^g = X_{k_1}^g + F(X_{k_2}^g - X_{k_3}^g) \tag{25}$$

  - Apply a crossover operator, initializing a randomly generated vector XC_k^g, whose dimensionality n is the same as that of X_k^g and each coordinate $xc_{k,i}^g$ follows a uniform distribution with outcome in $[0,1]\,\forall i \in \{1,\ldots,n\}$. In addition, generate randomly an integer index i^* in $\{1,\ldots,n\}$ from a uniform distribution to ensure that at least one coordinate from XM_k^g is exchanged to form trial individual XT_k^g, whose coordinates $xt_{k,i}^g$ are defined as follows:

$$xt_{k,i}^g = \begin{cases} xm_{k,i}^g & \text{if } xc_{k,i}^g \leq Cco \text{ or } i = i^* \\ x_{k,i}^g & \text{if } xc_{k,i}^g > Cco \text{ and } i \neq i^* \end{cases} \tag{26}$$

  - Evaluate the objective functions for the trial individual $\{f_1(XT_k^g),\ldots,f_{NF}(XT_k^g)\}$; if XT_k^g dominates X_k^g, i.e., $\{f(XT_k^g) \prec f(X_k^g)\}$, XT_k^g replaces X_k^g in POP^g, otherwise retain X_k^g in POP^g and save XT_k^g in the repository population $RPOP$
- Set a combined population $UPOP$ as $POP^g \cup RPOP$ and rank the individuals in $UPOP$, applying a fast non–dominated sorting procedure
- Compute and assign the crowding–distance value d_C to each individual in $UPOP$ and sort, in descending order with respect to d_C, the individuals belonging to the same non–domination–ranked group
- Set POP^g as the first NP individuals of the ranked and sorted population $UPOP$, $POP^g = \{X_k : X_k \in UPOP, k \in \{1,\ldots,NP\}\}$
- If the stopping criterion is reached return POP^g, otherwise set $g = g + 1$.

In correspondence to the nomenclature used in the proposed framework, the process of searching the set of non–dominated solutions carried out by the MOO–DE MCS–OPF is presented schematically in Figure 2.

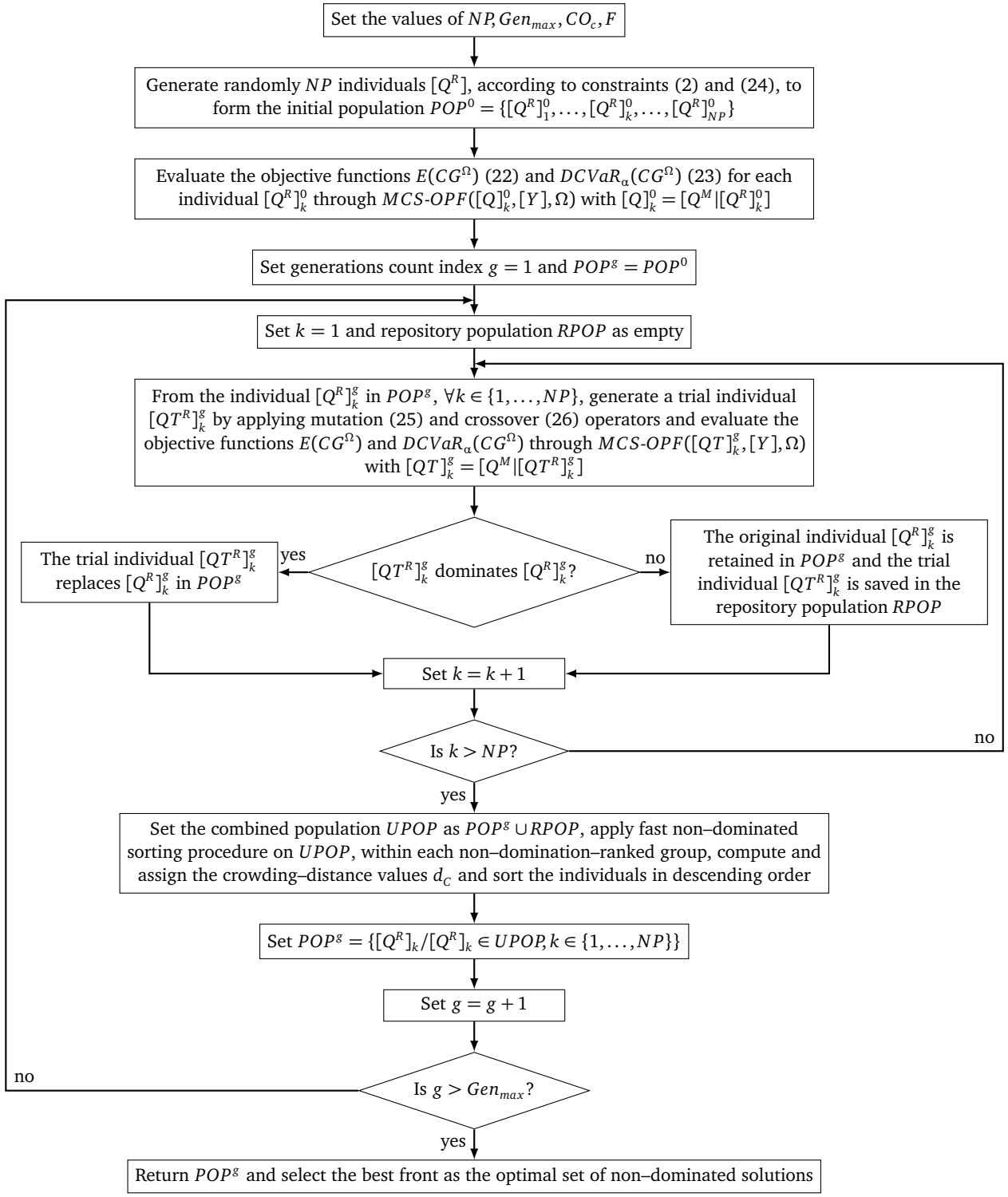

Figure 2: Flow chart of the proposed MOO–DE framework

4 Case study

We consider the IEEE 30 bus sub-transmission and distribution test system, a portion of the Midwestern U.S. electric power system which presents relevant characteristics of interest for the analysis, e.g., the presence of bulk–power supply spots different in type and with comparatively low and high nodal load profiles. An important consideration is that we neglect the synchronous condensers, given the DC assumptions made in the proposed framework for the resolution of the OPF problem.

4.1 Network description

The network consists of $n = 30$ nodes, a mesh deployment of 41 T&D lines and 2 transformers or bulk–power supply spots, as shown in Figure 3. The reference apparent power is $S_{ref} = 100$ (MV A).

