Luis Miguel Barbosa Proença

Towards a Comprehensive Methodology

DEEC
FEUP
1997

Faculdade de Engenharia da Universidade do Porto

Departamento de Engenharia Electrotécnica e de Computadores

Towards a Comprehensive Methodology for Power System Planning

Doctoral dissertation submitted to

the Faculty of Engineering of the University of Porto, Portugal

DEEC
FEUP

Luis Miguel Barbosa Proença

November 1997

Faculdade de Engenharia da Universidade do Porto

Departamento de Engenharia Electrotécnica e de Computadores

Towards a Comprehensive Methodology for Power System Planning

Doctoral dissertation submitted to

the Faculty of Engineering of the University of Porto, Portugal

Luis Miguel Barbosa Proença

Supervisor: Prof. Vladimiro Miranda, FEUP

621.3(043)/PROE/TOW

cat

200068

26.03.98

November 1997

Partially financed by JNICT (Project PRAXIS 3/3.1/TIT/43/94)

Basic research is what I am doing when I don't know what I am doing

Werner von Braun

Acknowledgements

Feeling somehow relieved for finally finishing the writing of my thesis, it comes the time to express my gratitude to all the wonderful people that, directly or indirectly contributed to the success of this work.

Particularly, I would like to show my deepest appreciation to my supervisor, Prof. Vladimiro Miranda for his stimulating guidance, his inspiring counseling and for giving me a hard time (for a good cause, certainly) during these last eight years. I know I will cherish that for all the years to come.

I am also grateful to Profs. Manuel Matos and Peças Lopes for providing me with a pleasant research environment, and for all the intelligent and insightful conversations we had during the course of my research. Obviously, I could not forget my colleagues and friends in the power systems group, whose support and advice were immensely appreciated. One would never find a better place to work!

Another special acknowledgment goes to Viktoria Bochkareva, Aleksandar Dimitrovski, José Antonio Domínguez, Prof. Erkki Lakervi and José Luis Bernal for all the interesting and useful discussions on planning and, of course, for their friendship.

I must also thank the people in Electricidade do Norte S.A., and especially to L. Oliveira and L. Carvalho for the excellent team work and cooperation provided. Thanks to João Ranito for the Genetic Algorithms platform, to João Germano for the clustering software and to Lou Chin Wang from the university of Macau for providing the results concerning capacitor placing studies.

The continuous support from my family was fundamental and I would especially like to thank my father for carefully reviewing the originals of this thesis.

Finally I would also like to thank INESC for providing me with excellent work conditions and JNICT for its financial support.

Miguel Proença ☺

Abstract

This thesis reports the results of a research project aimed at developing a new comprehensive methodology for the expansion planning of distribution systems. Realizing that this process is a matter of decision making and not of optimization, other than centering its concerns on the computational methods, the methodology focuses on the conceptual basis of planning, in an attempt to fully capture its basic nature.

Therefore, and grounded on innovative techniques such as genetic algorithms and fuzzy set models, the methodology takes into account several aspects usually neglected in previous approaches, like multiple criteria and a thorough representation of uncertainties.

In particular, the thesis proves that an approach based on evolutionary computing is both feasible and advantageous, offering more information and a better insight than classical optimization methods.

The thesis also presents new concepts to power system planning, such as the tree of fuzzy futures, fuzzy inadequacy and solution robustness. Decisions are taken in a explicit multicriteria environment under risk analysis policies, namely by minimizing future regrets.

The merits of the approach are discussed by analyzing its application to a case study based on a large distribution network.

Finally, the thesis presents some complementary studies on capacitor placing and a comparison between the paradigms considered to be fundamental in power system planning: *probabilistic choice/optimization* and *risk analysis/decision*.

Resumo

Esta tese apresenta os resultados de um projecto de investigação destinado a desenvolver uma nova metodologia integrada para o planeamento da expansão de sistemas de distribuição de energia eléctrica. Compreendendo o facto de que este processo está relacionado com decisão e não com optimização, em vez de centrar as preocupações nos métodos computacionais, a metodologia explora a base conceptual do planeamento, numa tentativa de capturar a sua essência.

Assim, e baseada em técnicas inovadoras tal como os algoritmos genéticos e fuzzy sets, a metodologia considera diversos aspectos geralmente ignorados nas abordagens tradicionais, tal como a consideração de critérios múltiplos e uma representação integral das incertezas.

Em particular, esta tese prova que uma abordagem baseada em computação evolucionária é útil e vantajosa, oferecendo mais informação e permitindo uma melhor compreensão do problema do que as abordagens baseadas métodos de optimização clássicos.

A tese também apresenta novos conceitos em planeamento de sistemas de energia, tal como os conceitos de árvore de futuros fuzzy, inadequação fuzzy e robustez de uma solução. As decisões são tomadas num ambiente multicritério explícito, sob análise de risco, nomeadamente minimizando os arrependimentos futuros.

Os méritos desta abordagem são discutidos pela análise dos resultados da aplicação desta metodologia a um sistema de distribuição de grande dimensão.

Finalmente, a tese apresenta alguns estudos complementares na colocação de baterias de condensadores, bem assim como uma análise comparativa dos dois paradigmas fundamentais em planeamento de sistemas de energia: *escolha probabilística/optimização* e *análise de risco/decisão.*

Résumé

Cette thèse présente les résultats d'un projet d'investigation destiné à développer une nouvelle méthodologie intégrée pour la planification de l'expansion de systèmes de distribution d'énergie électrique. Tenant compte du fait que le processus est lié à la décision et non à l'optimisation, au lieu de centrer les préoccupation sur les méthodes logicielles, la méthodologie exploite la base conceptuelle de la planification, tentant capturer son essence.

Ainsi, et basée sur des techniques innovatrices telles que les algorithmes génétiques et fuzzy sets, la méthodologie tient compte de différents aspects ignorés, en général, par les approches traditionnelles, tels que la considération de critères multiples et une représentation intégrale des incertitudes.

En particulier, cette thèse prouve qu'une approche basée sur le logiciel évolutif est utile et avantageux, offrant davantage d'information et permettant une meilleure compréhension du problème que les approches basées sur des méthodes classiques d'optimisation.

La thèse présente aussi de nouveaux concepts de planification de systèmes d'énergie, tels que les concepts d'arbre de futurs fuzzy, inadéquation fuzzy et robustesse d'une solution. Les décisions sont prises dans un environnement multicritères explicite, sous analyse de risque, notamment en minimisant les regrets futurs.

Les mérites de cette approche sont abordés par l'analyse des résultats de l'application de cette méthodologie à un système de distribution de grande dimension.

Enfin, la thèse présente quelques études complémentaires sur l'emplacement de batteries de condensateurs, ainsi qu'une analyse comparative de deux paradigmes fondamentaux sur la planification de systèmes d'énergie: Choix probabiliste/optimisation et analyse de risque/décision.

Index

1. INTRODUCTION

The wise man doesn't give the right answers, he poses the right questions.

Claude Levi-Strauss

1.1. Scope of the thesis

The systematic and efficient planning of the expansion of electrical distribution systems is, presently, object of universal interest, both in research groups and utilities. Considering that, in industrialized countries, around 50% of electricity investments are associated with distribution, it is obvious that even a small improvement in efficiency through better planning would have a considerable economic influence.

Since the 60s several models for distribution system planning have been proposed and discussed by the scientific community and within the utilities. However, unlike their counterparts on other areas of Power System planning (e.g. on hydrothermal operation planning, generation and transmission expansion planning, etc.), one should recognize that none of the models proposed in these three decades was very successful or had a wide practical dissemination.

The reasons behind this divorce between models and practical industrial applications seem to be the following:

1. The fact that the decisions in planning are strongly conditioned by **uncertainties**. Early models simply ignored this basic aspect of planning. Other models would use a representation of uncertainties that we could consider as crude or even naïf - as a general procedure, methodologies would try to find the "optimal" solution for a set of possible scenarios, each of these studied as if it were deterministic.

2. The fact that planning is intrinsically a **dynamic** process, in the sense that future decisions influence decisions that must be made in the present. Many models were proposed for static situations, then series of static models were chained in order to create a pseudo-dynamic approach.

Some models that had dynamic properties resulted both very heavy (only applicable to small systems) and as very simplified approaches to real problems, due to the use of "well-behaved" functions, needed because of computing limitations.

3. Targeting at optimization (the search for the Holy Grail of the "optimal solution") has obscured the real aim of planning - in general, costs (investment + power losses) were minimized and everything else was treated as constraints. Therefore, the sense of **multiple criteria** to judge the solutions was absent from the models.

4. Most models accept, without challenging, the traditional paradigm of planning based on probabilistic optimization, where the chosen solution would be the one that minimizes the expected cost over the set of futures considered. In recent years, however, this concept has been challenged and as an alternative the paradigm of **risk analysis** was proposed. The concepts behind this last paradigm (for example, the concept of hedging) and the way they deal with the problem are much closer to the way planners and decision makers think.

5. Finally the need to deal with **massive quantities of information**. In the past this was clearly one, if not the most important constraint in planning. The need of dealing with huge amounts of information is related, not only to the characteristics of the problem itself, but also to the recognition of the fundamental importance of the use of *Geographical Information Systems* in distribution network planning. These systems allow a greater identification between reality and models, presenting possible alternatives in a realistic manner. However, the impressive increase in computational, data processing and graphical representation power in the last few years clearly shows that this will no longer be a major concern for scientists and planners.

The main point here is that, above all, these models were not able to capture the real essence of planning, justifying the reduced interest in utilities.

This aim of this thesis, however, is not to provide the ultimate answer to all of these concerns. It will also not describe a full planning system intended for industrial use.

The problem of dealing with vast amounts of information (point 5) is already being addressed by researchers and utilities namely with the use of efficient GIS systems and the use of powerful computational tools. This thesis intends to address the conceptual part of the problem (points 1-4) with a strong focus on decision making.

In essence, the thesis states the following:

> *It is now possible to develop a full comprehensive methodology for distribution system expansion planning.*

The time has arrived to join all the new tools, all the imaginative and efficient methods for the search of solutions, and the now available computational power in an integrated tool for distribution planning. The global image of such a system is depicted in the diagram shown in Figure 1.

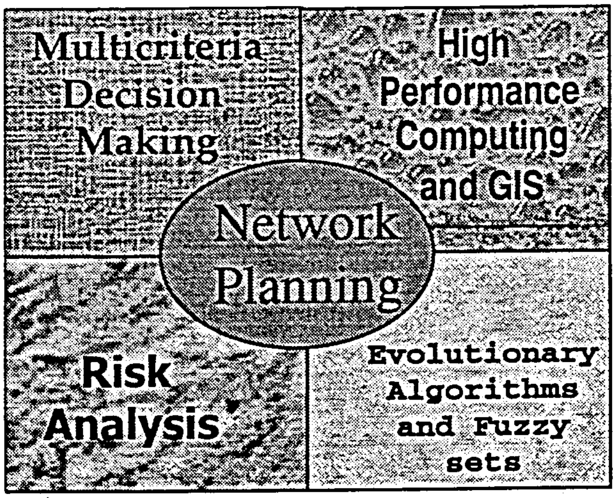


Figure 1 All tools fit together in a comprehensive methodology

The author believes that, in this moment, we are able to suggest a new methodological and instrumental perspective when approaching the problem. This new perspective is composed by the following main aspects:

- Abandoning the optimization paradigm, replacing it by an explicit multicriteria analysis.
- Concerns are centered on decisions, instead of solutions, by adopting a methodology based on risk analysis replacing the vision of optimizing for the average future.

- The representation of uncertainties is done in a plural and comprehensive manner, using simultaneously probabilistic techniques, scenario techniques and fuzzy set theory.

- The use of powerful tools like evolutionary algorithms for the determination of dynamic alternatives.

1.2. Structure of the thesis

The subject of distribution planning is a vast and complex one. It would be impossible to include in the thesis all the details on the developed methodologies and algorithms, and all the results obtained from a large number of tests and studies performed during the last 6 years. The author opted instead to concentrate on the essential matter and on the most interesting results, especially from already published papers and reports.

The following paragraphs summarize the material presented in each chapter. The author intended to write each chapter in a modular way, allowing the reader to concentrate on a specific topic on distribution planning. Therefore, parts of the material may be duplicated in different chapters.

Chapter 2 introduces the distribution expansion planning problem, reviewing its criteria, technical considerations and dynamics. It proposes a multiobjective analysis of the problem in a decision making environment, and the abandoning of the optimization paradigm.

Chapter 3 does a critical review of some existing planning models, classifying them according to their characteristics and scope. Recognizing the complexity and dimension of the planning problem, it sets the basis for the discussion on the divorce between academic models and industrial planning systems.

Chapter 4 proposes a general model for the representation of uncertainties and risk in power system expansion planning. In this model, all uncertainties are represented according to their nature by combining three different basic techniques: probabilistic models, scenario trees, and fuzzy models. The chapter introduces some new concepts derived from the fuzzy definition of loads: inadequacy, robustness and exposure.

Chapter 5 compares conventional algorithmic approaches generally used in distribution expansion planning with some new emerging techniques with particular emphasis on evolutionary algorithms. It also discusses the advantages and limitations of each technique and presents an overview on the use of evolutionary algorithms in power systems.

Chapter 6 presents a new comprehensive methodology for multistage distribution system planning based on genetic algorithms and on a thorough representation of uncertainties.

In order to discuss the merits of the approach proposed, chapter 7 analyses its application to a study on distribution expansion planning in order to highlight its most characteristic features.

Chapter 8 presents some complimentary studies on distribution network planning, namely on the area of reactive power planning (capacitor placing).

Chapter 9 proposes the Risk Analysis paradigm as an alternative to the traditional Probabilistic Choice approach to planning. It discusses and compares the fundamentals and underlying assumptions of the two basic paradigms, illustrating the discussion with a distribution network planning example.

Chapter 10 presents all the relevant conclusions from the thesis and shows possible directions for future work.

2. THE DISTRIBUTION PLANNING PROBLEM

All easy problems have been solved.

Anonymous

2.1. Summary

This chapter presents the distribution planning problem (DPP) as a large scale, dynamic and multiobjective problem subjected to a large range of uncertainties. In this perspective, it shows that the DPP is matter *of decision making* and not *optimization* and proposes the application of the principles of risk analysis to the DPP in order to replace the traditional single-criterion deterministic paradigm.

2.2. Introduction

The objective of this chapter is to go a little beyond the simple presentation of the distribution planning problem, by establishing the cornerstones of a discussion on the <u>conceptual part</u> of planning. In [Wil97], H. Lee Willis reflects on this question with clarity:

> *"**Identify goals and needs prior to selecting a tool.** Planners should explicitly study and document their goals and needs, and write down a prioritized list of their needs before picking a method, gathering data, or even setting a budget. This list should be used to evaluate potential planning tools."*

In addition to introducing the characteristics of the problem, this chapter <u>questions the traditional paradigm of power system planning</u>, a paradigm rooted in a certain cultural backdrop still present both in utilities and in research groups worldwide. This conventional view of the planning activity has the following characteristics that should be emphasized:

- <u>Single-criterion</u>: it adopts a dominant criterion (often unique) for the selection between alternatives for system expansion – the global cost criterion – reducing the importance of other criteria to simple technical constraints. The sense of multiple criteria to judge the solutions is simply absent from the models.

- Deterministic: most of the models based on the traditional view neglect the fact that we are dealing with a dynamic multitemporal problem under uncertainty.

- Optimistic: it believes that the future can be accurately predicted and that forecast errors will cancel or compensate each other. When the models consider uncertainty, they try to optimize for the so-called "average future". However, the future happens only once for every decision, and the average future, most of the time, will simply not occur.

It is curious, nevertheless, to notice how this paradigm still persists in utility practice, even though these aspects have been subjected to reflection and criticism for a long time now. Planning decisions are each time more subjected to conflicting multiple criteria. On the other hand, there has been a solid change in the culture and perspective of utilities, governments and public towards energy in general and several additional uncertainty factors.

Although the technical community recognizes that a simple exercise in optimization constitutes an unacceptable approach to planning, it appears odd when we notice that the models that have been presented in the literature and in scientific meetings are still very much attached to the conventional perspective.

The identification of different types of uncertainties and the fact that we cannot simply reduce them to probabilistic models, is an important step to the recognition of the limitations of the traditional paradigm, This specific and fundamental issue will be addressed in some depth throughout the dissertation.

This chapter, and the dissertation as a whole, will be devoted to present the distribution planning problem under a new light, coming as convergence of several contributions consisting of new ideas, concepts and techniques developed in the last few years. It will put a strong emphasis on the essential points of the new emerging paradigm: risk analysis and decision making. We will no longer be searching for optimal solutions but for strategies based on robust decisions, in an uncertainty and multiple criteria framework, driven by the consideration of the risks concerned.

The chapter will also present a historical introduction to distribution planning concentrating on the mentioned aspects. First, some basic definitions and terms utilized on the remaining of the thesis will be presented.

2.3. Basic definitions

This section introduces some basic definitions that will be adopted in the dissertation. Other definitions will be presented in the remaining sections of the chapter. Part of these definitions follows [Mer90].

Planning

- A utility chooses from a set of possible **Options**: build a new primary substation, upgrade an interconnection, build a new feeder, close a loop, etc.
- A **Plan** is a set of specified options – for example, "build a new primary substation 60/15 kV in 1997 and reinforce it in 50 MW in 1998".
- A **Strategy** is a set of coordinated plans to be implemented according to the future that actually occurs: "*if* load grows more than 15% until 1999 *then* build a 50 MW primary substation by 2000 *and* reinforce MV interconnection, *otherwise* build a 40 MW primary substation by 2002 *and* close loop L_5."

Multiobjective analysis

- **Alternatives** are instances (options, plans or strategies, for our purposes) that represent possible solutions to the problem in cause. Each alternative has several parameters to be specified.
- **Criteria** are general aspects about the problem that influence decisions. Examples of criteria: reliability, costs, environmental impact, security, etc.
- **Attributes** are measures of the goodness of a solution or alternative, i.e., the quantification of a given criterion. For example, if the criterion is *environmental impact* the attributes might be *total SO_2 emissions, km of*

roads in sensitive areas, etc. Other examples of attributes are: investments costs, operational costs (criterion: costs), energy not supplied, mean time between failures (criterion: reliability).

- **Objectives**, are goals concerning the attributes, in general, minimization or maximization (or, in some cases, reaching a target level). For example: *minimize SO_2 emissions, minimize investment cost.* Although the term *objectives* is often confused with *criteria*, the distinction is quite clear: sometimes, for the same criteria and attributes, the objectives may be different, depending on the decision maker. For example - criterion: car size. A decision maker's objective may be *minimize size* while another may want to *maximize* it or even another that may want to reach a certain target size for his car.

- **A Non-Dominated Alternative** is one alternative that, when compared to any other alternative, is better in at least one objective. That means that no other alternative is better than a non-dominated alternative in all of the objectives. If that happened, this alternative would be referred as *dominated.*

- **Ideal** is a fictitious alternative that would have, as values for its attributes, the best value found for each attribute.

- A **Decision Set** is a set of non-dominated alternatives.

- A **Conditional Decision Set** is a set of alternatives that may be taken as acceptable compromises between objectives and, therefore, will appear as good candidates to a final decision, in a given future.

Uncertainties and risk

- **Uncertainties** are factors that are not under the control of the utility or cannot be predicted with precision: load variations, economical background, etc.

- **Future** is a set of outcomes or realizations of uncertainties: "3% in load growth, 1% fuel price increase, project for large sports hall approved".

- **Scenario** is a complete set of specified options and uncertainties, i. e., a specific plan in a particular future. Not infrequently, the expressions *future* and *scenario* are used interchangeably and, in this dissertation, the word *scenario* will sometimes be used in the sense of *future*.

- **Risk** is the hazard to which a utility is exposed, if it selects one solution instead of another. Therefore, the concept of risk is associated to decisions and their consequences, in conjunction with the deviations that reality suffers relatively to the forecasted futures. One way of measuring risk is in function of attributes, comparing them to an ideal solution, if one could guess the future.

- **Regret** is the measure of what we lose in certain criteria when comparing it to the ideal solution. In a multiobjective problem, if there is more than one non-dominated solution, any decision will inevitably lead to regret in, at least, one of the criteria. It is important to notice that the possibility of a certain decision to generate regret, and its level, is not related to the expected value of any variable or attribute. While the expected value of a variable (for instance, average repair time) is calculated *a priori* from a given sample, regret is an *a posteriori* concept, since its assessment can only be made after the decision is taken and the occurrence of the future.

2.4. The distribution expansion planning problem

With the fast increase in electric energy consumption, the distribution planning problem has been revealed as one of the most important to utilities, considering the important investments that have to be made in that area.

Expansion studies are based on the existing system, load forecast, extensive economical and electrical calculations and on the planner's experience and judgment. However, the development of more elaborated studies, with a larger number of alternative projects and the use of better mathematical models and methodologies may improve considerably the traditional solutions obtained by planners. For a certain distribution system, the present load level is known and future loads may be forecasted for a single stage (generally one year) or for several stages. Consequently, the generic problem consists in planning the system expansion (decisions concerning construction and placing of lines, substations and various equipment) in order to satisfy consumption according to four basic and sometimes conflicting criteria: *Economy*, *Reliability*, *Security* and *Quality*.

Distribution planning is also conditioned by local geographic and demographic aspects: population evolution has a strong influence in future consumption and local geography also plays an important role in total investments [Lea95].

Until approximately the 70s, in the industrialized world, the main concern for utilities (generally, state-owned, in a protectionist market) was to foment consumption, bringing electrical energy to every region. In those decades, *expansion* meant *electrification* and, consequently, distribution expansion costs represented a high share of total investment costs. This tendency decelerated when the initial objectives (extensive electrification) were being met. By that time, consumers began to be aware of their rights, assumed a different attitude by demanding a better product, with the result of gradually shifting the focal point towards a new objective: quality.

This setting frames the traditional conflict: one that *opposes investment and operation costs* to *service quality*. However, circumstances have been changing rapidly and, more recently, other factors demand an independent and equally important consideration:

- The expansion in independent generation directly connected to distribution networks (mini-hydro, wind power, industrial co-generation, etc.) introduces additional uncertainty. On one hand, independent generation sources depend heavily on natural resources. On the other hand, this means that <u>external decisions</u> will have a growing importance in defining the adequacy of the distribution network.

- The increasing importance of the formation of an <u>energy market</u>, meaning that a certain utility will lose the absolute control of normal events in its network and in the connection and behavior of consumers [Pam95], [Sar96].

- The firm demand, from the consumers, of better service quality, including guarantees in the number and duration of interruptions, having as consequence, besides <u>new contractual practices</u>, the fact that network investments become more motivated by this factor than by load growth (as long as a minimum level of economic development is reached).

- The recent tendency of tightening design margins, lengthening equipment life times and increasing regulatory scrutiny.
- The emphasis on rational use of energy, emerging from its humble role in forecast tuning, to an important investment policy tool in utilities hoping to extract some benefits from it, namely by deferring future investments.
- The emergence of Demand Side Management (DSM) and of new incentive politics as a way of controlling peak load [Gel89].
- The growing importance of social and environmental impact of decisions, particularly concerning its future consequences.

Network dimension

Another conditioning aspect in expansion planning is dimension: the large number of injection nodes and distribution lines and, on the other hand, the dynamic nature of the problem leads to a large number of variables. Furthermore, since we are dealing with a finite number of possible solutions the problem may be considered as combinatory. These facts and the innumerous technical aspects (ranging from short-circuit concerns to switching policy), make any approach to the DPP a difficult and heavy task.

This discussion on the complexity of the problem leads us to another aspect in our analysis: the fact that planning is, in its essence, dealing with an uncertain future. This aspect will be detailed in the following section.

2.5. Uncertainties

As with most human activities, engineers have to face the fact that the activity of planning will always be encircled by a ubiquitous uncertainty.

Until the 80s, when planning was solely driven by the load growth/electrification logic, in a monopolistic environment, uncertainties were generally dealt as some kind of sensibility analysis around a previously chosen solution. However, as it was noticed in section 2.4, several added factors for uncertainty emerged due, not only to varying economic background, but also to a clear change in culture.

Thus, the planning engineers identify, in this rapidly changing environment, a vast range of uncertainties that they should incorporate in their models.

Uncertainties have two basic components: temporal (related to the question *"when?"*) and spatial (*"where?"*) and we may recognize three main types of uncertainties according to the sort and quantity of information available:

- Uncertainties related to equipment failures. In general, for this type of uncertainty, we have enough information and historical data to use traditional probabilistic models based on statistics. However, for some less common failures in some type of equipment, these premises do not hold and we have to deal with the associated uncertainty in a different manner.

- Uncertainties related to loads and costs. Planning is traditionally based on temporal and spatial load forecasts, and these are somehow unreliable and difficult to obtain. Especially in long term forecasts, the knowledge associated to loads is rather vague since it is related to areas with no history of past consumption. Generally, data related to loads and some type of costs (including equipment, fuel costs, energy not supplied, building duration, etc.) appears in the form of indicative values, typical values or, eventually, imprecise linguistic declarations [Lea95]. Assuredly, including these under a deterministic framework reduces the significance of the results. A probabilistic setting is also not adequate as it will be shown later in the thesis.

- Uncertainties related to unique events or to fundamental options. As example, we may refer: a sudden change in environmental legislation; a deregulation or re-regulation of the energy market; an international catastrophe; a regional burst of growth, etc. For this type of uncertainties, there is normally little information available and any precise forecast might be considered as simple futurology. Nevertheless, planners and decision makers have to be conscious of such type of uncertainty.

The following table summarizes these three basic types of uncertainty.

Type of uncertainty	Examples
Probabilistic nature	Equipment failure rates
Load and cost forecasts	Future loads Equipment costs Installation costs Interest rates Fuel Prices Potential supply of non-utility generation
Discrete nature	Economic background Legislation, regulations and incentives Major projects Technological Developments Customer response to DSM Energy market environment

Table 1 Basic types of uncertainty and possible application examples

All the questions related with uncertainty will be addressed in some detail in chapter 4: "Modeling Uncertainties and Risk".

2.6. Criteria

The previous sections have already established the DPP as a complex problem due to its large dimension and pervasive uncertainty. It was also referred that the problem was conditioned by multiple conflicting criteria. The following are the generally accepted criteria in distribution planning:

- **Economic and financial objectives.** The central objective of planning is to minimize costs, that may be divided in two basic types:
  - Investment costs (related to the construction or reinforcement of lines, substations and equipment);
  - Operation costs (Typically, power losses and maintenance).

We may also include in this analysis some financial attributes such as:

- Cash flow;
- Construction expenditures.

- **Reliability.** The planner will aim at maximizing distribution system reliability, by reducing indices like, for example, EPNS (Expected Power Not Supplied).

- **Security.** The distribution system should operate in a safe manner, for both consumers and operators. This factor is vital in distribution planning. However, security is generally evaluated *a priori* and considered as a technical constraint.

- **Quality.** Besides reliability aspects, there are other important quality factors in planning. Some of these factors may strongly influence a company's public image, an important criteria for decision makers.

  - Environmental Impact. This could include several components: visual impact, emissions, road construction, etc. In some sensitive areas, this criterion is determinant in planning.

  - Social impact. The use of territory for other uses, the need for diversification of energy sources, tax allocation, and finally the objective of energy efficiency could be included as general social concerns.

  - Voltage fluctuations. Apart from reliability matters, this is probably the biggest technical concern related to service quality.

  - DC component.

  - Frequency deviations.

  - 3-phase asymmetry.

  - Presence of harmonics and other types of wave pollution.

These criteria are, as mentioned before, basic in power system planning and have been addressed in several publications and research work.

However, in a closer look, a question arises: *is this all we should expect from a solution?*

The previous section mentioned that it is virtually impossible to eliminate uncertainty in planning and that there is a large range of different types of uncertainties. Most of the uncertainties referred (and, in particular, the ones related to load growth) can not simply be reduced to probabilistic models, since any probabilistic model is based in a scenario of "event repetition"; however, the future occurs only once and each decision, taken in any moment, is unique. Thus, an "optimal" decision based on an expected value of the objective function, may be revealed as catastrophic, when facing the future that actually occurs. Subsequently, it is very unlikely (and, somehow absurd) that this decision might be compensated by a favorable "repetition" of futures.

This leads us to recall the concept of risk, defined in 2.3, which is related to decisions and their consequences. From the risk analysis perspective, we may expect a solution not only to be economical, reliable and safe, but also to perform well in any plausible future or, at least, to be sufficiently flexible when facing an unpredictable future. In other words, we expect a solution to be robust.

In [Wil96], H. Lee Willis makes this point very clear concerning spatial/temporal planning:

> *"A clearly desirable goal of the planning process is to minimize risk due to uncertainty. Ideally, plans can be developed to confront any, or at least the most likely, eventualities."*

A solution (plan or strategy) is said to be robust if it is selected (in principle, with no regret) no matter what future occurs. We may, if necessary, define robustness degrees: for example, if a given plan is shown to be adequate in just 8 of 10 possible plausible futures, it would be 80% robust or with a robustness degree of 0.8.

Robustness corresponds to a measure of risk aversion, associated to a certain decision, by answering questions like, for example: "what is the possibility of the utility not regretting a certain decision?". Obviously, everyone would like to have 100% robust plans or strategies. However, that is not generally the case and, consequently, there will always be a future for which the decision taken will not be adequate and will generate regret.

The Exposure degree of a solution may be measured by assessing how much it differs (in the attributes) from the plan (supposedly "ideal") that we would have selected, if we knew, beforehand, the future that was going to occur. In simpler cases, the exposure degree could just measure the number of futures in which the plan would not be accepted, in favor of a set of adequate plans for that future.

Often, it is not common to find a 100% robust solution; however, in some cases, the risk may be low: if the futures for which the solution is exposed are too unlikely, or if the differences between the solution attributes and the ideal plan attributes (for that future) are unimportant and, consequently, the regret felt will be small.

If the preferred decision involves some risk, it is prudent to find ways of reducing it. This can be made by searching and generating new solutions, generally obtained by some kind of improvements in some of the aspects in the exposed plans. Usually, this will mean accepting some degree of extra investment. This process, known as hedging, may be considered as a sort of insurance against adversity and it will be detailed, along with the concepts of robustness and exposure, in the following chapters.

Summary

The following table summarizes the criteria identified in this section.

Type	Criteria
Economy	Investment Operation and maintenance Cash Flow Construction Expenditures
Security	Safe operation of the power system
Quality	Reliability Environmental Impact Social Impact Voltage quality DC component 3-phase asymmetry Harmonics

Risk related	Robustness Inadequacy Exposure

Table 2 Criteria for distribution system planning

2.7. Problem dynamics

Electrical power delivery is one of the most capital intensive industries, since all the investments involve substantial expenses. Arrangements for new or expanded facilities normally require several years, meaning that a utility must plan some years ahead, depending on the level of system.

Lead time means that the utility has to decide to construct or not to construct at that time, and waiting longer than the lead time means "do nothing" by default. Lead times set the basic planning horizon for utilities. For example, if it takes 5 years to plan, build and test a primary substation, it is absolutely necessary that the process is started 5 years before the substation is needed (according to the load growth forecast) or it will not be available when required (but there is no need to start the process more than 5 years in advance).

The following table shows typical lead times for distribution system equipment, according to [Wil96].

Level	Years Ahead
Distribution substations	5
Primary three-phase feeders	3
Single phase laterals	.5
Service transformers	.2

Table 3 typical lead times for distribution equipment

These values are just indicative, since conditions and requirements vary according to the situation. In some cases, the decision of building a primary substation has to be taken much more in advance due to the need of negotiating land rights for the substation site.

However, meeting lead times is not all there is to planning. Planners will also have to look beyond the lead time in order to assure that the short-term decisions provide lasting value. In long-range planning, we are concerned not only in ensuring that the system additions are made in time to meet needs, but in making certain that those additions have lasting investment and performance value.

For example, we know that we will have to install a three-phase feeder in order to meet forecasted load growth in a given area in three years. However, if load keeps growing, the feeder would have to be replaced just a few years later. Thus, it could be more economical to install a larger capacity feeder now, and avoid the additional cost later. This type of decisions has also to be made concerning primary substations and other type of equipment.

This means, and still according to [Wil96], that we will be planning for an extended period starting in the day the equipment is put into service and include a reasonable portion of its lifetime. Consequently, we will be looking around 20 years ahead for distribution substations and 15 years for primary three-phase feeders.

Unfortunately, it is common practice in utilities to neglect or minimize the importance of long-range planning, transmitting the perception that distribution system planning is a short-term matter, and that planners just need to sound warnings in order to meet the lead times and perhaps make some intuitive assessment on the requirements for the long term.

Furthermore, decisions will have to be taken in a dynamic framework. We will not be planning for a single stage but (since load grows continuously over space and time) for multiple stages within the planning horizon. Decisions made for the future will have a strong influence in decisions that have to be made today. Models for distribution network expansion will have to consider these facts and that will be the subject of discussion in chapter 3.

In conclusion, this section established the distribution planning problem as inherently dynamic and with a long-range planning horizon.

2.8. Multiobjective analysis

Engineering problems very seldom require the optimization of a single function with a single objective. Generally, several conflicting objectives must be optimized simultaneously. In some cases, the solution for these problems is obtained by combining the different objectives in a single one to be optimized, accordingly to *a utility function*. However, in most cases, the utility function is not known (or cannot be easily formulated) before the optimization process.

It is not hard to understand that is the case of the distribution planning problem. Besides the objectives related to traditional criteria (economy, reliability, quality and security), we identified some other objectives arising from the uncertain planning environment: maximize robustness and minimize exposure. It could be argued that some of the objectives could be combined in a utility function (for example, objectives related to costs) and some other objectives could be reduced to mere technical constraints (e.g., voltage quality, environmental impact). However, such kind of simplifications is, at least, questionable. It is time that we refuse, once and for all, this single-objective principle, since it indisputably obscures the real goals of the planning process.

The DPP (as any of this type of problems) must then be treated as multiobjective, with objectives that are not trivially measurable. This way, more than one solution may be found. These solutions provide the decision maker a deeper knowledge on the characteristics of the being addressed, allowing the articulation of the preferences, before a final solution is reached. Contrarily to single objective problems, the solution of a multiobjective problem is, in general, a set of points known as Pareto Optimal Set. Obtaining this set of non-dominated solutions is not always an easy task for conventional techniques and it is probably one of the reasons behind the persistence of the old paradigm. Another reason for this, is the planner not knowing how to deal with a bunch of points in a multiobjective space, instead of being delivered with a single, convenient answer.

<u>Generation of solutions</u>

One of the characteristics of the methodology proposed in the remaining of the thesis is to preserve, as far as possible, an explicit representation of the relevant attributes for assessing the quality of the alternative plans or strategies. Consequently, we will need a way of efficiently generating a set of non-dominated alternative solutions. The importance of calculating the set of non-dominated solutions is illustrated in Figure 2, representing a set of alternatives for placing of switching equipment in a 15 kV network [Mir91a], [Mir91c]. The alternatives are represented in a 2 attribute space: Investment (in PTE.10^6, dependent of the number and type of equipment) vs. *Average Annual Energy not-supplied* (dependent on the type and placing of equipment in the network).

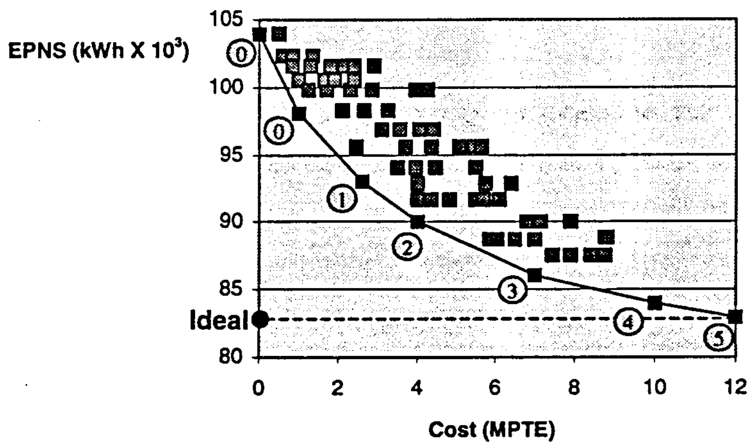


Figure 2 Representation of alternatives to switching equipment positioning in a distribution network, in a two-attribute space. Cost is shown in Millions of Portuguese Escudos (MPTE).

The thick line shows a convex border of this set, connecting a set of interesting non-dominated solutions. The numbers in the circles represent the number of automatic switching devices to include in the network.; the diversity in the values for the attributes derives not only from the number and type of installed equipment, but also from their positioning in the network.

We may notice that, even if the number of alternatives is relatively high, only a few deserve attention from the planner: the calculation of non-dominated alternatives is, therefore, a factor for effort economy and the introduction of discipline in planning tasks.

This fundamental aspect will be addressed in chapter 5, illustrated with some other interesting examples, when the thesis refers the use of Genetic Algorithms.

Another significant question that planners may ask is the following:

In a dynamic problem (meaning that decisions made in the future influence decisions that have to be made today), subject to uncertainties, can we avoid falling into unwanted or potentially catastrophic situations?

This essential doubt stresses the fact that we will need to introduce some kind of flexibility in solutions, making them more adaptable to changing futures. This means the development of adaptive planning strategies, instead of rigid plans.

A typical example in generation planning is well known - suppose that the decision maker has to choose between two options:

A. Build a large power station.

B. Build a series of smaller plants, to be implemented along time, according to the future that actually occurs.

Even if plan A is found to be less expensive in the expected future, it could have catastrophic consequences if the future doe not occur exactly as anticipated (which happens in most of the cases), creating a large regret.

Solution B (a strategy), being more flexible, is probably a better choice in an uncertain future, possibly leading to lower regret values.

Even though distribution expansion planning is intrinsically more flexible than generation planning (due to shorter building times *and* smaller investments at any time), it also involves important large and medium-term decisions that could lead to unwanted situations and potentially large regret values.

2.9. Decision making

The previous sections established the DPP as a problem of decision making, instead of simple optimization. This arises from the fact that we will have to deal with the DPP as a multiobjective problem, thus emphasizing the need for a way of generating alternative solutions.

However, generating a set of non-dominated solutions (a decision set) is just the initial phase in the decision making process. Subsequently, we will need to reduce this set to include only solutions that may be considered as good compromises between objectives, in a certain future (a conditional decision set).

Selecting Conditional Decision Sets

Several techniques may be adopted for the selection of a conditional decision set, namely:

- Methods based on trade-off analysis [Mat88]

Some methods rely on an interaction with the planner, allowing the definition of acceptable trade-off values between attributes. This concept allows utilities to define strategic values for trade-off, as guidelines for planning. This will allow a rational ordering of investments. The advantage of using such planning philosophy is that diversified investments, in different projects, will have an objective comparison basis.