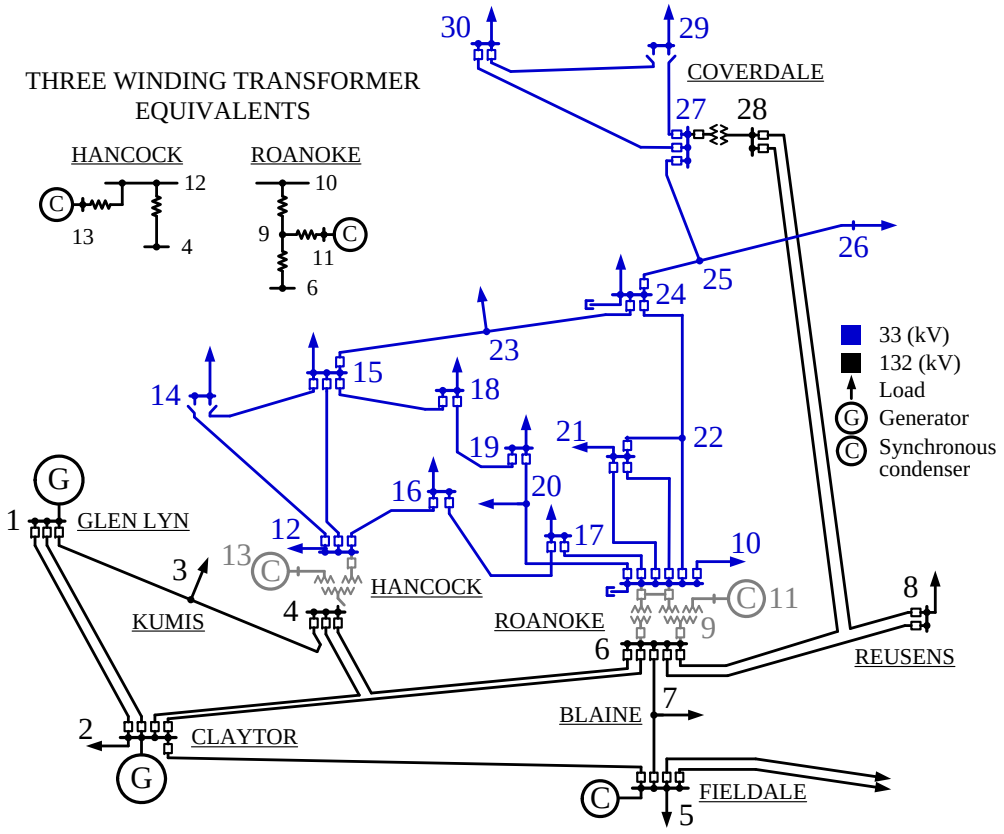

Figure 3: IEEE 30 bus sub–transmission and distribution test system diagram

Table 1 summarizes the characteristics and technical data of the T&D lines, specifically: the indexes of the pair of nodes that they connect (i, i'), the susceptance values $B_{(i,i')}$, power rating $P_{max_{(i,i')}}$, failure $\lambda_{F_{(i,i')}}$ and repair $\lambda_{R_{(i,i')}}$ rates and operating cost $COv_{(i,i')}$.

Table 1: T&D lines characteristics and technical data [40, 55, 56]

i	i'	$B_{(i,i')}$ (p.u.)	$P_{max_{(i,i')}}$ (p.u.)	$\lambda_{F_{(i,i')}}$ (n/h)	$\lambda_{R_{(i,i')}}$ (n/h)	$COv_{(i,i')}$ ($/MWh)
1	2	17.24	1.30	2.85e−04	6.67e−02	4.95
1	3	5.41	1.30	2.85e−04	6.67e−02	1.61
2	4	5.75	0.65	4.28e−04	1.00e−01	2.61
3	4	26.32	1.30	2.85e−04	6.67e−02	5.97
2	5	5.05	1.30	2.85e−04	6.67e−02	2.53
2	6	5.68	0.65	4.28e−04	1.00e−01	3.68
4	6	24.39	0.90	3.73e−04	8.72e−02	6.06
5	7	8.62	0.70	4.17e−04	9.74e−02	0.70
6	7	12.20	1.30	2.85e−04	6.67e−02	0.93
6	8	23.81	0.32	5.01e−04	1.17e−01	0.19
6	9	4.81	0.65	4.28e−04	1.00e−01	5.21
6	10	1.80	0.32	5.01e−04	1.17e−01	4.14
9	10	9.09	0.65	4.28e−04	1.00e−01	8.78
9	11	4.81	0.65	4.28e−04	1.00e−01	4.81
4	12	3.91	0.65	4.28e−04	1.00e−01	2.93
12	13	7.14	0.65	4.28e−04	1.00e−01	1.41
12	14	3.91	0.32	5.01e−04	1.17e−01	11.10
12	15	7.69	0.32	5.01e−04	1.17e−01	6.66
14	15	5.00	0.16	5.36e−04	1.25e−01	13.70
12	16	5.03	0.32	5.01e−04	1.17e−01	13.70
10	17	11.76	0.32	5.01e−04	1.17e−01	13.20
16	17	5.18	0.16	5.36e−04	1.25e−01	8.71
15	18	4.57	0.16	5.36e−04	1.25e−01	3.46
18	19	7.75	0.16	5.36e−04	1.25e−01	12.20
10	20	4.78	0.32	5.01e−04	1.17e−01	5.40
19	20	14.71	0.32	5.01e−04	1.17e−01	3.52
10	21	13.33	0.32	5.01e−04	1.17e−01	11.50
10	22	6.67	0.32	5.01e−04	1.17e−01	9.76
21	22	41.67	0.32	5.01e−04	1.17e−01	8.08
15	23	4.95	0.16	5.36e−04	1.25e−01	1.68
22	24	5.59	0.16	5.36e−04	1.25e−01	7.82
23	24	3.70	0.16	5.36e−04	1.25e−01	16.40
24	25	3.04	0.16	5.36e−04	1.25e−01	7.98
25	26	2.63	0.16	5.36e−04	1.25e−01	9.13
25	27	4.78	0.16	5.36e−04	1.25e−01	5.56
6	28	16.67	0.32	5.01e−04	1.17e−01	1.15
8	28	5.00	0.32	5.01e−04	1.17e−01	0.36
27	28	2.53	0.65	4.28e−04	1.00e−01	6.17
27	29	2.41	0.16	5.36e−04	1.25e−01	14.40
27	30	1.66	0.16	5.36e−04	1.25e−01	14.00
29	30	2.21	0.16	5.36e−04	1.25e−01	11.40

The nodal power demands are built from the load data given in [40] and reported in Figure 4 as daily profiles (accumulated according to the node indexes i), normally distributed on each hour t with mean $\mu_{i,t}$ and standard deviation $\sigma_{i,t}$. The technical data and uncertain model parameters of the different types of bulk–power suppliers and DG technologies, available to be integrated into the network, are given in Table 2. For the present case study, we consider that the number of photovoltaic cells per PV generation unit is $n_c = 20000$ and that the region covered by the system is such that the solar irradiation and wind speed conditions are uniform in the whole region, i.e., the values of the parameters of the corresponding Beta and Weibull distributions are taken equal for all nodes. The renewable power penetration factor is set to $PF = 30\%$ (constraint (24) in the MOO problem).

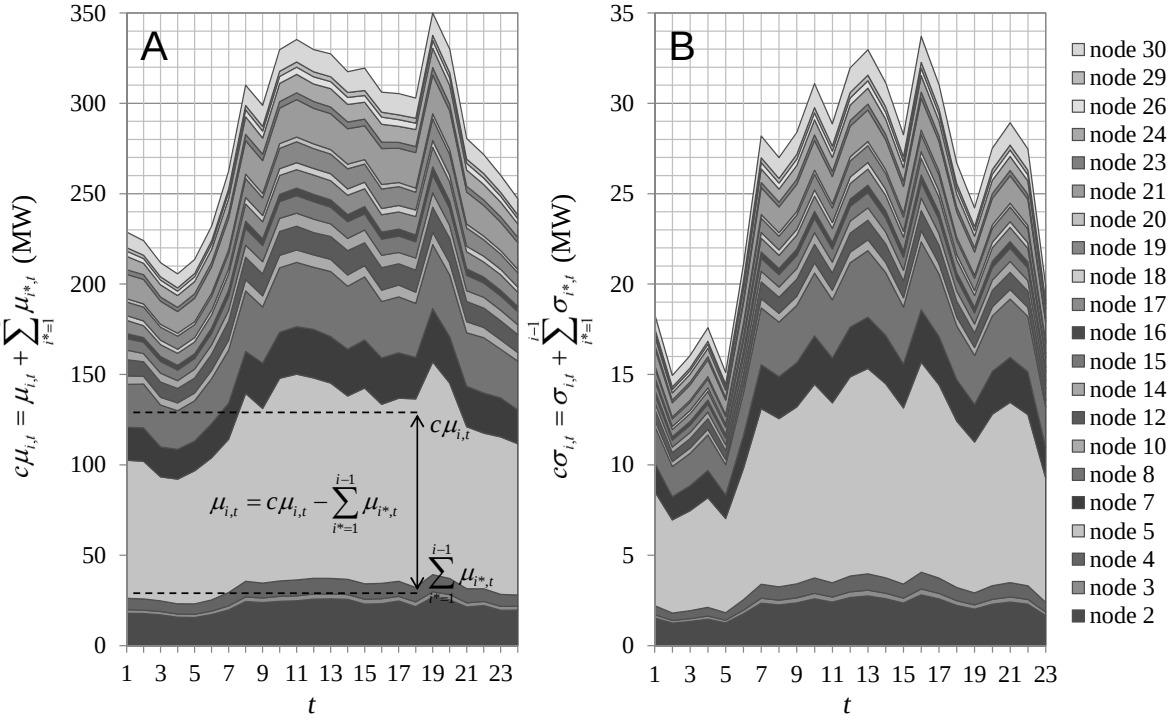

Figure 4: Accumulated mean (A) and standard deviation (B) values of nodal load daily profiles

Table 3 reports the failure and repair rates, λ_F and λ_R, respectively, the investment and operating costs $CI_j + COf_j$ and the variable operating cost of the different types of power generators.

Concerning the network economics, the maximum value of the energy price is $EP_{max} = 100$ (\$/MWh) [41, 42, 47] and the corresponding highest value of total demand $\Sigma Lmax_i$ (MW) is set to 445 (MW). The load shedding cost C_{LS} (\$/MWh) is considered as the maximum energy price. The horizon of analysis or lifetime of the project is 30 years, in which the investment and operating costs are hourly prorated. The confidence level or α-percentile considered to estimate the values $CVaR_\alpha$ is 75%, arbitrarily chosen.

Table 2: Power generators technical data and uncertain model parameters [20, 41, 46, 50, 57, 58]

Type		Technical parameters					Distribution parameters	
MG	G_1	P_{max_1} (MW)					Normal μ_1	Normal σ_1
		340					300	18.25
	G_2	P_{max_2} (MW)					Normal μ_2	Normal σ_2
		50					42.5	5
RG	PV	P_{AV} (MW)	P_{max} (W)	V_{OC}(V)	I_{SC} (A)	V_{MPP} (V)	Beta α	Beta β
		1.07	75	21.98	5.32	17.32	0.50	0.33
		I_{MPP} (A)	k_V (mV/°C)	k_I (mA/°C)	T_{No} (°C)	T_A (°C)		
		4.76	14.4	1.22	43	30		
	W	P_{AV} (MW)	P_R (MW)	U_{CI} (m/s)	U_A (m/s)	U_{CO} (m/s)	Weibull α	Weibull β
		0.93	1.5	5	15	25	15	2.2

Table 3: Power generators failure and repair rates and costs [20, 41, 46, 57–59]

Type		λ_F (n/h)	λ_R (n/h)	$CI + COf$ (M$/u)	COv ($/MWh)
MG	G_1	5.13e−04	2.77e−02	−	29.32
	G_2	6.84e−04	4.16e−02	−	8.92
RG	PV	6.27e−04	1.30e−02	2.20	9.69
	W	3.42e−04	9.00e−03	1.85	11.05

Finally, Table 4 summarizes the main parameters set for the general MOO DE and MCS–OPF framework.

Table 4: MOO–DE and MCS–OPF parameters

Parameter	Nomenclature	Value
Population size	NP	100
Maximum n° of generations	Gen_{max}	600
Crossover coefficient	COc	0.1
Differential variation amplification factor	F	1
N° of MCS–OPF scenarios	NS	24000
Scenario duration (h)	t^{Δ}	1

4.2 Results and discussion

The Pareto front resulting from the MOO–DE MCS–OPF is presented in Figure 5. The entire last generation population at convergence is shown by gray squares and the non–dominated solutions are the blue bullets. The base case (MG) in which no DG is integrated in the network is also shown as dark gray triangle. Each solution in the Pareto Front corresponds to a decision matrix $[Q^R]$ that indicates the number of units and locations of the different types of DG technologies integrated in the network.