In Figure 2, the trade-off between solution 1 and solution 2 is approximately 400 PTE/kWh, and between solution 2 and solution 3 is 800 PTE/kWh. If the strategic target had been set at 500 PTE/kWh, solution 2 would have been chosen, but not solution 3: each monetary unit invested would not allow a sufficient improvement in non-supplied energy, according to the strategic targets.

- Methods based on distance to the ideal [Mat88]

The concept of ideal corresponds to an imaginary solution that would have the best possible values in each attribute. Therefore, an acceptable solution should be as close to the ideal as possible, if we could define the concept of proximity.

To overcome this small difficulty, Zeleny [Zel82] defined *compromise set* as the set of solutions close to the ideal, that would be obtained by varying the metric used to evaluate this distance, namely from L_1 to L_∞.

A distance to the origin, in a system of n coordinates and in a L_i metric, is given by

$$d_i = \sqrt[i]{x_1^i + \ldots + x_n^i} \tag{1}$$

This issue will be further discussed in chapter 9, when analyzing different paradigms for network planning.

Another problem (which is not exclusive to this technique) is the establishing of a convenient relative scaling in each attribute, since scaling changes the perception of "distance".

In Figure 2, the ideal solution is represented in the graph, and solution 2 appears to be (at least visually) somewhat interesting.

- Holistic methods [Mat88]

The use of holistic methods is an alternative of progressive filtering in the space of attributes, in which the planner does not decide in function of any particular solution, but in function of generalized trends in solution clusters. These techniques allow the selection of a set of alternatives with reasonable close characteristics. Chapter 7 will introduce the use of a holistic methodology known as Successive Amplifications Method (SAM) [Mat92] in distribution system planning.

This phase will be concluded, no matter which method is used, with the definition of a set of solutions (and not a single solution) that are acceptable compromises between objectives – a conditional decision set for each possible future.

Robustness and exposure

Each solution in a conditional decision set associated to a given future must be examined in order to verify if it belongs to conditional decision sets relative to other futures. If a plan is acceptable in all possible futures (belongs to every conditional decision set), then it is 100% robust and a certain candidate in a final decision. However, this seldom happens.

When possible futures are finite or discrete, it is common to associate a numerical value expressing how plausible is the occurrence of such future. These values may be treated according to the axioms of probability theory and generally are referred as *subjective probabilities*, even if it must be stressed that they are not probabilities in the sense of repetition of events[1]. Consequently, it is possible to weigh the number of times a given solution appears in conditional decision sets by the plausibility of the respective futures, thus calculating a robustness degree.

At the same time, we have identified the futures in which a solution is not included in the respective conditional decision sets, which means that if that alternative was chosen and that future occurred, there would be a certain regret.

According to the regret value we may define the concept of exposure degree: for each attribute (e.g. investment cost) and for each future, it can be measured as the difference between the attribute value for that solution and a guideline value associated to the conditional decision set. It is common, but not necessary, to consider the guideline value as the best possible value (the *optimum*) if we knew beforehand which future would occur.

The global exposure value for a solution, expressing risk aversion, will be given by the largest exposure values obtained for each future. We may also calculate an average weighted exposure, from the subjective probabilities for each future in which the solution is exposed. However, this form of calculation, considered isolated, may be already interpreted (although with some caution) as expressing the concept of risk aversion.

[1] If I throw a fair die, most people would agree that the probability of a 6 coming up is 1/6. This means that, in the long run, a 6 will appear approximately once every six throws. However, if I ask 10 people "what is the probability of Brazil winning the World Cup in 1998?", I will get 10 different answers, since, in this case, we are talking about *subjective probabilities* or, more exactly, *degrees of belief*. This has to do with the fact that we cannot refer to a scenario of repetition of events (the '98 world cup will only be played once in 1998 – other world cups have been played but under different circumstances). The important here is that the assessment of the degree of belief is based on more than just frequency of events. This is also the case for most problems including some type of forecast in power system planning.

The computation of robustness and exposure indices may also be performed when futures are represented in a continuous and not discrete manner, e.g. by an interval or a fuzzy number. Chapter 4 will describe a way of calculating these indices when loads are represented by fuzzy numbers.

Hedging strategies

When we do not have a 100% robust solution and the exposure degrees are unacceptable, we must invest a little more in order to reduce risk. This attitude, paying a little extra to protect ourselves against an adverse future, corresponds to the development of a hedging (or insurance) strategy.

The building of hedging strategies is generally performed over solutions with a high degree of robustness, by adding some transformations that make them satisfactory in the futures on which they were exposed. As this introduces new solutions that have not yet been considered, the decision process may have to be restarted until it is possible to make a definitive decision.

The field of decision making is vast and is well established. It would be outside the scope of this thesis to fully analyze all the possible techniques for multiobjective decision making. The methodology proposed in chapter 6 for distribution expansion planning will follow the basic steps referred in the previous sections, with some variations and refinements, especially concerning the development of planning strategies.

2.10. Conclusion

This chapter introduced the distribution planning problem as a large, dynamic multiobjective problem, encircled in a series of complex questions. It promptly rejects the traditional single-criterion economy-driven paradigm and proposes a new approach to the problem, guided by a paradigm based on multicriteria risk analysis.

The first conclusion is that, for the practical translation of these ideas, we will need a set of efficient tools for the generation and analysis of a vast range of alternatives, in a set of possible futures. These tools will have to answer the following basic needs in distribution planning:

- Deal with real-sized networks;
- Allow fully dynamic planning;
- Allow multicriteria analysis, keeping criteria and tradeoffs explicit.

The assumption of risk criteria (robustness, exposure) will lead to the replacement of the concept of *optimal plan* by the concept of *planning strategy*, allowing more flexible and adaptable solutions in face of future uncertainty. This concept, curiously or not, comes close to the traditional practice in utilities, even if performed in a heuristic manner.

Ultimately, this chapter introduces the fundamentals of a new planning philosophy, based on the evaluation of the robustness of expansion strategies and centered on risk analysis. The following chapter will critically review the models proposed for distribution expansion planning and the remainder of the thesis will elaborate on the conceptual and instrumental part of the proposed methodology.

3. SYSTEM MODELING

As far as the laws of mathematics refer to reality, they are not certain, and as far as they are certain, they do not refer to reality.

Albert Einstein

3.1. Summary

This chapter seeks to give a panoramic view on the most representative models developed in the last years for distribution expansion planning. The discussion is centered on the operational and conceptual limitations of existing models and on the reasons behind the divorce between academic models and planning processes in utilities.

3.2. Introduction

Chapter 2 introduced the distribution planning problem (DPP) as highly complex due to its dynamics, multiobjective nature, and discrete characteristics. As consequence of this complexity, this problem can be considered as large dimensional, even for medium-sized networks. For this reason, and considering limitations in time and memory in computer systems, the direct resolution of the DPP is extremely difficult, if not impossible, for real-size systems.

In order to overcome these limitations, several approaches have been proposed in the last 30 years. These approaches, in general terms, can be distinguished by, on one hand, the number and type of simplifications introduced in the models and, on the other hand, by the methodologies and algorithms used. For now, we will not concentrate on discussing algorithmic techniques, since chapter 5 will be exclusively devoted to the discussion on this point.

This chapter will address the modeling issue by presenting a critical outlook on existing models, their characteristics and classification. In addition to referring the main questions related to modeling of the DPP, the main purpose of this chapter is helping to corroborate the statement made in chapter 1: that there is an effective divorce between three decades of models and practical industrial applications. Certainly, part of the reasons for this separation lays in the fact that traditional models tend to neglect some aspects that planning engineers consider to be essential.

In order to further debate these questions, some selected industrial models and new tools for planning will be presented.

3.3. What is, and why do we need a model?

The first step in planning is finding a suitable model for the problem in question.

A model is an abstraction or simplification of reality, since it is impossible to cope directly with the huge complexity of a real system for two main reasons: the limitations of numerical methods and computational power; and the quality of the data and forecasts available. Among all the possible modeling options, the planning engineer will have to include some and omit other. Therefore, a model will be no more than a combination of compromises, consequence of the mentioned limitations.

In power system planning there is a very large number of questions and options involved (as we will see in the remaining of the chapter) and any model will have to include at least the aspects considered fundamental by planners and decision makers. Other important aspects will have to be considered *a posteriori*, over the solutions found, since the direct inclusion of such aspects would introduce too much complexity in the models. Furthermore, some modeling is based on rules or on some type of technical constraints.

The following sections list most of the relevant matters that could be considered in distribution planning, including:

- Model dynamics
- Load and generation modeling
- Technical and quality issues
- Spatial planning

However, no matter how close to the physical reality our model may be, it will also have to reflect the needs and requirements of utilities. Therefore the model will have to contemplate the real objectives of planners and decision makers, and subsequently the way they reason and decide. This essential matter is fairly difficult to deal with, and it will be referred in other chapters of this thesis.

3.4. Planning horizon and model dynamics

Chapter 2 made clear that the distribution planning problem is intrinsically dynamic, which means that decisions taken in the present will influence decisions that have to be taken in the future. Nevertheless, most early models (and even some recent ones) simply ignored this basic fact.

The planning horizon may be divided in one or more planning stages. In general, one stage corresponds to one year in the expansion plan. Concerning the planning horizon, we may divide models in three basic types: static, pseudo-dynamic, and fully dynamic. The next paragraphs will explain the concepts behind each type of model and present some relevant examples.

a) Static models

Most models consider a single planning stage. Such models are known as *static models*. Frequently, in practical applications, the methodology may consider expansion planning as a series of annual expansions in a way that each expansion constitutes a different problem. However, this process will hardly result in a global optimum because, in general, partial optimal solutions will not lead to the optimal global solution. Some examples of early static models:

- In 1974, Masud [Mas74] presents a static model for the optimal placing of substations. This model considers the feeder system only in terms of load transfer between substations. The model was extended to a multitemporal universe in 1978 [Mas78].
- [Hin77] proposes a static model where radiality constraints are enforced by an interesting heuristic procedure of opening loops. This model

includes an interesting set of planning options and solves simultaneously the problem of placing and sizing of substations and the problem of planning distribution feeders. The model has been extended to include reliability evaluations [Oli79].

- [Bac79] is one of the most important static models presented to date, and it has been applied with a relative success in real system planning (see chapter 3). The objective of the model is to determine the optimal hierarchical structure of a distribution system and it includes most of the basic modeling options.

- In 1981, Gönem e Foote [Gön81] develop a quite comprehensive static model based on mixed integer programming and standard mathematical programming software. This model includes all the basic modeling options but it does not include aspects related to radiality and voltage drops. In 1982 [Gön82] the model is extended to a fully dynamic model.

b) Pseudo-dynamic models

In order to overcome the problem of partial optimization, some more complex models were developed in order to solve the expansion problem in several stages instead of a single stage. Pseudo-dynamic models consider the global duration of the plan ignoring loads in intermediate stages. In this methodology, each annual expansion is obtained by an optimized concatenation process after the global solution is attained. Some examples:

- [Sun82] presents a model where the global planning is divided in two phases. In the first phase a static model is applied in order to obtain one solution that satisfies the constraints corresponding to the final of the planning horizon, considering the complete set of network components (feeders, substations, etc.) that would be built in the period being studied. So, in this phase the whole planning horizon is considered as a single stage. In the second phase, the static model is applied to each intermediate network expanded from the previous stage, being the components chosen from the set determined in the first phase. This way, the final solution is obtained in a process of optimization by concatenation. The static model is based on mixed integer programming.

- In a publication from 1991, Nara et al. [Nar91] also propose an approximated method based on a pseudo-dynamic decomposition methodology for large networks. The global problem is divided in several stages that are solved using an algorithm proposed in an early publication [Aok90]. Next, the results of these sub-problems are coordinated in order to minimize the overall cost and the cost of constraint violations. This model proposes interesting solutions, and can be considered superior when compared to other pseudo-dynamic methodologies.

c) Fully dynamic models

In this case, building decisions are obtained simultaneously for each expansion stage, leading to dynamic solutions. For example, [Kag93] (see section 3.12) and [Mir94b] (the first application of Genetic Algorithms to distribution planning) are fully dynamic models. Also, in 1986 Gönem e Ramirez-Rosado [Gön86] presented a rather complete multitemporal model based on mixed integer programming techniques that solves simultaneously the temporal and spatial problems in planning and includes several interesting modeling options. However, the inclusion of radiality constraints seems to limit the use of this model to small systems.

The methodology presented in chapter 6 may be included in this category of models, but in a somehow different perspective.

3.5. Load and Generation modeling

A detailed study on load forecasting and models would be outside the scope of this thesis. As it has been referred in the previous chapters, the main focus of the thesis lays in discussing the conceptual basis of planning and providing a few clues on the instrumental questions in planning, <u>assuming that we have adequate information</u> about future loads and generation.

Nevertheless, since these issues are basic in the planning process, the next sections will present the most important aspects related to load and generation modeling, as well as load forecasting.

3.5.1. Load forecasting

The essential objective of planning is to achieve an expansion of the existing system in order to satisfy future load. Therefore, the first and indispensable step in the planning process is to obtain information about the location, timing and amount of future load growth in a suitable way both for short and long term planning purposes.

Load forecasting is therefore an important decision making tool, the first and essential part of planning. The main objective of the forecast is to lead the planner to make the right decisions since the quality of the results obtained is intimately related to the quality of the forecast.

The following points summarize some important aspects related to load forecast:

- The system has to perform adequately under extreme situations (and not just in average conditions) and for peak situations (in general annual maximums). Therefore, the forecast has to expect the most adverse conditions, for example, in terms of weather (e.g. the coldest winters or the hottest summers in ten years).
- Accuracy. Temporal and spatial correctness in forecasts is a determining factor for distribution planning, since little accuracy will contribute to inferior quality in planning. Some studies may forecast correctly the amount of load growth but if these methods fail to forecast with accuracy where this growth will happen (spatial accuracy), the planners will be lead to make the wrong decisions.

Load forecasting methods

Several studies on for load forecasting have been published, and they basically differ in the following aspects:

- Methodology. There have been several types of computational techniques used in forecasting, from linear regression to pattern recognition methods. Some methods require some interactivity with the planner. Questions related to computational capacity are also relevant. Please refer to [Wil97] for a comprehensive analysis on this and other subjects related to forecasting techniques.

- Quality of data. This is an interesting aspect since it could lead us to discuss the *worth of information*, i. e., how much would we be willing to pay for having higher quality data [Cle91]. Chapter 7 discusses some aspects related to the value of having information that is more accurate.
- Type of results. All forecast methods for power industry provide results for annual peak load, since most planning is based on this value. However, some studies provide other type of important forecasts [Wil96]: off season peak, total annual energy, load factor, power factor, customer class evolutions, DSM impacts, etc.

It is also important to realize the existence in utilities of two types of forecast results, arising from two different contexts and based on two different paradigms: technical (engineering) forecasts and economical (corporate) forecasts. The latter is generally performed by the financial department of a utility and its function is forecasting future economical revenues for the utility, on which it will base all type of commercial and financial planning. While both types of forecast may be accurate from the modeling perspective their basic objectives are essentially different. In [Wil96], the author strongly recommends that technical forecasts should always be based upon, or at least adjusted to conform to corporate forecasts.

Spatial forecasting

The need for information on the location of loads is essential in distribution planning and any forecast has to have the necessary precision to allow the placing of distribution equipment such as substations and main feeders.

The forecast for elementary regions (Small Area Load Forecast) is the most used methodology for long term forecasting. This method is based on the division of the area to be studied in small square areas (or other type of polygons), allowing the integration of all or part of these small areas for global studies on a given region. This type of studies is very efficient and allows a large flexibility and accuracy in spatial forecasting. Some computer programs can handle large numbers of these small areas, sometimes up to a few million.

A detailed study on spatial electric forecasting may be found in [Wil96].

Planning horizon and temporal detail

An adequate forecast should provide information for both short and long term planning. As it was referred before, forecasts are usually done for annual peak loads and the planning will be done for these conditions along the planning horizon, with higher detail in the short-term (usually one-year stages) than in long-term planning, where a lower detail is required. For example a planning horizon of 25 years could be divided in 6 stages for 1, 2, 3, 5, 10 and 25 years ahead. Other possibilities for the same planning horizon would be a more detailed set [1,2,3,4,5,7,10,15,20,25] or a less detailed set [1,3,10,25]. Some kind of interpolation could be used for years in between. By interpolation, we mean a detailed analysis of the periods in between, in order to obtain a better scheduling of investments that would otherwise be allocated to the starting year of this intermediate period.

The planning horizon may also be adapted to the type of system being planned.

Another aspect to be considered in forecasts is the temporal detail. For some type of planning purposes it will be necessary to have other types of forecasts like, for example, a 24 hour load diagram for an average or peak day. This type of forecast could be used for reactive power planning or for assessing the effects of Demand Side Management. It could also be useful in situations where the system is functioning near its limits and we could accept some equipment to function in overload for a limited time.

Forecasts for different times of the year may also be relevant, since weather conditions will be varying along the planning horizon (e.g., some type of equipment have increased stress in high temperatures). Moreover, it is common that load will have a different type of geographical distribution in different times of the year.

Occasionally, for planning purposes, forecasts should also provide the value for the total annual energy and some aspects related to the load diagram, such as load factor in order, to perform some type of economical or technical calculations (for example related to power losses).

Uncertainty in forecasting

No matter how good our method of forecasting may be there will always be a certain amount of uncertainty in data. Unfortunately, it is well known both by planners and decision makers that planning is very sensitive to varying forecasts.

Short-term load growth for the existing system may be predicted based on historical information and statistical methods. However, for new development areas or for regions going through structural changes these forecasts are somehow harder to obtain due to longer time frames. In this case, uncertainty arises from diversified economical, technical and even political sources, and it is shaped in rather different ways.

Consequently, it becomes important to know how these factors affect the forecast. Sensitivity analysis, i.e. the assessment of how varying conditions influence load forecasts is essential in most planning situations. For this type of analysis the most used (and probably indispensable) approach is based on *multi-scenario techniques,* by assigning values to a set of uncertain parameters reflecting a possible future state of these non-controllable factors.

In distribution planning, where spatial analysis plays a large role, the capacity of performing studies on varying conditions (for example, the building of a large sports hall or bridge) is crucial.

The discussion on aspects related to uncertainty in forecasts could easily be extended to include several other issues omitted in this section. However, this question is not central to the thesis. Chapter 4 will be entirely devoted on way of translating uncertain forecasts to mathematical models.

3.5.2. Consumer modeling

As it was noticed in the previous paragraphs, in order to fulfill the objectives of long term planning an adequate forecast must be performed. However, this task is complicated by the need of the use of economical models, and by the existence of complex daily and annual cycles. These cycles will be different for consumers with different types of activities (commercial, residential, industrial, etc.)

Forecasting consumer behavior from the information about their type and activity leads to the definition of consumer classes. This classification, well-known to planning activities, refers to groups of consumers with similar behavior, being impossible, however, to quantify with precision the exact form of the variables involved. [Mir89] proposes a methodology for the definition of such classes, using processes of aggregation and analysis based on fuzzy set concepts, in order to preserve the imprecise character of the original definitions.

3.5.3. Disperse generation

Among the most determinant factors in distribution system planning we have now to count on the growing penetration of disperse generation, initially favored by an environment concerned public attitude, encouraged in many countries by favorable legislative framework, and sustained presently by a constant decline in costs, making these new types of energy competitive with classical generation forms.

The presence of disperse generation (independent producers or co-generation) has a strong influence in utilities namely by the aggravated degree of uncertainty, as it was referred before in this thesis. Apart from the economic point of view, disperse generation also has a technical influence in network modeling, namely in:

- Voltage levels (see section 3.8.2)
- Cold Load Pickup (see section 3.5.5)
- Presence of Harmonics
- Topology (even if the operational topology remains radial, there is now more than one injection point)

The main consequence of the existence of disperse generation is that the planner has to design the network in order to cope with the presence of new injection points in the grid. However, the planner has to realize that the distribution network has to be robust enough to perform adequately in case of contingency or loss of power.

A study on distribution system planning with disperse generation based on the methods proposed on this thesis may be found in [Mir96c].

A detailed study on the modeling of small independent producers (hydro, co-generation and wind power) may be found in [Lea95].

3.5.4. Demand Side Management (DSM)

At first sight, the idea of managing load in the consumer end of the electrical distribution system seems appealing to utilities. Demand Side Management techniques allow, among other things, the reduction of peak load by flattening the load diagram, and consequently diminishing the load factor $P_{average}/P_{peak}$ [Yau90]. This operational capacity of the utilities leads to a considerable reduction of costs by allowing the utilities to reduce total energy in expensive peak hours and by deferring investments to a later time in the future.

Nevertheless, this type of action can affect the quality of service for consumers and have adverse effects in utility image. Some studies have been made concerning the tradeoff *consumer comfort/cost reduction*, but these models have been criticized for assuming the utility's point of view and because of the difficulty of measuring such subjective criterion as comfort.

In conclusion we may state that, since plans are generally performed for peak situations [Wil96] and DSM is intended to reduce peak load (but not total energy), it is very important for planners to take in consideration the effects of this technique. Furthermore, it is also crucial for decision makers to carefully ponder the advantages and drawbacks of DSM and to transmit the conclusions and expectations of the utilities to the planner.

3.5.5. Cold Load Pickup (CLP)

Another aspect that both planners and decision makers have to be aware of, is related to the phenomenon of Cold Load Pickup [Del96]. The simultaneous switching of a large number of loads (e.g. air conditioning, refrigerators, water heaters, etc.) after a power failure can have a pernicious effect on the network, by increasing dramatically the peak load by overshoots with time constants of a few seconds, destroying the concept *of simultaneity factor* commonly used in planning studies. The uncertainty associated to independent generation and its consequences (including what we could call "cold dispersed generation pickup") increase considerably the problems derived from CLP.

Therefore, planners must design their systems in order to allow an orderly and controlled reconnection of loads, reducing the potential damaging effects of CLP.

3.6. Reliability

One of the main concerns in distribution networks is their reliability. Low reliability indices will result in high loss of load, leading to higher indirect operation costs. Thus, there is a growing conscience that it is necessary to include this factor in planning models. However, reliability has different aspects and their mathematical formulation is of high complexity, generally involving some kind of tradeoff analysis (see Figure 2).

Consequently, reliability has mostly been ignored in planning models (at least its detailed representation). However, some authors have been proposing models where reliability considerations are considered. [Mir91a] proposes, after obtaining the optimal radial network, the building of new feeders (according to several economical and technical criteria) for the closing of loops in order to allow network reconfigurations in case of contingency, increasing the network's reliability. However, this procedure will not lead to an optimal meshed configuration.

Another aspect that strongly influences the reliability of distribution systems is the switching policy, i. e., the optimal placing of switching equipment. This problem is rather complex and basically involves some kind of tradeoff analysis (cost/reliability). Generally, this problem is considered *a posteriori*, i.e., after the solution for the distribution network is obtained. This way, very few models contain reliability related elements in their formulations. Still, we may refer a study published in 1979 [Oli79] as the first attempt at including reliability aspects directly in the model.

A common policy is the placement of switches in most lines in urban distribution networks. This, however, does not happen in rural networks. In addition, having manually operated or remotely operated switches has a large impact in distribution reliability and on network design. A convenient mix must be searched, and this is a complex problem in itself.

Besides, connected with switching policy is the definition of open loops or open switcher location, which also has a major impact in reliability.

The methodology proposed in this dissertation, based on the *minimal cut set* method[1] (using simple contingency analysis), is similar to the one used in [Mir94b], and it will be detailed in chapters 4 and 6, including some aspects on the basics of a proposed fuzzy reliability theory. The methodology tries to search for a compromise between computation complexity and quality of results in determining possible meshed solutions. However, it does not fully answer the need for a method that determines meshed networks including reliability considerations. This aspect has to be further studied.

3.7. Repressed demand

The concept of repressed demand has to do with the situations when the utility is simply not able to satisfy all the requests for connection to the electrical network. This situation in general arises from financial or technical difficulties in the utilities or from an unexpected load growth, and it should not be confused with the aspect of Energy Not Supplied, due to contingencies. In general this situation is not common in the industrialized world (where emergency actions can be taken at a certain – in general high – cost) but occurs systematically in developing and underdeveloped countries, sometimes reaching levels of 40-50% of total possible future expected load.

The costs (economical, social, public image, etc.) of repressed demand have not yet been completely studied and quantified, but it is more or less clear that this is one of the factors limiting the economical growth in certain regions of the world.

The methodology proposed in this thesis is concerned about the aspect of repressed demand, by allowing the inclusion of the costs of possible emergency measures (in order to avoid equipment overload) in case the load

[1] A *cut set* is a set of components which, when removed from service causes a lack of continuity in power flow, resulting on a interruption. A *minimal cut set* is a cut set (of components) that does not contain any other cut set as a subset [Wil97].

grows more than expected and the lead times for equipment construction have been passed. This analysis is done in the perspective of minimizing risk.

3.8. Technical considerations

3.8.1. Radiality

The planner may be interested in finding a solution that possesses radial characteristics, depending on the localization of the network and on the criteria followed by the utility, generally related to service security.

For reliability reasons, it is common that distribution networks (especially in urban areas) present meshed structures, being the system operation performed radially. This way, system reconfigurations are allowed in case of contingency (for example the loss of a transformer or distribution line).

This fact reveals an important distinction between a network's *structural topology* (usually meshed) and *operational topology* (usually radial). For this reason, any planning model should not strictly enforce structural radiality on the solutions. What generally happens is that the planning process will produce an operational topology (radial) and then the planner will place new branches forming open loops, in order to increase network flexibility. Generally, this process is guided by reliability considerations and tradeoff analysis and will most certainly lead to sub-optimal solutions. No model has ever been able to address this issue in a complete manner. The methodology presented in [Mir94b] and later shown in chapter 6 tries to find some kind of compromise solution, allowing the inclusion of open loops under specific conditions.

In addition, the inclusion of radiality related constraints in the formulation of the distribution planning problem increases considerably the complexity, introducing undesirable properties in the problem domain like, for example, non-convexity and non-connectivity.

Due to the difficult inclusion of these constraints, most models do not include radiality in their formulations, which reduces considerably the interest of such models. Some other models use special rules or heuristics in order to force radiality in their solutions [Hin77], [Faw83]. However, these heuristic rules lead, in general, to sub-optimal solutions. On the other hand, some authors refer the possibility of inclusion of such constraints in their models without presenting any application to a real-size system [Gön86], [Kag90].

In [Mir94c], radiality constraints are included directly in the model in a straightforward manner, facilitated by the use of genetic algorithms as search mechanisms. In fact, radiality is probably the single most important aspect that limits the applicability of traditional algorithmic techniques (see chapter 5).

3.8.2. Voltage drops

Consumer voltage level is determined by voltage in the primary substation and by voltage drops in lines and transformers, and it varies with load levels [Jor85].

Generally, we may divide voltage variations in two groups: fast and slow variations. From the planning point of view, we are only interested in slow voltage variations.

Slow variations cannot be noticed immediately, and they are caused mostly by gradual load variations and by voltage variation in the primary substations. This type of variation in voltage has a considerable effect in the efficiency and life span of consumer equipment. Usually, a maximum limited is established for variations from the nominal voltage level in distribution networks.

Slow voltage variations have a considerable importance in service quality and, consequently, are a major factor to be considered in planning the expansion and the operation of a distribution system.

Most models reviewed do not include voltage drop considerations in the optimization process itself. Some of the models verify voltage drops in the computational process but do not include them directly in the models [Hin77], [Faw83]. However, it is common the verification of voltage drops *a posteriori*, i.e., after the optimization process. Some models make some adjustments (transformer tap, line reinforcement) in order to force the constraints [Bac79]. More recent models include these constraints directly in the formulation of the model [Nar91], [Gön86], [Kag93], [Mir94c]. This last model includes voltage drops in the model as an independent criterion in a multiobjective analysis and will serve as base for the methodology proposed in chapter 6.

It is important to notice that whenever dispersed generation is present, the planner faces the problem of eventual "overvoltages" and not only classical "voltage drops". This is becoming quite an important issue in several European distribution utilities, and the enforcement of upper voltage limits must be taken as seriously as lower limits.

3.8.3.Power losses

Costs related to energy losses constitute the easiest calculable parcel of operation costs in planning models. Often, and according to the present value of power losses, it is possible that the capacity of lines be determined by economical considerations instead of thermal limits.

This cost is modeled in several different manners and degrees of precision according to the authors. Basically, the models may be divided in three different groups, according to the way they deal with this factor:

a) The models that simply ignore costs related to Power Losses [Mas78], [Hlt76].

b) The models that do not include power losses cost in the problem formulation but consider them *a posteriori*, for the calculation of the system's global cost.

c) The most general group of models that include power losses costs as an additional cost in distribution lines and primary substations: [Nar91], [Kag93], [Gön86], [Mir94c]. This last model considers power losses as a distinct criterion in a multiobjective analysis. This possibility is somehow relevant, for it allows the consideration of independent uncertainties in future costs of energy.

The methodology proposed in chapter 6 is based on the model presented in [Mir94b] and [Mir94c], and therefore power losses are considered as an independent criterion.

3.8.4. Additional modeling options

Several other factors have to be considered when planning a system's expansion. It would be outside the scope of this thesis to elaborate on every aspect that might be considered during planning a distribution system. The following list shows some of the these aspects, which are generally studied over a small set of alternatives, and not included directly on the optimization process:

- **Harmonics** are multiples of the base frequency appearing in equipment that breaks the base frequency, in most cases devices containing thyristors [Jor85]. The additional disturbance of harmonic current depends on the impedance of the circuits. Harmonics occupy some capacity of the cables, provoke extra heating in appliances and finally, there is the possibility of resonance in LC circuits. The significance of this quality factor has been increasing considerably in operation planning of distribution networks due to the increase in electronic control devices.

- **Fast voltage fluctuations** are primarily caused by the connection and disconnection of individual devices in the network. These fluctuations in voltage affect the planning of the network mainly as a technical condition.

- **Short-circuit power** is a determining aspect in choosing the adequate solution for a distribution network. In fact in several utilities (e.g. in the People's Republic of China) this factor has become the most important limitation to network expansion due to the equipment presently installed. The inclusion of short-circuit power related constraints and/or criteria in planning models is not a particularly difficult task. However, the planner will face conflicting objectives: high P_{sc} (short-circuit power) values demand expensive equipment and low P_{sc} values limits the connection of dispersed generation and provokes low values for short-circuit currents I_{sc}, causing problems with the activation of protective switching equipment. In other cases, existing breakers limit the number of expansion alternatives due to their limited ability to deal with high short circuit currents.

- **Three Phase Symmetry** is also an important quality factor that planners must take in consideration by evenly loading the different phases in the network. Load Asymmetry can cause asymmetry of the voltage and has considerable effect on the efficiency of motors, not to mention the inadequate occupation of the distribution network capacity.

- **Dynamic security** studies are generally performed only to a certain number of network alternatives, in order to know if the configurations perform correctly under some predetermined situations. These studies are particularly important in networks with a large share of dispersed generation, in particular wind power. [Lop95] presents an interesting application of Genetic Algorithms to dynamic security assessment of distribution networks.

- **Environmental and visual impact** has already been referred in this thesis and has become, in some situations, the most important criteria in planning. Models for the assessment of environmental impact of distribution networks are starting to come to light, mostly based on Geographical Information Systems.

3.8.5.AC/DC Modeling

Most authors consider DC models sufficient for distribution system planning. However in [Mir94b] and in the results shown in this thesis an AC model [Mir83] was used in order to show the versatility of the genetic algorithm approach in dealing with complex problems. An extra benefit is the higher precision in the results from the models, avoiding questionable simplifications and the possibility of planning a mixed network of aerial/ground cables. The disadvantage is related to a higher computational effort.

3.9. Subsystem optimization

In network expansion planning models it is common to divide the global problem in two basic sub-problems, to be solved successively:

a) The sub-problem of determining the optimal capacity and/or optimal location of substations. In some approaches, the formulation considers the existing network in terms of *load transfer capacity* ([Mas74], [Mas78], [Ada74] ,etc.) or in terms of *load X distance* [Hlt76]. However, the representation of the network is ignored. This simplification is acceptable but, in general, leads to low quality results in terms of planning.

b) The sub-problem of determining the optimal location and/or capacity of distribution feeders. The models for the solution of this problem will consider the full network but will not consider the previous problem.

However, this division will lead to sub-optimal solutions, since each system is optimized separately.

Some models ([Kag93], [Gön86], [Nar91], etc.) search for the global solution of the problem, but the additional computational weight makes them feasible only to small systems.

3.10. Other planning issues

Planning models may also include address some other topics:

a) Determination of optimal paths for distribution lines. This problem involves a high level of complexity and is generally solved *a priori*.

b) Calculation of optimal capacities for lines and substations [Gön81], [Hin77]. Again, this problem is generally solved before the optimization process itself. However, in practical real-life problems, the number of options for line and substation capacities is rather small, so it is acceptable that projects are defined beforehand.

c) Possibility of including two or more lines in the same ditch with consequent reduction in costs. This possibility is difficult to be formulated in a conventional linear programming model: if we have two Boolean variables a_1 and a_2 representing the existence of lines ℓ_1 and ℓ_2, the simultaneous presence of ℓ_1 and ℓ_2 would be represented by $a_1.a_2$. This representation introduces a non-linearity in the formulation. [Bur85] overcomes this difficulty with some ease since this model uses simulation techniques and not a direct linear programming method. Methodologies based in Evolutionary Algorithms, like the one presented on chapter 6, are immune to this problem.

3.11. Dealing with uncertainty

The way models deal with uncertainty has been referred throughout the thesis as a central aspect in power system planning. In fact, and especially in long term planning studies, factors related to uncertainty (in particular related to load forecast) are extremely relevant, and this lead us to the conclusion that planning is foremost dealing with uncertainties.

Unfortunately, early planning models and most new models simply overlook this basic fact. However, the so-called *optimal solution* is extremely sensitive to variable load forecasts and this fact has been recognized for a long time now.

Nevertheless, some later models try to include factors related to uncertainty in several different ways in their formulations. Some examples:

- [Bur85] shows the way we can convert the obscure effects related to uncertainties (modeled using probabilities) in a clear priority list of building measures. This model uses simulation techniques allied to high degree of interactivity with the planner.
- In [Kag93] the distribution planning problem is formulated on a fuzzy mathematical programming environment.
- [Sar91b], [Sar94] and [Kau93] also propose the modeling of uncertain data in loads with the use of fuzzy sets. The methodology is proposed for application in operation problems.
- In [Mir95] the authors propose a comprehensive modeling of uncertainties, using scenario trees, probabilistic models and fuzzy sets. This work was the first publication regarding the methodology proposed in chapter 4.

The most common way of dealing with uncertainties is the scenario technique, in which several plausible scenarios are studied for each possible solution. This technique is been used since earlier models. Probabilistic methods and fuzzy modeling have also been used in different degrees.

This thesis proposes a comprehensive way of modeling uncertainties, where each type of uncertainty is modeled according to its type and the amount and quality of information available. In this manner, chapter 4 will be solely devoted to an extended analysis of uncertainty modeling in planning.

3.12. Multiple objectives

As it was stated before (see section 2.8), the recognition of the multiobjective characteristics of the distribution planning problem is essential for the development of a truly comprehensive methodology for this problem.

Again, most early models consider just one objective (in general investment costs, sometimes combined with operation costs) and treat the others as mere technical constraints.

However, newer models recognize the multiobjective essence of the DPP, by accepting the principles of multiobjective analysis in planning. Some examples:

- In 1979, Oliveira and Miranda [Oli79] include for the first time aspects related to reliability directly in the objective function.
- [Kag93] extends the capacity of existing models to multiobjective analysis, including uncertainties modeled with the use of fuzzy sets. The final solution is chosen with the participation of the decision maker through an effective interface, by searching the best compromise solution between conflicting objectives. However, the complexity of the procedures used in this model limits its application to small systems.
- In [Mir94b], V. Miranda, J. Ranito and the author of this thesis propose a model where multiple objectives are combined in a utility function (used as fitness function in the genetic algorithm – see chapter 6 for more details). An extension of this model to full multiobjective analysis was presented in [Mir95] and serves as basis for the methodology presented in chapter 6.
- [Boc97] rejects the concept of optimal planning and embraces the idea of multiobjective analysis as a tool providing an extra insight on the problem. The model is based on Dynamic Programming and uses a technique known as Optimal Initial States [Dal90].

- [Fer96], formulates a model intended to be as realistic as possible, clearly adopting the principles of multiobjective analysis.

3.13. New tools

In the last few years, some new computational tools have appeared, together with the fast increase of computational power and memory. This section will present the instruments that the author believes will have a larger impact in power system planning: Geographical Information Systems and High Performance Computing.

3.13.1. Geographical Information Systems (GIS)

Industry is starting to fully understand the importance and potential of the information contained in a Geographical Information System. In fact, GIS, conjugating effectively the relation between time and space, have become a decisive tool in planning, administration and management of resources and, consequently, crucial in any information system.

Geographical Information Systems is much more than a map drawing system, it is a powerful tool for geographic analysis. The concepts behind GIS integrate hardware and software instruments for the gathering, processing, analysis, storage and visualization of spatial data.

Considering that the problem of electrical energy distribution has clear geographical characteristics, it is obvious the efficiency of GIS in this type of applications. Looking at the capabilities of GIS in the several stages of power system planning, we may better understand the potential of GIS [Mon96b]:

<u>Data Collecting and Processing</u>

The first phase of any planning process is the gathering and manipulation of data. Most data in distribution planning has geographical characteristics. The secret for the success of a GIS is projecting all data in a common platform, since the data exists in several forms: maps, photographs, satellite images, tables, databases, files, etc. Thus, the best tool for normalizing all the information is surely a GIS.

<u>Data analysis</u>

GIS is an open and modular tool, with its own language and capacity of using programming modules from other systems. GIS also include internal or user-programmed modules in order to perform powerful data analysis.

GIS are able to operate with geographical grids and thematic layers in which, for each graphic element corresponds a geographical entity. The flexibility allowed by GIS in the operation of these grids and layers, facilitate the formulation and implementation of methodologies.

Presentation of results

Since GIS originated in cartography it is understandable their aptitude for the presentation of results as maps and charts. This type of presentation is probably the clearest form of showing the most important aspects among the global results. This way, by simple inspection, it is easy for the planner to take his conclusions and detect important details that could be missed in a mere result table.

For all these reasons, there is a clear trend in industry when it tries to move in direction of a total realization of the potential of the new tools by integrating SCADA, DMS (Distribution Management Systems) and GIS.

Several studies have been performed in INESC related to network planning using GIS [Mon96a] and the objective now is to combine the methodology proposed in this thesis for distribution planning with a GIS for spatial planning.

3.13.2.High Performance Computing

Power systems simulation, optimization and control problems may be included in the category of highly computer intensive problems found in practical engineering applications [Dja96].