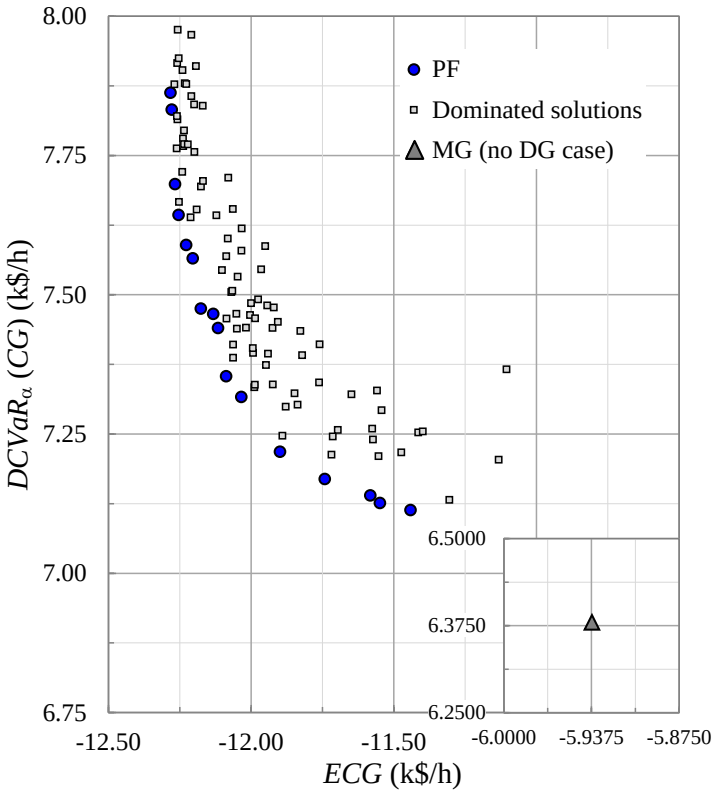

Figure 5: Set of non–dominated solutions: Pareto front

Recalling the relation given by equation (21), $DCVaR_\alpha(x) = CVaR_\alpha(x - E(x))$, it is possible to draw a map of iso–CVaR curves in the plot of non–dominated solutions and so, to include risk into the trade–off between expected performance and uncertainty, represented by ECG and $DCVaR_\alpha(CG)$, respectively. Then, we can find the compromised solution, namely $[Q^R]^{min}_{CVaR(CG)}$, that minimizes risk as reported in Figure 6(A). The reciprocal case, i.e., a map of iso–DCVaR curves is drawn in the distribution of non–dominated solutions plotted as ECG vs $DCVaR\alpha(CG)$, and it is shown in Figure 6(B). We look to three representative non–dominated solutions for the analysis: those with minimum values of ECG and $DCVaR_\alpha(CG)$, denoted $[Q^R]^{min}_{ECG}$ and $[Q^R]^{min}_{DCVaR(CG)}$ respectively, and $[Q^R]^{min}_{CVaR(CG)}$ which, as mentioned earlier, minimizes risk.

In Figure 6, we can observe that the three solutions of interest, $[Q^R]^{min}_{ECG}$, $[Q^R]^{min}_{CVaR}$ and $[Q^R]^{min}_{DCVaR}$, lead to considerable improvements in expected performance and risk with respect to the base case MG, in which no renewable generation is integrated into the network. However, the level of uncertainty in the ECG estimation is increased, in all DG–integrated solutions, because of their stochasticity. Even so, in comparison to the MG case, the increase in the level of uncertainty for all DG–integrated cases (on average 1.067 (k$/h)) is much less than the gain in both expected performance and risk (on average −6.035 and −4.967 (k$/h)), respectively). This fact can be seen also in the empirical CG probability density functions (*pdf*) shown in Figure 7.

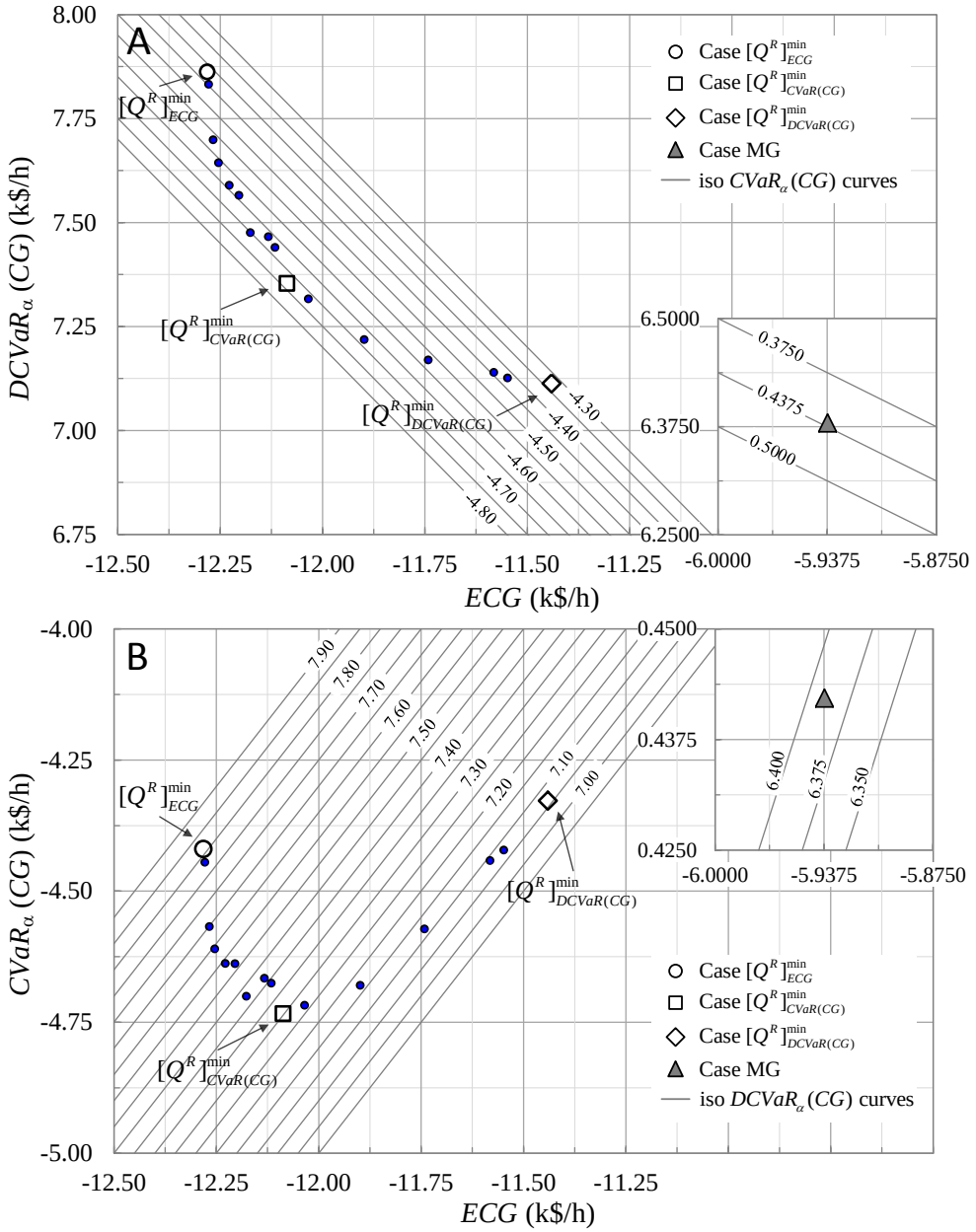

Figure 6: Set of non–dominated solutions with iso–CVaR (A) and iso–DCVaR (B) curves

Furthermore, it can be inferred from Figure 7 that, in general, all CG empirical *pdf*s show three main peaks. This is due to the characteristics of the daily load profiles (Figure 4) which present three important ranges: a low power demand range during the night, between 23 and 6 hours and two high ranges of load taking place in the intervals 10 to 13 and 18 to 21 hours, respectively. Thus, the left peak of the distributions corresponds to the highest range of loads (18 to 21 hours), because higher levels of power demands imply more energy sold and, therefore, more profits or negative values of CG. Following the same logic, the central peak is due to the second