The processing capabilities of single processor computers (despite the substantial growth attained in the last years) has been shown not to be able to cope with the increasing computing requirements of some of these applications. High Performance Computing (HPC), a general term for parallel, distributed, vector and other type of processing techniques has achieved a stage of industrial development which allows economical use in Power

System applications. HPC relies on the exploitation of concurrent tasks in the programs that can be executed in parallel in computational systems with multiplicity of hardware components. This type of implementations allow reducing drastically run times, increasing efficiency. Some different types of architectures are available:

A. *Superscalar processors* are single processors able to execute concurrently more than one instruction per clock cycle.

B. *Vector processors* are processors designed to optimize the execution of arithmetic operations in long vectors.

C. *Shared memory multiprocessors* are machines composed of several processors that communicate among themselves through a global memory shared by all processors.

D. *SIMD (Single Instruction Stream Multiple Data Stream) Massively Parallel Machines* are composed of hundreds of thousands of relatively simple processors which execute, synchronously, the same instructions on different sets of data under the command of a central control unity.

E. *Distributed Memory Multiprocessors* are machines composed of several pairs of memory-processor sets, connected by a high speed data communication network.

F. A *Heterogeneous Network of Workstations* may be used as a virtual parallel machine to solve a problem concurrently by the use of special developed coordination and communication software.

Several potential areas for the application of HPC in Power Systems have been identified: control, simulation, optimization, probabilistic assessment, intelligent systems, distributed databases, etc.

Evolutionary Algorithms, algorithmic-mathematical techniques described in chapter 5 and central in the methodology presented in the dissertation, due to their structure, are particularly adapted to parallel or distributes implementations. In INESC, we have implemented a Genetic Algorithms platform based on *sockets* (an interface program from UNIX standard library) over a cluster of Heterogeneous Workstations (as described in F) connected in a local network. This platform has been used with some success, in special in determining optimal Neural Network topologies, but also in distribution planning. Presently, we study other types of implementation of the distributed platform, eventually using personal computers running on Windows NT. Other possibility being studied is the use of a programming environment known as PVM (Parallel Virtual Machine) that has all the potentialities of *sockets*, and allows the synchronization and control of processes in multiprocessor systems.

3.14. Industrial applications

Until a few years ago, distribution planning in the industrialized world was mainly performed with the help of calculations deriving from individual judgment or from a set of practical rules dictated by utilities, based on experience and regard of regional characteristics. The computer was used mostly for dealing with specific calculations such as network drawing, power flow calculations, stability analysis, short circuit studies, etc.

Gradually, the industry became aware of the importance of computer aided design for distribution networks in order to free the planner from routine tasks. Therefore, some industrial computer applications for planning have increased productivity, improved design, and helped reducing overall planning costs.

However, due to the complexity of the process, the vast amount of data, and the constraints and uncertainties involved, attempts at some kind of optimization of the planning process (or part of it) have had a very limited success in utilities.

With the shift of paradigm from *optimization* to *decision making*, a few of the models proposed in utilities have been moving toward a different approach: the system would present the planner and the decision maker with a set of suggested alternatives for expansion of the distribution network. The possibilities could then be analyzed individually before the final decision is taken.

Nevertheless, although some of the methodologies are moving in the right direction, so far none of them has had a definitive success in industrial planning.

The next sections will present some interesting examples of distribution applications in different utilities in the industrialized world. We can distinguish the proposed applications in two basic groups:

- The more general group of applications designed as planning aid tools i. e., these applications do not suggest expansion plans but simply assist the planner in this task. In this group, we may find very interesting and comprehensive systems that are having a large impact in terms of efficiency in utilities. These include systems like EDF's PRAO, ABB's CADOPS, and DINIS (used by the Portuguese Power Company EDP). These last two systems are described in the following sections.
- The applications designed to propose plans (or partial sets of options) for network expansion. These systems are mostly experimental and, although promising, are having (for the time being) little impact in utilities. The focus of the following sections will be on this type of systems, since that is what this thesis is concerned about.

By the time this thesis is being written, the author's group in INESC is concluding the details on a specification of a module for expansion planning to be integrated in a DMS system, based on the principles exposed on this thesis. This module is proposed to have several interesting features such as Risk Analysis, switching policy studies, capacitor bank planning, reliability studies, etc.

3.14.1.Planning Aid Tools

<u>CADOPS (ABB)</u>

CADOPS is an integrated system for Power System planning developed by the Swiss company ABB (Asea Brown Bovery). It is probably the most advanced system in this area, provides all the necessary tools for analysis of distribution systems, and includes an extraordinary graphical user interface based on a GIS.

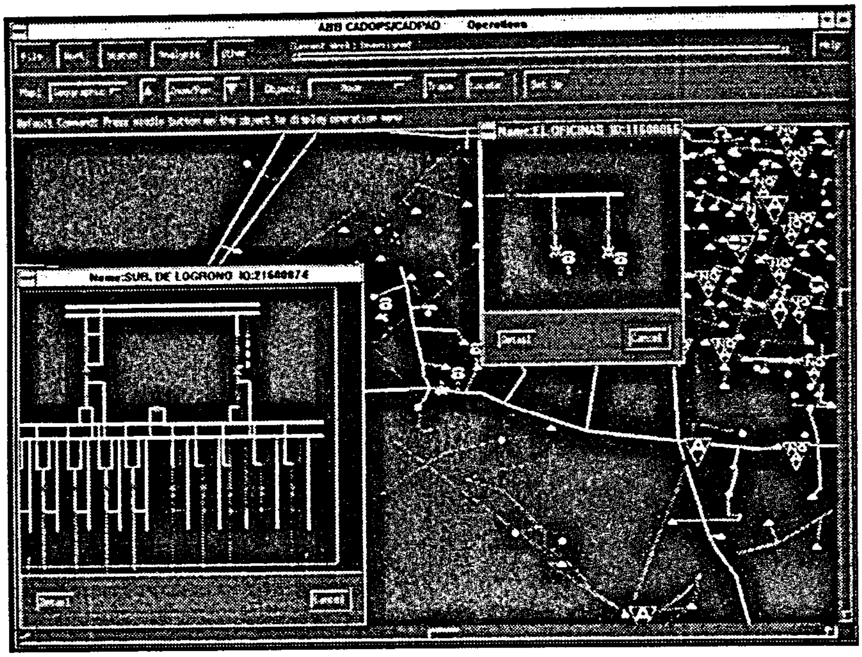


Figure 3 ABB's CADOPS/CADPAD system (picture from INTERGRAPH®)

This type of system, eventually integrated with a SCADA system, answers many of the requirements of the utilities especially related to the operation of power systems (for example outage management) but is used for expansion planning as well.

<u>DINIS (ICL)</u>

The Portuguese power utility EDP (Electricidade de Portugal) adopted a system from ICL (International Computers Limited) for the analysis of distribution networks named DINIS [ICL97]. It provides an interactive, graphics-based way to input network designs and carry out a certain number of studies including load flow and fault level analysis.

DINIS records and stores a network design in digital form using a geographic information system allowing, in different manners, the modification of a basic network design and then the performance of studies and comparisons between alternative designs. According to the specifications, the system is designed in order to provide the planner with the following possibilities:

- Supporting decisions relating to the design of networks, by identifying the effects of changing the network.

- Confirming that the base model of the network accurately reflects the way the real network operates.
- Evaluating the effects of adding new loads to the system.
- Fault analysis
- Transient stability studies
- Load allocation studies

Some other major users of DINIS are:

- The Southern Company who, using AutoCAD as a basis, developed Georgia Power's Automated Mapping/Construction package for the maintenance of existing digitized maps.
- NORWEB and South Wales Electricity
- London Electricity who use DINIS to maintain an up to date record of their master supply network.
- Yorkshire Electricity

3.14.2. Expansion planning applications

SwedPower (Sweden)

In 1991, SwedPower presents a report for the development of planning methods for distribution networks. The methodologies presented in the report are similar to the ones proposed by Backlund e Bubenko [Bac79]. The general proceeding follows two basic steps:

- A general planning of energy distribution for new areas;
- A detailed network planning based on regional development plans and forecasted load demand.

The methodology also stresses the importance of the use of Geographical Information Systems emphasizing the necessity of an interaction with the planner.

Hydro-Quebec (Canada)

Hydro-Quebec has developed a system [Bla96] for planning major investments in distribution networks. According to the authors, the system allows the planner to identify, from load forecasts, data on the existing distribution system and all future projects, together with the commissioning year, in such a way that the investment, operation, and maintenance costs are minimized over the planning horizon (typically, 20 years). The system is based on a dynamic model of network planning solved by a heuristic method comprising five phases. The software presents the planner with a group of alternative solutions for expansion of the system (including costs), that are analyzed in detail before the final selection. The methodology includes some interesting modeling options and allows the possibility of sensitivity analysis. The mathematical model used in the system is a non-linear mixed-integer programming type. Still according to the authors, part of the system is already in use in Hydro-Quebec with a significant impact on savings and planner productivity.

Although it involves some interaction with the planner, the methodology does not consider aspects related to uncertainty - the system is basically an optimization program for major investments without any type of risk analysis.

<u>VDEW (Germany)</u>

In Germany, VDEW (Vereinigung Deutcher Elektrizitätsweke) has developed in 1989 a set of programs together with university researchers (University of Darmstadt and University of Saarland). The program ODIN is based on work presented in [Bur85] and [Fre89], following the work of Backlund and Bubenko in Sweden [Bac79]. This program is a practical application of this research effort. The initial work [Bur85] included uncertainties in forecasts using probabilistic models, considering confidence intervals for future loads. However, as the authors noticed, the optimal solution responds very sensitively to variable load forecasts. Therefore, the results from the system would be an ordered list of building measures.

In 1996, researchers from the University of Saarland presented an evolution of the methodology, introducing a tool for planning with a strong focus on reliability evaluation [Back96]. In this work, a software package **ZuBer** is used for the calculation of reliability indices, using fuzzy numbers for the modeling of uncertainties. The final objective of the project is to develop a planning system with a graphical interface for Windows, in a multiobjective environment (reliability and investment and operation costs), presenting the planner with several non-dominated network alternatives. The system is not yet fully implemented, but it is clear that it has the potential of becoming an interesting decision aid tool in distribution planning, for it tries to answer some of the planner's concerns.

3.15. Conclusion

Some authors blame the divorce between academic models and industrial practices on the little efficacy in the articulation between science and industry. However, the reasons for that separation seem to lie much deeper than that.

The distribution planning problem is extremely complex and vast, involving a large number of questions that planners want to be answered in the methodologies proposed by researchers and scientists. Planners certainly want a methodology that includes all the fundamental modeling options and demand a user-friendly graphical interface, probably over a Geographical Information System. They want to be given the possibility of performing different studies on different planning conditions and they want the answers to be fast, comprehensive and accurate. Planners want a system that comes as close to reality as possible.

Decision makers want to have a deeper insight of the problem by knowing the possible tradeoffs between the multiple conflicting objectives involved. They need the system to provide them with different alternatives and to help them in the process of decision. Decision Makers want to know the risks and consequences of a decision.

Therefore, and in conclusion, any planning methodology will only have an ultimate success when it is able to fulfill two main objectives:

- Include all or most of the modeling options required by planners;
- And, certainly more important than that, help and guide the decision maker in the process of decision.

This chapter left an important point open: the modeling of uncertainties and risk, two basic concepts in this dissertation. This will be the function of the next chapter.

4. MODELING UNCERTAINTIES AND RISK

If we don't succeed, we run the risk of failure.

Dan Quayle

4.1. Summary

This chapter proposes a general model for the representation of uncertainties and risk in power system planning. All types of uncertainties are modeled according to their nature: discrete scenarios by the means of a tree of futures, qualitative imprecision by fuzzy descriptions of data and parameters related to system components by probabilistic models.

4.2. Introduction

In the previous chapters we discussed the fact that, in the process of planning, there comes a time when the planner must face factors that influence decisively the planning activity and that are not under the planner's control or cannot be forecasted with precision: that's what we referred to as uncertainties. Therefore planning is, essentially, *dealing with uncertainties.* This fact has been, in many cases, camouflaged behind dense mathematical formulations and allegedly objective economic criteria.

Planning is also, as stated in chapter 2, intrinsically dynamic, in the sense that decisions to be taken in the future influence the decisions that must be made in the present.

When facing an uncertain future there are two basic attitudes in planning.

One attitude is to plan for what is perceived as the most realistic expectation. This concept lead to planning approaches where decisions are driven by the calculation of the average future outcomes of the decision or the attribute variables, weighting the futures by some perceived probabilities. These probabilities could result from some rational model (for example, a reliability model) or from subjective way of perceiving them (for example, when allocating a subjective probability to some load growth scenario in the future). The paradigm related to this attitude will be referred as *Probabilistic Choice.*

Another attitude is to avoid decisions that, although with high potential for a probable future, could lead to unpleasant or even catastrophic consequences, in an improbable future. This line of reasoning lead to formalizing concepts such as risk aversion and to the development of a planning approach based on the paradigm of *Risk Analysis*.

This chapter will not discuss these two paradigms, for they will be object of a detailed analysis in chapter 9.

Being aware that the concept of uncertainty is present in any planning problem, the following sections will discuss how to model factors related to uncertainty in the perspective of its use in power system planning.

A probabilistic model is certainly adequate if one has enough information to make it valid and under some assumptions: the phenomenon has to be quantifiable and there should be enough information for such analysis. Probabilities relate to a scenario of repetition of events, according to a law that supposedly governs the phenomenon. Probabilistic models that rely on average outcome values lie on a hidden principle that time enough will be allowed for a system to digress through its state space. This is, most of the time, the case in reliability models for repairable systems, or for inflow models in hydro systems.

But, in many cases, such stability conditions cannot be verified, or one cannot admit that the lifetime of a system will be so large as to cancel out exceptional negative events with positive events, or when the information needed to build a probabilistic model just does not exist. In this case, probabilities are no longer an adequate tool to work with, and other representation of uncertainties must be used. An example of a wrong use of a probabilistic model would be the use of a Gaussian distribution to represent uncertainty in load 10 years from now.

Until recently, apart from probabilistic models, the only widespread alternative technique for the representation of uncertainties was known as *Scenario Techniques*: uncertainties are represented as a discrete set of deterministic scenarios. However, the development of the *Fuzzy Set Theory*, proposed by Lofti Zadeh [Zim85] showed the existence of an alternative for the modeling of uncertainties, where these could be represented by continuous intervals according to degrees of "possibility".

Which approach to adopt, then? As it is usually the case, the most pragmatic attitude is to combine the three, using each one according to the type of uncertainty to be modeled. In the following sections, there will be no allegations of superiority of the fuzzy set approach over the other two. The purpose is just to show the contributions of each technique to the planning activity, with some emphasis on fuzzy set concepts, because it is still a novel technique with its own application space, not because it might be superior to any previous knowledge or able to replace any known techniques. However, it must replace them in those cases where they are not being adequately applied [Mom95].

The following sections will present a review of these three ways of modeling uncertainties and concentrate on the representation of uncertainty and how it may influence planning decisions or how it may contribute to the decision making process in utilities. An example of a practical application of decision making based on fuzzy concepts will be presented in chapter 9.

4.3. Probabilistic models

"When the only tool you have is a hammer, all your problems will look like a nail"

Probabilistic models have been the main technique for modeling uncertainties in power systems. There have been innumerous attempts at modeling uncertainty in loads by probabilistic models. Consequently, several approaches on probabilistic load flow (PLF) models appeared in the literature. The first DC models appeared in 1974 [Bor74], [All74], followed by the development of AC models by Allan et al. [All76a]. Also in 1976 the same author applied PLF in operational decision making [All76b]. Both Sullivan [Sul77] and Leite da Silva in 1990 [Lei90], used PLF for expansion planning purposes.

However, the application of this type of models to expansion planning in power systems may be questionable in some cases for two main reasons:

- The extreme mathematical complexity introduced by these models, leading to dubious simplifications;
- The fact (referred in the introduction of this chapter) that, in this case, it is at least arguable that the basic assumptions for the use of probabilistic models can be verified. Furthermore, the existence of some type of uncertainties that cannot be modeled by probabilistic models, clearly suggests that the use of these subjective probabilities (or degrees of belief) for loads is inadequate.

However, as we have already seen in chapter 2, there are some cases in expansion planning where probabilistic models may be used, although with some care. This would be the case of some type equipment reliability indices, only when there is sufficient information and historical data to establish probabilistic models.

Still, it is important to realize that the conditions in which equipment is installed are never the same twice and that, for some types of failures, the historical data is minimal. In those cases, the planner must be careful on using this probabilistic assessment of reliability.

4.4. Fuzzy set theory

4.4.1. Basics

A simple way of understanding fuzzy sets is to consider them as an extension of traditional (crisp) sets. A fuzzy set $\tilde{A}$ is characterized by a membership function $\mu_{\tilde{A}}(x)$ defined in [0, 1], relating each element X_1 to its compatibility degree with X_1. The transition between extreme situations of full and complete lack of membership is gradual, while in classical crisp sets the membership degree to a set would be defined by a value in {0,1} - an element either would belong or not to a given set. A fuzzy set $\tilde{A}$ defined in a Universe X_1 is

$$\tilde{A} = \{ (x_1, \mu_{\tilde{A}}(x_1)), x_1 \in X_1 \} \tag{2}$$

An α-level set or an α-cut of a fuzzy set **Ã** defined in $\mathbf{X_1}$ is the hard set $\mathbf{A_\alpha}$ obtained from **Ã** for $\alpha \in [0,1]$ such that:

$$A_\alpha = \{x_1 \in X_1 : \mu_{\tilde{A}}(x_1) \geq \alpha\} \tag{3}$$

The usual set operations can be readily extended to fuzzy sets using the Extension Principle formulated by Zadeh [Zim85]. In particular, given the fuzzy sets **Ã** and **Õ**, where **Õ** is defined as in (2), the intersection **Ã∩Õ** is most frequently calculated by

$$\tilde{A} \cap \tilde{O} = \{(z_1, \mu_{\tilde{A}\cap\tilde{O}}(z_1)), z_1 \in X_1\} \tag{4}$$

$$\mu_{\tilde{A}\cap\tilde{O}}(z_1) = \min\{\mu_{\tilde{A}}(z_1), \mu_{\tilde{O}}(z_1)\} \tag{5}$$

and the union **Ã∪Õ** is given by

$$\tilde{A} \cup \tilde{O} = \{(z_1, \mu_{\tilde{A}\cup\tilde{O}}(z_1)), z_1 \in X_1\} \tag{6}$$

$$\mu_{\tilde{A}\cup\tilde{O}}(z_1) = \max\{\mu_{\tilde{A}}(z_1), \mu_{\tilde{O}}(z_1)\} \tag{7}$$

Although the **max** and **min** operators are the classical ones, other operators have been proposed, having the advantage of introducing compensatory characteristics that the former do not exhibit. However, the ones above are those which have been more generally accepted.

<u>Fuzzy numbers</u> - A fuzzy set **Ã** is said to be a fuzzy number if it is <u>convex</u>, such that its membership function is piecewise continuous, and <u>normal</u>, meaning that there is at least a value $\mathbf{x_i}$ for which $\mu_{\tilde{A}}(x_i) = 1$.

Fuzzy numbers (FN) may be represented by the intervals defined by each α-cut, such as with $\mathbf{A_\alpha = [a(\alpha), b(\alpha)]}$, like in [Kau88]. Convexity implies the following nesting property:

$$\alpha' < \alpha \;\Rightarrow\; [a'(\alpha) \leq a(\alpha),\; b'(\alpha) \geq b(\alpha)] \tag{8}$$

It should also be emphasized that the extension of arithmetic operations to FNs is much simpler than with random variables and therefore very efficient in computer calculations. In fact, they may be seen as an extension of interval arithmetic, applied at every α-level. In fact, if the FNs are described by rectangular membership functions, fuzzy mathematical models become analog to interval arithmetic models.

Some FNs, for their simplicity, are commonly used, such as the triangular (Tr) and the trapezoidal (Tp) fuzzy numbers. They can be represented by their characteristic points at α-levels 0 and 1; for instance, the FN in Figure 4 could be described as **A = [1,2,3,5]**.

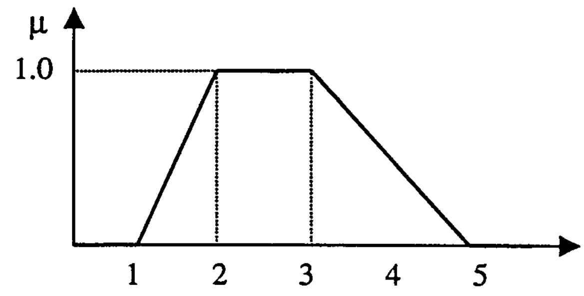


Figure 4 - A fuzzy trapezoidal number, expressing a range of possible values [1,5], at membership 0, and a range of likely values [2,3] at membership grade 1.

Adding Tr or TpFNs becomes very simple: one just adds the characteristic points of the operands to get the result. Let

B = [-2,1,2,3]

then

C = A+B = [-1,3,4,7]

Subtracting Tr or Tp FNs needs some care: A - B must be seen as A + (-B), meaning that

A - B = [1,2,3,5] + [-3,-2,-1,2] = [-2,0,2,3]

This means that there is no neutral element, and if one performs (A - A) one does not get zero, but instead [-4,-1,1 4]. The implication is clear: *it is not possible to reduce uncertainty just by operating with uncertain data.*

Adding or subtracting Tr or Tp fuzzy numbers gives Tr or Tp fuzzy numbers, but that no longer happens with multiplication, inverse or division (besides, these operations only have an easy definition in IR^{+}). However, in many cases and for practical purposes, taking the result as Tr or Tp will be a reasonable approximation (Figure 5).

For instance, the inverse of C = [0.5, 1,4] will be approximately the TrFN

$$C^{-1} = [0.25, 1, 2]$$

There are, of course, other ways to represent the membership functions of FNs, but their description would fall out of the scope of this introduction to fuzzy systems –refer to [Zim85].

In some circumstances, a fuzzy set can be associated to a possibility distribution. In such cases, the membership function corresponds to a possibility distribution function and may be interpreted as such.

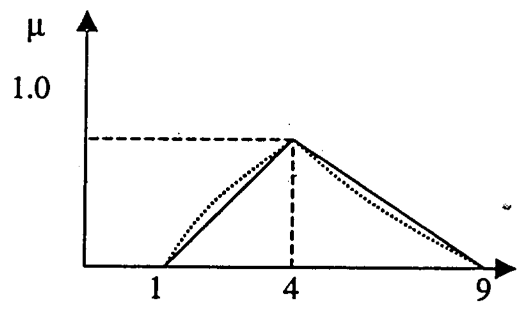


Figure 5 Fuzzy number resulting from squaring the TrFN [1,2,3] - it can be approximated by the TrFN [1,4,9].

Linear ordering of fuzzy numbers - It is not possible to define a unique linear order in fuzzy numbers contrary to real or crisp numbers. However, for practical purposes or in some practical applications, a sequence of criteria may be applied such as the ones proposed by Kauffmann [Kau88], where the concept of *removal* is introduced. This is important within optimization frameworks, where one would like to compare fuzzy results and define a coherent "gradient" of preferences.

Defuzzification

There are several ways of mapping fuzzy numbers onto a crisp number. This section will refer two of them: Removal and Center of Gravity (COG). The remainder of the thesis will use the technique of removal for defuzzification, when needed.

The COG defuzzification index, applied to a fuzzy set A defined over $\Re$ and having a discrete membership function

$$\mathbf{A\ \{x_i,\ \mu_A(x_i))\},\ for\ x_i \in \Re\ and\ i=1,2,\ldots,N} \tag{9}$$

With N being a finite positive integer, can be expressed as

$$G(A)=\frac{\sum_{i=1}^{n} x_i \mu A(x_i)}{\sum_{i=1}^{n} \mu A(x_i)} \tag{10}$$

When A is a continuous fuzzy number

$$\mathbf{A\ \{x_i,\ \mu_A(x)\},\ for\ x \in \Re} \tag{11}$$

then we have the following expression:

$$G(A)=\frac{\int_{-\infty}^{\infty} x\mu_A(x)dx}{\int_{-\infty}^{\infty} \mu_A(x)dx} \tag{12}$$

The *Removal*[1] value maps each fuzzy set to a crisp number by the integration along the membership axis of the arithmetic mean value of the α-level of the fuzzy set considered [Saa96].

$$\mathbf{F(A)=\int_{0}^{1} M(A_\alpha)d\alpha} \tag{13}$$

where $M(A_\alpha)$ is the arithmetic mean of the α-level set of A.

Figure 6 illustrates the concept of removal in terms of areas. The value of removal is equal to the sum of the two areas A and B divided by two (corresponds to the integration of the mean value if A_α).

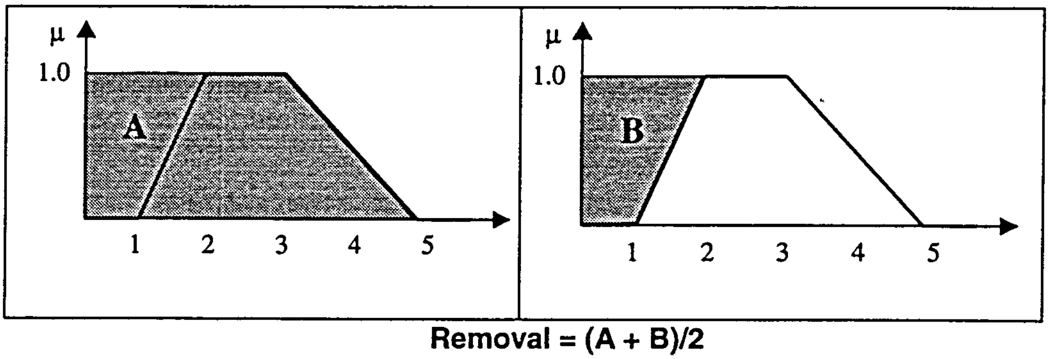


Figure 6 Illustration of Removal.

4.4.2. Fuzzy loads

The application of FNs in power system modeling becomes clearer if one considers the concept of Fuzzy loads.

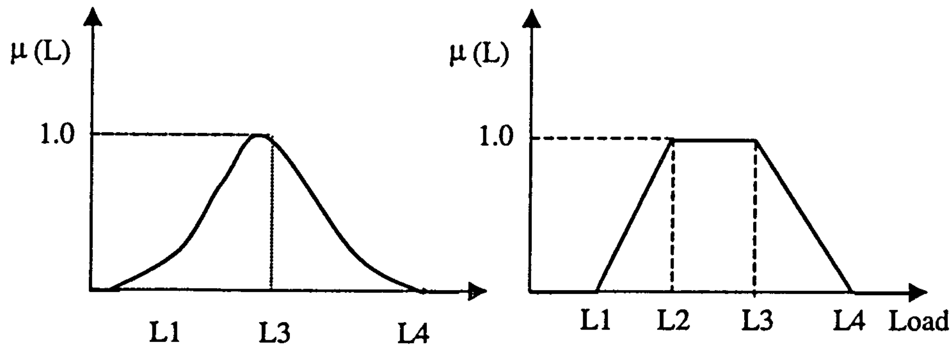


Figure 7 Fuzzy numbers modeling a fuzzy load; left - general; right - trapezoidal.

A fuzzy load is in general associated with a qualitative description, usually through a linguistic declaration [Mir91a]. Simple representations use Tr or TpFNs, but more general membership functions may be adopted (Figure 7).

An uncorrelated fuzzy load, given by estimates of active and reactive power, is represented as in Figure 8. One may see that there is no simple conversion to a representation by fuzzy power module and angle; this may introduce some complication in models.

[1] [Saa96] refers this value as Total Distance Criterion (TDC). Although there is no consensus on this matter, the term *removal* will be used in the remainder of the thesis.

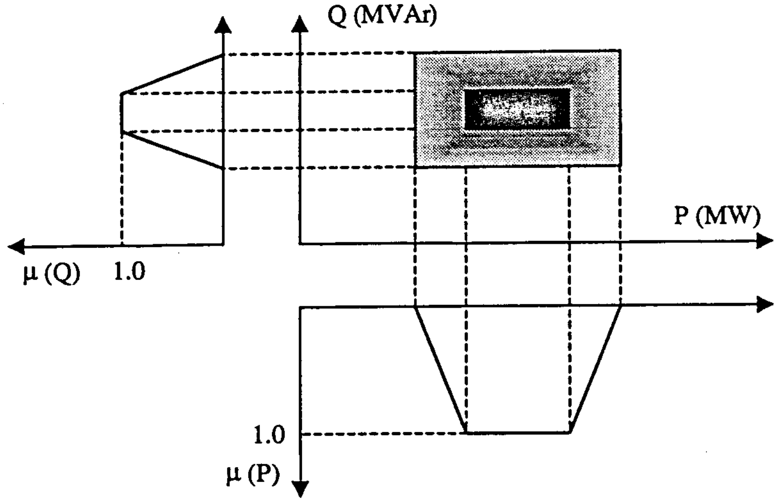


Figure 8 Fuzzy load, in terms of P and Q with their trapezoidal membership functions. In black: region for which membership has maximum value 1.

4.4.3. Linguistic interfaces

One good thing about fuzzy numbers is that they are able to represent many qualitative aspects of the engineering knowledge (or lack of knowledge).

An uncertain load may simply be declared as "*more or less L_2*", and be represented by a fuzzy number such as in Figure 7 (left). Alternatively, through a declaration such as: "*load may occur between L1 and L4 MW but it is likely to be between L2 and L3*"; this sentence is translated into the trapezoidal fuzzy number in Figure 7 (right). The same would apply, for instance, to a reliability index such as a failure rate.

A general approach departing from a basic estimate and changed by means of linguistic modifiers seems to be adequate for most purposes. Some sets of modifiers ("little", "some",...) qualifying the type of uncertainty have been proposed. In some cases the trapezoidal shape of membership functions is kept, with calculus effort advantages; in another cases [Mir91a], these modifiers have been derived from fuzzy set operations such as contraction and dilation [Zim85], leading to non-linear membership functions.

Let us again stress that the type of uncertainty we are talking about is not the probabilistic one: rather the linguistic declaration may be produced by an expert, in a situation where data that are more specific may be missing.

4.5. Fuzzy Power Flow

DC Fuzzy Power Flow Model

The DC FPF model [Mir90b] allows one to obtain fuzzy descriptions of the bus angles and branch active power flows, departing from uncertain (fuzzy) loads and generations, given as possibility distributions (ΔP) of the deviations from specified active injected powers.

The basic model derives possibility distributions of both bus angles [$\Delta\theta$] and active power flows [ΔP_{ik}] using the known DC model matrix [B] and sensitivity coefficient matrix [A], and the rules of fuzzy arithmetic

$$[\Delta\theta] = [B]^{-1}.[\Delta P] \tag{14}$$

$$[\Delta P_{ik}] = [A].[\Delta P] \tag{15}$$

The uncertainty of nodal injection in the slack bus depends on the uncertainties of the other injections. The results obtained are the widest distributions obtainable from the combination of all possible scenarios implicit in ΔP. Different chosen slack buses will in general lead to different results, because equations (14) and (15) do not condition the uncertainty at the slack bus.

One important thing to bear in mind, and that the DC FLF technique clearly demonstrates, is that the extreme value of a branch flow can be obtained at a solution where some generators or loads are not at any of their extreme possible values.

AC Power Flow Model

The AC fuzzy power flow allows the imprecision of bus data to be reflected on the voltages, angles, active and reactive flows and losses, generated active and reactive powers and currents. The algorithm, fully presented in [Mir90b, Sar91a, Mir92a], can be summarized considering the following points:

a) the possibility distributions are built departing from a previous deterministic Newton-Raphson (NR) load flow giving crisp values of voltages (V_d), angles (θ_d), active and reactive flows ($P_{d.ik}$ or $Q_{d.ik}$), generated powers and losses.

b) the fuzzy $\begin{bmatrix} \dfrac{\Delta\theta}{\Delta V} \end{bmatrix}$ angle and voltage deviations are evaluated using the Jacobean matrix [J] built in the last NR iteration and the fuzzy deviations $\begin{bmatrix} \dfrac{\Delta P}{\Delta Q} \end{bmatrix}$ of the injected active and reactive powers. The possibility distributions of voltages and angles $\begin{bmatrix} \dfrac{\Delta\theta}{\Delta V} \end{bmatrix}$ are obtained by combining the deterministic values $\begin{bmatrix} \dfrac{\Delta\theta d}{\Delta Vd} \end{bmatrix}$ and the distributions of their deviations.

$$\begin{bmatrix} \dfrac{\Delta\theta}{\Delta V} \end{bmatrix} = \begin{bmatrix} \dfrac{\Delta\theta d}{\Delta Vd} \end{bmatrix} + [J]^{-1} \begin{bmatrix} \dfrac{\Delta P}{\Delta Q} \end{bmatrix} \tag{16}$$

c) Branch flows and generated powers are non linear functions of V and θ which can be replaced by the first terms of their Taylor series expansions around the crisp point obtained in a). As an example, for the active power flow in branch i-k, expression (17) is used to obtain ΔP_{ik} from ΔV_i, ΔV_k, $\Delta\theta_i$ and $\Delta\theta_k$ fuzzy descriptions. Fuzzy P_{ik} is finally given by (18).

$$\Delta P_{ik} \approx \frac{\partial P_{ik}}{\partial V_i}\Delta V_i + \frac{\partial P_{ik}}{\partial V_k}\Delta V_k + \frac{\partial P_{ik}}{\partial \theta_i}\Delta\theta_i + \frac{\partial P_{ik}}{\partial \theta_k}\Delta\theta_k \tag{17}$$

for $V_i=V_{di}$; $V_k=V_{dk}$; $\theta_i=\theta_{di}$; $\theta_k=\theta_d$

$$P_{ik} = P_{d.ik} + \Delta P_{ik} \tag{18}$$

where

$$\begin{bmatrix} \dfrac{\Delta\theta}{\Delta V} \end{bmatrix} = [J]^{-1} \begin{bmatrix} \dfrac{\Delta P}{\Delta Q} \end{bmatrix} \tag{19}$$

d) For branch currents and losses, the above is not a satisfactory method in lightly loaded lines or for lines where reversing of power flows may occur. [Mir90b] presents an algorithm to be applied in these situations.

Fuzzy flows with nodal dependencies

The fact that the generation is driven by the consumption imposes that for every possible scenario of loads there must be at least one scenario of generations that meet the demand. There are thus several ways to define the fuzzy power flow problem; for instance:

(a) Declare uncertainties at all nodes but the slack bus;

(b) Declare uncertainties at all nodes but the slack bus and admit that total uncertainty in generation equals total uncertainty in consumption;

(c) Declare uncertainties at all nodes but the slack bus and declare for this node the max and min admissible generation values; this is equivalent to describing by a rectangular membership function the feasible power injection at this node.

(d) Declare uncertainties at all nodes including the slack bus.

Details for solving these problems can be found in [Mir92a] and [Mir92b] concerning optimal power flow.

Radial Fuzzy Load Flow

In distribution system planning, load flow calculations are performed in radial networks which makes them much easier. Therefore, a simplified version of the model described above was used. Basically, the fuzzy load flow used in the methodology described in chapter is a fuzzy extension of the algorithms described in [Mir83].

4.5.1. Fuzzy voltage drops

A solution is generally considered unfeasible if the voltage drop exceeds a given voltage threshold (in the example in chapter 7, we have considered 8%). As we are dealing with fuzzy voltage drops resulting from the Fuzzy Power Flow, the fuzziness of node voltage values may be translated into a fuzzy index as follows (Figure 9).

if $V(\alpha)$ is a fuzzy node voltage, expressed through a membership function associated to membership or possibility level α,

$I(v)$ is a fuzzy description of the acceptable voltage levels and unacceptable thresholds. $I(v)$ is a fuzzy interval with the properties of a fuzzy number. $C[I(v)]$ is the complement of this number: $C[I(v)] = 1 - I(v)$.

then VQ is a fuzzy voltage quality index given by the functional composition

$VQ(\alpha) = \text{Max}\ \{ C(I) \circ V \}$, at every α level

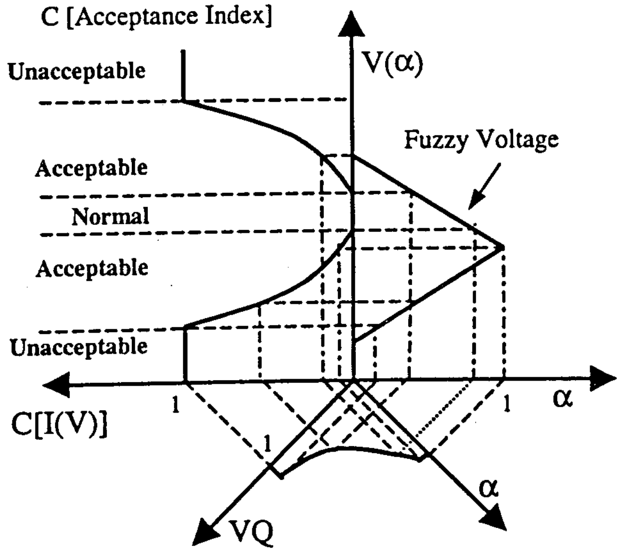


Figure 9 Building a fuzzy voltage quality index VQ(α) from a fuzzy nodal voltage V(α) and a quality index I(V)∈[0, 1] shown by its complement C[I(v)].

4.6. Fuzzy reliability

This section will offer a small introduction on the subject of fuzzy reliability theory. A full discussion of the subject may be found in [Mir96b] and the modeling used in the proposed methodology is presented in chapter 6.

Component fuzzy indices, as data

There are several reasons why reliability indices should be taken as fuzzy. By this statement, we do not mean replacing probability distributions (usually exponential) by possibilities: we just mean that the expectation of such distributions should not be a crisp value.

In fact, in many cases reliability indices used in planning studies are values taken from data bases and adopted by similarity, considering the type of the devices, their conditions of installation, the efficiency of the utility in maintenance actions, etc. Therefore, an expert will often propose an index such as "a failure rate of *more or less 0.001*", which is after all a fuzzy description! Fuzzy models will then allow the analysis of the impact on system indices from this uncertainty in component data, such as on the unavailability of supply or on the average annual energy not supplied (which become associated also with possibility distributions).

Fuzzy Reliability Indices (FRI) are not in competition with probabilistic indices: they retain the properties of these kind of indices. However, they also share properties of qualitative uncertainty, and they may be described by what sometimes mathematically is called a random *walk of a fuzzy number.* In simple terms, this implies that, among other things, the mean value of a probability distribution is not a crisp real value, but it is described by a membership function.

The general idea of considering fuzzy reliability indices has been proposed, for instance, by Kaufmann et al. [Kau88]. In power system studies, Fuzzy Reliability Indices (FRI) have been proposed for power system reliability evaluation using the general min cut set method [Mir90a]; the design of distribution systems has incorporated FRI in some proposals, such as in [Mir91a] and [Sei94]. Since the methodology proposed in chapter 6 uses the min cut set method for reliability assessment, it will be presented in section 4.6.2. The next section will present an introduction to FR models.

4.6.1. Fuzzy reliability models for system components

Classical reliability studies are based on two assumptions: *the probabilistic assumption*, which states that the system behavior can be fully described through probabilistic models (the laws and axioms of probability apply, and the behavior is characterized in the context of probability measures); and *the binary state assumption*, which requires that the state of an equipment, of failure or functioning, shall be completely defined.

However, the consideration of some fuzzy concepts leads to several reliability models [Cai91]; we have therefore:

- the PROBIST model: assuming the probabilistic assumption and the binary state assumption;
- the PROFUST model: keeping the probabilistic assumption, but introducing a fuzzy state assumption;
- the POSBIST model: introducing a possibilistic assumption as to the description of the events and the laws governing their repetition, together with the binary state assumption; and
- the POSFUST model: with the possibilistic assumption and the fuzzy state assumption.

This section will focus only in the PROFUST model for component reliability; but it is necessary to clarify the meaning of the fuzzy state assumption: one cannot precisely define the state of a component, having therefore a fuzzy success and a fuzzy failure state; or one can define exactly the success and the failed states, but one is unable to define precisely how the transition occurs, namely how often. Therefore, the probability of finding a particular component at one state is given by a fuzzy number.

However, because the probability assumption is retained, particular probabilities of instantiated outcomes must add up to 1 in the whole state space. This introduces a dependency between fuzzy probability values that is not found in other fuzzy models.

The raw data are usually the failure rate λ and the mean repair time r. We may have, for a λ value, instead of a crisp number such as 0.01 fl./year, a fuzzy description such as a best estimate of (0.01) and an interval of confidence of [0.008, 0.012] fl./year. Therefore,

$$\lambda_{\alpha}= [0.01- (0.002\,(1-\alpha)),\ 0.01+(0.002(1-\alpha))] \tag{20}$$

will represent a triangular fuzzy failure rate, with

$$\lambda_{\alpha} = [\lambda_{\alpha}^{-}, \lambda_{\alpha}^{+}] \tag{21}$$

being the interval of confidence at level

$$\alpha \in [0 ; 1] \tag{22}$$

Fuzzy exponential distribution

We can now consider the extension of the Reliability function R(t) to the fuzzy case. For every **t**, the boundaries of the λ interval define

$$R_\alpha(t) = [R_\alpha^+ = e^{-\lambda_\alpha^+ t} ; R_\alpha^- = e^{-\lambda_\alpha^- t}] \tag{23}$$

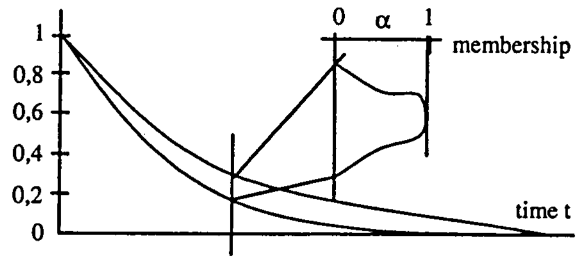


Figure 10 Fuzzy reliability function, and an illustration of its membership values at a given t

We have, at each α, an interval of confidence delimited by a lower survival law R_α^+ and an upper law R_α^-. This defines a fuzzy reliability function $R_\alpha(t)$, such as suggested in Figure 10. The mean time to failure of a classical Reliability function is given by MTTF = $1/\lambda$. Taking in account (23), one may define a fuzzy MTTF for the fuzzy reliability function $R_\alpha(t)$, based on the interval of confidence representation

$$MTTF_\alpha = [\frac{1}{\lambda_\alpha^+} ; \frac{1}{\lambda_\alpha^-}] \tag{24}$$

This definition is straightforward. We define an interval mean time to repair r_α in relation with repair rates μ_α^- and μ_α^+ such as

$$r_\alpha = [\frac{1}{\mu_\alpha^+} ; \frac{1}{\mu_\alpha^-}] \tag{25}$$

This means that the fuzzy numbers **r** and **μ** are calculated as the inverse of each other: **r = 1/μ**. This operation is valid because the fuzzy numbers **r** and **μ** are defined on the positive axis of the real line.

Series and parallel systems

The FR of a serial system of n components is given by

$$R^{ser} = R_1 . R_2 . \dots . R_n = \prod_{i=1}^{n} R_i \tag{26}$$

We have here the product of fuzzy numbers, all defined in the positive axis. In terms of intervals of confidence, if each

$$R_\alpha = [R_\alpha^- ; R_\alpha^+] \tag{27}$$

we have

$$R_\alpha^{ser} = [R_{\alpha 1}^- . R_{\alpha 2}^- . \dots R_{\alpha n}^- ; R_{\alpha 1}^+ . R_{\alpha 2}^+ . \dots R_{\alpha n}^+] \tag{28}$$

The result of such operation over triangular R_i does not give a triangular result; however, a triangular approximation is usually acceptable.

The FR of a n-parallel system is given by

$$R^{par} = 1 - \prod_{i=1}^{n} (1 - R_i) \tag{29}$$

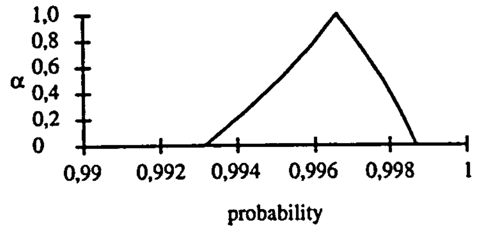


Figure 11 FR of a 3-component parallel system

Admit a parallel of three similar components, each of them with a reliability described as a triangular fuzzy number [0.81;0.85;0.89]. Applying (29) one reaches a result depicted in Figure 11. again, it is not a triangular fuzzy number but still, a reasonable approximation.

4.6.2. Min Cut Set method with fuzzy indices

The min cut set method is a well-known approach to the assessment of the reliability of general systems. It is not an exact method, but within the relative values for component reliability indices usually found in power systems, it gives a satisfactory accuracy. It is appropriate to find nodal or local indices, and is especially adequate for continuity assessment, in the design of power stations or substations, and in sub-transmission or distribution systems. A relevant paper referring to this technique is [All76]. The fuzzy model has been first discussed in [Mir90a].

The min cut set method demands one to determine which component failure mode (a "cut") causes the supply to a node to be interrupted; then, reliability indices are calculated for each cut - applying, if necessary, formulas for components in parallel; and finally, the reliability calculations are performed as if the cuts were independent entities in series - namely, applying the fuzzy formulas for component series.

Although one will have, during calculations, repeated entries of the same fuzzy indices (for instance, two 2nd order cuts may share one component), we can apply directly the "fuzzyfied" formulas of the crisp method because, in series or parallel components, we only find additions and multiplications - therefore, a variable will never divide itself or subtract from itself.

We can get also a fuzzy index for the average annual energy not supplied ENS. The fuzzy ENS value results from the multiplication of a crisp number PNS for a fuzzy number U:

$$\mathrm{ENS} = \mathrm{U.PNS} \tag{30}$$

The equations adopted in the fuzzy min cut set method are similar to the ones used in the "crisp" traditional approach (except that fuzzy numbers are used instead). However, as those equations include products and divisions, a special care must be taken in the order of performing them, so that no unnecessary uncertainty is added to the results.

4.6.3. Frequency and duration models

We now present some approximate formulas for fuzzy calculations in parallel and serial systems for two components, arranged so that the fuzzy calculations may remain valid.

Parallel systems

System failure rate $$\lambda = \lambda_1 \lambda_2 (r_1 + r_2) \qquad (31)$$

System unavailability $$U = \lambda_1 r_1 . \lambda_2 r_2 \qquad (32)$$

System mean repair time $$r = \frac{1}{\frac{1}{r_1} + \frac{1}{r_2}} \qquad (33)$$

Serial systems (no overlapping failures, or neglected)

System failure rate $$\lambda = \lambda_1 + \lambda_2 \qquad (34)$$

System unavailability $$U = \lambda_1 r_1 + \lambda_2 r_2 \qquad (35)$$

The calculation for system mean repair time is not straightforward. One can derive exact formulas, but it is easier to follow a de-convolution approach. In fact, as the following fuzzy equation must hold:

$$U = \lambda\, r \qquad (36)$$

then an easy way of obtaining **r** from (36) is by finding the fuzzy number **r** that multiplied by the fuzzy λ would give the fuzzy **U**; this is not the same as solving

$$r = \frac{U}{\lambda} \qquad (37)$$

by applying the rules of fuzzy arithmetic.

The impact in the results of optimization procedures from the imprecision in reliability data has also been dealt with. For instance, in the problem of deciding on the location and type of switching equipment to be included in a distribution network, considering the investment and the quality of supply [Mir91a], the cost of equipment and also the cost of the kW disconnected and the kWh not supplied had also fuzzy definitions. The power system is assumed repairable and the traditional probabilistic concepts keep their validity, but the mean values of the probability distributions are now fuzzy.

4.6.4. Fuzzy System Indices, as analysis results

Another way of obtaining system fuzzy indices, departing from the traditional probabilistic component indices, is by combining a probabilistic simulation of the failure-repair cycle with a possibilistic representation of the uncertainty in generations and loads.

The Monte Carlo simulation methodology has been a well known and used technique in power system reliability evaluation. In classic models, loads are represented by crisp values so that each sampled state is run through an OPF model (usually a DC approach) in order to derive, among others, indices such as the expected power not supplied. The integration of load uncertainties, namely non-probabilistic ones, may be especially important in long term planning.

This approach was first sketched in [Mir91b] and fully developed in [Sar92]. The DC Fuzzy OPF model, referred to above, can be used to analyze sampled states in a Monte Carlo process, so that one can derive estimates of reliability indices given by fuzzy numbers. In fact, for each sampled state one can get its PNS membership function.

These membership functions can be aggregated according to the fuzzy rules so that fuzzy estimates of the corresponding expected value of PNS become available. A fuzzy LOLP index may also be derived.

The advantage of this approach is that a whole continuous set of uncertainties (in loads) is added to the problem, but the Monte Carlo simulation will only take about the same number of iterations and some 40% longer in time to converge, therefore rendering feasible this process.

As a final comment, it should be emphasized that, in this process, no load values are sampled: imprecision in load forecasting is dealt within the fuzzy load model. Besides, it must also be stressed that this simulation process integrates, in an innovative way, two conceptually different uncertainties as random information related to the probabilistic nature of the component outages is combined to load qualitative knowledge represented by fuzzy numbers.

4.7. New concepts

Since the possible values of branch flows could exceed branch limits, we have to define some new criteria: the *robustness* of a solution (based on the α-cuts above which the fuzzy flows are within branch limits) and the *inadequacy* of a network (based on the sum of the fuzzy flow subsets that exceed branch limits). In this new framework, the concepts of fuzzy regret and fuzzy dominance also have to defined.

4.7.1. Robustness and Exposure

The concept of robustness in risk analysis was introduced in chapter 2. Robustness is one measure of the safety or lack of risk in a decision and, in the proposed methodology, a non-fuzzy criterion. We define technical robustness of a solution as an index deriving from a FPF study. As it is clear in Figure 12, for some scenarios of loads the power demand will have the possibility of exceeding the limit P_{max} in branch capacity. This happens below membership level α; so,

$$RO=(1-\alpha) \tag{38}$$

is an index of how much uncertainty the system is able to cope with - it is a Robustness index, and

$$EX = \alpha \tag{39}$$

is an Exposure index. Therefore, minimizing α, in terms of planning, means that the planner wishes to accept solutions that cover or are technically sound in a wider range of possible future scenarios.

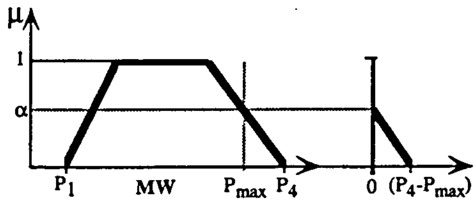


Figure 12 Possibility distribution for a line Power Flow. Below level α, load may exceed the thermal limit P_{max}. On the right side the definition of branch inadequacy.

4.7.2.Inadequacy

The right tail of the branch power flow distribution will be used to define the fuzzy concept of *branch inadequacy* $\mathbf{IN_{br}}$ (Figure 12). It may be interpreted as the consequence of some adverse demand scenarios, for which the distribution system will not be able to cope with. Facing this structural constraint, the utility will not authorize more loads to connect to that particular line or should take costly emergency measures. Therefore, the actual load allowed will be less than the forecasted load - it is an event of a nature different from disconnecting load that was previously being supplied.

One is facing here a situation of possible repressed demand, which incurs in social costs of a nature different from those related to power not supplied as a consequence of failures. Therefore, we should keep this analysis independent of reliability evaluations, by making explicit a new criterion.

We have thus defined a measure of *system inadequacy* IN_{sys} of a distribution system as the (fuzzy) sum of the possibility distributions of branch inadequacies:

$$IN_{sys} = \sum_{br=1}^{n.\ branches} IN_{br} \tag{40}$$

One of the planning objectives will be to minimize inadequacy. It is also noticeable that in the comparison of two alternatives they may share the same robustness index and, nevertheless, have different inadequacy values.

4.7.3. Hedging

If a given decision involves risk, the planner must find a way to limit its effects. In risk analysis, *hedging* occurs when one is willing to pay an extra in order to reduce or even avoid the consequences of an adverse future.

In [Mer90], H. M. Merrill and Allen Wood present a simple but clear example to illustrate the concept of hedging based on a problem with a single attribute and a single uncertainty. In the example presented there is, initially, no robust solution:

Plan A: Buy a car

Plan B: Ride taxis

Uncertainty: Will I have an accident?

Attribute: cost

FUTURE 1 (no accident)		FUTURE 2 (accident)	
PLAN	COST	PLAN	COST
A	Low/Medium	A	High
B	Medium/High	B	Medium/High

It is clear that no plan is robust and a hedge is needed. So Plan C is introduced:

Plan C: Buy a car *and* buy insurance

FUTURE 1 (no accident)		FUTURE 2 (accident)	
PLAN	COST	PLAN	COST
A	Low/Medium	A	High
B	Medium/High	B	Medium/High
C	Medium	C	Medium

Plan C is a robust plan, hedging against possible unwanted outcomes from uncertainties.

In fact, there is nothing new about hedging in Power Systems planning. Consciously or not, engineers and decision makers have always used some kind of hedging policy in planning. Typically, engineers would project the systems in order to satisfy the load in an economical manner, and then add some extra equipment or reinforce lines to make the system more robust and/or reliable. This "just in case" policy is clearly a good example of hedging, where the planner is willing to spend a little more for improved security in facing the possible futures.

In order to formalize these concepts in power system planning, we may define hedging as *reducing the value of the exposure index* referred on the previous sections (thus generating a solution more insensitive to uncertainties). Within the framework of fuzzy set models, some hedging approaches have already been proposed [Sar95, Sar94], related to investments in the generation or transmission system, that lead to a ranking of investment decisions associated with increased robustness of design. This can also be seen in Figure 13, if one imagines that moving from option A (that permits a maximum load P_{Amax}) to B (P_{Bmax}) reducing exposure to adverse load scenarios, has been done with some extra expenditure on a given system. The alternative leading to α_A is less robust than the other one: for this option, only up to P_{maxA} the system will have no load disconnected. Consequently, moving from A to B will also reduce inadequacy of the system design ($In_A - In_B$). Improving system design in order to cope with extra load may be considered, therefore, a hedging strategy.

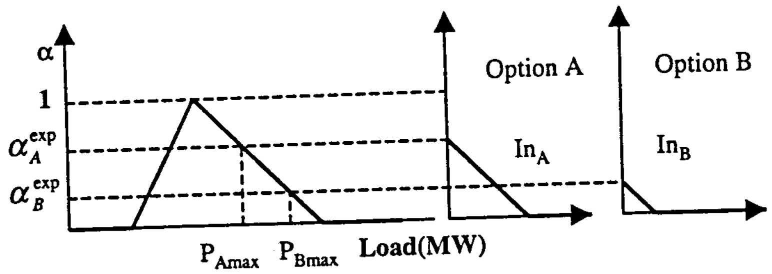


Figure 13 Load uncertainty and power disconnected for two system designs A and B. Design B copes with more load than design A and, consequently is less inadequate.

4.7.4. Comparing solutions

During the planning activity, comparison of alternative solutions or configurations will occur. If crisp models are used it is possible to compare, for example, the value of the Power Not Supplied (PNS) for two system designs and establish a preference order.

However, when dealing with fuzzy definition of variables and values, this comparison is no longer trivial. Certainly, for two situations ranked at the same robustness value (1-α), a comparison is still possible and meaningful.

Nevertheless, two configurations with different exposure levels and different inadequacies cannot be compared directly, because the assumption that they serve the same load is no longer valid (since we admitted situations of repressed demand). A multicriteria comparison is therefore necessary: the alternatives should be compared on the basis of fuzzy reliability indices evaluated at the same α level, but at the same time, increased solution robustness should also be taken in consideration.

Two solutions with robustness associated with α_1 and α_2 are only comparable for the levels above which they are both robust, which means that they would both operate normally and fail from time to time. Below Max $\{\alpha_1 , \alpha_2\}$, one of them would still be able to meet the loads, but the other one would fall into a situation of possible repressed demand.

Comparisons between more than two solutions must be made by successive distillations guided by the (1−α) index.

4.7.5. *Fuzzy regret*

The concept of Regret is essential to this thesis and the concept of fuzzy regret has thus to be defined.

Figure 14 displays the typical decision making dilemma, in comparing the cost of two alternative plans A and B. If crisp calculations (at α=1) were made, it would seem that plan B should be preferred, because it is less expensive; however, a fuzzy modeling of data uncertainties makes one hesitate. One realizes that, below α_1, B may exceed A, and this becomes obvious below α_2; i.e., if the decision maker has uncertainties in data larger than the ones defined at the α_1 level, then he must weight the risk of making a regrettable choice [Mir90a].

If uncertainties could be constrained to remain above α_1, then his decision in favor of plan B would be the best whichever instantiation of the data values would occur.

There are some methods to defuzzify and rank fuzzy numbers, as we have seen. However, they should not be applied directly to rank costs. One important question that direct ranking of fuzzy costs may not answer is: "how much will one regret, for choosing one alternative instead of another?".

Let us assume a context of minimization. We now introduce the new concept of **Fuzzy Regret** Reg(A|B), the possible regret felt by a decision maker when he chooses A instead of B, and then B happens

$$\mathbf{Reg(A \mid B) = max\{0,(A_\alpha^- - B_\alpha^+)\}} \qquad \textbf{(41)}$$

Of course, if one chooses A, and A results better than B, no regret will be felt - that is why expression (41) does not allow for negative regrets.

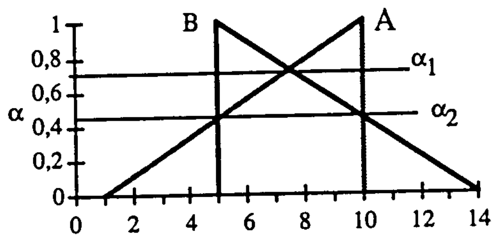


Figure 14 Triangular fuzzy costs of two planning alternatives. Plan B seems to cost less than plan A, but is perhaps riskier.

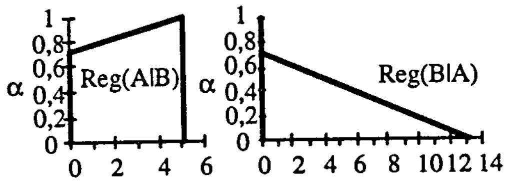


Figure 15 Fuzzy Regrets, referring to Figure 14 Left - regret for choosing A instead of B; right - regret for choosing B instead of A

Figure 15 shows Reg(AIB) and Reg(BIA) in relation to Figure 14 The decision dilemma lies now between these two regrets. We can now apply the Center of Mass technique and rank the two fuzzy regrets: we get Reg(BIA)<Reg(AIB) - therefore, decision B would be preferable to decision A, in this comparison (however, applying the Center of Mass technique directly to the fuzzy costs A and B would have ranked A<B; this shows the importance of changing the point of view from objective function values to decisions).

4.7.6.Fuzzy dominance in multicriteria problems

This section will discuss the concept of dominance in multi-criteria problems.

The concept of efficient, dominant or Pareto-optimal set of solutions, in a multi-criteria environment was already defined before. The concept of non-dominated solutions is clear in Figure 16 which represents three solutions in a two-attribute space: C is dominated by A and B, and these are non-dominated, in a context of attribute minimization.

The fuzzy case is somewhat more complicated. In Figure 17 we have three solutions D, E and F, with fuzzy values on each attribute (say, a fuzzy cost and a fuzzy power not supplied, as a consequence of uncertainties in data).

To make it more evident, admit that this fuzziness is just represented by intervals at α=**0** (giving a rectangle) and a point at α=**1**, and suppose that these points would represent the Center of Mass of each fuzzy solution.

If a decision maker evaluates the solutions only based on the Centers of Mass, he will conclude that D dominates E and that both dominate F. However, the fuzzy representation shows that the uncertainties in data do not allow such a simplistic conclusion: for example, at level α=0, there is the possibility that E may dominate D (while still both dominate in every case F).

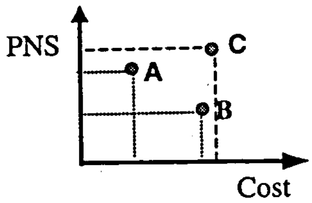


Figure 16 Three crisp multi-attribute solutions (Power Not Supplied v. Investment Cost)

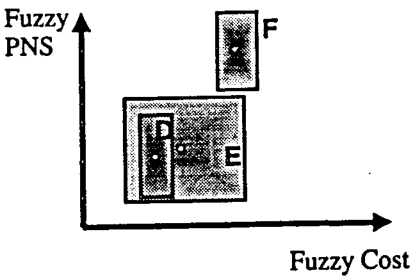


Figure 17 Three fuzzy multi-attribute solutions

This illustrates very well one of the important contributions of fuzzy modeling to the understanding of the problems dealt with, in Power Systems: contrary to crisp models, where the concepts of solutions and decisions are confused, in fuzzy models there is a clear separation between the (fuzzy) solution values and the decisions that must be made, and this distinction derives from the evaluation of risk, resulting from the simultaneous consideration of multiple possible scenarios.

In a crisp model, solutions may be ranked for decision according to their solution values, because the ordering of regrets, in pairwise comparisons, may be considered the same as the ordering of solution values.

However, in a fuzzy model this is not necessarily true: the ordering of central points or of centers of mass does not reflect the possible regrets, as we have seen before. This happens also in a multi-criteria environment (one might in theory even reach situations of intransitivity of preferences).

The discussion of decision making in fuzzy context is a very interesting topic of research in itself. Related to the subjects discussed in this thesis, the fuzzy modeling allows the identification of risks and the evaluation of regrets, when opting for one solution instead of another. For the multi-criteria case, an extension of the Fuzzy Regret concept, developed in section 4.7.5., must be considered.

4.8. Scenario Trees

One technique that tries to capture the implications of the uncertainty related to trying to "guess the future" is known as *tree of futures*, representing possible scenarios at different time stages and the admitted possible transitions among them.

This approach has been adopted in Power System Planning, for example in [Gor93] regarding generation expansion planning. The tree of futures is a scenario technique, where the scenario is a trajectory or path in the tree.

It was mentioned, in the previous sections, that some type of continuous uncertainties should be modeled using probabilistic models while other should be modeled as fuzzy numbers. We will use the tree of futures technique to model events or possibilities perceived as discrete - for instance different economic or legislative background scenarios leading to different load levels, globally or locally. Also, uncertainty related to the realization of large projects (e.g., a large sports stadium or a shopping mall) can be modeled using this technique, since it becomes clear that none of the other techniques is adequate for modeling this type of discrete uncertainty. We cannot plan, for example, for the construction of *half a stadium*, if the subjective probability of the project being implemented is 50%. Willis in [Wil96] makes this point very clear, using mainly spatial examples.

However, until recently, each scenario was studied as if it were deterministic

Figure 18 tries to illustrate a new concept, the *tree of fuzzy futures*: it combines a discrete understanding of the possible scenarios in the future with a fuzzy definition of each scenario. The tree in the figure covers for time steps and six future stages, and the possible trajectories through them. For pictorial reasons, it is only possible to represent intervals of uncertainty in this figure. Considering that the *yy* axis represents total load values, then within each interval we have a fuzzy definition of the load forecast.

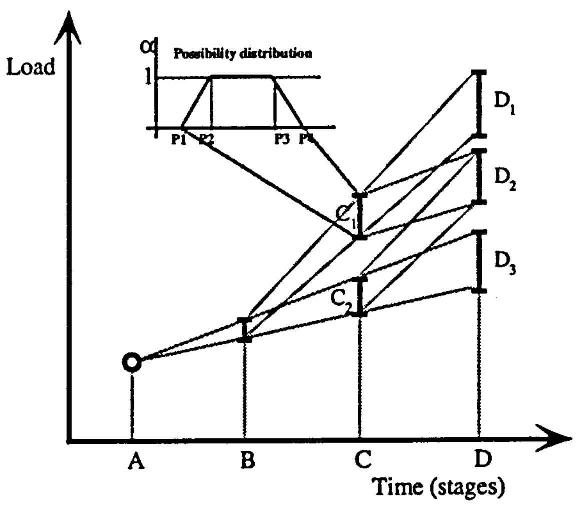


Figure 18 Tree of fuzzy futures and possibility distribution for future C1

This technique represents a way of introducing granularity in the information about possible futures - the paths form a discrete set, but each path is represented as "blurred", also contaminated with continuous uncertainty modeled by fuzzy numbers. In the sense of the tree of futures, each path is a possible future (a future is not only a static *picture* in some stage).

4.9. A general model under risk analysis

After reviewing some tools that add a fuzzy dimension to data and results in power system modeling, this section will examine, in a generic fashion, how these tools may be used in power system planning studies (not restricted, for the time being to distribution planning).

The first relevant question to ask is: how would a planning methodology based on these models work, under the general paradigms of decision making and risk analysis?

1. First, one should try to find the (supposedly) "best" expansion plan for each trajectory in the tree. Each of these "best" plans is then determined conditioned to an *a priori* knowledge that the future will occur in a pre-determined sequence of stages. The calculation of each "best" plan would require a dynamic model that might deal with the multi-temporal nature of the problem,

However, because each trajectory is fuzzy defined, one must make use of fuzzy analytical tools such Fuzzy Power Flow or Fuzzy Optimal Power Flow, fuzzy reliability and fuzzy economic evaluation.

If one defines several criteria to qualify alternatives to system expansion (investment, reliability, losses, voltage quality - and, of course, solution robustness and inadequacy, because the underlying load data are fuzzy) then multi-criteria decision methods should be applied. The notion of "best" plan may then be replaced by the concepts of conditional decision set and ideal values of this set, in the multi-criteria sense.

2. After having the ideals for each trajectory, one proceeds to define a risk aversion strategy. The decision dilemma, at each node of the tree of futures, is that in general one will have different optimal decisions or plans depending on the futures departing from a node, and the option for one of those plans might prove disastrous if by chance the "wrong" future occurs. The risk aversion strategy will therefore try to determine which decisions minimize the possible future regret, i.e., which plan minimizes the sum of the differences between it and all the ideals departing from the node (this sum could be weighted, for instance by some subjective probabilities).

If the values of the several planning criteria are expressed as fuzzy numbers, then this new objective (of minimizing the "distance" between the actual decisions and the ideals on each trajectory within the tree of futures) will also be expressed in fuzzy terms.

This general model could seem huge and leading to unmanageable computer calculations, making it unrealistic to use in the power industry. However, the next chapter of the thesis will present new emerging techniques that open promising pathways into the ability of actually dealing with such a complex model. Chapter 6 will elaborate on the principles sketched in this chapter and propose a full methodology for the expansion of distribution systems.

4.10. Conclusion

This chapter presented the basis for a complete representation of uncertainties in Power System planning, according to the type of uncertainty and amount of associated information.

The focus was on fuzzy sets since it is a relatively novel technique and there are some interesting consequences of including fuzzy set theory contributions in Power System planning [Mir94a]:

- Fuzzy sets may be used to model some type of uncertainties where probabilistic models do not seem to be adequate or when there is no sufficient information.
- We may also use fuzzy sets to model soft constraints in mathematical programming models.
- In decision making models, fuzzy sets may be used to account for the compromise among objectives, as a consequence of preferences and priorities being usually vaguely defined.
- Fuzzy numbers allow the continuous representation of future scenarios. Therefore, fuzzy models seem adequate to examine the process of reaching a decision in system planning through new approaches, such as risk analysis.
- Dealing with fuzzy variables implies in fact an extra computing effort, when compared with crisp models; however, the increase in computing time is usually not too large, so the adoption of the technique is perfectly feasible.

So, in fact, accepting the use of fuzzy concepts in planning seems to be just a matter of time: time enough to build a scientific culture leading to the intuitive perception of the new dimensions of analysis and synthesis this tool is able to bring.

The consideration of uncertainties, and their modeling using fuzzy numbers lead to the revision (within the framework of expansion planning) of concepts such as solution robustness and system inadequacy. These concepts, together with the notion of hedging, are well known to both planners and decision makers and were not necessary in crisp models. They will be central on the development of the dissertation, since they allow us to move one step closer to a realistic representation of reality and to modeling the way planners think.

5. ALGORITHMS AND MATHEMATICAL TECHNIQUES

Natural selection is a mechanism for generating an exceedingly high degree of improbability.

Sir Ronald Aylmer Fisher (1890-1962)

5.1. Summary

Now that we have a potential model for the distribution planning problem, this chapter reviews the possible algorithmic approaches. Reflecting on the complexity of the problem, it realizes the limitations of conventional methods and presents some innovative mathematical-algorithmic techniques based on natural evolution, known as Evolutionary Algorithms. These algorithms are introduced as robust problem-solvers in a multicriteria environment.

5.2. Introduction

The previous chapters set the basis for a possible model for the distribution planning problem. They analyzed the complexity of the questions involved and the previous chapter proposed a comprehensive model for the representation of uncertainties and risk. This chapter will discuss some aspects related to the techniques and algorithms used to solve the problem. Naturally, it is impossible to separate the two concepts (models and algorithms) so this chapter will review the techniques keeping in perspective the problem we are dealing with.

In the last 30 years, several methodologies were proposed for distribution system planning. These techniques have evolved together with the development of scientific knowledge and benefiting from the remarkable increase in computational capacities.

We may state that the history of these methodologies is the history of the conflict between two crucial factors in planning models:

- On one hand, computation capacity and precision.

- On the other hand, the network model (number and type of simplifications) and the planning model (questions related to, for example, model dynamics and uncertainties)

It is clear that the common feature to most, if not all, of these models is the use of classical operation research techniques for the determination of an optimal solution (or at least a solution considered satisfactory).

This chapter will not discuss all the variety of algorithms that have been used to solve the problem, but only the most relevant in the short history of planning models.

First, it will refer the main conventional algorithmic approaches for distribution expansion planning: numerical optimization, dynamic programming, mixed integer programming, decomposition algorithms and heuristic methods.

Then it will review some less conventional algorithmic approaches like simulated annealing and TABU search. These techniques have been used with some degree of success in Power System planning.

Finally, it will introduce Evolutionary Algorithms as powerful search and optimization mechanisms, with exceptional characteristics for the application to the distribution planning problem. Specifically, these algorithms have shown to be remarkably useful when dealing with multiobjective problems, where several alternatives are needed. A small survey on the use of Evolutionary Algorithms in Power Systems will also be presented in order to justify the worth of these methods.

In each section, there will also be a critical discussion on the advantages and limitations of each methodology.

5.3. Conventional algorithmic approaches

This section will refer the traditional approaches to the network expansion problem in Power systems. Curiously (or maybe not), these cover approximately all the standard techniques in the operation research field. Unfortunately, these techniques inherited from XIX century mathematicians the idea of a perfect, differentiable and linear universe that, as everyone knows, does not correspond to reality. The actual world is made of "bad-behaved" functions that are non-linear and non-differentiable. The real universe is mostly non-convex, non-contiguous and full of local optima. That is probably the reason of the relative little success of these techniques in the field of power system expansion planning.

5.3.1.Numerical optimization

Numerical optimization may be considered the traditional approach for optimization, with the advantage of, at least in theory, converging to the optimal solution and not just to a "good" solution. However, methodologies based on this type of optimization cannot, in practice, be applied to real dimension cases in distribution planning, due to the extreme mathematical and computational complexity introduced by the discrete and non-linear nature of the problems being considered.

5.3.2.Dynamic programming

Methods based on Dynamic Programming are apparently attractive, for they naturally allow the representation of the dynamic nature of the planning process. Another advantage of these methods is that we do not need to linearise the objective function used on the optimization process. Furthermore, the objective function may contain the present value of costs, reducing the influence of investments made in the future. This way, the decisions made for a near future will be correct, and the decisions for a distant future may be corrected when better forecasts are available. [Afu82] and [Ada73] are examples on the use of dynamic programming in the DPP.

Nevertheless, methods based on Dynamic Programming are applied with difficulty to real-size problems due the famous "curse of dimensionality", for they demand high computational resources. However, methods based on Dynamic Programming were used with some success in problems related to reinforcements of existing network [Dal90, Par90].

In order to overcome the difficulties related to problem dimension, there were some attempts to reduce computational needs on the use of Dynamic Programming. For example, [Dal90] describes a method based on DP known as "Optimal Initial States" . The idea behind this algorithm consists in, during the dynamic optimization process, maintaining only the states that might lead to the optimal solution, reducing the needs in terms of computation time and memory.

5.3.3. Mixed Integer Programming

The basic mathematical technique for the optimization of the distribution planning problem is Mixed Integer Programming. This technique is well adapted to this problem because the decisions of building/not building may be easily represented by binary variables 0/1. However, this technique suffers from the same problem as dynamic programming, with the exponential increase of computation time with the growth of the number of variables. On the other hand, it is very difficult to include in this type of models some aspects considered fundamental in planning, e.g.: multiobjective analysis, uncertainties, strategic planning, etc.

It is also common in recent models the use of standard libraries of mathematical programming [Gön81], [Kag90].

5.3.4. Decomposition Algorithms

In the early 90s there was some interest in the use of decomposition techniques, namely Benders decomposition, in network planning [Kag90], [Nar91]. However, these techniques were also limited by the need of "well behaved" functions and by the dimension of real-life problems. Decomposition techniques are also not adequate for obtaining non-dominated solutions in a multiobjective problem.

5.3.5. Heuristic methods

A great number of methods in distribution and transmission planning are based on heuristic enumeration. According to the literature [Che96], these heuristic approaches can be classified according to the method of exploring the solution space: constructive or destructive (greedy search) or approaches of the type *branch exchange* [Aok90]. Some variations of these basic types follow the principles of an optimization technique known as TABU search [Che96]. This less conventional methodology will be referred with a little more detail in a following section. The common principle of all these approaches is the following: from a reasonable initial plan, the configuration is replaced by one of the neighbor configurations, obtained from elementary changes in the current configuration. Other examples on the use of heuristic methods may be found in [Nar91] and [Bac79].

However, these approaches have some difficulty in dealing with the dynamic nature of the problem. The recent development of new heuristics for distribution planning suggests different types of *forward/backward* procedures [Kuw96, Bla96, Bia95]. The multi-temporal dynamic problem is decomposed in several smaller sub-problems for a single stage and then each problem is solved separately, leading to pseudo-dynamic solutions.

Some researchers suggest the inclusion in this class of methodologies of a methodology known as Simulated Annealing. However, this technique will be referred separately in the next section.

5.4. Other algorithmic approaches

5.4.1.Simulated Annealing

This technique is based in the analogy between the simulation of annealing of solids and the problem of solving large combinatorial problems. The objective function is known as *Energy Function*. The system to be optimized is initially at a high temperature and is cooled until it reaches a zero energy value in the global optimum [Jon96]. One basic characteristic of this technique is the fact that the final configuration does not depend on the initial configuration. It is proven that the algorithm converges asymptotically to the global optimum with probability one. Even if it may be heavy in computational terms, this is an interesting characteristic of this approach [Lea95].

5.4.2.TABU search

TABU search (TS) is based on the use of flexible memory of search history in order to guide the search process in overcoming local optima [Glo89,90].

In TABU search, for each configuration, one defines a set of possible moves that may be applied to a solution in order to produce a new one. Among all the possible moves, TS looks for the one that improves the most the objective function. In some situations, no move will produce a better solution, which means the problem is at a possible local optimum. Then, TS permits the substitution of a configuration by another configuration that is worse than the first one. Generally TS chooses the one that least degrades the objective function.

However, if this occurs, the risk of cycling becomes important. Thus, the reverse move must be forbidden. This is done by storing the move in a data structure, called *tabu list*. The elements in this list are called *tabu moves*. Due to tabu moves, we can keep the search bias toward points with lower objective function values and escape from local optima. As soon as a trial solution is generated, it is checked against the tabu list. This allows the reduction of the search space. It is very important to keep the size of the tabu list within a reasonable range, to avoid closing paths to valuable solutions.

TABU search has already been applied to distribution planning [Che96] and to other areas of Power systems like, for example, capacitor planning [Hua96].

5.5. Evolutionary Computation (EC)

The complexity introduced by the new concepts in planning such as the ones referred in previous chapters (uncertainties, multiobjective analysis, etc.), associated to the combinatorial complexity of the problem, lead to the perception of the limitation of the traditional algorithmic approaches. These methods, in general, are clearly insufficient to deal with problems with the following characteristics:

- Fuzzy variables and constraints
- Integer and linear variables
- Non-linearities
- Multiple objectives
- Large dimension networks
- Dynamic Planning

- Non-convex spaces

After trying several conventional algorithmic approaches (namely Benders Decomposition) with limited success, the author, in 1991, came across with some new computational techniques based on the principles of natural evolution. These techniques are known as Evolutionary Algorithms (EA) [Hol75, Gol89] and had already been applied to several areas of engineering and control with some success. Still, in 1993, there was only a handful of research papers on the use of Evolutionary Algorithms in Power Systems and none on distribution planning.

Inspired by the promising results of these new techniques, one type of EA, known as Genetic Algorithms, was tried and tested for solving the Distribution Expansion Planning Problem with excellent results. These research results were published in 1994 [Mir94b].

Since then, the interest in these algorithms has been rising exponentially for they provide robust and powerful adaptive search mechanisms. The interesting biological concepts on which EA are based also contribute to their attractiveness.