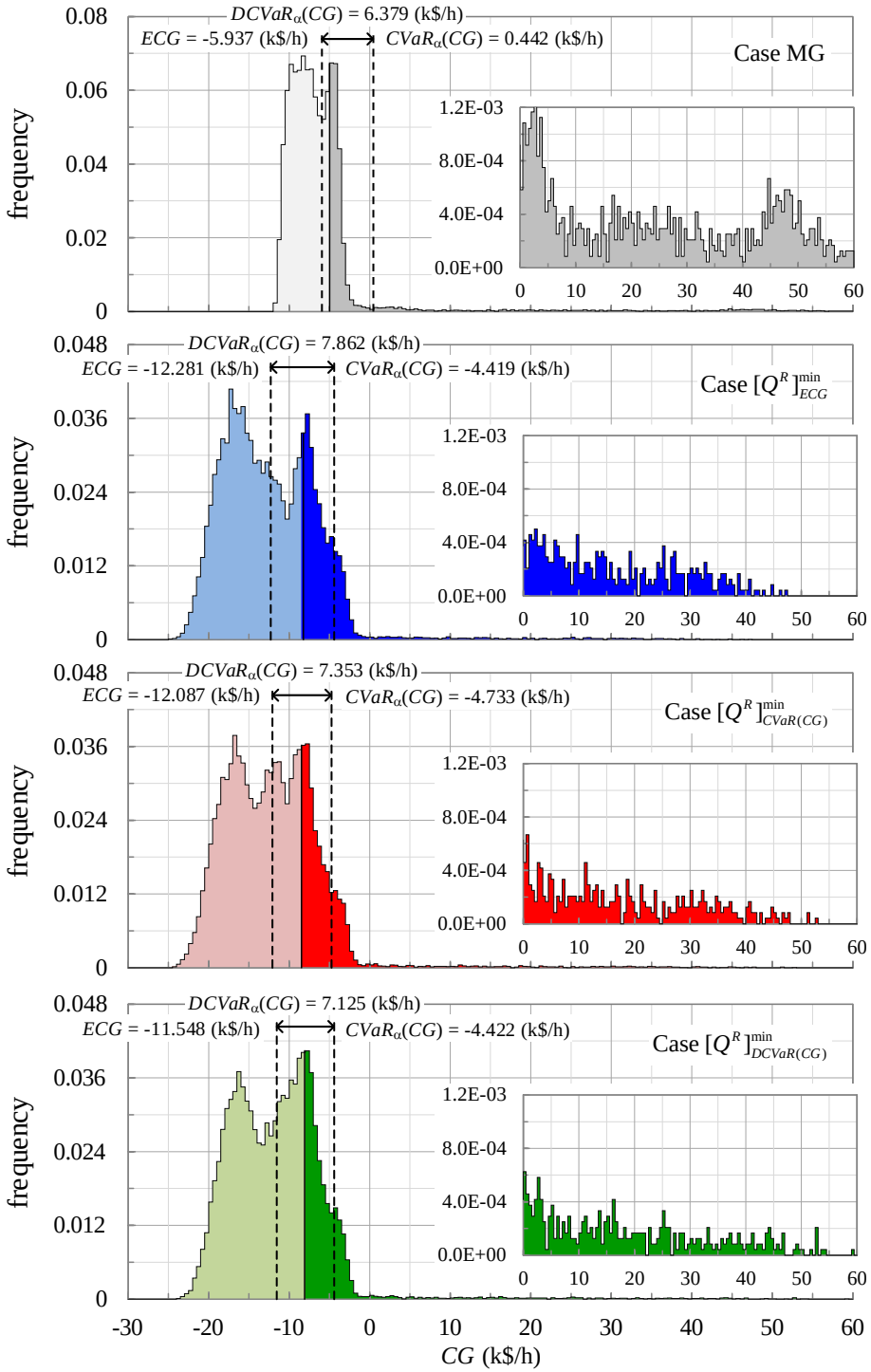

Figure 7: Empirical CG probability density functions

highest range of loads (10 to 13 hours) and the right peak to the low range of power demand. As mentioned before, the three DG–integrated network cases improve the cost profiles because the usage of power is transferred from the bulk–power suppliers (MG) to the renewable generators (RG) as summarized in Table 5, presenting the ratio of power usage defined as the proportion of power used to satisfy the loads, determined by the MCS–OPF, over the power available. According to the IEEE 30 bus test systems information, the bulk–power supply arriving at node 1, G_1, comes from a conventional coal–fired power plant whereas G_2 is supplied by a hydro–power plant. The nature of the source of power supply conditions the operating costs, in particular, G_1 that presents the higher variable operating costs (Table 3). Nevertheless, the ratio of power usage associated to G_2 is in all cases 100%, even though its operating cost is higher than the renewable. This is because no investment is being paid for G_2, contrary to the DG technologies (PV and W). In this view, considering investment, PV and W are more convenient than the coal–fired power supply G_1 but not more than the hydro–power supply G_2.

Table 5: Ratio of power usage by type of generator

Case	G_1	G_2	PV	W
MG	77.93%	100.00%	–	–
ECG^{min}	41.88%	100.00%	99.96%	99.94%
$CVaRCG^{min}$	40.38%	100.00%	99.95%	99.42%
$DCVaRCG^{min}$	40.71%	100.00%	100.00%	98.73%

The extreme CG scenarios encountered in the tail of the distributions are mainly produced by the occurrence of failures in the components of the system, power generators and T&D lines. In Figure 7, we notice the stability of the CG to these non–desirable events. Since the cost objective function to be minimized in the OPF considers a load shedding cost, the occurrence of failures in the components, interrupting the power supply and/or the ability of distribute it, will impact the CG function depending on how much centralized is the power supply. Focusing on the MG base case, the power supply and its distribution depend on two generators and the T&D lines connected to them: the reliability of power supply is determined by few components and so, the eventual losses of functionality of these components can lead to high amounts of non–satisfied demands. This is precisely the effect observed in the tails of the CG distributions and, in particular, we see that distributing and diversifying generation helps to improve the risk impacts from multiple failures in the network, even if the number of generation units in the system is increased and with it, the overall absolute likelihood of occurrence of failures.