This way, over the last few years we have been witnessing a considerable increase in research based on EA that has been documented in a large number of conference proceedings, journal articles, research reports and working papers. For that reason, it is now almost impossible to keep track on all the published works based on EA, both in Power Systems and in other areas of engineering. Section 5.8 refers to some of the applications of EA in Power Systems.

The next point will present an introduction to Evolutionary Algorithms principles.

5.5.1.Introduction to Evolutionary Algorithms

Evolutionary Algorithms are computer based problem-solving systems based on principles of evolution theory. A variety of EA has been developed and they all share a common conceptual base of simulating the evolution of individual structures via processes of Selection, Mutation and Recombination.

The processes depend on the perceived performance of the individual structures as defined by an environment.

It is important to notice that the field of Evolutionary Computation is not more than a small part of a greater, more complex scientific universe that incorporating Fuzzy Systems and Artificial Neural Networks, is referred by some authors as Computational Intelligence (CI) and by others as Soft Computing (SC). Some hybrid approaches within CI have been conceived with interesting results.

The most popular EA developed so far are the following:

- Genetic Algorithms
- Evolution Strategies
- Evolutionary Programming
- Genetic Programming
- Classifier Systems

Some researchers [Ebe90] would also include Simulated Annealing in this list, but it is preferably to treat it separately.

The next sections will introduce these forms of Evolutionary Algorithms. All of the techniques have already been applied to Power System design problems with different degrees of success. However, the focus will be on the two more promising techniques for distribution network planning: Genetic Algorithms and Evolutionary Programming. The methodology proposed in chapter 6 uses Genetic Algorithms as the technique for the search of planning alternatives.

5.6. Genetic Algorithms

5.6.1. Introduction

Genetic algorithms (GA) are an attempt at the implementation of strategies based on evolution in a large range of information systems. The idea of using evolutionary methods in automatic learning systems appeared in the 60s, and was discussed, for example, in [Fog66]. In 67 the term *Genetic Algorithm* is introduced, but only in 1975 [Hol75] the scientific community considers the use of genetic algorithms.

The concept behind genetic algorithms is attractive: for a period of over 4 billion years nature has produced, applying the process of evolution, organisms that no engineer could even dream creating. These organisms have a quasi-perfect capacity of adaptation to any situation. In other words, evolution led us to almost optimal solutions to a large number of complex problems.

The basics on the model of the process are well known: Darwin's evolution theory. Nowadays, we have computers that operate infinitely faster than nature. Why not try to emulate the natural processes?

That is how genetic algorithms emerge, applied mostly in search and optimization problems.

5.6.2. The Evolution Mechanism

Before we try to emulate nature, we need to talk about populations of individuals.

Each of these individuals defines a partial or potential solution for the problem being considered. These individuals may represent just anything: approximate functions for the solution of a mathematical problem, cluster prototypes in a classification system, rules in an expert system, etc.

The individuals are codified in strings – chromosomes – built upon a given alphabet or code (for example, the binary code {0,1}), in order to be univocally mapped in the domain of decision variables. The whole search process happens at coding level.

After calculating the representation of the current population in the decision variable domain, the performance of each individual is obtained according to an objective function that characterizes the problem being considered. This function is generally referred as fitness function.

Each individual is then characterized by the pair

(chromosome, fitness)

Example

In order to illustrate the concept of individual, suppose we want maximize the following function:

$$f(x) = 10 - (x-2)^2$$

in the interval **[0,31]**

We could codify **x** as a binary number, using five bits. Each solution should then be evaluated by the objective function, producing a given value. The chromosome $\mathbf{00100_2}$ ($x \leftarrow 4$) would have a fitness value of

$$f(00100_2) = 10 - (4-2)^2 = 6$$

The pair

$$(00100_2,6)$$

represents an individual.

This leads us to the next stage in the evolution process:

<u>Reproduction</u>

In this stage, the fitness of an individual will guide the selection process. Individuals with a higher fitness value will have a greater probability of being selected to the next evolution stage than individuals less fit and, consequently, is it expected the average fitness of this intermediate generation to be higher.

Therefore, there will be a larger number of copies of the best individuals, although there is the possibility of existing a few copies of the worst.

This process is performed by one of the genetic operators known as selection.

The different ways of implementing this operator will be referred in the next section.

The selected individuals are then modified through the application of other genetic operators in order to obtain the next generation. This genetic operators manipulate directly the genes that compose the chromosome, assuming that some genes codify, in average, better individuals than others. These genetic operators are divided in two categories:

Recombination. This type of operator makes that pairs (or larger groups) of individuals interchange genetic information.

Mutation. Mutation provokes modifications in the genetic representation of an individual, according to a probabilistic rule.

5.6.3. Canonical Genetic Algorithm

Although there are many forms [Gre91] for Genetic Algorithms, we will only refer to the *canonical* algorithm. This means that we will be dealing with three genetic operators (selection, crossover and mutation) and linear, binary, fixed-size chromosomes. Canonical GA use a fixed-size, non-overlapping population scheme and each new generation is created by the selection operator and altered by crossover and mutation. The first population is generated at random .

Selection

As it was referred in the previous section, the *selection* operator creates a new population (or *generation*) by selecting individuals from the old population, biased towards the best. This operator can be implemented in a variety of ways, although in the proposed methodology we use a technique known as Stochastic Tournament [Gol91a]. This implementation is suited to distributed implementations and is very simple: every time we want to select an individual for reproduction, we choose two, at random, and the best wins with some fixed probability, typically 0.8. This scheme can be enhanced by using more individuals on the competition [Gol91b] or even by considering evolving winning probability, eventually leading to *Boltzman Tournament* [Gol91b], generalizing the *Simulated Annealing* paradigm [Kir93].

Crossover

Crossover is the main genetic operator and the engine of genetic algorithms. It consists in swapping chromosome parts between individuals. The simplest crossover operator is implemented by selecting a random crossover point in the chromosome, and swapping the genes that reside between the crossover point and the end of the chromosome. For example, if we have two individuals:

A=01000 ; B=01011

and choose a crossover point C=3 (indicated by 'I')

A=010 I 00 ; B=010 I 11

the resulting individuals after crossover would be:

A'=010 I 11 ; B'=010 I 00

Crossover is not performed on every pair of individuals, its frequency being controlled by a crossover *probability*. This probability should have a large value, typically P_C=0.8.

Crossover allows the exchange of genetic material between two *parent* chromosomes, allowing beneficial genes from the parents to be combined in their descendents.

<u>Mutation</u>

The last genetic operator is *mutation* and in its simplest form consists in toggling a random bit in an individual. This operator should be used with some care, with low probability, typically P_m=0.001, for normal populations.

The selection and crossover operators effectively search and recombine existing chromosomes. However, they do not introduce any new genetic material in the population. Thus, the mutation operator is used in order to guarantee the possibility of searching in any particular subspace of the problem space, preventing the search of finishing in a local optimum. Nevertheless, too much mutation can be harmful: a mutation probability of 0.5 always leads to random search [Gol89], independently of crossover probability.

<u>Parameters</u>

Similarly to other stochastic methods, GA have a certain number of parameters, such as:

Population size;

Probabilities related to genetic operators;

Number of individuals in the tournament.

These parameters should be selected with care, since the performance of a GA depends heavily on the values utilized. Usually, it is recommended the use of relatively low population sizes, high crossover probabilities, and low mutation probabilities (inversely proportional to population size). [Gol89] makes a in-depth sensitivity analysis of GA in function of several parameters.

5.6.4.How does a genetic algorithm work?

A canonical GA is a very simple process: we first generate a random initial population, evaluate it and start creating new populations by applying genetic operators. This high-level behavior can be depicted on the following piece of pseudo-code:

```
1. START
2. GENERATE(OLDPOP)
3. Repeat until iteration limit
EVALUATE(OLDPOP)
NEWPOP=SELECT(OLDPOP)
CROSSOVER(NEWPOP)
MUTATION(NEWPOP)
OLDPOP=NEWPOP
4. END
```

Obviously, there is the need for some bookkeeping functions, for statistics and so on, but they are not central to this explanation.

This very simple behavior hides a powerful processing, done by the GA. In fact, the combination of selection and crossover leads to a proliferation of individuals that possess small, tightly coupled *blocks* of bits leading to good performance. These blocks, usually called *schemata* [Hol75], are replicated through selection and combined or separated by crossover.

So, GA tends to select individuals with good performance and recombine some of their building blocks, creating more and more copies of good schemata, simply by the use of selection and crossover. This hidden processing is called *implicit parallelism* because the number of schemata processed in each generation is typically $O(N^3)$, being N the population size. This compares very well with the number of fitness function evaluations, N. This characteristic is distinctive of Genetic Algorithms and leads to their excellent performance.

5.6.5. Variations of Genetic Algorithms

The canonical genetic algorithm has been improved in several ways apart from the ones referred above. Different selection methods have been proposed in order to reduce stochastic errors related to roulette selection. In [Bak85] Baker proposes a ranking system as alternative to proportional attribution of fitness, showing how to avoid premature convergence and accelerate search when the population is approaching convergence [Whi91]. Some recombination operators have also been proposed, namely multiple crossover [Spe91].

The mutation operator has been more or less stable, but the use of real-coded (non-binary) chromosomes requires the use of alternative mutation operators. This factor has also motivated the use of other crossover operators [Bäc91].

Other parameters have been introduced, such as the *generation gap* concept. The canonical genetic algorithm approach implies non-overlapping populations that, as we know, is not the case in natural systems. The concept of generation gap establishes how many descendants are produced in each generation.

Operators such as inversion (a form of mutation) and cyclic crossover [Gol89] have also been proposed.

Finally, some genetic models implement the concept of geographical isolation [Gorg92], an important principle in natural evolution. These models present very interesting performances.

5.6.6. Advantages of Genetic Algorithms

Genetic algorithms present several when compared to traditional techniques:

- Since genetic algorithms develop their search from a population of points and are based in probabilistic transition rules, they are less likely to converge to local optima than traditional techniques based on *hill climbing* deterministic methods.
- As it was referred before, traditional optimization techniques require, in general, "well-behaved" functions, restricting their application. GAs support non-continuous, non-convex functions and the existence of *noise* in the objective function. GAs do not need any *a priori* information about gradients of the objective function.
- GAs do not assume any a priori information on the problem, besides the one needed to define the decision variable space and the problem itself (the objective function).
- GA are, like Evolutionary Algorithms in general, very well adapted to distributed or parallel implementation.
- GAs are robust search mechanisms and lack any connection to specific heuristics of a given problem domain (this cannot be said about other types of Evolutionary Algorithms).
- Contrarily to some traditional techniques, Simulated Annealing and some Evolutionary Algorithms, GAs have memory about the search space (the schemas)

Nevertheless it is important refer that, besides their limitations, the power of traditional algorithms is recognized – GAs should be used when it is difficult or impossible to obtain efficient solutions using conventional approaches.

5.6.7.Special topics

Discrete and real variables

Discrete variables may be treated directly through binary (or n-ary) codification. Binary codification is always preferable, since it carries more information[1]. Due to their nature, GAs are particularly well adapted to discrete variables.

Real variables can be approximated, with the necessary precision using, for example, fixed-point binary representation. It is also possible to represent real variables in floating-point, expanding the range of possible values.

<u>Constraints</u>

Most search and optimization problems are subjected to constraints. GAs can deal with constraints in two distinct ways. The most efficient way is to embed these constraints directly in the coding, eliminating unfeasible solutions. When this is not possible, the fitness of unfeasible individuals should be calculated according to some kind of penalty function that assures that these individuals have, in fact, a low fitness value. The fitness should also reflect the proximity to the feasible domain.

However, appropriate penalty functions are not always easy to establish, since they may affect considerably the efficiency of the genetic search.

5.7. Other Evolutionary Algorithms

5.7.1. Evolution Strategies

Evolution Strategies (ES) were conceived in 1960 by I. Rechenberg to solve parameter optimization problems. ES employ real-coded variables and, in its original form, it relied on Mutation as the search operator, and a Population size of one. Since then, it has evolved to share many features with Genetic Algorithms [Hof91]. The major similarity between these two types of algorithms is that they both maintain populations of potential solutions and use a selection mechanism for choosing the best individuals from the population. The main differences are at three levels:

- Different representation methods: ES operate directly on floating point vectors while classical GAs operate on binary strings.

[1] For example, the binary forms of numbers two (010_2) and three (011_2) show more similarities (being adjacent integer values) that their representations in decimal form 2_{10}, 3_{10}.

- Different operators: GAs rely mainly on recombination to explore the search space, while ES uses mutation as the dominant operator. However, some variants of ES also use recombination as a search operator.
- ES are an abstraction of Evolution at individual behavior level, stressing the behavioral link between an individual and its offspring, while GAs maintain the Genetic Link.

An additional feature of ES is the self-adaptation of mutation variances by incorporating these parameters in the solution itself.

ES are also capable of solving high dimensional, multimodal, nonlinear problems subject to linear or nonlinear constraints. The objective function can also be, for example, the result of a simulation; it does not have to be given in a closed form (although this is common to EC models).

Evolution Strategies have already been introduced to solving some Power System problems with very exciting results. For instance, these algorithms are very promising optimization methods for solving complex parameter optimization problems. It is likely that this technique will be used, eventually, in Distribution System Planning with considerable success.

5.7.2.Evolutionary Programming

Evolutionary Programming, originally conceived by Lawrence J. Fogel in 1960 [Fog66, Fog65], is a stochastic optimization strategy similar to Genetic Algorithms, which places emphasis on the behavioral linkage between parents and their offspring, rather than seeking to emulate specific Genetic Operators as observed in nature. Evolutionary Programming is similar to Evolution Strategies, although the two approaches developed independently.

Like both ES and GAs, EP is a useful method of optimization when other techniques such as gradient descent or direct analytical discovery are not possible. Combinatorial and real-valued function optimization in which the optimization surface or Fitness landscape is "rugged", possessing many locally optimal solutions, are well suited for Evolutionary Programming.

In EP, like in GA, we assume that it is possible to assess fitness characterized in terms of variables and that there is an optimal solution in terms of those variables. In Multistage Network Planning, each solution is an expansion plan (with investment and operation decisions). The overall cost associated with the plan (including different criteria) is the solution's fitness. The objective is to find the plan with the minimum overall cost.

Its main distinctive features are the following:

1. In EP there are no constraints to the representation of the solution. In a GA approach, it is necessary to encode the solution in a binary string known as chromosome. In EP, the representation of the solution derives naturally from the problem: for example, an expansion plan for a network is freely represented, because the *mutation* operation does not require linear encoding. One just requires the definition of "what is a mutation" for each case, which depends on the decision variables chosen to represent the problem. The *crossover* operation (i.e. recombining parts of the network) is not used in this EP approach.

2. The mutation operation changes the solutions according to a statistical (Gaussian) distribution, in which minor changes in the offspring (for example, exchanging one line) are highly probable, while major changes (for example, complete branch deletions) are less probable. In a typical EP, the severity of mutations is reduced as we approach the end of the process.

The optimization algorithm involves three steps that should be repeated until the iteration (generation) limit is reached (or until an acceptable solution is obtained). EP has already been experimented with a relative success in distribution system planning [Mir96a]. The following description of the algorithm is one possible approach to the problem of multistage expansion planning:

1 - Choose an initial population of solutions (expansion plans). The initial solutions for the multistage network planning problem are created at random. Nevertheless, the algorithm that is used for generating the initial solutions may always find topological feasible solutions (by simply adding branches at random until all the nodes are covered in all stages).

2 - Each solution is copied into a new population. Each offspring solution is mutated according to a distribution of mutation types ranging from minor

(swapping two lines over the same node, for example) to extreme (deleting branches, for example - this mutation might be useful is the nodes served by these branches have no load).

3 - The fitness of each solution is assessed. To create the example presented in [Mir96a] the authors used only costs of lines and substations to assess fitness. However, one could use a much more elaborated scheme (similar to the GA approach proposed in chapter 6). The fitness function in this example also includes strong penalties for electrically and/or topologically non-feasible solutions, considering radiality and maximum line flow constraints.

In a typical EP algorithm there is no requirement that the population size be held constant or that each parent only generates one offspring, but in the example, we chose the simpler approach of having a fixed-size population.

Recently, the author has been applying with some interesting results, the technique of Evolutionary Programming to the extremely complex combinatorial problem of hydrothermal operation planning [Pro97], with innumerous advantages already shown by Genetic Algorithms in this application.

5.7.3.Classifier Systems

Classifier Systems [Hol86] are rule-based machine-learning systems that are capable to learn by examples. A CFS takes a set of inputs and produces a set of outputs which indicates some classification of the inputs. An example might take inputs from on-line measurements in a power system, and classify them in terms of "Stable", "Unstable", "Emergency". These systems are based upon Genetic Algorithms' modules and sometimes they are not referred independently. Based on the results obtained so far, it is likely that these systems will have an interesting future in Power Systems and, in particular, in the areas of Security Assessment [Lop95] and State Estimation.

5.7.4. Hybrid approaches

Evolutionary algorithms have also been hybridized with knowledge-based systems, Artificial Neural Networks and Fuzzy Systems in several applications and, in particular, in Power systems.

For example, in our research group at INESC we have been using Genetic Algorithms to determine optimal topologies of Neural Networks used in power system stability assessment, with success.

5.8. Evolutionary algorithms in Power Systems

As it was referred in Section 5.5, in the beginning of the 90s there were very few references on the use of Evolutionary Algorithms in Power Systems. In fact, a small survey performed by the author and his colleagues in order to find some references for a research paper in 1993 found only a handful of publications ranging from Optimal Capacitor Location [Ajj91] to Reactive Power Control [Iba91].

In 1996, another survey [Mir96a] presented in the PSSC'96 in Dresden, found over 130 (!) research papers and reports (only on major international publications) on the use of several types of Evolutionary Algorithms in different areas of Power Systems. Table 4, taken from this paper, summarizes the applications surveyed.

As an example, the mentioned survey also describes and presents some interesting results on a GA approach to the planning of the operation of a hydrothermal system and of the application of an EP strategy to a simple distribution network planning problem.

Why this interest in the use of Evolutionary Algorithms, in Power Systems?

Because these approaches are very well suited to deal with all those kinds of problems that usually represent nightmares for researchers and developers: integer variables, non convex functions, non differentiable functions, domains not connected, badly-behaved functions, multiple local optima, multiple objectives, fuzzy data, etc. Furthermore, they are not necessarily restricted to deal with numerical models, allowing the natural building of hybrid models including knowledge, under the forms of rules or other.

AREA	FIELD
A. Expansion or structural planning	A1. Generation-transmission
	A2. Transmission-distribution
	A3. VAr planning, capacitor placement
B. Operation planning	B1. Unit commitment, generator scheduling
	B2. Load dispatch
	B3. Reactive power dispatch, voltage control
	B4. Maintenance Scheduling
	B5. Security Assessment
C. Generation/transmission and distribution operation	C1. Loss minimization, switching
	C2. Alarm processing, fault diagnosis
	C3. Service restoration
	C4. Load management
	C5. Load forecasting
	C6. State estimation
	C7. FACTS
D. Analysis	D1. Power Flow
	D2. Harmonics
E. Control	E1. Parameter estimation and tuning
F. Surveys	F1. Surveys

Table 4 Applications of Evolutionary Algorithms in Power Systems

This complexity is what is required, in order to build larger Power System models with more adherence to reality. In very complex situations, they seem to be the only practical tool available to reach global optimization. In addition, they can make use of all consolidated knowledge in Power Systems - Fitness evaluation may call up any kind of tools, including specific heuristics that will perform the role of "learning".

Of course, in simplified models, when there are available specialized tools (say, for instance, in Linear Programming), EA models will surely be outperformed. However, in many cases, the simplified models were used just because there was no other reasonable way to get an answer (though approximate) to a problem. The point here is that EA provide robust search mechanisms that perform well for a large range of problems, and that is of extreme importance in Power Systems.

The most important conclusion is that all of the techniques offer immense potential, in spite of not being completely mature (notice that the vast majority of papers was published in the last three years!).

However, there has not been still a major breakthrough of Evolutionary Computation in practical applications in Power Systems, for most researchers are still concentrating in test problems that do not reflect the complexity of practical applications. Nevertheless, the excellent results being produced by Evolutionary Algorithms will inevitably lead to more practical applications and to a better understanding of the advantages and weaknesses of these type of approaches. It will also enhance the opportunity for the use of hybrid systems combining the potential of these versatile tools with conventional problem-solving algorithms, as well as with emerging techniques like, for example, Fuzzy Systems and Artificial Neural Networks.

5.9. Multiobjective analysis

The DPP, within the framework of the new planning paradigms, is clearly one problem that must then be treated as multiobjective and this fact has already been mentioned in some detail in chapter 2. This means that, in the multi-objective approach that is proposed throughout this dissertation, it is essential to have methods that allow the efficient generation of an extended range of alternative solutions within a fully dynamic model.

Genetic Algorithms, among other Evolutionary Algorithms, by maintaining a population of solutions, may search (in conditions of implicit parallelism) for non-dominated solutions. This unique characteristic makes Evolutionary Algorithms particularly attractive for solving multiobjective problems [Fon93].

The relevance of multiobjective analysis in distribution planning will be elucidated in the example presented in chapter 7.

5.10. Conclusion

As we have seen in chapter 3, early models for distribution system planning neglected several aspects now considered to be fundamental by planners and utilities: the fact that we are dealing with a dynamic, multiobjective problem; the need for more than one planning alternative; and the fact that the planning problem is profoundly conditioned by uncertainties.

The need to include all these aspects in the models (to make them closer to reality) and the increase in computational power uncovered the fact that traditional algorithmic approaches were no longer capable of dealing with this type of problems.

Consequently, we needed more effective, robust and powerful techniques in order to overcome the weaknesses of conventional algorithms. Evolutionary Algorithms proved to be the instrument we needed, for they provide flexible and easy to use search mechanisms in a vast range of problems, allowing the utilization of optimization software widely available. This way, the interest in the application of this type of algorithms in Power Systems has grown exponentially in the last few years.

In particular, Evolutionary Algorithms show two characteristics that are particularly interesting for the type of problems we deal in planning: the adequacy to multiobjective problems (generating a set of alternatives) and the easy implementation in distributed and parallel systems.

The next chapter will propose a methodology that will integrate most of the concepts and techniques referred in this and the previous chapters.

6. PROPOSED METHODOLOGY

Everything should be made as simple as possible, but not simpler.

Albert Einstein

6.1. Summary

This chapter comes as a convergence of most of the concepts, techniques and models proposed in previous chapters in a methodology for expansion planning. Risk analysis, a thorough representation of uncertainties and multicriteria decision making will be the core to the development of the methodology.

6.2. Introduction

Chapters 3 and 4 discussed questions related to the modeling of the expansion planning problem. Considering that we need to consider a large number of aspects in the planning process, the complexity of the problem was clearly demonstrated. Chapter 5 tried to address the issue of finding efficient algorithms in order to overcome the deficiencies of conventional techniques, showing the interest of using some form Evolutionary Algorithms in the network expansion problem.

This chapter will seek to merge some of the material referred in previous chapters in a methodology that, on one hand, addresses a large part of the technical concerns referred in chapter 3 and, on the other hand, includes some important concepts, as follows:

- Multicriteria environment. The methodology should take in account several different objectives when assessing the merits of possible plans. This objectives will include not only the traditional ones (investment, losses, etc.) but also new criteria associated to the concepts presented in chapter 4 (inadequacy, robustness, exposure).
- Strategic planning. Instead of an optimal expansion plan, one aims at defining a strategy that may lead to acceptable decisions in face of the uncertainties in the future.

- Risk analysis. The methodology intends to build risk adverse strategies for system expansion planning. The central idea is not to plan for the average future, but to take planning decisions such that, whatever plausible future occurs, the regret felt is kept as low as possible. Plans developed under such principles usually display a diversity of solution resources (which assure their flexibility) while the more classical "optimization for the average future" tends to propose less flexible solutions. This discussion will be advanced in chapters 7 and 8.

- Alternative generation. Genetic algorithms will be used to generate multiple dynamic solutions, in order to allow the use of multicriteria decision methods in the decision process.

- Comprehensive uncertainty modeling. A combination of fuzzy sets, probabilistic models and scenario trees will be used to model uncertainties and decision making.

It is important to become aware that the methodology does not intend to be a complete model for planning or, even less, a complete planning system. Even though we have included a great deal of planning options seldom considered in traditional models, the main purpose of this methodology is just to suggest the basic procedures for a planning process under the new paradigms.

Accordingly, the proposed methodology involves three basic steps:

Phase I

This is what we could call the analysis phase. In this phase, the planner, based on adequate forecasts, should establish possible scenarios and uncertainties, leading to the definition of a tree of futures. It is also necessary, in this phase, to delineate all the data required, to establish the network representation and finally to define the criteria that will guide the planning process.

Phase II

For each possible future (defined as a path in the tree of fuzzy futures) a Genetic Algorithm is executed in order to find expansion plans. The algorithm should try to find as much representative solutions as possible in the universe of non-dominated solutions.

Next, for each path in the tree of futures, and using MCDM (multicriteria decision making) methods, the decision maker should define a conditional decision set, determining the (conditional) ideal plan, under the assumption that we know precisely this path will occur.

Phase III

Considering all conditional ideals, a new algorithm will determine robust strategies that minimize the regret felt (deviations to the ideal values), in all possible futures. The final decision will be taken by selecting a strategy, again using MCDM. This decision is taken in an interactive manner and may eventually be supported by Geographical Information Systems. The possible strategies could also be analyzed according to other criteria not defined in phase I and eventually be added some reinforcements to improve some quality factor.

These phases will be detailed in the following sections, but first we will elaborate on the discussion initiated in chapter 2 on the need for a strategy.

6.3. The need for a strategy

Solving for a path in a given tree of futures (see chapter 4 for more details) means that one has anticipated that this particular path would be followed. However, decisions to be taken today are influenced by decisions later. Therefore, even if two paths share some nodes, it is just natural that optimal decisions at the very same node would be different, according to the path studied.

In many aspects of the planning activity, engineers face decisions that are not only irreversible, but also that have consequences that escape the law of large numbers (which roughly states that the frequency of occurrence of an event will only approximate its probability for a large number of trials).

Therefore, assessing the merits of long term investment policies on the basis of average returns is meaningless in many cases: the future will happen only once, and it is unlikely that there will be any repetition of events or circumstances, such that a bad decision could be later compensated by a lucky one. This explains the success of risk analysis approaches to system planning, in many fields of activity; in power systems, several authors have also stressed its importance [Mer90]. Chapter 9 will present a detailed analysis of this question and its consequences.

An acceptable strategy, for a decision maker, will be one in which decisions are not largely regrettable, no matter which plausible (or feared) future occurs. Certainly, decisions will not be optimal, for the actual path followed along but they will not (if possible) have catastrophic consequences. In fact, the best strategy would be the one that would minimize the regret felt for any decision, no matter what future becomes actual.

The concept of regret is central in this risk analysis approach. It is deeply related with other two concepts presented in chapter 4: robustness and exposure. A solution would be 100% robust if it would be considered good in no matter which future. Exposure is associated with futures for which regret would be felt, related to some decision taken.

Introduction to the methodology

The following points will present the three-phase methodology proposed for distribution expansion planning. It is important to notice, however, that there might be some feedback loops through the phases, just like in any normal planning activity. For example, after defining a good strategy, the planner might want to add some extra equipment in order to improve some aspects of the expansion project (e.g. reliability). Subsequently, the process could be resumed in phase II in order to determine new ideals under the new conditions.

6.4. Phase I - Analysis

In any decision process – and we have established the planning as a decision making process - the first phase is devoted to the analysis of the problem, the identification of objectives and uncertainties. This is probably the most important phase, since it is determinant for the success of any planning activity. Any errors during the analysis phase will be irremediably reflected in the posterior phases, by introducing serious errors and leading the planner to wrong decisions. We may establish the main steps in this phase, each one related to the definition of the following items:

1. Load forecast. Scenarios and uncertainties
2. Representation of the network
3. Preprocessing
4. Criteria

The next points will detail on each one of these analysis steps. Again we will stress that these steps, although fundamental, will have to be taken before the search process itself and are mostly related to the problem of spatial forecast and modeling, which is somehow marginal to the scope of this thesis.

Furthermore, it is important to be aware that the several options in terms of modeling taken during the development of the example, concern exclusively the example itself and are not limitations of the methodology. Naturally, there will always a tradeoff between the number of modeling options and the computational capacity (especially CPU time). However, the author believes (from contact with utilities and planners) to have included an interesting number of modeling options in this example.

6.4.1. Scenarios and uncertainties

A clear identification of the uncertainties affecting the problem and the adequate modeling of these uncertainties are critical elements in the analysis phase as well as in the whole decision process. These uncertainty models will have the immediate effect of allowing the planner to find insights that are not apparent on the surface.

Planning horizon

Initially, the planner will have to decide, from considerations already referred in chapter 2, the planning horizon for the expansion problem. It will also have to decide on the number of stages on which we are going to divide the planning horizon.

Load and network modeling

As it was mentioned in chapter 2, the analysis on expansion planning will be performed for extreme situations in terms of load (for each planning stage) taking into account all the considerations made in the chapter 2 (simultaneity factor, load factor, DSM effects, etc.). The planner will have to consider the results from spatial load forecast techniques such as the one referred as "Small Area Load Forecasting" (by mapping its results on a simplified network) and define plausible scenarios with maximum care.

In the example proposed, for each node we will consider a possible *fuzzy* load that will grow according to the scenario under consideration. The fuzzy value of the load could derive from some imprecise assessment of the possible load or from subjective judgement deriving from the planner's experience (perhaps using statistical values).

After aggregating the load in the several nodes and considering the possible growth scenarios the result will be a fuzzy tree of futures as represented in Figure 18.

Costs

The costs for equipment items (lines, substations) have to be clearly defined before the search process, and may also be defined as a fuzzy number in order to represent the uncertainty in this basic attribute. Again, this value may come from an imprecise assessment of future costs. Figure 19 could represent the assessment of the cost of a primary substation in a given stage as "It is expected that the cost of the substation will be between C_2 and C_3 but it could also be as low as C_1 or eventually reach C_4"

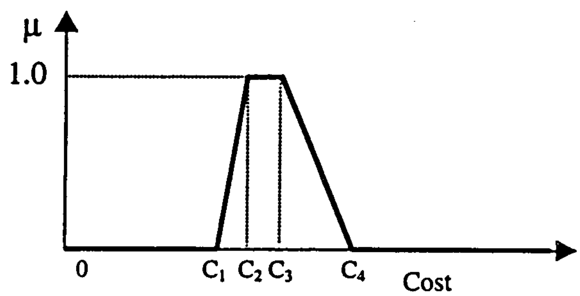


Figure 19 Imprecise assessment (represented by a fuzzy number) of a future cost of a primary substation.

In the example, we assume that each cost for equipment will be given as a fuzzy number. Consequently, the total cost of the expansion will also be given as a fuzzy number.

6.4.2. Network representation

In the example designed to demonstrate the methodology, we assume that some studies on the placing of equipment have already been made. This means that the planner should know, beforehand, the possible alternatives for lines and primary substations. In the example, the algorithm will not determine the best routing for lines or the placement of a primary substation. It will solely choose between alternative lines and primary substations and determine the year in which they should be put into service. Again, this is only a limitation concerning the example.

Therefore, the network will be represented by its single-phase equivalent and there will be two sub-systems:

- The existing system composed by all the lines and substations in the present system.
- The system in project, consisting on all the possible alternatives for lines, primary substations, reinforcements in line capacity and reinforcement in primary substations.

Since they are not central to the description of the methodology, details on the representation of the network may be found in Appendix A.

The objective of the algorithms is to determine which equipment is going to be put into service and in which stage. However, the system will have to perform some kind of preprocessing of the network before the Genetic Algorithm starts the search of expansion plans in phase II. This will be referred in the next section.

6.4.3. Preprocessing

In order to improve the performance of the genetic algorithm search procedure, the network will have to go through a preprocessing phase where the system eliminates some unneeded variables in the network representation.

This phase follows the general principles:

- If a line has to be built (put into service) under any conditions, the system will not include the line in the search procedure. For example, if a node with load in a certain stage is in the extremity of a line, this line will certainly have to be built and will be eliminated from the network as far as concerns the search algorithm. Obviously, it will be present in the solution as presented to the user.
- The same happens if the line does not have to be built under any circumstances. In the situation mentioned above, if a node with no load assigned in a specific stage (and in the stages before that) is in the extremity of a line, this line will not have to be built (unless it could be put in the same trench as another line and that would be considered to be economical interesting).

The preprocessing phase has three main advantages in terms of the overall performance:

- Reduces the number of variables in the problem;
- The decoding process becomes faster;
- The efficiency of the fitness function increases.

6.4.4. Criteria

The step of clearly identifying objectives is an important one in decision making, involving some kind of introspection: what are the issues and what is important? What are the real objectives? What do we mean exactly with risk? Is risk related to the fear of a monetary loss or to conditions that could lead to security or environment problems? A careful identification of all aspects of a problem, including pertinent objectives can lead, during the search process, to the discovery of alternatives that were not obvious in the beginning [Cle90].

Chapter 2 already offered an extended discussion on the criteria relevant for the distribution planning problem: we want the system to be economic, safe, reliable, but also to perform in a robust manner in an uncertain environment. The criteria chosen in order to illustrate the methodology presented in this thesis were the following:

- **IC** Fuzzy investment cost
- **VQ** Fuzzy voltage drop quality
- **PL** Fuzzy power losses
- **RB** Fuzzy Reliability (Related to energy not supplied due to contingencies)
- **RO** Robustness
- **IN** Fuzzy inadequacy

One must realize that other criteria could have been included in the fitness evaluation, without any major problems.

Why using fuzzy criteria?

Chapter 4 made the point for the use of fuzzy modeling of uncertainty in loads and reliability indices. The use of fuzzy modeling for costs appears logic since it would make no sense using crisp values for costs, when uncertainty in costs may be more determinant than technical uncertainties.

Note that most of the values in the following expressions are fuzzy. Robustness derives directly from the fuzzy definition of loads and is a non-fuzzy criterion. However, the notation will not reflect this fact for reasons of simplicity.

These values have been included in the assessment of the quality of each solution in the following way:

Investment costs (IC)

The value for the Fuzzy Investment Costs coefficient IC is obtained from the capitalized summation of investments for all the stages within the planning horizon. This value will include the fuzzy cost for the following items:

- Lines

- Primary substations
- Reinforcements in lines or substations

Investment costs will also have to consider the reduction in costs related to the simultaneous construction of two items (for example, two lines in the same trench built in a given stage).

IC will therefore be calculated using the following expression - the first term refers to investment costs and the second to possible savings as mentioned in the last paragraph:

$$IC = \sum_{i=1}^{ns}\sum_{j=1}^{k}\frac{C_j}{(1+\frac{D_r}{100})^{i-1}} - \sum_{i=1}^{ns}\sum_{m=1}^{Nv}\frac{S_m}{(1+\frac{D_r}{100})^{i-1}} \quad (42)$$

n_s Number of stages

k Number of items built

C_j Present cost of item j

D_r Discount rate

N_v Number of situations where savings could be considered from the simultaneous construction of two or more items

S_m Saving achieved from the simultaneous construction of two or more items for case **m**.

IC is then the cost referred to the first stage of all the investments to be made in the planning horizon.

Power Losses (PL)

The value of power losses in lines for the peak situation studied is obtained using an AC load flow in the fitness function during the GA search process. The fuzzy value PL is obtained adding the fuzzy values of power losses all the planning stages in the planning horizon, considering a fixed discount rate, similarly to the calculations for investment costs.

$$PL = \sum_{i=1}^{ns} \sum_{j=1}^{l_i} \frac{p_{ji}}{(1+\frac{D_r}{100})^{i-1}} \tag{43}$$

p_{ji} power losses in line **j** in stage **i**

n_s number of stages

l_i number of lines in stage i

D_r Discount Rate

<u>Voltage quality (VQ)</u>

Voltage quality can be accessed in many different ways. In this example, we used a scheme similar to the one proposed in chapter 4. A solution is considered unfeasible if the voltage drop exceeds a certain voltage threshold (e.g. 8%). As we are dealing with fuzzy voltage drops, each solution will have a certain degree of inadequacy, if there is the possibility that the voltage in any node should exceed this threshold. Voltage drops inferior to another threshold (5%) will be considered acceptable. Between these two values, for node **j**, there will be a penalty for the solution proportional to the square of the deviation.

$$Q_j = (\Delta V - 5)^n \Leftarrow \Delta V > 5\% \tag{44}$$

being ΔV the voltage drop (in percentage) relative to the nominal voltage V_{nom}.

$$\Delta V = \frac{|V_{nom} - V_i|}{V_{nom}} \cdot 100(\%) \tag{45}$$

and **n** a penalty factor (in the example in chapter 7 → n=2) depending on the importance of the node in terms of voltage drop.

However, as we will see later in this chapter, a solution is considered unacceptable if the removal value of the voltage drop in a certain node exceeds the maximum limit. Again, VQ will be calculated for the set of stages (**i**...n_s) using capitalization, as follows:

$$VQ = \sum_{i=1}^{ns} \sum_{j=1}^{I_i} \frac{Q_{ji}}{(1 + \frac{D_r}{100})^{i-1}} \tag{46}$$

This is the basic process for assessing the voltage quality of an expansion plan. However, since the voltage drop will be defined as a fuzzy number, refer to Figure 9 for detail on the calculation of a fuzzy voltage quality index.

Reliability (RB)

Reliability evaluation is based on the analysis proposed in chapter 4. The reliability parameters are probabilistic values modeled by fuzzy numbers. This means, for example, that the probability of a given line being down at a given time will be modeled by a fuzzy number. The procedure follows a general process based on the min cut set method and on the evaluation of the system's EPNS (Expected Power Not Supplied, a fuzzy number) for the peak load considered. The analysis includes branch failures, switching device location and load transfer through open loops.

The major problem is that it is impossible to calculate an exact value for a system's reliability without considering the switching device location, which can also be defined *a posteriori*. Since this is a quite complex problem and dependent on the type of network (urban or rural), as referred in a previous chapter, we tried to obtain an approximated value for the EPNS, through the following scheme:

(a) An upper limit U_r is calculated for the EPNS value, assuming that there is no switching equipment in the system. In this case, all the load cuts are processed at primary substation level. Consequently, no reconfigurations are permitted in case of contingency. This value is fairly simple to obtain, by considering the power not supplied for all the possible single branch failures. This value is the power at the corresponding output in the primary substation.

(b) A lower limit L_r is obtained for the EPNS assuming that all the branches in the network are switchable, allowing the isolation of the branches in case of contingency and service restore (taking into account line capacity). The process followed for this calculation is considerably more complex than for the previous limit. In case of branch failure, the algorithm has to verify all the reconfiguration alternatives, putting into service lines that may transport load to the affected nodes. Next, the process will calculate an approximate value for the EPNS considering the average repair times and line capacities.

Finally, the value for the reliability index is obtained using the following expression:

$$RB = EPNS = U_r + (1-\nu)\, L_r \tag{47}$$

Where

$$\nu \in [0,1] \tag{48}$$

is defined as a switching coefficient, whose objective is to simulate the effect of a compromise solution in the switching politics.

The overall reliability index is obtained by the capitalized summation of the reliability for each stage, as follows:

$$RB = \sum_{i=1}^{ns} \frac{RBi}{(1+\frac{D_r}{100})^{i-1}} \tag{49}$$

n_s number of stages

IN_i System Inadequacy for stage **i**

Dr Discount Rate

It is important to notice that there would be several other (eventually better) possibilities for obtaining a reliability index based on different methodologies or resulting in different attributes such as EENS (Expected Energy Not Supplied), LOLP (Loss of Load Probability), SAIFI (System Average Interruption Frequency Index), SAIDI (System Average Interruption Duration Index) or even MAIFI (Momentary Average Interruption Frequency Index). However, the author believes this process is a good compromise between accuracy and computational complexity considering that the objective is mainly to compare and rank alternatives for a multiobjective analysis. A better and more accurate method to assess the reliability of the system would need more information, namely related to load diagrams and other. This type of analysis would be too time consuming to be included in the search process, but it could be used a posteriori, i. e., for a more precise analysis of expansion alternatives.

Robustness (RO)

Robustness is a measure of risk in a certain decision, in this case related to an expansion design and is a non-fuzzy criterion already defined in section 4.7.1. Maximizing the robustness of the solution means, in terms of planning, that the planner wishes to accept solutions that cover or are technically sound in a wider range of possible future uncertainties. Robustness for a network branch **b** where the power could only be supplied under an α_b level was defined as:

$$RO_b = 1 - \alpha_b \tag{50}$$

We define the system's robustness for a given stage **i** - RO_i as the minimum value for branch robustness in the whole system:

$$RO_i = \min (RO_b) = 1 - \max (\alpha_b) \quad (b=1\ldots \text{ number of branches}) \tag{51}$$

The overall value for robustness is not capitalized. It will be the minimum robustness considering each stage:

$$RO = \min(RO_i) \quad (i=1\ldots\text{number of stages}) \tag{52}$$

This criterion has to be considered with some care, for the existence of a single "bottleneck" in a given branch would reduce the robustness of a whole network in some degree. The planner should then consider the possibility of increasing the capacity of this particular branch.

In this sense, the fuzzy criterion Inadequacy is probably better in terms of assessing the quality of the solution when facing uncertain futures.

Inadequacy (IN)

As it was referred in chapter 4, one of the objectives of the planning process is to minimize the total inadequacy of the system, given by the following expression:

$$\text{IN}_\text{i} = \sum_{b=1}^{n.branches} IN_b \text{ (for the network in stage i)} \tag{53}$$

Inadequacy is a fuzzy criterion and its consideration corresponds to the need of minimizing structural constraints to load growth. The total inadequacy is the capitalized summation of inadequacy for each stage:

$$\text{IN} = \sum_{i=1}^{ns} \frac{\text{INi}}{(1+\frac{D_r}{100})^{i-1}} \tag{54}$$

n_s number of stages

INi System Inadequacy for stage i

Dr Discount Rate

<u>Why use a Discount Rate?</u>

One could question the use of a discount rate for some of the criteria defined above, since the methodology proposes the use of a Discount Rate for all of the criteria, with the exception of robustness.

In a fully dynamic model, the correct attitude is to capitalize the values of all the criteria since they are somehow related to economical aspects. An investment in a given stage corresponds to an improvement in other objective in that same stage and in the stages after that. The same investment in a later stage will cost less and we could increase quality investing less. The use of a Discount Rate for non-investment attributes could appear at a first glance as if we were accepting a decline in quality along time, which is not the case.

If, for instance, the planner decides not to use the capitalized value of a reliability attribute such as ENS (Power Not Supplied) he is accepting the principle of minimizing the overall value of ENS in the whole planning horizon, but betraying, in a sense, the principles of dynamic modeling.

Note that the fact that we are dealing with fuzzy numbers has no influence in this analysis.

6.5. Phase II – Obtain ideals

Phase II may be divided in two basic steps:

- First, for each possible trajectory in the tree of fuzzy futures (represented by gray lines in Figure 18), a group of non-dominated solutions or expansion plans is obtained using a genetic algorithm approach.
- In a second stage, a *conditional decision set* is chosen for each trajectory, based in Multicriteria Decision Making Methods. From this set, we may define the ideal and anti-ideal for each possible future, if we knew exactly that this future would happen.

The following sections will specify the particulars of these two steps.

6.5.1. Expansion plans

For each trajectory in the tree of futures, a GA search process is executed with the objective of finding a set of alternatives, corresponding to a expansion plan for the distribution system.

<u>Assessment of the quality of a solution</u>

Within the GA we have to assess the quality of each possible solution using a fitness function.

A solution is said to be <u>topologically feasible</u> if it is radial[1] and connected for all expansion stages.

A solution is said to be <u>electrically feasible</u> when there are no violations on the constraints of power flows for all the lines or on the voltage limits in the nodes.

The quality of a feasible solution is assessed according to the criteria defined in a previous section in this chapter.

[1] In this case we are only accepting radial (in terms of operation) structures. However the methodology may consider other type of structures in its formulation.

Chromosome coding

The coding process in phase II is similar to the one presented in [Mir94b] and its objective is to lead to a minimum amount of unfeasible solutions and at the same time provide very fast decoding. Since the explanation of chromosome coding is rather lengthy and not central to the method, it is presented in Appendix B.

The essential about this codification is that we have managed to conceive an efficient coding that, by reducing the number of unfeasible solutions generated, we are able to increase the efficiency of the Genetic Algorithm by handling directly topology constraints.

Fitness function

The fitness function must reflect both the desired and the unwanted properties of a solution, rewarding the former and penalizing the latter. In the electrical distribution problem, desired properties are, for instance, low cost, high reliability and flexibility under uncertain futures, while unwanted features are non-radial configurations (open loops are accepted, but not closed loops), violations of thermal cable limits or of voltage drop constraints.

The general trend is to maximize fitness. Figure 20 presents the general scheme of the evaluation of the fitness of a solution, represented by a chromosome, at any generation. The functions g(), h() and v(), referred to in this figure, must be chosen so that, for no matter what solutions **x** are evaluated, one always obtains

$$g(x_i) < h(x_j) < v(x_k) < f(x_m), \qquad \forall x_{i,j,k,m} \tag{55}$$

where the notation **A < B** stands for *"B is preferred to A"*, **A** and **B** being fuzzy numbers.

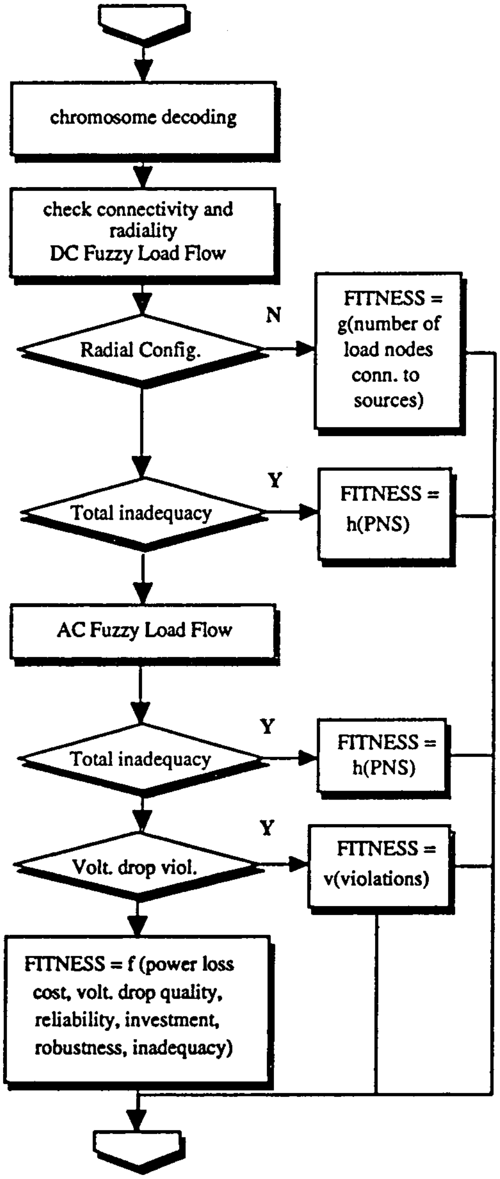


Figure 20 Fitness function evaluation scheme for phase II

Decoding

The chromosome is decoded following the algorithm corresponding to the coding process presented in Appendix B. From the decoding process we will obtain the whole expansion plan for the distribution system, including lines, substations and reinforcements. Thus, the network is defined for every stage in the planning horizon.

Topological analysis

After the decoding process, the algorithm will try to eliminate all the topologically unfeasible solutions. This analysis takes a very small amount of computation time, allowing the elimination of these solutions by attributing to them a very small fitness value.

The algorithm is also searching for radial connected solutions in all stages. If the solution does not have these characteristics, the fitness function returns a value inversely proportional to the summation (extended to the number of stages) of the number of nodes (with attributed load) not connected to the source node. The idea is that solutions with a higher number of not connected nodes will have a lower fitness value thus leading the algorithm to feasible solutions from the topological point of view. On the other hand, the non-radiality of the network is also penalized with low fitness values. The fitness is therefore assessed by a function g()

$$\textbf{Fitness} = \textbf{g(number of nodes not connected to sources, radiality)} \qquad (56)$$

At the same time, the algorithm recognizes the network structure in order to ease future calculations (electrical analysis). This process has revealed to be very efficient in tests developed over the model.

Electrical Analysis

If the solution is topologically feasible, the algorithm immediately passes to the electrical analysis. This analysis is composed by two different steps as follows:

DC radial fuzzy load flow - For each stage, a DC fuzzy load flow is performed. The result is the total fuzzy load not supplied, due to violations in the constraints in thermal limits in lines or in substation capacities. Since the loads are defined as fuzzy numbers, the algorithm uses the removal values of the flows in lines in order to verify the constraints. It is considered that, in a given line, if the removal value of the flow is higher than the thermal limit, then this situation is considered to be too inadequate and thus unfeasible. A network is considered unfeasible if there are such violations in any of the lines.

If the network is considered feasible then the algorithm passes to step 2.

If the network is unfeasible, the fitness function returns a value inversely proportional to the summation of the total Power Not Supplied in the network. Thus, the fitness function is calculated using the following expression:

$$\textbf{Fitness} = \textbf{h(Power Not Supplied)} \qquad (57)$$

In order to guarantee that solutions topologically feasible (even if not electrically feasible) always get a fitness value higher than topologically unfeasible solutions we have the following condition:

$$h(\,) > g(\,) \qquad (58)$$

This step is rather important since it allows a very fast filtering of excessively inadequate solutions (A DC radial fuzzy load flow is a considerably fast process) by attributing them a rather low fitness value.

AC fuzzy load flow - the step mentioned previously allowed the elimination of most unfeasible solutions from the electrical point of view. This second step involves the execution of a AC fuzzy load flow for radial networks. This load flow is the extension to fuzzy numbers of an algorithm described in [Mir83], applied to balanced systems. This iterative algorithm was developed for radial networks allowing precision adjustments according to the planners needs. For this reason, it is well adapted to the problem being considered.

As results of this procedure, we will have the following information for each planning stage:

- Fuzzy power flows in lines
- Fuzzy node voltages
- Fuzzy active power losses in lines

If this procedure detects violations on the thermal limits or on substation capacity (that were not detected in the previous step), the fitness function returns a value calculated by a function **h()** similarly to the previous step.

Otherwise, the algorithm will proceed to verify the voltage limits in each node. If the removal value of a node voltage in any stage is higher than the imposed limit (for example 8%) the solution is considered to be unfeasible and the fitness function returns a value proportional to the number of nodes violating these limits:

$$\textbf{Fitness} = \textbf{v(number of violations)} \qquad \textbf{v( ) > h( )} \qquad \textbf{(59)}$$

This last condition guarantees that the violation of voltage limits is considered less important than thermal limits violations, leading the algorithm to fully feasible solutions.

If there are no violations to voltage limits, the solution is considered feasible for each stage in the planning horizon. The fitness evaluation passes to the next phase, the assessment of solution quality.

<u>Solution Quality</u>

After passing the previous stages of the fitness function, the solutions that arrived so far are feasible from the topological and electrical point of view. Therefore, they may be analyzed according to the criteria defined previously. The versatility of this methodology lays precisely here, since we may assess the quality of the solution in the best manner according to our objectives.

The fitness fuzzy value **f** of a solution **x** is obtained from the following fuzzy equation:

$$f(x) = M - c_1 \cdot (IC+PL) - c_2 \cdot VQ - c_3 \cdot RB - c_4 \cdot (1-RO) - c_5 \cdot IN \qquad (60)$$

where

M - Large (enough) constant value.

c_i - Constants externally fixed

IC, PL, VQ, RB, RO, IN - Fuzzy values (except for RO) for each criterion defined in the previous section.

Constant **M** and the negative signs in the function are needed because we are trying to maximize fitness.

The comparison between two solutions is based on the concept of Removal presented in chapter 4 (which corresponds to a defuzzification process). This comparison is also needed during the selection procedure.

Varying constants c_i allows one to obtain a picture of a non dominated region of the domain of feasible solutions (in the sense of multi-criteria decision making). Encouraging *"geographical isolation"* in the space of criteria, by benefiting (in terms of individual fitness) solutions that present good characteristics in one criterion or that are non-dominated, also helps a larger coverage of the criteria domain.

6.5.2. Conditional Decision Set

A *Conditional Decision Set (CDS)*, as it was referred in chapter 4, is a set of alternatives or plans that may be considered as acceptable compromises between the objectives and, therefore, will be good candidates as a final decision, in a given future. A conditional decision set is generally constituted by a subset of non-dominated solutions. The selection of this set can be done using well-known multicriteria methods. This part of the methodology will be presented later in this chapter. To each of these conditional decision sets, one may define a *conditioned ideal* associated to the best value in each criterion and a *conditional anti-ideal*, corresponding to the worst value for each criterion (see Figure 21).

The use of Multicriteria Decision Making (MCDM) methods will be detailed in section 6.7.

6.5.3. Results

To summarize, the results from phase II are:

- A large set of (fuzzy) non-dominated solutions for each trajectory, obtained by the genetic process described above.
- A conditional decision set for each trajectory in the tree of fuzzy futures. From this set, we are able to determine the ideals for each future, in each possible trajectory. These values will be used in phase III to calculate possible regrets from the decisions taken.

Figure 21 depicts a simple example of a conditional decision set in a problem with just two attributes, *Investment Cost* (in switching equipment) and *Expected Power Not Supplied* (EPNS), both to be minimized. The picture represents several non-dominated solutions (**a-n**) and it is clear that an increase in investment will lead to higher reliability, by decreasing EPNS. In this case the determination of the CDS could be fairly simple and would not require elaborated MCDM methods: the planner (or maybe the decision maker) is not wishing to invest more than 8 MPTE (Millions of Portuguese Escudos) and will not be satisfied with a system where EPNS exceeds 95 kWh.10^3. Therefore, the planner will only consider solutions **e-j** as possible compromise solutions for his problem. This set of solutions will then be designated as the *Conditional Decision Set* (CDS) for this particular problem. However, a more detailed evaluation of the selection set is always possible, before choosing a CDS.

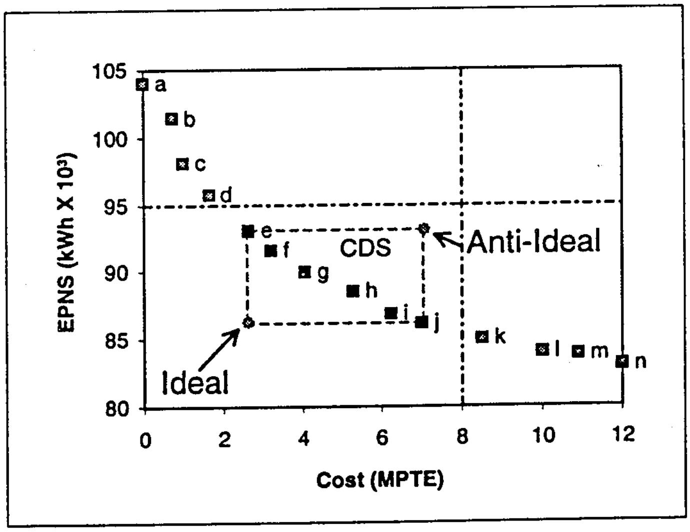


Figure 21 Example of a Conditional Decision Set (CDS, solutions e-j) in a bi-attribute space. The conditional ideal and anti-ideal are also shown.

6.6. Phase III – Obtain a robust expansion strategy

In <u>phase III</u>, we try to obtain a robust expansion strategy, according to the principle of risk minimization. This means that the network expansion strategy should contemplate every evolution trajectory as a possibility to be taken in account.

6.6.1. Compute possible strategies

In this phase another genetic algorithm is used. This time we have to codify and try to find an *expansion strategy,* so all the possible future decisions (depending on the future that actually occurs) have to be codified.

The risk aversion strategy is defined as the one that minimizes regret, measured as the overall distance between a strategy and the ideals (in the conditional decision set defined in phase 1) in all futures:

$$\min\{ \text{Max} [(f_{ik}^{x} - f_{ik}^{opt}), k=1 \ldots n] \} \text{ (for each attribute)} \qquad \textbf{(61)}$$

being:

n: number of futures

f_{ik}^{opt} value of attribute **i**, in the conditional ideal, for future **k**

f_{ik}^{x} : value of attribute **i** in future **k**, of solution **x**.

The equation above implies the linearity of the regret function, and that may not be the case in several problems, especially if we want to avoid "catastrophic" situations. For now, we will assume the linearity of the regret, but in chapter 9 some other possibilities will be explored.

Variables

The variables used to represent a multi-stage strategy are the same used in phase one to represent a network solution for one possible trajectory, but a different set of variables has to be considered for each stage where the decision depends on the future that corresponds to the situation.

Chromosome coding

In phase III the chromosome coding strategy is essentially identical to the process used in phase II. The main difference is that now we are codifying a strategy, so all the possible future decisions, (depending on the future that actually occurs) have to be codified. Figure 22 presents a graph depicting the possible futures in a given time horizon A-D. There is no uncertainty in terms of scenario for stage B. However, there are two possible situations for C and three situations for D. In this case the following decisions would have to be codified:

- Decision to be taken in time stage A.
- Decision to be taken in time stage B.
- Decision to be taken in time stage C if future C_1 occurs.
- Decision to be taken in time stage C if future C_2 occurs.

This leads to a larger chromosome, and consequently to a larger computation time due to extra analysis (an extra decision stage corresponding to an extra network) and added combinatorial complexity for the Genetic Algorithm. One way of reducing the complexity in the chromosome is to codify only the differences between decisions C1 and C2, which are generally small.

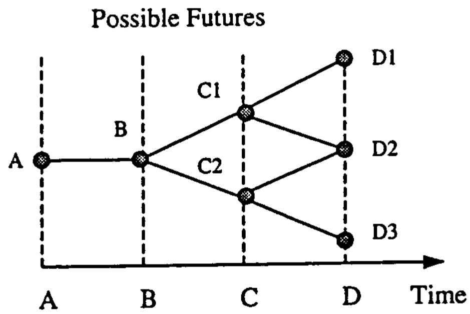


Figure 22 Graph representing possible paths in a tree of futures. Decisions have to be taken in A, B, and C.

Criteria

Phase III considers the same criteria as phase II for the evaluation of the fitness function. This means that each strategy is evaluated for all the stages considering exactly the criteria defined for phase II. The difference between these two phases concerning evaluation lays basically in the fitness function.

Fitness function

The fitness function in phase III follows exactly the same steps as phase II, until a feasible strategy is found. To be feasible a strategy has to be suitable for all the possible futures represented in the tree of fuzzy futures.

The main difference is that now we have to evaluate the quality of this strategy. The values for all the criteria have to be calculated for each possible trajectory in the tree of futures. The final objective is to determine the strategy that minimizes the distances between the values for each criterion and the ideals calculated in phase II, as seen above. This corresponds to minimize overall regret on the decisions to be taken.

6.6.2. Choose an expansion strategy

After the completion of the search procedure, the result will generally be a large number of possible expansion designs (strategies). Therefore, the planner will have to find a way to reduce the number of possible strategies in order to obtain one of the following, which will constitute the final result of the procedure:

(a) A robust expansion plan (or a set of robust plans – a strategy) consisting on a flexible set of investment decisions.

(b) A small set of possible expansion strategies to be further analyzed and tested according to other criteria or technical matters.

If the desired result would be the one referred in (b), these solutions could be expansion designs representative of the whole set of alternatives or designs with just minor differences between them. Among the large number of possible additions and improvements to each expansion design, we could stress the following:

- Improve reliability by adding new lines in order to form loops or by installing new switching equipment.
- Improve voltage profiles and reducing power losses by an adequate planning of reactive power, for example, by installing capacitor banks. This issue will be referred in chapter 8.
- Short circuit power analysis. Some measures could be taken in order to improve this aspect.

- Research on the possibility of reducing environmental or visual impact. This could be done, for example, by passing from aerial lines to underground cables.

Considering we would have only a small number of expansion strategies to analyze, the comparison between this set of possibilities and the final decision (by the utility management) would then be relatively simple tasks.

Whatever result is preferred (a single alternative or a set of alternatives) the planner will have much difficulty in order to reduce the set of solutions, since we are dealing with a problem with multiple criteria and a large number of alternatives. This is precisely the point where Multicriteria Decision Making Methods prove their efficacy. This important aspect will be referred in the next section.

6.7. Multicriteria decision making (MCDM)

Chapter 2 presented the distribution planning problem multicriteria, and mentioned the importance of obtaining not a single solution, but an extended set of non-dominated solutions. The existence of several alternatives allows the carrying out of interesting and valuable exercises on comparisons and tradeoffs, helping the planner to gain insight on the problem he or she is faced with, and opening the doors for better decisions to be taken. This type of understanding could never be provided by a methodology based on single objective optimization.

On the other hand, the fact that more than one solution is available enhances the opportunity for MCDM methods to be explicitly applied, which appears to be an important step in direction of a correct planning.

In this perspective and considering the methodology proposed in this chapter, the following paragraphs will detail on the use of these methods in phase II and phase III.

For reasons related to completeness and in order to present an example of a possible decision process, this section will also present a description of one MCDM method known as Successive Amplification Method as well as an example of its application in distribution system planning.

6.7.1.MCDM in phase II

If we have only two criteria, the determination of the conditional decision set is generally straightforward as we have seen in section 6.5.2.

However, if we have several criteria, which is the case described in this chapter, the problem of obtaining a CDS is a much more complicated matter. In this case, for phase II, and for each possible future we have two basic hypothesis:

a) If the number of non-dominated solutions is relatively small (something like less than 8-10 solutions), the planner will be able to analyze all the possible alternatives and choose the ones that seem to be good compromises between the objectives. That group of alternatives would constitute the Conditional Decision Set and the ideals would be determined from it. It is important to notice, however, that we are dealing with a multicriteria problem, and even with just a few solutions, a CDS could be hard to obtain.

b) The general case where the number of non-dominated alternatives is very large. In this case, the planner or the decision maker is confronted with a number of solutions which is excessive for his limited processing capacity [Mil56]. This is where MCDM methods may be useful in the determination of a CDS.

There are several possible approaches to this problem based on well-known MCDM methods. Good material on this subject may be found in [Zel82].

One of the possible strategies for non-prescriptive decision-aid on this type of problems consists on gradually reducing the set of solutions without surpassing, at any stage, the processing limits of the Decision Maker. In section 6.7.2 we will present a method first proposed by M. A. Matos in 1988 [Mat88] known as Successive Amplification Method or SAM. A full description of the method including that of a graphical interface developed together with the author of this thesis may be found in [Mat92].

Other interesting methods that could be applied in this methodology have been proposed. These methods could be used to determine possible Conditional Decision Sets or make the final decision [Mat88], [Zel82].

MCDM in phase III

In phase III, the planner faces a slightly different problem. After obtaining a (generally large) set of possible planning strategies, the planner may have to do one of the following:

a) Choose one of the strategies from all possible solutions proposed by the search algorithm.

b) Reduce the set of alternatives in order to present a smaller solution set (just very few solutions) to the decision maker. The DM will analyze the possibilities concerning the proposed strategies and take the final decision.

In both cases the planner will have to use some decision-aid method in order to reduce the set of alternatives to one solution (a) or to just a small number of solutions (b). Again, there are a large number of decision-aid methods that could be used. Due to its characteristics, the Successive Amplification Method represents a very good choice in this case as well.

6.7.2. Successive Amplification Method (SAM)

First, it is important to notice, that several other methods could be used in the decision making processes described in this chapter. However, the SAM is an interesting technique originating from the school of thought and culture in INESC's Power Systems research group and, besides, it does not diverge from the predominant line in the thesis. Specifically, the use of fuzzy clustering in the method stresses the questions related to uncertainty modeling.

Description

The Successive Amplification Method uses fuzzy clustering techniques to define a limited number of macro-solutions, which represent holistically the decision set. At each stage, the DM selects the most promising macro-solutions, and the actual decision set is reduced accordingly. This corresponds to a progressive amplification of the preferred area of the attribute space, until an adequate Conditional Decision Set is reached. In the process, the DM has only to decide between alternatives (real or virtual).

For a given multiobjective problem, we define a set **Z** of **n** non-dominated alternatives:

$$Z=\{z_1, z_2,\ldots,z_n\} \tag{62}$$

in a space of dimension **m**:

$$z_i = (z_{i1},z_{i2},\ldots,z_{im}\} \tag{63}$$

The objective is to consider a set of macro-solutions that represent **Z** in a structured and aggregated manner. The number of macro-solutions **s** should not exceed the number of solutions the DM is willing to consider simultaneously n_{DM}. A macro-solution M_i is defined as a set of fuzzy solutions:

$$M_i = \{ (z_k , u (z_k , M_i)) \mid z_k \in Z)\} \quad i = 1\ldots s \leq n_{DM} \tag{64}$$

Where u: Z→ [0,1] is the degree of belonging of alternative z_k in relation to macro-solution Mi, generally represented by:

$$u_{ik} = u (z_k , M_i) \tag{65}$$

In this method, we also define the characteristic points of each macro-solution in the space of attributes (these points do not necessarily belong to **Z**):

$$s_i \quad i=1\ldots s \tag{66}$$

The set of characteristic points of macro-solutions represents the set **Z**:

$$S = \{s_1,s_2,\ldots,s_s\} \tag{67}$$

The identification of macro-solutions from the set **Z** allows a reduced presentation of **Z**, from its representative points. Successive applications of this procedure over sub-sets of **Z**, from partial decisions, allows the focusing on preferred regions of the search space, until the definition of a final set of solutions of reduced dimension. The identification of the macro-solutions has the following natural conditions:

(a) The sum of the degrees of belonging of any solution z_k in **Z** equals the unity:

$$\sum_l u_{lk} = 1 \qquad u_{lk} \in [0,1] \tag{68}$$

(b) There are no empty macro-solutions:

$$0 < \sum_k u_{lk} < 1 \tag{69}$$

This conditions correspond to a fuzzy partition of set Z. Consequently, the cluster prototypes will be the characteristic points of macro-solutions. We may use, for example, the fuzzy clustering algorithm known as *fuzzy c-means* developed by Bezdek [Bez81] in order to obtain of $\mathbf{u}_{ik}$ and $\mathbf{s}_l$.

Due to the use of fuzzy clustering analysis, the choice regarding the number of clusters has less importance than in a general case; the impact of any "wrong choice" is attenuated by the progressive character of decisions and by the possibility of controlling the speed of progression through the parameter u_0 (see below). We may set **s** equal to $\mathbf{n}_{DM}$ or, to a value defined in each stage of the process. The basic procedure is the following:

Successive Amplification Method

(i) Obtain a fuzzy partition **U** of set **Z**, using a fuzzy clustering algorithm. Present solution prototypes $\mathbf{s}_i$ to the DM, as representation of macro-solutions;

(ii) Decision: the DM selects the most promising macro-solutions. Let **P** be the set of indices of the selected clusters;

(iii) Determine the reduced set:

$$Z^R = \{z_k \mid z_k \in Z \textit{ and } \sum_{i \in P} u_{ik} \geq u_0\} \tag{70}$$

Where u_0 is a pre-specified threshold.

(iv) If the number of elements in Z^R is higher than desired then do **$Z=Z^R$ and return to (i)**

(v) Decision: The DM chooses the favorite alternative. Alternatively, Z^R may be considered as the Conditional Decision Set for the problem in cause.

Some characteristics of this method are quite interesting, in what concerns the Decision Maker:

(a) The DM has a global perspective during the whole process

(b) The DM does not have to consider, at any time, more alternatives than desired.

(c) It is always possible to make multiple decisions.

(d) All decisions are taken between (real or virtual) alternatives. There is no need to define weights, tradeoff values or aspiration levels.

(e) The method is not prescriptive, but it prevents "wandering", due to the progressive reduction of the alternative set.

The presentation of the prototypes should be made in an adequate manner, namely with the use of histograms, which allow a global perspective to be maintained. The method is also favorable to reconsideration (backtracking), backing a few steps and restarting the process.

Example in distribution planning

The SAM has proven to be very interesting in applications related to decision making in Power System Planning. In the example in chapter 7, it was the method used to define the Conditional Decision Sets in phase II and was also used to simulate a possible decision process in phase III. This section will present a simplified example of the use of the SAM in a decision process concerning the choice between 53 possible expansion plans for a 3-stage distribution network expansion problem in a 4-criteria environment. Obviously, the example refers only the last phase of the planning process when the DM has to choose between a large set of non-dominated alternatives, generated by a genetic algorithm over a distribution network. Naturally, the simultaneous consideration of a decision set of 53 alternatives would be an impossible task for the decision maker and that is the reason that lead to the application of a MCDM method, in this case the SAM. The supporting software was developed by the author of this thesis and is presented in [Mat92].

The deterministic attributes (to be minimized) considered were:

- Total Investment cost (including discount rate, in MPTE)
- Power Losses (considering load diagram factor and discount rate, in kWh)
- Voltage Quality Index (non-dimensional)
- Expected Power Not Supplied (in kWh, for a peak load)

A full description of the aspects related to the attributes was not considered necessary in order to illustrate the method but the way these attributes are calculated is essentially identical to the description supplied above in this chapter.

The values for the attributes for the 53 non-dominated alternatives are also not going to be presented. Nevertheless, the ideals (non-conditional) for this decision set are:

Criterion	Ideal	Anti-ideal
Investment Cost (MPTE)	826	1069
EPNS (kW)	281	542
Voltage Quality Index	0	22.4
Power Losses (kWh)	1.12	1.89

Table 5 Ideals and Anti-ideals for the global decision set

The DM will have to indicate how many solutions he or she is willing to consider simultaneously (4 in the example) and the acceptation threshold $\mathbf{u_0}$, which parameterizes the successive reductions of the set.

The method will perform a fuzzy partition of the initial alternative set (53 alternatives) presenting the prototypes to the DM (see table). Note that these are virtual alternatives and do not correspond to any of the initial alternatives.

Macro-solution	s_1	s_2	s_3	s_4
Investment Cost (MPTE)	952	1025	827	1009
EPNS (kW)	363	202	538	349
Voltage Quality Index	2.69	7.98	5.31	1.44
Power Losses (kWh)	1.41	1.44	1.86	1.36

Table 6 Macro-solutions for first iteration

In this moment, the DM will indicate that the most promising solutions are s_1 and s_4. Alternative s_3 is considered to be low cost but not very reliable and s_2 has very good reliability and reasonable power losses, but the DM considers its cost too high.

In this point the DM has **30** alternatives left and decided to proceed with the reduction process, since this Conditional Decision Set is not considered satisfactory. Otherwise, the DM could stop the process here and use these 30 alternatives as the CDS. In this case the conditional ideals are:

Criterion	Conditional Ideal	Conditional Anti-ideal
Investment Cost (MPTE)	909	1055
EPNS (kW)	314	420
Voltage Quality Index	0	13.2
Power Losses (kWh)	1.12	1.63

Table 7 Conditional Ideals and anti-ideals for iteration 1 (30 solutions)

However, we will suppose that the DM chose to continue, by grouping these solutions again in 4 clusters. The following table presents the result:

Macro-solution	s_1	s_2	s_3	s_4
Investment Cost (MPTE)	923	958	997	1026
EPNS (kW)	386	335	359	334
Voltage Quality Index	1.12	1.38	1.4	0.28
Power Losses (kWh)	1.36	1.38	1.4	1.37

Table 8 Macro-solutions for iteration 2

In this case, the DM preferred the cheapest alternatives s_1 and s_2, which lead to a CDS of **13** alternatives, which seemed reasonable. The conditional ideals for the second iteration were the following:

Criterion	Conditional Ideal	Conditional Anti-ideal
Investment Cost (MPTE)	909	964
EPNS (kW)	314	392
Voltage Quality Index	0	13.2
Power Losses (kWh)	1.12	1.63

Table 9 Conditional Ideals and anti-ideals for iteration 2 (13 solutions)

It is curious to notice that the ideals for the second iteration are exactly the same as for the first iteration (although the anti-ideals are not). This happens because of the choice of solutions in this second step. Macro-solutions s_1 and s_2 are better in terms of cost (s_1) and EPNS (s_2) and Power Losses (s_1) - although none is dominant! It is natural that the ideal values for these three criteria will be in the reduced set from these macro-solutions. On the other hand, most solutions have a voltage index of 0 (meaning all nodes have a voltage drop inferior to 5%, in this case). Again, it is natural that there will be at least one solution in the reduced set with a perfect voltage index.

The DM could then continue the process until reaching a single solution (as in phase III, if the alternatives are planning strategies) or accept the set of 13 solutions as the CDS (for phase II, where the objective is to obtain a reasonable CDS).

6.8. The use of Evolutionary Algorithms

Finally, this section will refer some topics related to the use of Evolutionary Algorithms and in particular, Genetic Algorithms, as the tool for search in this methodology. The extraordinary characteristics of Genetic Algorithms were already referred in chapter 5:

- They provide robust search mechanisms immune to complex and "bad-behaved" objective functions.
- They naturally give a set of solutions and not a final single solution. Using a simple technique of weight variation and geographical isolation, the GA will cover the maximum possible of the search universe.

However GA, just like other stochastic search algorithms have some limitations and should be used with some attention to some important questions:

- GA are considerably sensitive to variations in the parameters (crossover and mutation probabilities for example) and these have to be tuned for each type of problem (although we may use some typical values). Therefore, a thorough sensibility study on this question should be performed before the methodology is applied.

- GA are based on probabilistic transition rules and random number generators, and therefore they will not always provide the same solutions for different search processes, even if they have exactly the same parameters. Therefore, it might happen that a particular good solution will not appear in all the runs of the genetic procedure. It is advisable that a few runs of the search procedure are executed in order to better ensure a good coverage of the search space.
- Even if GA perform better than other search algorithms in face of the "curse of dimensionality", a system that is too large could lead the algorithm to perform deficiently or take to much CPU time before reaching a good solution. Also, if the tree of futures is too complex and there is a large number of possible paths, the codification of the strategies will need too many bits and the analysis time will explode in phase III. The use of specific heuristics within the GA or the use of distributed or parallel systems could help mitigating this problem.

In spite of these limitations, GA have performed consistently and efficiently in network expansion planning as we will refer in chapter 7.

6.9. Conclusion

This chapter presents a new methodology for electric distribution network planning. This methodology leads to the substitution of the *single criterion /optimal solution* concept by a new integrated notion of *expansion strategies* in order to obtain solutions that are more flexible and adaptable to changing futures. It was developed based on two points presented in previous chapters: a technique grounded on genetic algorithms; a comprehensive representation of uncertainties for the generation of solutions and expansion strategies; and a planning philosophy based on the robustness evaluation of these strategies, guided by a paradigm of multicriteria risk analysis.

The main purpose of the methodology is to respond to the fundamental requirements in power system planning:

- Be able to deal with real sized networks
- Allow a multitemporal representation

- Generate sets of solutions
- Permit multicriteria analysis, keeping the criteria and their respective tradeoffs explicit
- Allow the planner to obtain indices that measure the distance between the strategies to implement and the ideal solutions he would choose if he could have a perfect knowledge about the future
- Be well suited to distributed or parallel implementations
- Cover a large range of modeling options and be flexible enough to respond to other requirements

The planning philosophy adopted is the one that minimizes the *regret* in the decisions taken - and this implies that the consequences of these decisions are evaluated within the large set of uncertainties and futures defined.

Finally, the use of an adequate Multicriteria Decision Making method for the choosing of an alternative is an indispensable step in the type of analysis for planning proposed in this thesis.

The next chapter will present an example of an application of the proposed methodology to an expansion planning problem, on a distribution network.

7. CASE STUDY

Computers are useless. They can only give you answers.

Pablo Picasso

7.1. Summary

This chapter presents the application of the proposed methodology to a large-size network and draws some relevant conclusions from the results. The aim is to investigate the applicability and usefulness of such approach in the framework of the new paradigms of decision making and risk analysis.

7.2. Introduction

The approach presented in chapter 6 and some of its variants were applied to several test systems of different dimensions and characteristics. One of the most relevant studies was based on a real system in a Portuguese utility EN (Electricidade do Norte, a power distribution company operating in the north of Portugal). The problem reported in this study is centered in the possible building of a new primary substation in the city of V.N. de Gaia, in Portugal. The simplified system included 20 load nodes (resulting from the clustering of more than 200 nodes), 5 primary substations and a planning horizon of 15 years. The results of this study were published in [Mir95] and show the potential of the methodology when applied to a practical system.

Moreover, some of the tools developed during the study, including the genetic algorithm platform for network expansion, were also integrated in prototype for a distribution management system being developed by a Portuguese company.

In order to fully illustrate the methodology, this chapter presents the results concerning a more general, large-size distribution system, based on real networks. Although not related to a specific practical problem, it combines several interesting characteristics and uncertainties usually found in practical problems.

The approach follows the general principles of the three-phase method mentioned in the previous chapter: analysis, calculation of ideals, and finding a robust strategy.

The search algorithms were implemented in language C, run on an UNIX operating system, using the genetic platform referred in [Ran93]. Some of the data processing related to the multicriteria decision making methods was done using a Personal Computer under Windows 95, and the code was implemented in Visual Basic for Applications for Excel 7.0 and Excel 97. Some experiments were carried out on a distributed platform, based on a network of workstations under UNIX.