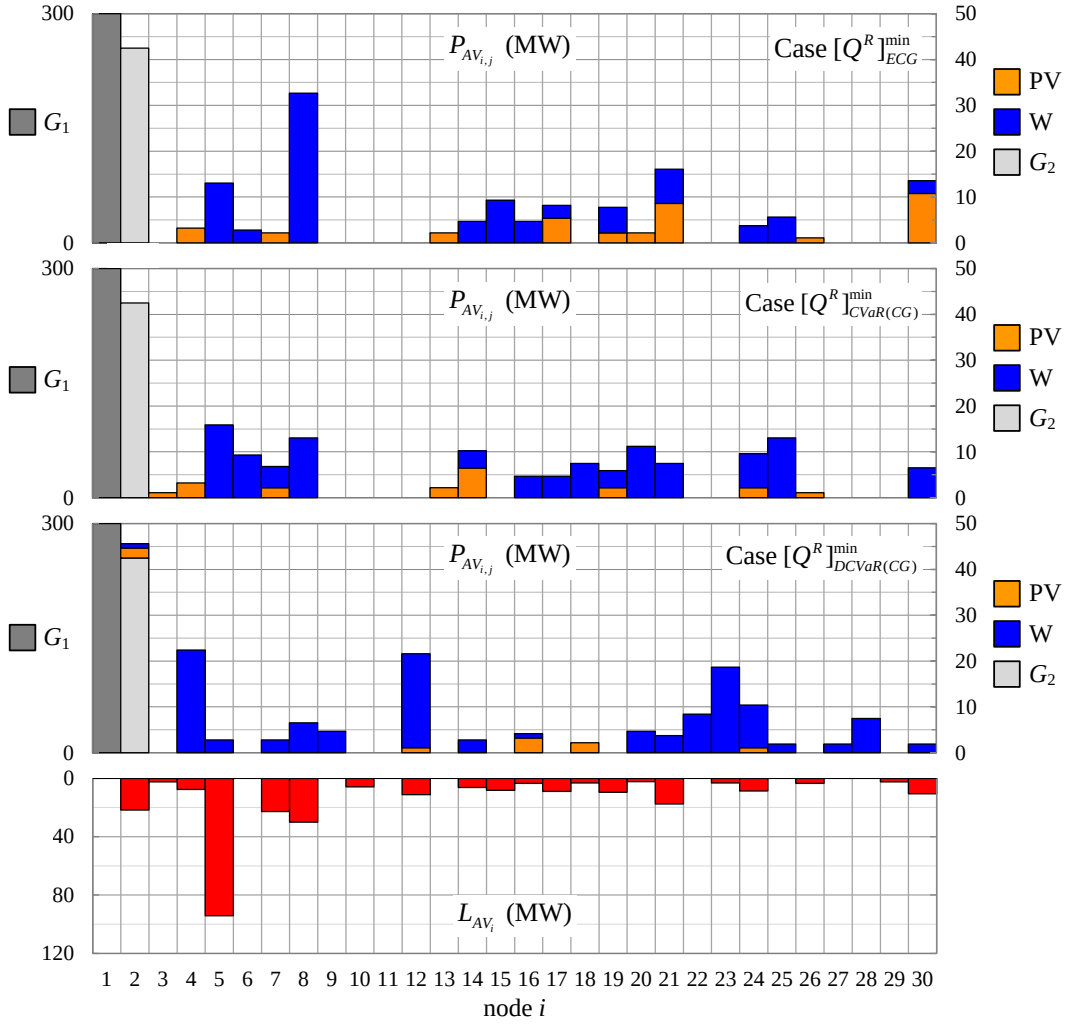

Figure 8: Total average renewable power installed. Cases ECG^{min} (A), $CVaRCG^{min}$ (B) and $DCVaRCG^{min}$ (C)

Figure 8 shows the total average DG power integrated in the network for the three solutions under analysis and the corresponding obtained values of ECG and $CVaR_\alpha$. It can be pointed out that in all cases the DG power installed is almost equal to the limit value set by the penetration factor PF. In the present case, a maximum of 30% is accepted, of average DG power installed over the maximum aggregated load in the system. This leads to an approximated limit of 135.5 (MW). Furthermore, in general terms moving along the non–dominated solutions, starting from the one that minimizes ECG ($[Q^R]^{min}_{ECG}$), passing to the compromising one that minimizes $CVaR_\alpha(CG)$ ($[Q^R]^{min}_{CVaR(CG)}$) and ending with ($[Q^R]^{min}_{DCVaR(CG)}$) that minimizes $DCVaR_\alpha(CG)$, the amount of PV power integrated in the network decreases progressively, being replaced by W power. Then, we infer that the more PV power generators are integrated, the better expected global cost performance is achieved. This seems physically coherent, taking into account that the average power outputs delivered by one generation unit of PV and W technologies are 1.07 and 0.93 (MW), respectively,

and the cost performance benefits comparatively more from the PV MW h sold during the daylight interval within the two peaks ranges of power demand. Nevertheless, the amount of average integrated PV power is invariably smaller than W power, this is because PV generation units do not supply during the night interval, making convenient to integrate always a certain amount of W power and, thus, to avoid resorting to the more expensive coal–fired power supplier G_1. Moreover, it is precisely this lack of PV generation during the night interval that strongly conditions the trade–off between expected performance and uncertainty, ECG vs $DCVaR_{\alpha}(CG)$, i.e., the more PV power is integrated, the better expected global cost performance is achieved but the more uncertain is its estimation.

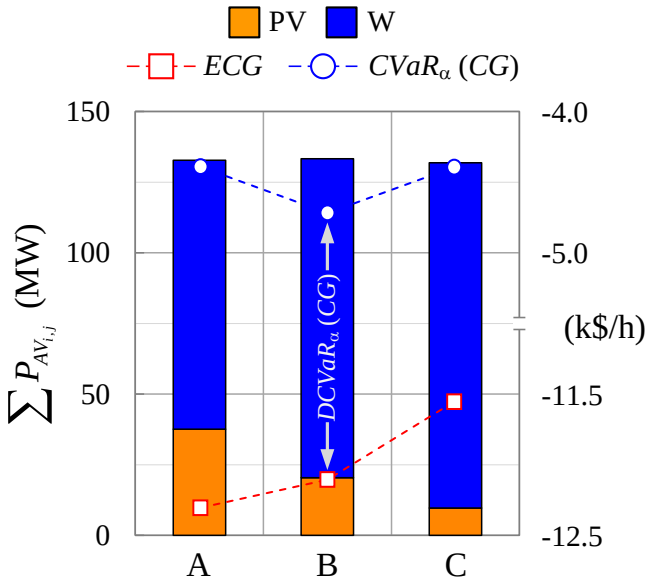

Figure 9: Nodal average power by type of generator

Concerning the compromising solution $[Q^R]^{min}_{CVaR(CG)}$ that minimizes risk, as it was derived from the empirical CG distributions, the risk associated to a solution depends mainly to the occurrence of extreme non–desired events. Improvements in risk performance can, then, be achieved by solutions for which the allocation of DG power generation units decentralizes and diversifies to a large extent the supply. This insight is noticeable in Figure 9 that reports the nodal average power by type of generation for the three solutions of interest. For both extreme solutions, $[Q^R]^{min}_{ECG}$ and $[Q^R]^{min}_{DCVaR}$, the tendency is to integrate localized sources of renewable DG at two identifiable portions of the network, in the region close to nodes 2, 5, 7 and 8 of the sub–transmission portion of the network, and nodes 17, 19 and 21 in the distribution part, favoring the sub–transmission portion which presents higher and non homogeneous nodal load profiles. In a different manner, the solution $[Q^R]^{min}_{CVaR(CG)}$ presents a more homogeneous deployment of DG power, allocating comparable generation capacities in both sub–transmission and distribution parts of the network.

5 Conclusions

We have presented a multi–objective optimization framework for the integration of renewable distributed generation into an electric power network. Multiple uncertain operational inputs are taken into consideration: the inherent uncertain behavior of renewable energy sources and power demands, as well as the occurrence of failures of components. For managing the uncertainty and risk associated to the achievement of a certain level of expected global cost performance, we have introduced the conditional–value–at–risk deviation measure, which allows trading off the level of uncertainty and, given the axiomatic relation to the conditional–value–at–risk, enables conjointly the trade–off of risk by constructing an iso–risk map in the non–dominated set of solutions. The proposed framework integrates the multi–objective differential evolution as a search engine, Monte Carlo simulation to randomly generate realizations of the uncertain operational scenarios and optimal power flow to evaluate the network response. The optimization is done to simultaneously minimize the expected value of the global costs and the respective conditional–value–at–risk deviation.