7.3. Phase I – Analysis

As it was referred in the previous chapter, the analysis phase is devoted to the establishing of the following general points:

1. Representation of the network
2. Load scenarios and uncertainties
3. Preprocessing
4. Criteria

7.3.1. Data

The development of an example involved a very large set of data, since several modeling options had been included. In order to simplify and keep the clarity in the presentation of the results, the author opted for keeping only the essential data and the most relevant results in this chapter.

Distribution system

The methodology proposed in the thesis was applied to the system shown in Figure 23, where thick lines represent branches in the initial system and thin lines represent possible sites for the expansion of the system. The system is constituted by 54 load nodes, 16 lines and 2 substations in the initial system (S1 and S2), 45 lines and 2 substations in project (S3 and S4). There is also the possibility of expanding the capacity of substations S1 and S2 (data shown below).

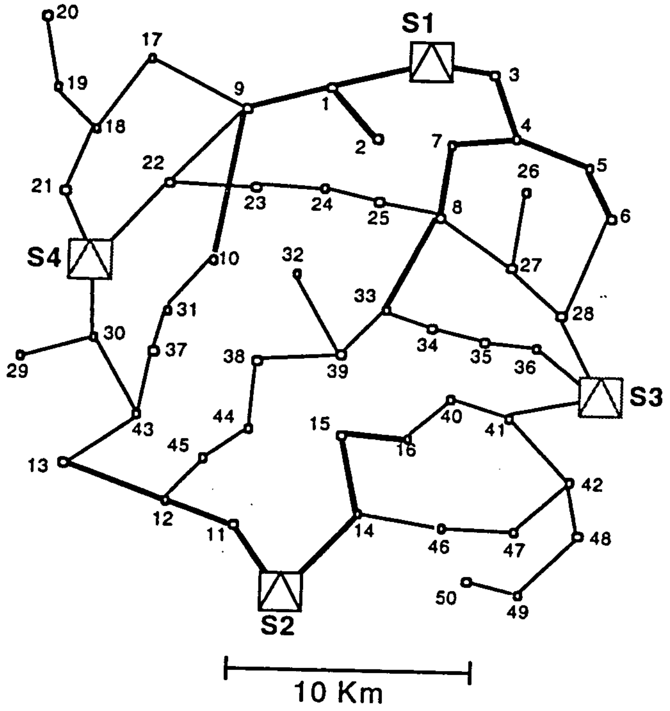


Figure 23 Initial System (thick lines) and system in project (thin lines)

Planning Horizon

The example was developed for a planning horizon of 6 years divided in three stages:

B 2 years ahead

C 4 years ahead

D 6 years ahead

A present time (first decisions have to be made here)

Note: The choice of a short planning horizon and stage steps seems unduly simplistic but it was made for the sake of clarity. Other studies were performed with larger planning horizons (up to 20 years divided in 4 stages) [Mir95]. However, the main objective in this chapter is to illustrate the methodology and too much temporal detail could shift the focus away from the important issues in the methodology.

Discount Rate

A discount rate has been established per stage:

$$Dr = 10\% / \text{stage} \quad (71)$$

Investment costs, as well as some other criteria (as mentioned in chapter 6) will be capitalized according to this value.

Loads

Load is defined as a trapezoidal fuzzy number for all the nodes in all planning stages as shown in Figure 24. Table 10 presents the values for total load corresponding to the sum of fuzzy loads in each stage.

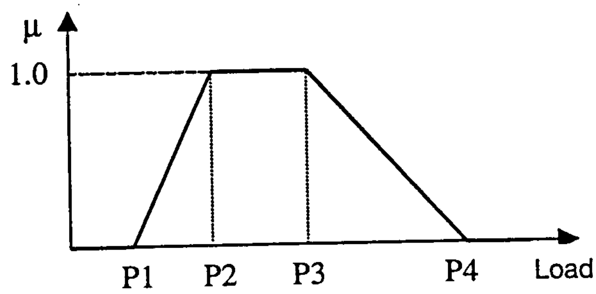


Figure 24 Trapezoidal fuzzy load in a node, as defined for the example

Stage	Total load (MVA)			
	P1	P2	P3	P4
B	43.0	43.5	46.5	47.0
C1	59.4	61.0	65.0	66.6
C2	50.1	51.8	56.2	57.9
D1	81.2	82.4	94.0	95.2
D2	70.2	71.9	79.3	81.0
D3	58.0	59.7	66.3	68.0

Table 10 Total fuzzy loads for each planning stage. Notice the large range of uncertainty covered by these values.

The load presented in the table leads to the definition of a tree of fuzzy futures as shown in Figure 25. As it has been referred before in the thesis, decisions on expansion options have to be made in A, B and C.

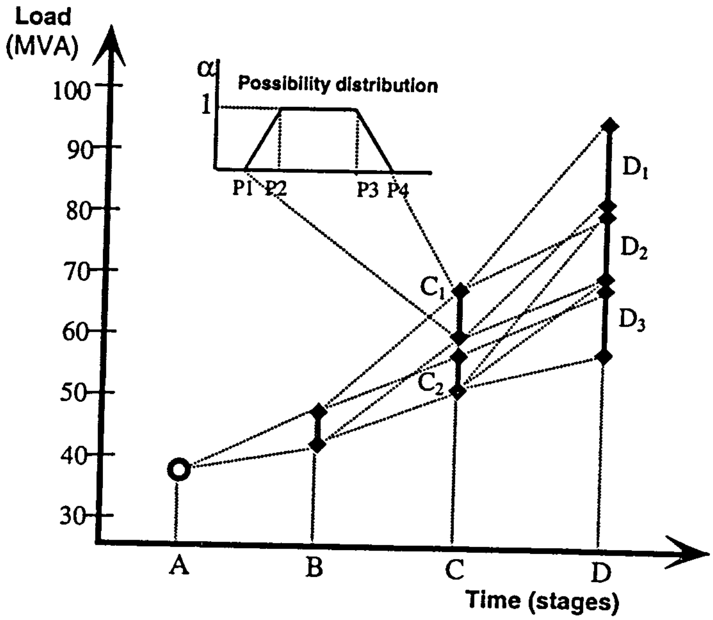


Figure 25 Tree of Fuzzy Futures. Decisions have to be taken in A, B and C.

For the whole set of loads and for all the planning stages, a fixed load factor **cos ϕ** was defined:

$$\cos\phi = 0.9 \quad (72)$$

Substations

In the initial system there are two primary substations S1 and S2. There are two primary substations in project S3 and S4. The total capacity of these substations as well as the values for possible reinforcements in capacity of existing substations are shown in Table 11 and Table 12. The tables also present the trapezoidal fuzzy costs corresponding to each one of these options and the values of failure rates for each one of the substations. Costs for substations S1 and S2 correspond to capacity reinforcement (expansion).

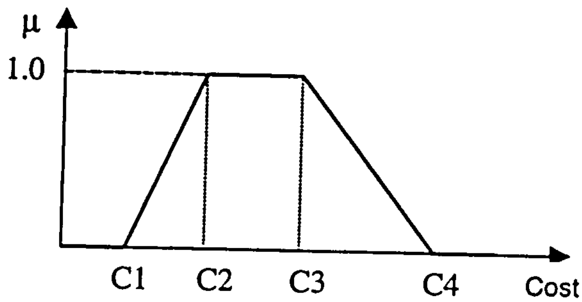


Figure 26 Trapezoidal fuzzy cost for a piece of equipment (substation or line).

Substation	Initial Capacity (MVA)	Projected capacity (MVA)	Fuzzy Cost (MPTE - stage 1) - TrFN			
			C1	C2	C3	C4
S1	16.7	16.7 (exp.)	90	100	110	120
S2	16.7	16.7 (exp.)	110	120	130	140
S3	-	22.2	190	200	210	220
S4	-	22.2	210	220	240	250

Table 11 Primary substation data

Failure Rate (failures.year^{-1}) - TrFN				
S1	0.009	0.0010	0.0010	0.0011
S2	0.0011	0.0012	0.0012	0.0013
S3	0.0011	0.0012	0.0013	0.0014
S4	0.0013	0.0014	0.0014	0.0015

Table 12 Failure Rates for Substations

Average repair time (ART) for all the substations is given by the fuzzy number:

$$\mathbf{ART = [0.008\ ;\ 0.01\ ;\ 0.01\ ;\ 0.015]\ year} \qquad (73)$$

Lines

The cost for lines is generally assessed using the following formula:

$$\mathbf{C_l = C_{fl} + l_l \cdot C_{vl}} \qquad (74)$$

$\mathbf{C_l}$ Investment cost for line **l**;

$\mathbf{l_l}$ Length of line **i**;

$\mathbf{C_{fl}}$ Fixed investment cost for line **l**;

$\mathbf{C_{v_l}}$ Variable investment cost for line **l**.

In this example, and for the sake of simplicity, the fixed cost for a line is considered to be zero, and the variable cost is approximately:

$$\mathbf{C_{v_l} = 4\ MPTE/km} \qquad (75)$$

The actual values are given in the following table in form of a Trapezoidal Fuzzy Number.

Line		Tr. Fuzzy Cost			
Node A	Node B	C1	C2	C3	C4
37	31	7	7.2	7.4	7.6
31	10	8	8.7	9	9.7
13	43	10	11	12.2	13.2
12	45	8	8.7	9	9.7
45	44	8	8.7	9	9.7
44	38	10	11	12.2	13.2
38	39	12	12.5	13	13.5
39	32	14	15	15.4	16.4
39	33	8	8.7	9	9.7
33	8	14	15	17	18
33	34	7	7.2	7.4	7.6
34	35	8	8.7	9	9.7
35	36	8	8.7	9	9.7
53	36	10	11	11.4	12.4
53	28	10	12	13	15
53	41	10	12	13.8	15.8
41	40	9	9.1	9.8	9.9
40	16	9	9.1	9.8	9.9
41	42	10	12	13.8	15.8
42	48	8	8.7	9	9.7
48	49	10	12	13	15
49	50	8	8.7	9	9.7
42	47	10	11	11.4	12.4

Line		Tr. Fuzzy Cost			
Node A	Node B	C1	C2	C3	C4
19	20	9	10	10.6	11.6
18	19	8.3	9	9.8	10.5
17	18	12	12.5	13	13.5
9	17	14	15	17	18
18	21	10	11	11.4	12.4
54	21	8	8.7	9	9.7
54	22	12	13	14.6	15.6
9	22	15	16	17	18
22	23	14	14.2	14.6	14.8
23	24	7	8	9	10
24	25	7	8	9	10
25	8	7	8	9	10
8	27	12	13	14.6	15.6
27	26	10	11	12.2	13.2
27	28	9	9.1	9.8	9.9
6	28	14	15	17	18
54	30	9	9.1	9.8	9.9
30	29	10	11	11.4	12.4
30	43	12	13	13.8	14.8
43	37	8	8.7	9	9.7
47	46	10	11	11.4	12.4
46	14	10	12	13	15

Table 13 Trapezoidal fuzzy costs for lines (line between Node A and Node B)

Note: these costs vary according to the line considered.

All these costs (in MPTE - 10^6 Portuguese escudos) are referred to the first stage of expansion (stage B) and capitalized for the other stages according to the defined discount rate.

<u>Electrical data</u>

The following table presents the electrical data for all the lines in the system. For reasons related to simplicity, the lines are considered to be of a single type, which is common in distribution systems. Reliability parameters are given as the number of failures per year for each line and the Average Repair Time (ART) in years. The overall reliability of the system will be an approximation (since it is calculated for peak load) to the value of the annual energy not supplied (ENS).

Nominal Voltage	15 kV
Resistance (R)	0.3 Ω/km
Inductance (X)	0.3 Ω/km
Maximum capacity	20 MVA
Failure Rate	[0.007 ; 0.008 ; 0.01 ; 0.012] failures.year^{-1}.km^{-1}
Average Repair Time	[0.002 ; 0.003 ; 0.003 ; 0.004] year

Table 14 Electrical Data for lines

Voltage Constraints

The following values correspond to the preferable and maximum limits for voltage drops in all network nodes for all the planning stages. Again, it is important to stress that the consideration of different constraints and penalties for each different node would not increase the complexity of the model.

Preferable limit: 5%

Maximum limit : 8%

Since loads are defined by fuzzy numbers, voltage values for nodes and the voltage quality index VQ will also be given as fuzzy numbers. The model considers that all the solutions where voltage (removal) exceeds the maximum value in any node, will be considered as unfeasible. For the other solutions, a VQ index will be assessed.

Genetic Algorithm

The following parameters are the most important used in the genetic process in phase II. These parameters were obtained after careful sensibility studies on the performance of the genetic algorithm.

Population	50
Crossover Probability	0.8
Mutation Probability	0.004
Approximate number of generations	300

7.4. Results for Phase II

In **phase II** the genetic algorithm procedure described is applied for every possible trajectory on the tree of fuzzy futures represented on Figure 25. By varying, within a fixed range, the weights for each criterion and benefiting non-dominated or isolated solutions (in the space of criteria), the result will be a set of efficient solutions (defined as possible expansion plans) for each future. From each one of these decision sets, a Conditional Decision Set may be assessed and from it we can derive the corresponding ideals.

The next point will present, as an example, one possible expansion plan for a particular trajectory in the tree of futures.

7.4.1. Example of an expansion plan

Figure 27 presents one of the efficient alternatives (expansion plans) found in the search process. This particular expansion plan was determined for future ABC_2D_3 in the tree of fuzzy futures and had an overall investment cost of

[584 ; 631 ; 685 ; 732] MPTE - Stage B

[207 ; 220 ; 231 ; 244] MPTE - Stage C

[15 ; 16 ; 17 ; 18] MPTE - Stage D

[806 ; 867 ; 933 ; 994] MPTE

Which corresponds to a defuzzified value of

900.3 MPTE (removal)

The values for each stage are already capitalized. It may be considered as an average solution in terms of cost as well as in terms of other criteria.

However, its importance in the overall presentation of the results is very small. It does not represent any solution to the expansion problem as defined in chapter 6. It is presented here just as an illustration of a possible plan for a particular fuzzy scenario.

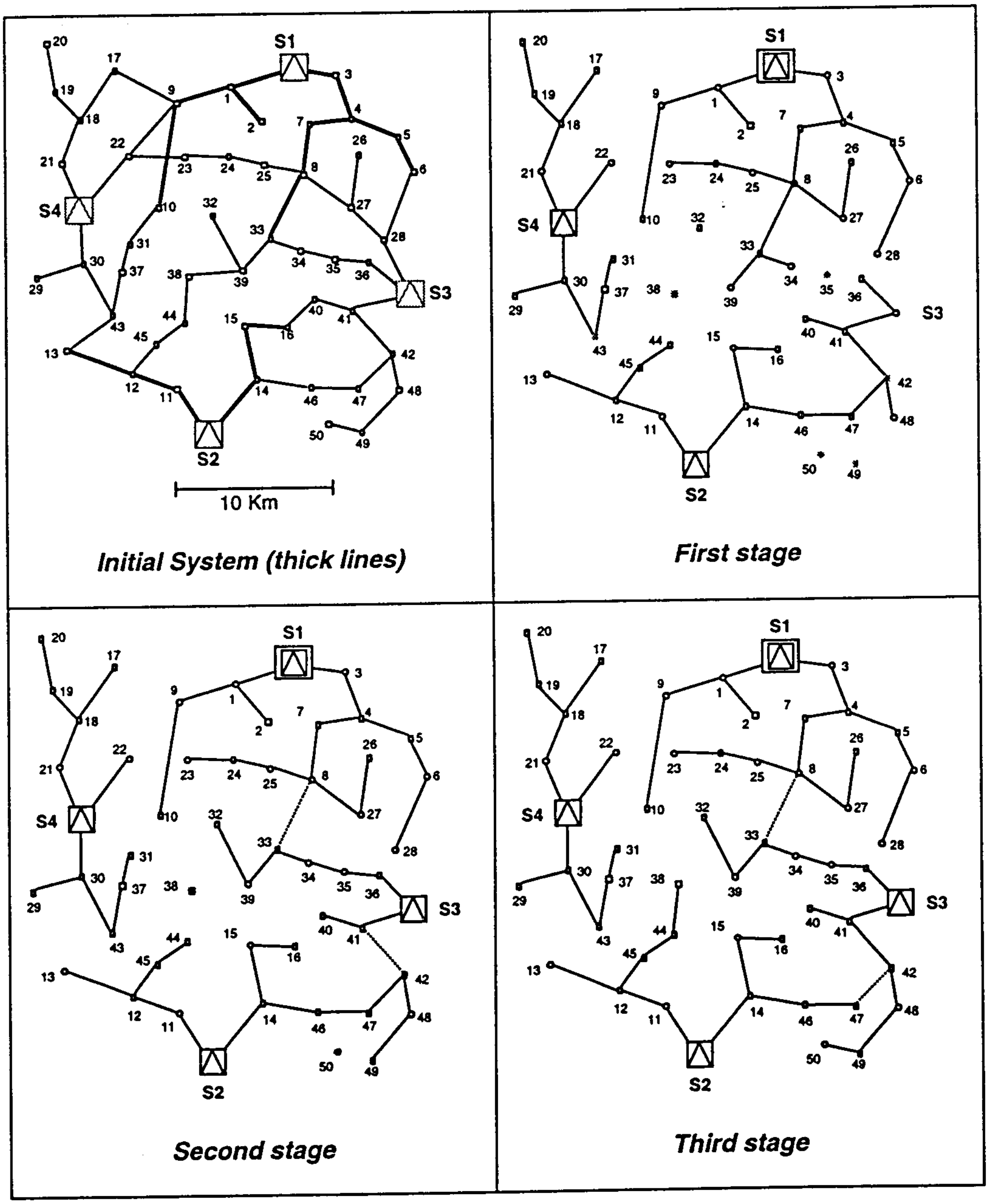


Figure 27 Network expansion plan

The plan presented is one of the possibilities for the expansion of the system in a specific future. In the figures shown, a substation framed in a square represents a reinforcement in its capacity. A dotted line represents a branch that exits in a given stage, but it is not used. However, these lines may be used to close loops in case of a contingency, and therefore are considered in reliability calculations. At last, a node in gray represents that it has no load attributed in a given stage. This particular alternative has some interesting characteristics that we could notice:

- In the first stage, nodes 42 and 43 do not have load attributed. However, the network goes through these nodes to allow flow to other parts of the network.
- Line 41-42 is built in the first stage, declassified in the second stage and reclassified in the third.
- There is an interesting reorganization of the network in the second stage after S3 is put into service.

7.4.2. The importance of multicriteria analysis

The application of the GA model for each trajectory in the tree of futures immediately allowed us to recognize a very important result, with direct implications in future model developments: we were able to identify projections of domains of feasible solutions, in the attribute space, that were clearly <u>not convex</u>.

Figure 28 and Figure 29 depict one of such cases: the projection of non-dominated solutions for future **ABC_2D_2**, on two planes, formed by pairs of criteria. The indices corresponding to the criteria are represented by their removal values. An approximation of the non-dominated border is also drawn.

- Figure 28, *investment vs. reliability*, seems quite classical, even if it is showing a slight concavity.
- In Figure 29, *investment vs. losses* clearly displays a non convexity.

The fact that the set of solutions is not convex (with special incidence on the non dominated border) makes it evident that some classical approaches of "pseudo-multicriteria", that would just perform an optimization using as single criterion a linear function of all criteria in the problem (adding them up multiplied by some weights) are misused: they can never detect "hidden" solutions in the concave parts of the solution set which, in fact, might become the most attractive to the decision makers. Chapter 9 explores this aspect even further in a slightly different perspective.

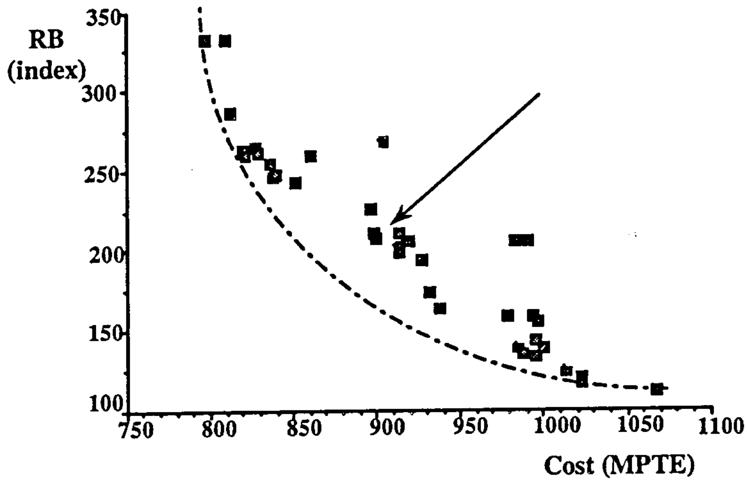


Figure 28 Investment costs versus RB with an approximation of a convex non-dominated border. The alternative presented in the last section is shown by an arrow.

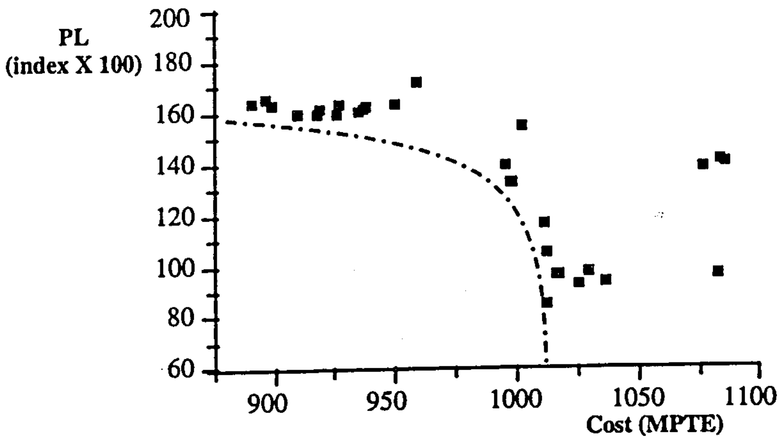


Figure 29 Investment versus Power Losses: the same solutions as in Figure 28 clearly display a non convexity in this domain projection

This result gives even more importance to the use of GA, because we have demonstrated that they deal adequately with this feature of the planning problems.

Figure 30 displays an interesting plot of a group of non-dominated solutions for another trajectory or path **(ABC_1D_2)** in the tree of futures. Solutions are plotted in an attribute space defined by two axes (*cost vs. robustness* - cost is fuzzy and therefore is represented here by *removal* values of the alternative possibility distributions). It becomes obvious that a growing investment does not contribute, in this case, to increase the ability of a system design to be acceptable in a larger set of futures. In fact, due to bottlenecks imposed in substation design, we were not able to find solutions (for this specific trajectory) with α lower than approximately 0.45, which means that the distribution system, due to some project constraints imposed, will not be able to cope with the whole range of uncertainties defined in the load forecast.

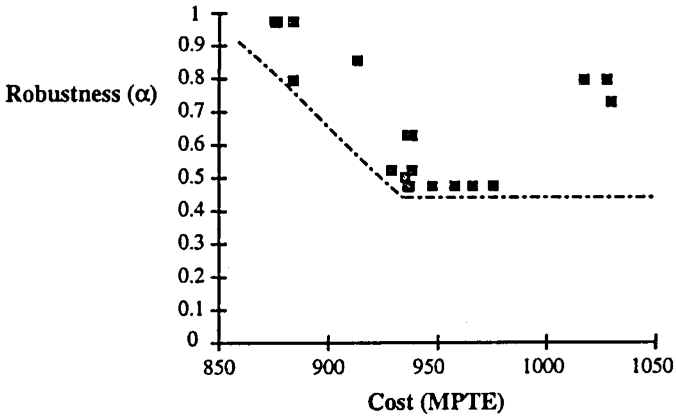


Figure 30 Cost v. robustness for a group of non-dominated solutions (maximizing robustness corresponds to minimizing α).

7.4.3. Evolution of the genetic process

Figure 31 represents the typical evolution of a Genetic Algorithm for the best individual in the population and the population median (fuzzy numbers represented by their removal values). This evolution is typical in the search for an expansion plan in phase II. Some areas may be distinguished in the figure (for the best individual):

A solutions topologically unfeasible

B solutions unfeasible in terms of load supplying

C solutions unfeasible in terms of node voltage

D fully feasible solutions

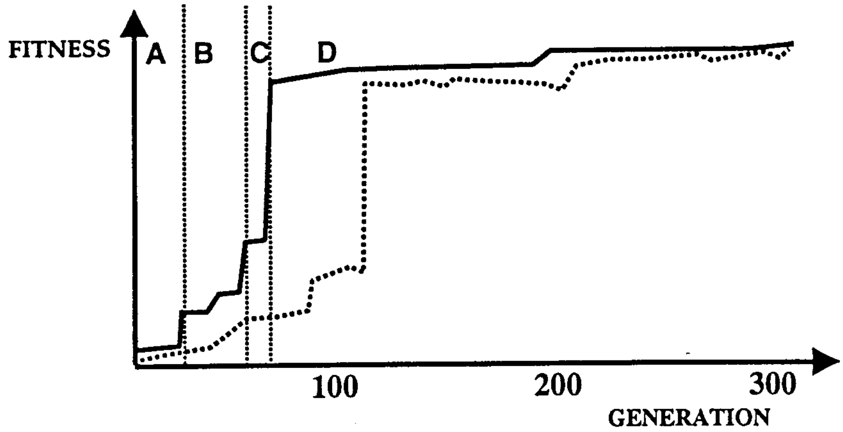


Figure 31 Typical Evolution of the Genetic Algorithm. Thick line - Best Individual; Dashed line - Median of the population

By analyzing the chart, one may verify that in this case the GA reached feasible solutions (zone D) in less than 80 generations.

It is also common that, in some genetic search processes, zone C does not appear in the chart. This happens when the first feasible solution found (in terms of load supply) does not violate voltage limits.

<u>Execution times</u>

The approximate CPU running time for a genetic process for phase II is t=180 sec. This value is not excessive since during the genetic process several feasible solutions are found and tested. The fact that we are dealing with fuzzy variables increases somehow the time necessary to obtain a solution. Distributed implementation reduce to an extent these running times. Improved implementation of the algorithms could also improve on these values.

7.4.4. Ideals

After the search process and the determination of expansion plans for each possible scenario, the next step corresponds to the use of multicriteria analysis in order to determine the Conditional Decision Set (CDS) for each one of the scenarios. From these sets, we will have the ideals for future. These ideals will correspond to a virtual solution that has the best possible values within the CDS for each criterion.

It is important to notice, however, that the choice of a CDS is necessarily a subjective process that depends on the judgment of the planner or the decision maker. Therefore, the results presented in this section are simple simulations of a possible choice process, eventually with the help of multicriteria decision making methods. Each CDS could also correspond to some kind of expectations from the planner concerning the attributes in each possible future.

Figure 32 shows one possibility for the choosing of a CDS in the example presented. This would correspond to solutions considered good compromises between the objectives. Notice that we are dealing with a problem with more than two objectives and the figure shows only the projection in a plane formed by the objectives *Investment Costs* and *Reliability* both to be minimized. Obviously, the choice of CDS for each trajectory would involve more complexity (in a 6-D space) and this graph does not correspond exactly to the ideals presented in the next section.

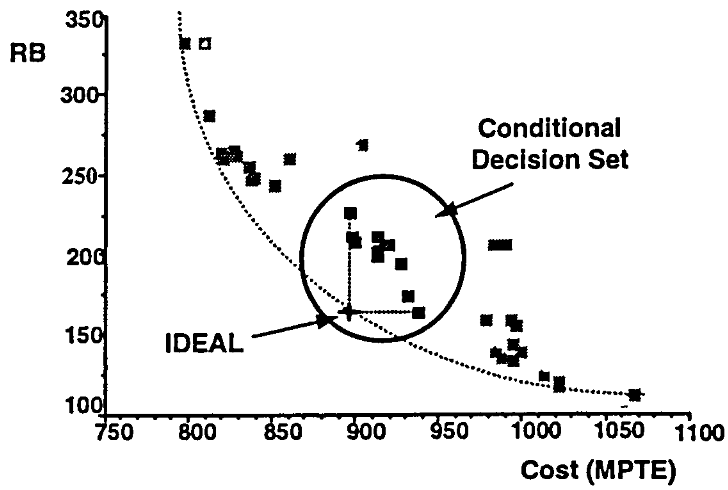


Figure 32 Possible choice of a Conditional Decision Set in a bi-objective plane. The conditional ideal is also shown as a virtual solution combining the best values for the two objectives.

Ideals

The following table presents, for each trajectory, the ideal values for two of the criteria, investment and reliability (represented by their *removal* values).

Trajectory	Investment (10^6 PTE)	RB (index) (~MWh/year)	VQ (index)	PL (index)	RO (α)	IN (index)
ABC_2D_3	707	150	0.0	0.95	0.0	0.0
ABC_2D_2	854	166	1.15	1.00	0.32	2.4
ABC_1D_2	885	183	5.15	1.05	0.45	8.3
ABC_1D_1	932	202	7.61	1.13	0.45	13.6

Table 15 Ideals for all the possible trajectories in the tree of futures

Again, it is important to stress that, apart from investment costs, all the other criteria are mere indices related to the quality of the solution. However, we should make some remarks on these results:

- As it was expected, the cost for the solutions increases and reliability decreases for trajectories with higher load. For low values of the load the ideal solution within the CDS would have no inadequacy. For higher values of load a certain inadequacy is inevitable as shown in the previous section.
- Curiously, the ideal for trajectory ABC_2D_3 has a very low value for investment costs because the algorithm was able to find at least one solution where the primary substation S3 did not have to be built. However, the solutions not including S3 had a high value for inadequacy (especially considering voltage drops) and low reliability. Nevertheless, this solution was included in the CDS for this trajectory. It seems clear that a robust expansion plan will have to include this primary substation.

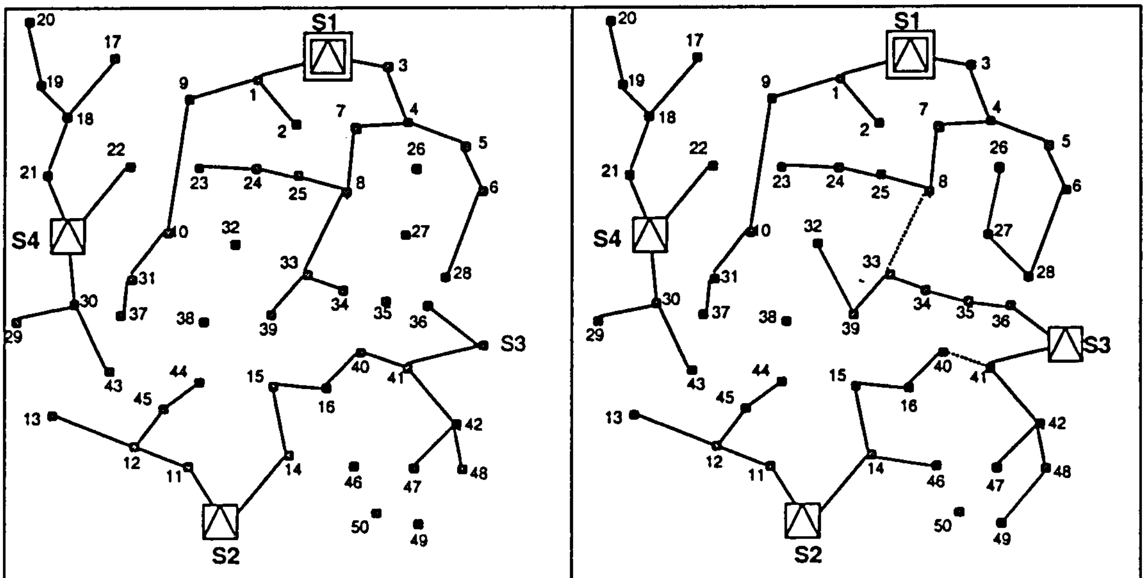
7.5. Results for Phase III

In **phase III** the objective is to find a robust expansion strategy, according to the principle of risk minimization. This means that the network expansion strategy should contemplate every evolution trajectory as a possibility to be taken into account.

The genetic algorithm search procedure is in many ways identical to the one followed in phase II. However, and as it was referred in the previous chapter, now the GA is looking for a robust expansion strategy based on the principle of regret minimization. Since the number of variables will increase, the CPU time will be considerably higher than in phase II (around 1 hour in a SPARCstation Alpha). The final result will be a set of possible strategies for the expansion of the network. In the example, one would typically find 40-50 efficient solutions.

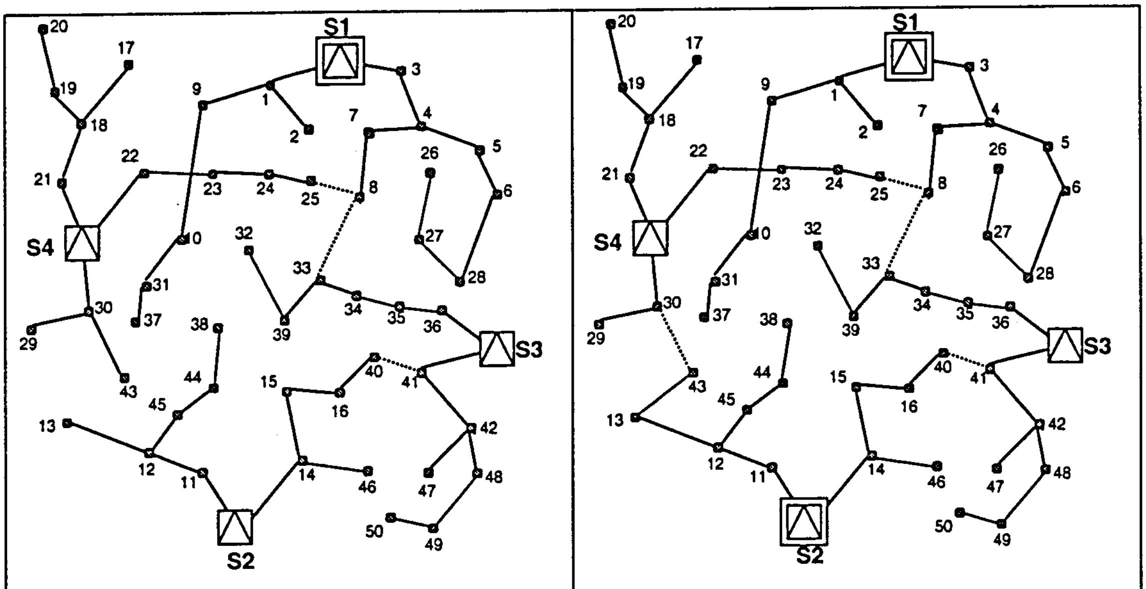
Again, the final choice will have to be done with the help of Multicriteria Decision Making methods. It is significant to stress that the final result is a single expansion strategy to be implemented. A possible simulation of a decision process will not be shown. The decision maker will have to choose between a few possible expansion strategies presented to him according to his judgement and other type of information.

Figure 33 shows one of the possible final strategies to be adopted for the network expansion. It is important to notice that this strategy is not a plan, but a flexible set of plans, to be implemented along time according to the actual evolution of the previously considered uncertainties (and, of course, considering future uncertainties). Again, a substation within a square means that there has been a reinforcement in capacity of that substation; shadowed lines represent feeders already installed that are not used in a given stage, remaining as open loops. Notice that the main difference between C1 and C2 is the need of reinforcement in S2 in C1. In this case, S2 will be able to supply extra load in node 43 (thus, line 13-43 has to be built).



A [551 ; 593 ; 644 ; 686] MPTE

B [233 ; 248 ; 262 ; 278] MPTE



C2 [19 ; 20 ; 22 ; 23] MPTE | C1 [123 ; 134 ; 144 ; 155] MPTE

Figure 33 Decision plans for stages A, B and C. C1 and C2 represent the decisions that have to be taken if the corresponding futures occur. The Tr Fuzzy Investment cost is also shown.

7.5.1.The cost of information

The following table presents the values for each criteria for both the ideals (ID) and the robust solution (RS) presented in the previous section. The difference between these values is what is defined as regret. Since we are dealing with a multiobjective problem, it is easy to see that there will always be some regret at least in one of the objectives (since, for the vast majority of the cases, there are no ideal solutions).

Trajectory	Investment		RB		VQ		PL		IN		RO	
	ID	RS	ID	RS	ID	RS	ID	RS	ID	RS	ID	RS
ABC_2D_3	707	895	150	188	0.0	5.35	0.95	1.01	0.0	0.0	0.0	0.0
ABC_2D_2	854	895	166	207	1.15	6.12	1.00	1.07	2.4	5.2	0.32	0.32
ABC_1D_2	885	1013	183	211	5.15	8.76	1.05	1.14	8.3	12.5	0.45	0.45
ABC_1D_1	932	1013	202	233	7.61	18.3	1.13	1.24	13.6	18.7	0.45	0.45

Table 16 Objective values for ideals (ID) and the proposed robust strategy (RS)

The fact that we are dealing with a large degree of uncertainty has a consequence that a robust solution will always correspond to a large regret especially in terms of investment costs (even if are trying to minimize regret!).

A graphical analysis of the results is even more interesting: Figure 34 shows, in a bi-dimensional space, global deviations that are obtained with the best risk aversion strategy, relatively to the ideal solutions that would be possible to adopt, if one could perfectly know beforehand what future would occur. These deviations are, as we have seen, potential regrets, from the risk analysis point of view. Both investment cost and ENS are fuzzy and therefore they are represented here by their *removal* values of the possibility distributions. Notice that, obviously, the investment values are similar for futures $\mathbf{ABC_2D_3}$ - $\mathbf{ABC_2D_2}$ and for futures $\mathbf{ABC_1D_2}$ - $\mathbf{ABC_1D_1}$. This is not the fact for reliability values because load is different in each pair of possible scenarios.

The values of these regrets are rather high due to the high range of uncertainties translated in the tree of fuzzy futures. For example, the regret in terms of investment if future $\mathbf{ABC_2D_3}$ occurs is considerably high, because we are not sure of the load in stage D and we have to build substation S3 (not building it would be "ideal"...) to hedge against the possibility of future D_2 occurring. A similar analysis could be made regarding future $\mathbf{ABC_1D_2}$. This consideration illustrates an interesting aspect related to planning under uncertainty. In fact, these values may be seen as the opportunity cost for a *crystal ball* – the utility management has the option of paying for hedging or buying such ancient technological device and guessing the future.

The crystal ball is obviously a metaphor for an essential aspect already referred in chapter 3: the <u>cost of information</u>. This type of analysis provides insight on the costs related to the uncertainty in data and the worth of having better and more accurate forecasting tools. Any improvement in forecast, allowing the narrowing of the range of uncertainties would result in lower regret values and considerable savings, especially in investment.