A case study has been analyzed, based on the IEEE 30 bus sub–transmission and distribution test system. The results obtained show the capability of the framework to identify Pareto optimal solutions of renewable DG units allocations. Integrating the conditional value–at–risk deviation into the framework has shown effectively the possibility of optimizing expected performances while controlling the uncertainty and risk, analyzing, in addition, the contribution of each type of renewable DG technology on the level of uncertainty associated to the outcome performance of the optimal solutions and the importance of the deployment of the renewable generation capacity to lower the risk of incurring in non–desirable extreme scenarios. In this view, a complete and comprehensible spectrum of information can be supplied in support of specific preferences of the decision makers for their decision tasks.

References

[1] R. Viral and D. Khatod, "Optimal planning of distributed generation systems in distribution system: A review," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 7, pp. 5146–5165, 2012.

[2] C. L. T. Borges, "An overview of reliability models and methods for distribution systems with renewable energy distributed generation," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 4008 – 4015, 2012.

[3] Y.-C. Chang, T.-Y. Lee, C.-L. Chen, and R.-M. Jan, "Optimal power flow of a wind-thermal generation system," *International Journal of Electrical Power & Energy Systems*, vol. 55, pp. 312 – 320, 2014.

[4] A. M. Imran, M. Kowsalya, and D. Kothari, "A novel integration technique for optimal network reconfiguration and distributed generation placement in power distribution networks," *International Journal of Electrical Power & Energy Systems*, vol. 63, pp. 461 – 472, 2014.

[5] G. Allan, I. Eromenko, M. Gilmartin, I. Kockar, and P. McGregor, "The economics of distributed energy generation: A literature review," *Renewable and Sustainable Energy Reviews*, vol. 42, pp. 543 – 556, 2015.

[6] J. J. Justo, F. Mwasilu, J. Lee, and J.-W. Jung, "AC-microgrids versus DC-microgrids with distributed energy resources: A review," *Renewable and Sustainable Energy Reviews*, vol. 24, pp. 387 – 405, 2013.

[7] S. Javadian, M.-R. Haghifam, M. F. Firoozabad, and S. Bathaee, "Analysis of protection system's risk in distribution networks with DG," *International Journal of Electrical Power & Energy Systems*, vol. 44, no. 1, pp. 688 – 695, 2013.

[8] E. Zio, M. Delfanti, L. Giorgi, V. Olivieri, and G. Sansavini, "Monte carlo simulation-based probabilistic assessment of dg penetration in medium voltage distribution networks," *International Journal of Electrical Power & Energy Systems*, vol. 64, pp. 852 – 860, 2015.

[9] M. Samper and A. Vargas, "Investment decisions in distribution networks under uncertainty with distributed generation - Part I: Model formulation," *Power Systems, IEEE Transactions on*, vol. 28, pp. 2331–2340, Aug 2013.

[10] S.-Y. Lin and A.-C. Lin, "RLOPF (risk-limiting optimal power flow) for systems with high penetration of wind power," *Energy*, vol. 71, pp. 49 – 61, 2014.

[11] A. Alarcon-Rodriguez, G. Ault, and S. Galloway, "Multi-objective planning of distributed energy resources: A review of the state-of-the-art," *Renewable and Sustainable Energy Reviews*, vol. 14, no. 5, pp. 1353 – 1366, 2010.

[12] V. Martins and C. Borges, "Active distribution network integrated planning incorporating distributed generation and load response uncertainties," *Power Systems, IEEE Transactions on*, vol. 26, pp. 2164–2172, Nov 2011.

[13] S. Ganguly, N. Sahoo, and D. Das, "A novel multi-objective PSO for electrical distribution system planning incorporating distributed generation," *Energy Systems*, vol. 1, no. 3, pp. 291–337, 2010.

[14] M. Gómez-González, A. López, and F. Jurado, "Optimization of distributed generation systems using a new discrete PSO and OPF," *Electric Power Systems Research*, vol. 84, no. 1, pp. 174 – 180, 2012.

[15] J. Aghaei, K. M. Muttaqi, A. Azizivahed, and M. Gitizadeh, "Distribution expansion planning considering reliability and security of energy using modified PSO algorithm," *Energy*, vol. 65, pp. 398 – 411, 2014.

[16] J. H. Zhao, J. Foster, Z.-Y. Dong, and K. P. Wong, "Flexible transmission network planning considering distributed generation impacts," *Power Systems, IEEE Transactions on*, vol. 26, pp. 1434–1443, Aug 2011.

[17] H. Hejazi, M. Hejazi, G. Gharehpetian, and M. Abedi, "Distributed generation site and size allocation through a techno economical multi-objective differential evolution algorithm," in *Power and Energy (PECon), 2010 IEEE International Conference on*, pp. 874–879, Nov 2010.

[18] H. Hejazi, A. Araghi, B. Vahidi, S. Hosseinian, M. Abedi, and H. Mohsenian-Rad, "Independent distributed generation planning to profit both utility and DG investors," *Power Systems, IEEE Transactions on*, vol. 28, pp. 1170–1178, May 2013.

[19] L. Arya, A. Koshti, and S. Choube, "Distributed generation planning using differential evolution accounting voltage stability consideration," *International Journal of Electrical Power & Energy Systems*, vol. 42, no. 1, pp. 196 – 207, 2012.

[20] M. Raoofat, "Simultaneous allocation of DGs and remote controllable switches in distribution networks considering multilevel load model," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 8, pp. 1429 – 1436, 2011.

[21] M. Shaaban, Y. Atwa, and E. El-Saadany, "DG allocation for benefit maximization in distribution networks," *Power Systems, IEEE Transactions on*, vol. 28, pp. 639–649, May 2013.

[22] Y. Zhang and G. Giannakis, "Robust optimal power flow with wind integration using conditional value-at-risk," in *Smart Grid Communications (SmartGridComm), 2013 IEEE International Conference on*, pp. 654–659, Oct 2013.

[23] J.-H. Wu and Y.-H. Huang, "Electricity portfolio planning model incorporating renewable energy characteristics," *Applied Energy*, vol. 119, pp. 278 – 287, 2014.

[24] A. Bhattacharya and S. Kojima, "Power sector investment risk and renewable energy: A japanese case study using portfolio risk optimization method," *Energy Policy*, vol. 40, pp. 69 – 80, 2012. Strategic Choices for Renewable Energy Investment.

[25] C. González-Pedraz, M. Moreno, and J. I. Pena, "Tail risk in energy portfolios," *Energy Economics*, vol. 46, pp. 422 – 434, 2014.

[26] D. Ramos, L. Camargo, E. Guarnier, and L. Witzler, "Minimizing market risk by trading hydro-wind portfolio: A complementarity approach," in *European Energy Market (EEM), 2013 10th International Conference on the*, pp. 1–8, May 2013.

[27] P. Vithayasrichareon, J. Riesz, and I. F. MacGill, "Using renewables to hedge against future electricity industry uncertainties - An Australian case study," *Energy Policy*, vol. 76, pp. 43 – 56, 2015.

[28] R. Fagiani, J. Barquín, and R. Hakvoort, "Risk-based assessment of the cost-efficiency and the effectivity of renewable energy support schemes: Certificate markets versus feed-in tariffs," *Energy Policy*, vol. 55, pp. 648 – 661, 2013. Special section: Long Run Transitions to Sustainable Economic Structures in the European Union and Beyond.

[29] S. Ahmed, M. Elsholkami, A. Elkamel, J. Du, E. B. Ydstie, and P. L. Douglas, "Financial risk management for new technology integration in energy planning under uncertainty," *Applied Energy*, vol. 128, pp. 75 – 81, 2014.

[30] L. Kitzing, "Risk implications of renewable support instruments: Comparative analysis of feed-in tariffs and premiums using a mean-variance approach," *Energy*, vol. 64, pp. 495 – 505, 2014.

[31] R. Rockafellar, S. Uryasev, and M. Zabarankin, "Generalized deviations in risk analysis," *Finance and Stochastics*, vol. 10, no. 1, pp. 51–74, 2006.

[32] M. Nazir and F. Bouffard, "Risk-sensitive investment in renewable distributed generation under uncertainty due to post-feed-in tariff policy," in *Developments in Renewable Energy Technology (ICDRET), 2012 2nd International Conference on the*, pp. 1–5, Jan 2012.

[33] J. Francis and D. Kim, *Modern Portfolio Theory: Foundations, Analysis, and New Developments*. Wiley Finance, Wiley, 2013.

[34] S. Martin, Y. Smeers, and J. Aguado, "A stochastic two settlement equilibrium model for electricity markets with wind generation," *Power Systems, IEEE Transactions on*, vol. 30, pp. 233–245, Jan 2015.

[35] S. Ghosh, J. Kalagnanam, D. Katz, M. Squillante, and X. Zhang, "Integration of demand response and renewable resources for power generation management," in *Innovative Smart Grid Technologies (ISGT), 2011 IEEE PES*, pp. –, Jan 2011.

[36] I. G. Moghaddam, M. Nick, F. Fallahi, M. Sanei, and S. Mortazavi, "Risk-averse profit-based optimal operation strategy of a combined wind farm-cascade hydro system in an electricity market," *Renewable Energy*, vol. 55, pp. 252 – 259, 2013.

[37] M. A. Tajeddini, A. Rahimi-Kian, and A. Soroudi, "Risk averse optimal operation of a virtual power plant using two stage stochastic programming," *Energy*, vol. 73, pp. 958 – 967, 2014.

[38] R. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *Journal of Banking & Finance*, vol. 26, no. 7, pp. 1443 – 1471, 2002.

[39] R. Storn and K. Price, "Differential evolution - A simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.

[40] University of Washington, College of Engineering, Department of Electrical Engineering, Available at http://www.ee.washington.edu/research/pstca.

[41] H. Falaghi, C. Singh, M.-R. Haghifam, and M. Ramezani, "DG integrated multistage distribution system expansion planning," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 8, pp. 1489 – 1497, 2011.

[42] H. Ren and W. Gao, "A MILP model for integrated plan and evaluation of distributed energy systems," *Applied Energy*, vol. 87, no. 3, pp. 1001 – 1014, 2010.

[43] W.-S. Tan, M. Y. Hassan, M. S. Majid, and H. A. Rahman, "Optimal distributed renewable generation planning: A review of different approaches," *Renewable and Sustainable Energy Reviews*, vol. 18, pp. 626 – 645, 2013.

[44] E. Zio, *The Monte Carlo Simulation Method for System Reliability and Risk Analysis*. Springer Series in Reliability Engineering, Springer London, 2013.

[45] D. Kroese, T. Taimre, and Z. Botev, *Handbook of Monte Carlo Methods*. Wiley Series in Probability and Statistics, Wiley, 2013.

[46] Y. Atwa, E. El-Saadany, M. Salama, and R. Seethapathy, "Optimal renewable resources mix for distribution system energy loss minimization," *Power Systems, IEEE Transactions on*, vol. 25, pp. 360–370, Feb 2010.

[47] H. Ren, W. Zhou, K. Nakagami, W. Gao, and Q. Wu, "Multi-objective optimization for the operation of distributed energy systems considering economic and environmental aspects," *Applied Energy*, vol. 87, no. 12, pp. 3642 – 3651, 2010.

[48] A. Thornton and C. R. Monroy, "Distributed power generation in the United States," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 9, pp. 4809 – 4817, 2011.

[49] Y. Li and E. Zio, "Uncertainty analysis of the adequacy assessment model of a distributed generation system," *Renewable Energy*, vol. 41, pp. 235 – 244, 2012.

[50] Y.-F. Li and E. Zio, "A multi-state model for the reliability assessment of a distributed generation system via universal generating function," *Reliability Engineering & System Safety*, vol. 106, pp. 28 – 36, 2012.

[51] D. Van Hertem, J. Verboomen, K. Purchala, R. Belmans, and W. Kling, "Usefulness of DC power flow for active power flow analysis with flow controlling devices," in *AC and DC Power Transmission, 2006. ACDC 2006. The 8th IEE International Conference on*, pp. 58–62, March 2006.

[52] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans, "Usefulness of DC power flow for active power flow analysis," in *Power Engineering Society General Meeting, 2005. IEEE*, pp. 454–459 Vol. 1, June 2005.

[53] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *Evolutionary Computation, IEEE Transactions on*, vol. 6, pp. 182–197, Apr 2002.

[54] J. Branke, K. Deb, K. Miettinen, and R. Slowinski, *Multiobjective Optimization: Interactive and Evolutionary Approaches*. LNCS sublibrary: Theoretical computer science and general issues, Springer, 2008.

[55] P. Buijs, D. Bekaert, D. Van Hertem, and R. Belmans, "Needed investments in the power system to bring wind energy to shore in Belgium," in *PowerTech, 2009 IEEE Bucharest*, pp. 1–6, June 2009.

[56] O. Alsac and B. Stott, "Optimal load flow with steady-state security," *Power Apparatus and Systems, IEEE Transactions on*, vol. PAS-93, pp. 745–751, May 1974.

[57] K. Zou, A. Agalgaonkar, K. Muttaqi, and S. Perera, "Multi-objective optimisation for distribution system planning with renewable energy resources," in *Energy Conference and Exhibition (EnergyCon), 2010 IEEE International*, pp. 670–675, Dec 2010.

[58] U.S. Energy Information Administration, Available at http://www.eia.gov/.

[59] Z. Liu, F. Wen, and G. Ledwich, "Optimal siting and sizing of distributed generators in distribution systems considering uncertainties," *Power Delivery, IEEE Transactions on*, vol. 26, pp. 2541–2551, Oct 2011.