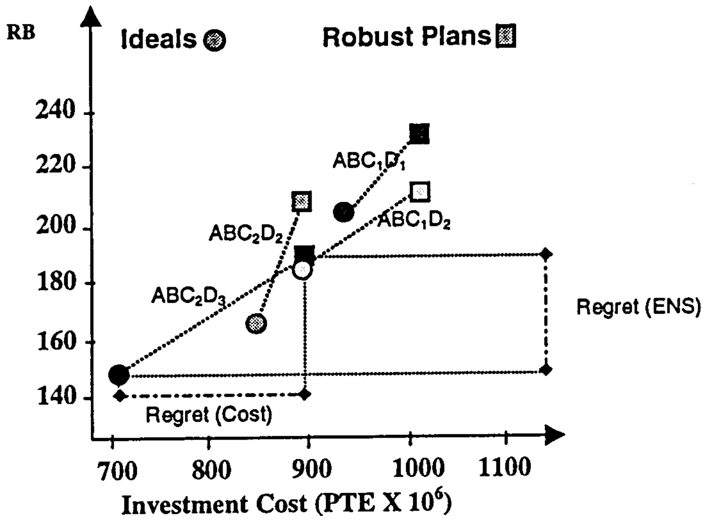


Figure 34 Distance between ideals and robust plans for each possible trajectory, shown in a bi-dimensional space (investment cost versus ENS). Regrets are shown for trajectory ABC_2D_3.

7.6. Conclusion

This chapter presented an application of the methodology proposed on chapter 6 to a distribution system planning problem.

The analysis is based on a three-phase methodology:

- An initial analysis phase, consisting of an adequate identification of criteria and uncertainties and the development of an appropriate model for the system.
- A phase where the ideals for each possible scenario are determined, as if the planner knew which future would happen.
- Finally, phase III with the determination of robust expansion strategies for the distribution network based on the principles of risk analysis.

The combination of Genetic Algorithms, Fuzzy Sets and Scenario techniques concepts for the instrumental part of the problem and, on the other hand, risk analysis and multiobjective decision making tools for dealing with the conceptual part, proved to be both suitable and resourceful, enhancing the representation ability of the models and the flexible interpretation of concepts such as risk and regret.

Furthermore, one could argue that probably the most important benefit from this innovative approach is the added insight and awareness that planners and decision makers gain in the process.

Therefore, the author believes that, undoubtedly, this type of approach will be retained for future work, due to its maturity and to the potential economical benefits arising from it. Nevertheless, the models and algorithms will eventually have to be perfected on the base of experience and improved techniques. The introduction of new modeling options and the full integration with a GIS platform will be one of the directions to follow.

Chapter 9 will present another case study in order to further analyze the consequences of this enhanced perspective on planning. Finally, chapter 10 will equate all the aspects related to the evolution of the model.

8. COMPLEMENTARY STUDIES

Facts do not cease to exist because they are ignored

Aldous Huxley

8.1. Summary

This chapter presents some complementary studies developed for solving the problem of optimal sizing and placing of capacitor banks including tap control schemes, meeting the objectives of power losses and investment minimization, as well as keeping node voltage within adequate bands. The methodology uses genetic algorithms as search tools.

8.2. Introduction

From the tests and studies performed in distribution expansion planning, we may reach the following conclusions:

- One cannot reasonably expect to include every detail in a planning model.
- Solutions emerging from either phase II or III can be further improved.

Considering this two facts, we may want to perform some complementary studies on the solutions found for the system's expansion. These studies could include:

- Placing of capacitor banks in order to reduce losses and improve voltage quality.
- Placing of switching equipment, considering reliability aspects.
- Placing of new branches in order to close loops and improve reliability.

All these aspects have already been studied with some detail in the author's power system group in INESC. This chapter will present some interesting studies on capacitor placing with the use of Genetic Algorithms.

We may start by referring the general objective of reactive power planning, which consists in finding the optimal adjustments of transformer tap positions, generator voltage levels and the placement of capacitor banks with the objective of reducing total real power losses and improving the voltage profile [Swa96], [Tom92].

This highly complex problem may be included in the field of operational planning, with the possible exception of the problem of optimal Capacitor/Reactor placement. Therefore, and similarly to the switcher placing problem, this problem is rarely included in the expansion planning process itself. However, studies related to reactive power planning are occasionally considered after the network expansion is determined.

In [Wan96] the authors (Lou Chin Wang, V. Miranda, and the author of this thesis) propose a new method for the placement and sizing of capacitor banks in distribution systems based on Genetic Algorithms. The model tries to minimize investment and operation costs, keeping node voltage within adequate bands. It also presents the application of the methodology to the distribution network in Macau. The matter presented in this chapter is a summary of what published and the next sections will briefly present the problem formulation and the solution technique developed in this work.

8.3. Problem formulation

The formulation of optimal capacitor placement planning problem as a constrained, non-differentiable, multi-objective optimization problem is given as follows where three objectives are considered: the cost associated with capacitor placement (i.e. purchase, installation and maintenance); the cost for energy losses; and the one related to voltage quality.

8.3.1. Cost function

The cost function includes the cost of capacitor placement and the cost of total energy losses in a distribution system. Generally, costs for both fixed and switchable types are considered.

<u>Cost of fixed capacitors</u>

The cost of fixed capacitors is associated with the cost of installation and purchase of capacitor banks. Mathematically, it can be given as follows:

$$C_F(u) = \sum_{k \in N_C} C_k(u_k^0)$$
$$= C_c(N_A) + \sum_{k \in N_c, u_k^0 \neq 0} C_I(u_k^0) \qquad (76)$$

where

u_k^0	: Size of capacitor at bus k during peak load level
$C_c(\cdot)$	: Fixed installation and maintenance cost
$C_I(\cdot)$	: Cost of capacitor banks
N_c	: Number of candidate buses
N_A	: Number of bus where capacitors are added
k	: bus number index

Cost of switchable capacitors

The cost of installing switchable capacitors can be represented by $\mathbf{C_s(u)}$. This cost function depends on the relative values of the k^{th} bus capacitor control setting u_k^i (where u_k^i is the control setting of the k^{th} bus capacitor during i^{th} load level). If all u_k^i are equal for all **i**, it means only fixed capacitors are needed to install at bus **k** and this portion of cost can be considered to be zero. Thus, the total cost of capacitor can be represented by

$$C_{cap}(u) = C_F(u) + C_S(u) \qquad (77)$$

Cost of energy and power losses

One of the objectives of the problem is to minimize the total real power losses arising from transmission lines, which can be calculated as follows:

$$P_{loss} = \sum_{k=1}^{N_l} G_{k(i,j)} \left[V_{ii}^2 + V_j^2 - 2V_iV_j \cos(\delta_i - \delta_j) \right] \qquad (78)$$

where

N_l : Total transmission lines on the system

$G_{k(i,j)}$: Conductance of transmission line k connected between bus i and j

V_i : Voltage magnitude at bus i

δ_i : Voltage angle at bus i

and can be directly obtained from load flow results. Now, let P^i_{Loss} represent total real power loss in the distribution system during i[th] load level and T_i be the time duration of i[th] load level. Then, total energy loss of the distribution system is:

$$E_{loss}(x,u) = \sum_{i \in N_T} T_i . P^i_{Loss}(x,u) \tag{79}$$

where

N_T : Total number of load levels to be studied

The cost of energy losses can be calculated by multiplying the total energy loss by a cost-to-kWh conversion factor K_E.

$$C_{e.Loss}(x,u) = K_E \cdot E_{Loss}(x,u) \tag{80}$$

Here, the factor K_E reflects a capitalized value of the kW of losses. In practice, it must reflect the marginal cost of energy generation.

There is another effect of losses which is the marginal gain in capacity of the power system, because of reactive compensation. This gain must reflect marginal investment costs both at generation and at transmission level (lines, transformers).

In our model this is calculated as a function of the maximum value of P^i_{Loss} obtained for a particular configuration, at the peak value of the load curve. If marginal costs of power are known (at least an estimation), then the following expression may be added to the costs incurred due to the existence of power losses:

$$C_{p.loss}(x,u) = \sum_{i \in N_G} K_{p.i} \cdot P^o_{Loss}(x,u) \tag{81}$$

where

$P^{o}_{Loss}(x,u)$ - Losses at peak hour

$K_{p.i}$ - Marginal investment cost coefficients for the generation and transmission system

N_G - Number of different marginal costs to be considered; for instance, for generation, for substation transformers and for transmission lines.

<u>Voltage quality</u>

Bus voltage is an important index for security and leads us to use the following deviation function (deviation of bus voltage from a specified voltage) as one of the objectives function to the problem in order to force bus voltage magnitude to approach specified voltage magnitude V^{spec} which is normally set to unity.

$$C_{dv} = \sum_{k=1}^{N_b} \left(\frac{\left| V_k - V_k^{spec} \right|}{\Delta V_k^{max}} \right)^2 \qquad (82)$$

where

N_b : Total buses of the system under study

V_k : Voltage magnitude at bus k

V_k^{spec} : Specified voltage magnitude at bus k

ΔV_k^{max} : Maximum allowable voltage deviation limit at bus k

8.4. Overall problem formulation

The objective of the overall problem is to find

$$\text{Min } C(x,u) = C_{cap}(u) + C_{e.Loss}(x,u) + C_{p.Loss}(x,u) + k \cdot C_{dv}(x) \qquad (83)$$

subject to

(1) Load constraints:

The load constraints are the real and reactive power balance described by a set of power flow equations that can be expressed in compact form

$$F(x,u) = 0 \tag{84}$$

(2) Operational constraints on bus voltage:

$$V_{min} \leq V_k \leq V^{max} \qquad k = 1,2,3,\ldots,N_b \tag{85}$$

(3) Constraints on fixed capacitors:

$$u_k^i = u_k^0 \leq m_c u_s, \qquad \text{for all } k \in N_c, i \in N_T \tag{86}$$

where

m_c : maximum number of capacitor banks to be installed at bus k

u_s : standard capacitor size of one bank

(4) Constraints on switchable capacitors

$$u_k^i \leq u_k^0 \leq m_c u \qquad \text{for all } k \in N_c, i \in N_T \tag{87}$$

It should be pointed out that if more information about active and reactive production limits of generators is provided, the operational constraints such as

$$\begin{aligned} P_{gj}^{min} &\leq P_{gj} \leq P_{gj}^{max} \qquad j = 1,2,\ldots,n_g \\ Q_{gj}^{min} &\leq Q_{gj} \leq Q_{gj}^{max} \qquad j = 1,2,\ldots,n_g \end{aligned} \tag{88}$$

can also be included without major difficulties.

8.5. Solution Technique

In the study, a Genetic Algorithm (see chapter 5 for details on Genetic Algorithms) has been applied to solve the problem because it has several attractive features:

- It can handle different kinds of equality and inequality constraints. Differentiable and non-differentiable functions are also easily dealt with.
- GA have no difficulty in dealing with discrete variables such as the number of capacitor units.

The only disadvantage of this optimization technique is that it might take considerable amount of computation time to achieve a (near) global optimal solution, if an advanced model (and some care in its development) is not adopted.

The details about the use of genetic algorithm will be omitted in this brief explanation. However, we may summarize the algorithm as follows:.

1. The Genetic Algorithm proposes a specific configuration and sizes for capacitors.
2. The system finds the optimal capacitor settings in order to minimize operation costs and maintaining the voltage in adequate levels.
3. The overall cost of the solution is assessed, including investment and operation costs.

The output of the above solution algorithm gives the optimal capacitor size vector $(u_1, u_2, \dots, u_{N_c})$, where $u_k = \left(u_k^1, u_k^2, \dots u_k^{N_T}\right)$ and u_k^j is the control setting of bus k at load level j. This piece of information can be used to determine the number, location, sizes of either fixed or switchable capacitors to be placed. The types of capacitors to be installed at bus k can be determined by control setting u_k, using the following rule:

if $u_k^1 = u_k^2 = \dots = u_k^{N_T}$, then a fixed type of capacitor is to be installed at bus k, otherwise a switchable type of capacitor is to be installed at bus k.

For the purpose of illustration, a 3-node distribution system will be used. Let us suppose that the output of the solution algorithm gives the following control setting vector:

$$u = \begin{Bmatrix} \{3 & 5 & 0\} \\ \{3 & 4 & 0\} \\ \{3 & 2 & 0\} \end{Bmatrix} \quad (89)$$

This indicates that three banks of fixed capacitors are needed to be installed on node 1 for all load levels, and five banks of switchable capacitors are needed to be placed on node 2 (a fixed base of 2, and 3 switchable) and no capacitor is needed at node 3.

The application of this method to the distribution network in Macau (China, under Portuguese administration) operated by CEM (Macau Power Company) has proved to be efficient and has provided robust and reliable solutions. [Wan96] makes public some of the results obtained. Nevertheless, we may present some results referring to one of the solutions for capacitor placing obtained for the simplified 12-bus network of Macau (the system presents a set of alternatives with different costs and improvements). This results are illustrated in the figures.

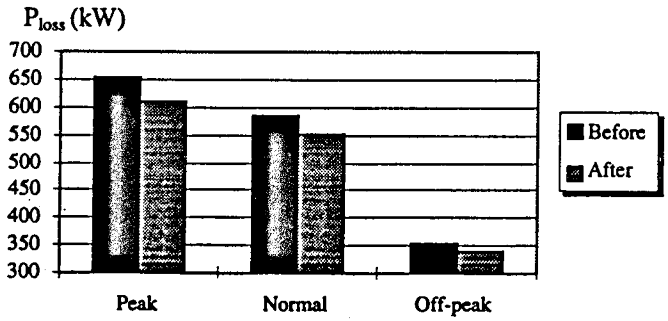


Figure 35 Power losses at three different load levels before and after capacitor installation

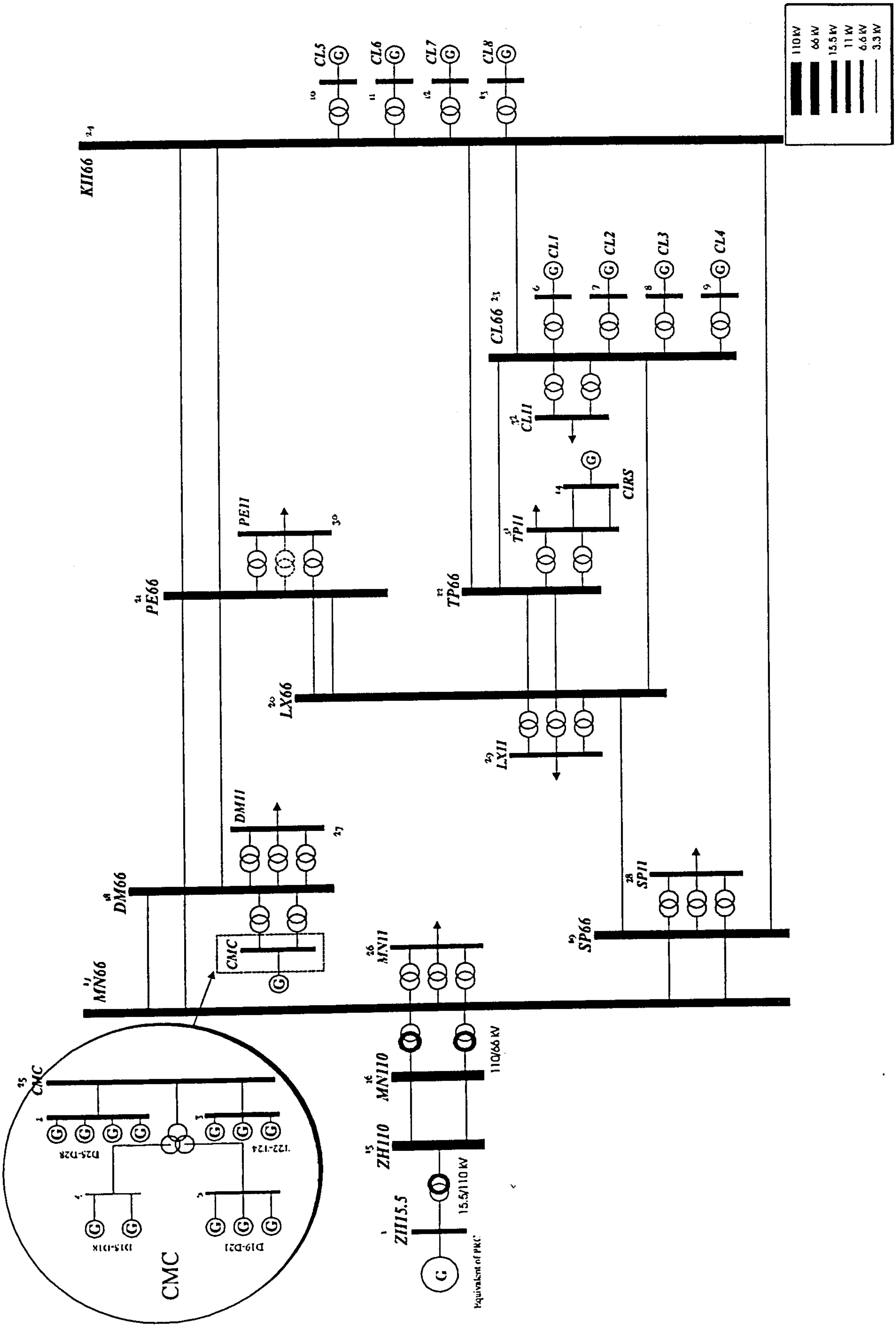


Figure 36 Network of Macau. This network was simplified to a 12-bus system for the example.

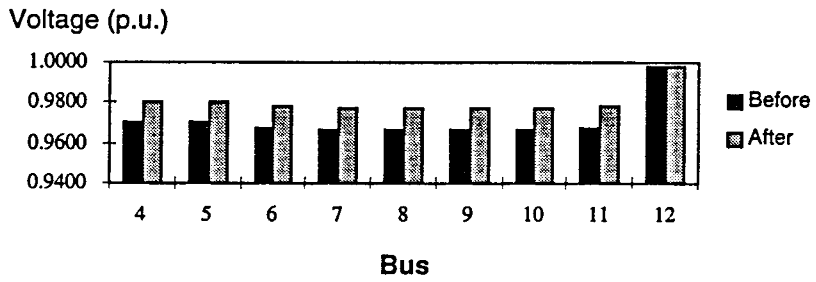


Figure 37 Voltage profiles at peak load before and after capacitor installation

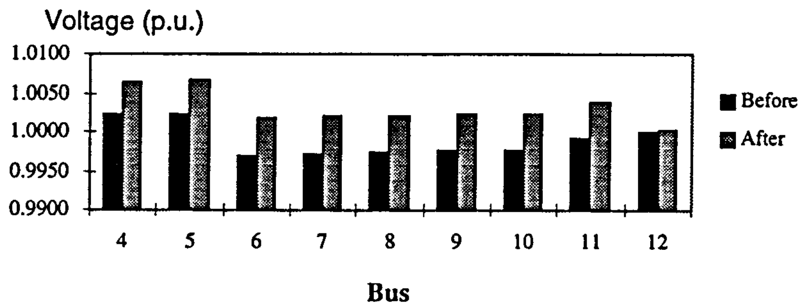


Figure 38 Voltage profiles at off-peak, before and after capacitor installation

We may appreciate from the figures that, in fact, the installation of capacitor banks may improve both voltage profiles and reduce losses in a considerable manner in any load level considered. Obviously, some analysis on cost/benefits from the solutions should be performed before making a final decision.

8.6. Conclusion

This chapter illustrates how the optimal capacitor placement may be solved based on Genetic Algorithms. Both fixed or switchable capacitor can be considered, demonstrating the flexibility of the modeling.

Obviously, including capacitors in a power system would require other types of studies to be performed, namely about possible resonance with harmonic frequencies at some load levels or system configurations, and also their contribution to the dynamic stability maintenance of the system. These types of studies, presently, are made after some scenarios about sizing and location of capacitors have been set: it is still not possible, today, to integrate everything in the same search process.

The algorithm is capable of finding robust, reliable solutions for system planners. However, it is important to notice that, for a real case study, the following factors must be taken in account:

- A comprehensive modeling of system components (Generators and Transformers);
- An accurate enough load modeling;
- Precise (as possible) information installation and maintenance costs of both fixed and switchable type capacitor;
- Adequate information about site conditions (such as the space available to install capacitors) when considering candidate buses in the problem. Their restriction or consideration may result in more costly investment (since the problem is more constrained), but it may reflect better the physical realities;
- Information about load growth and network planned expansion. This will allow to establish a dynamic policy on the investment in capacitor, taking into account future scenarios and not only a static load profile.

The model developed by the author's group in INESC is able to take in account all these aspects and qualitative rules that a utility may wish to consider, conditioning the search of a solution.

9. RISK ANALYSIS AND PROBABILISTIC MODELS

9.1. Summary

This chapter, based on results already published in [Mir97a] and [Mir97b], demonstrates that a classical stochastic optimization is, in many cases, not convenient for power system planning. Instead, a risk analysis approach is proposed. The underlying assumptions of each paradigm are reviewed and the capacity of each to find acceptable solutions is analyzed. The comparison is illustrated by an example where the subjective evaluation of the possible future scenarios is expressed by fuzzy probability assessments. The chapter also shows how a decision can be reached if fuzzy probabilities are assigned to each possible future.

9.2. Introduction

The general expansion planning problem in Power Systems has been developed traditionally under the Probabilistic Choice (PC) paradigm, which may be described as following: admit that one has defined a "cost function" that measures the goodness of a solution (*latu sensu*: we refer as "cost" any criterion that may be considered and adequately transformed into an objective of minimization); given a set of futures, each one with a probability assigned, the optimal solution should be chosen among those that minimize the expected cost over the set of futures considered.

In recent years, however, the generalized adoption of this planning paradigm has been increasingly challenged. As an alternative, a Risk Analysis approach has been suggested. This approach applied to Power systems was clearly defended, with enlightening examples, by H. M. Merrill and A. J. Wood in 1990 [Mer90]. Since then, this approach has been adopted, in greater or lesser degree, in a growing number of papers published, as for example [Gor93] and [Mir95]. Also in some recent books this approach has been defended [Wil96], with some interesting spatial examples:

> *"One of the worst mistakes that can be made in transmission and distribution planning, integrated or otherwise, is to try to*

circumvent the need for multi-scenario planning by using 'average' or 'probabilistic forecasts'. Generally, this approach leads to plans that combine poor performance with high cost."

The Risk Analysis (RA) paradigm indicates a preferred solution as one that minimizes the regret felt by a Decision Maker (DM) *aftèr* verifying that the decisions he had made were not optimal, given the future that in fact has occurred.

Why has been the RA approach so attractive (although not many models have included it explicitly yet)? One of the reasons may well be that it reflects with much more accuracy the way people think. In fact, the traditional models in planning, namely those working under the PC paradigm, concentrate their analysis on the *solutions* of the problem, while the RA paradigm is mainly about *decisions*. Factors as risk aversion or risk attraction associated with DMs, the concept of *hedging* (paying an extra to avoid adverse futures) and the measurement of regrets felt, all are reflected in the RA way of dealing with a problem, and are very well understood by those planners that have a daily contact with real and practical problems.

However, models developed by academics and scientists have insisted one way or the other in a kind of stochastic optimization, where none of these concerns are correctly or completely addressed.

Recognizing that Power System planning is a matter of Decision Making, and not of Optimization, has been one major step into a new perspective on this activity, with deep practical consequences.

However, recognizing that the optimum on the average of futures may not be the best decision, from the point of view of Decision Making, has been surely harder.

So far, this discussion, in the Power System area, has been limited to some philosophical arguments, from which sometimes one may suspect that an unbiased appreciation may be absent in some degree. People used to "stochastic thinking" will need some time to adapt to the meaning and use of other paradigms, while people that do not fully understand probabilistic models will tend to over-evaluate risk analysis as a new universal tool. Neither position seems to us as deserving to be sustained.

This chapter is devoted to, giving a clear and rigorous demonstration of why a PC paradigm may be the wrong choice, in Power System planning, and why the RA approach may be superior, from a decision making point of view, in a number of cases. It also presents a sketch of a method to reach a safe decision, even if the subjective assessment of probabilities for distinct scenarios is given, not by numbers, but by fuzzy declarations - such as "the probability of scenario A is *more or less* 0.4", or "the probability of scenario B is *most likely* between 0.3 and 0.4, and surely not below 0.2 or above 0.5". In this perspective, it offers some innovative discussions.

9.3. Risk analysis *versus* Probabilistic choice

In a previous chapter, it was shown that a planning model should present, as a result, not only one "optimal" solution, but a set of planning alternatives.

Admit that one has a set of **m** planning alternatives. Admit also that the planning uncertainty is represented by a set of **n** futures, each one having assigned a subjective probability. Admit that we have evaluated, for each alternative **i**, its regret in each future **k**. We can therefore imagine these alternatives represented by their regrets in an n-dimensional space of scenarios.

We will compare the PC and RA paradigms under a set of assumptions, namely that for a planning problem we have been able to define a finite set of possible futures, having each future **k** assigned a *subjective* probability value w_k.

Usually a future will denote a particular dynamic sequence of events uncontrolled by the Decision Maker, and not only a static image. The dynamic nature of future was referred on a previous chapter, illustrated by the concept of "fuzzy tree of futures"; *Figure 39* depicts the evolution of load for 3 stages (plus the initial state) in the planning horizon, and we have considered 3 possible futures, where loads are represented by fuzzy numbers; a future will in this case be a path in the tree of futures, and not a "frozen" state at a time

T. Even so, a path k can have assigned a probability w_k. The data presented in the figure will be used in the example presented in the following sections.

A scenario is usually considered as a given set of decisions in a particular future; sometimes in the literature these concepts are not clearly distinguished, and the words "future" and "scenario" are used equivalently.

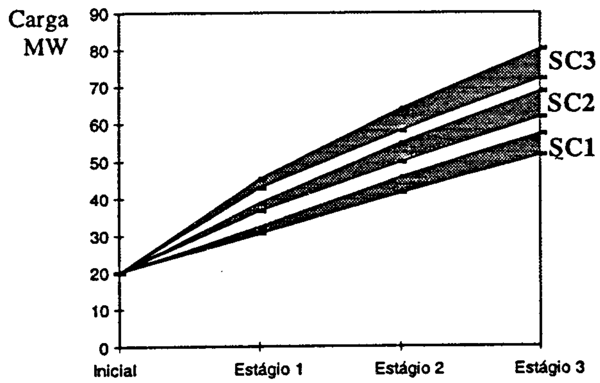


Figure 39 Tree of futures with fuzzy loads

9.3.1. Applicability of the PC paradigm

The PC paradigm depends on a linear composition of the future costs: if the cost incurred in future k by an alternative i is f_{ik}, then the PC paradigm is characterized by:

$$\min_i \sum_k w_k f_{ik} \tag{90}$$

This does not depend on the evaluation of the consequences of the decisions, compared to what could have been decided if one knew the future in advance. This is why we say that the PC paradigm deals with solutions and not with decisions.

If the future k were known in advance with no uncertainty, we will admit that one would be able to calculate an optimal solution for the expansion of a Power System, having cost f_k^{opt} - this is conditional optimum (for future k). For the sake of making concepts clear, let's define the *Regret* (**R**)felt for a

Decision Maker, for having chosen a given alternative i and then seeing future k happening, may be defined as a function.

$$\mathrm{Regret}_{ik} = f_{ik} - f_k^{opt} \tag{91}$$

If we define a stochastic optimization of the possible regrets felt, we have:

$$\begin{aligned}
&\min_i \sum_k w_k \left(f_{ik} - f_k^{opt}\right) \Leftrightarrow \\
&\min_i \sum_k (w_k f_{ik} - w_k f_k^{opt}) \Leftrightarrow \\
&\min_i \sum_k w_k f_{ik}
\end{aligned} \tag{92}$$

because $w_k f_k^{opt}$ is a constant for every k. In this case, therefore, the stochastic optimization of the solution costs is equivalent to the stochastic minimization of the decision regrets.

This, however, is not true in the general case of having **R** as a non linear function of the deviation from the conditional optima. The non-linearity of the regrets may perhaps be neglected, if the real costs are relatively close to the conditional optima; but the perspective of unwanted or even catastrophic events surely gives to the function **R** a highly non-linear characteristic - which means that scenarios with a very low probability of occurrence may in any case have a decisive influence in the final decision, namely when they lead to the acceptance of higher costs to avoid their consequences (this attitude is called *hedging*).

The concept of *unwanted* events is very important, because it contradicts the basic assumption behind the Probabilistic Choice paradigm: that bad situations will be compensated by good ones along time, so that one can evaluate a solution by its average behavior. However, if a catastrophic event or scenario occurs, no reasonable recovery will ever be possible and the PC assumption cannot be verified - the PC paradigm is not an useful context for decision making in planning, in this case.

As a general rule, the PC paradigm is not adequate to problems where an enough significative repetition of events does not occur; this may result either from a too low frequency of the event cycle (compared with the planning time horizon) or from the occurrence of unwanted or catastrophic situations that constitute a disruption in the stochastic process of repetition of *experiences* and compensation of consequences.

When can we use linear Regrets? Basically, in any problem context where the cost of any solution f_{ik} is not very different from an ideal cost f_k^{opt}, and where no unwanted events are likely to occur. Figure 40 displays an example of a Regret function with linear characteristic near the zone of no regret and increased risk aversion for extreme scenarios.

$$Regret = \begin{cases} f - f^{opt} \Leftarrow (f - f^{opt}) < 1 \\ \left(f - f^{opt}\right)^2 \Leftarrow (f - f^{opt}) \geq 1 \end{cases} \tag{93}$$

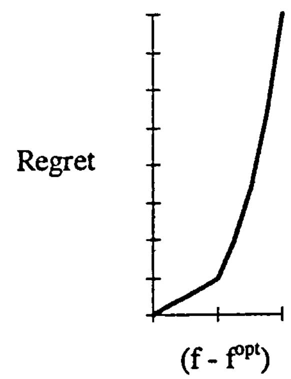


Figure 40 Regret function with aversion to extreme cases

9.3.2.Applicability of the RA paradigm

The RA paradigm may be characterized by the following expression:

$$\min_{i}\left\{\max_{k}\left(w_k \text{Regret}_{ik}\right)\right\} \tag{94}$$

By minimizing the maximum regret, one really tries to avoid selecting a solution that may have a bad performance in any future considered. This is not equivalent to the PC paradigm, not even in its "updated version" where the solution costs are replaced by the associated regrets, such as in:

$$\min_{i}\sum_{k}\left(w_k \text{Regret}_{ik}\right) \tag{95}$$

The application of the RA paradigm is associated with the concept of *robust* solution, meaning a solution that will be acceptable in all futures considered. *Exposure* relates a solution with adverse scenarios, and the concept of *hedging* results therefore naturally as a mean of reducing exposure or increasing robustness.

The RA paradigm is applicable whenever one cannot assume a probabilistic compensation of bad and good results, because there will not be enough frequent repetition of events or situations (validating a PC paradigm through the law of large numbers); this may result either from a too low frequency of the event cycle (compared with the planning time horizon) or from the occurrence of unwanted or catastrophic situations that constitute a disruption in the stochastic process of repetition of "experiences" and compensation of consequences.

The RA paradigm has a mathematical formulation that is not adequately tamed by classical optimization algorithms, which may yet be another reason for its late popularity. However, the recent emergence of innovative optimization methods, like the family of evolutionary computing algorithms [Mir96a], referred in chapter 5, opened a clear path to deal with these types of non linear models in a hopefully efficient way.

9.4. Comparison

The last section has so far justified why the PC and the RA paradigms may lead to different planning decisions, and why *the PC paradigm may result unjustified due to the non-linearity of the Regret concept and the possibility of occurrence of unwanted or even catastrophic scenarios*. This section will now explain why the PC paradigm may also be inconvenient for indicating a good compromise solution in a planning environment including uncertainty.

9.4.1. The PC paradigm misses compromise solutions

It is possible to define the planning problem as the simultaneous minimization of **m** objective functions, each one being the minimization (weighted by the probabilities) of the regret in each future **k**:

$$\begin{aligned} &\min \ w_1 \, Regret_1 \\ &\dots \\ &\min \ w_m \, Regret_m \end{aligned} \qquad (96)$$

We recall that an alternative is said to be non dominated or Pareto optimal if any other alternative that is better than the former in one criterion is also worse in at least one other criterion - this means that we cannot improve in one criterion without losing somewhere else. These non dominated solutions are universally accepted as being the natural candidates for a final decision. We will admit then that the set of solutions has already been screened and that we have retained only those that are non dominated.

Equation 90 or its regret-equivalent 95 represents a linear combination of the criteria. Therefore, varying the weights w_k, we will only discover those solutions that lie on the convex hull of the non dominated set. However, the surface joining all non dominated solutions is not necessarily convex and, in fact, in problems with integer variables, this non-convexity most often happens. Therefore, many non-dominated alternatives that may possibly be interesting compromise solutions will be missed, if the search is conducted by a PC paradigm.

Certainly, an approach that allows "invading" the concave parts of the non dominated border will be necessary to uncover new alternatives that may constitute good compromise solutions. The RA paradigm, translated by equation 94 has precisely this ability.

9.4.2. The PC paradigm is riskier

The last appreciation of the relative merits of the PC and RA paradigms refers to the consequences of the decisions produced by both approaches.

It is easy to demonstrate the following*: the PC paradigm consistently tends to propose riskier decisions.* Of course, this is only bad if one cannot recover from the negative consequences of an adverse event by compensating it with a number of favorable events. We are stating therefore that, not only for philosophical reasons based on its assumptions but also for operational reasons related with the emergence of adverse conditions in the future, the

PC paradigm is not an adequate tool for problems or environments where compensation is not likely to occur.

The PC solution is the best on the average of the futures; however, if the future does not happen close to that average or to the most probable forecasted future (which often happens), the regrets incurred are in general higher than those that would be generated by decisions proposed by the RA paradigm.

Of course, if the RA minimizes risk, it is natural that we may find less risky decisions with it. What is impressive is that the PC paradigm tends systematically to push decisions to one of the extreme futures while the RA denotes more capacity for compromising.

If we compare the PC and the RA paradigms in the space of scenarios, equations 94 and 95 may receive the interpretation that one is trying to find the solution that is closest to the *ideal* one. The concept of *ideal* has been referred in a previous chapter as a virtual solution having the best value in all criteria, chosen from the values found in the non dominated set - this would be the solution chosen by the DM, if it were feasible.

In fact, these equations represent the minimization of the distance to the ideal (in this case, represented by the origin, which would mean a solution with no regrets in all futures, a fully robust solution).

The PC paradigm measures this distance in a L_1 metric; the RA paradigm measures this distance in a L_∞ metric. Zeleny defined *compromise set* as the set of solutions close to the ideal, that would be obtained by varying the metric used to evaluate this distance, namely from L_1 to L_∞.

A distance to the origin, in a system of n coordinates and in a L_l metric, is given by

$$d_l = \sqrt[l]{x_1^l + \ldots + x_n^l} \tag{97}$$

The Euclidean metric is L_2, of course. The L_∞ metric converts into the expression

$$d_\infty = \max(xj;\, j=1,\ldots n) \tag{98}$$

which is reflected in the RA paradigm. Because of this definition of the L_∞ metric, the distance$_\infty$ value will not be affected by any simultaneous transformation of the coordinates by a increasing monotonous function; when applied to equation 94, this means that the RA paradigm may be represented just by the difference $f-f^{opt}$, and that a Regret function such as the one in Figure 40 will not affect the decisions proposed under this paradigm. This is why the RA paradigm may be seen as the most conservative one, in terms of risk aversion.

Any intermediate metric, such as the Euclidean, will therefore tend to propose solutions in between the ones offered under the riskier PC paradigm and the conservative RA paradigm.

This cannot easily be demonstrated mathematically, but will be clearly illustrated with the example in the next section.

9.5. Worked example

We will describe the main results from a distribution planning exercise where we have considered 3 possible load futures, referred as SC1, SC2 and SC3, with subjective probabilities assigned. In each future path, the load has a fuzzy definition, leading to the fuzzy future tree shown in Figure 39. The example was developed for a planning horizon divided in 3 time stages.

The initial distribution system will be the same as in the case study presented in chapter 7 and is composed by 54 nodes, 16 lines and 2 substations and there are 45 lines and 2 substations in project. The data on this system is the same as presented in chapter 7.

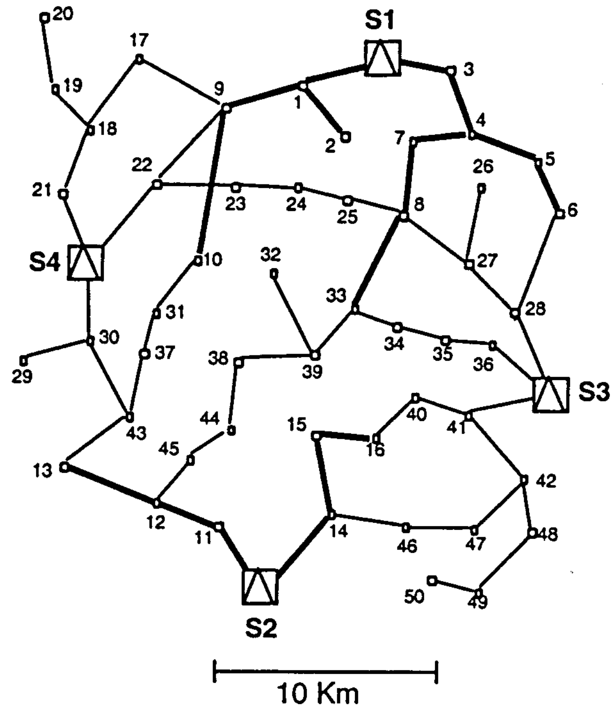


Figure 41 Thick lines represent existing branches and thin lines represent possible sites for the expansion of the network. Substations S3 and S4 are in project.

In determining feasible alternatives, three criteria were considered:

- Investment (+power losses).
- Solution inadequacy. This criterion was already defined in chapter 4 and derives from the fuzzy definition of loads, and it may be interpreted as the consequence of adverse scenarios in which the system capacity remains below the forecasted load, leading to the need for corrective measures or to repressed demand. A way to measure this attribute is also described in chapter 4.
- Reliability - evaluated through the assessment of energy not supplied as referred in chapter 6 and 7.

In order to find a set of non-dominated solutions, we have applied a genetic algorithm tool as described in chapter 5. The 3 criteria above were combined to a single value for each scenario, by the use of appropriate fuzzy tradeoff weights, which were varied to present a picture of the non-dominated border of the problem.

The fuzzy tradeoffs reflected an uncertain cost of the kWh not supplied (reliability) and of the peak kW not suppliable (inadequacy).

As it can be seen, we have incorporated uncertainty in the model, through fuzzy modeling. This aspect is not described in detail in this chapter, because it is not essential to the discussion; a full description may be found in chapter 4. The results presented hereafter result from the defuzzification of fuzzy values obtained.

Note: For the sake of clarity, weights have been used for the criteria, in a utility function. This was done in order to simplify the analysis and regard the problem as multiobjective *in the space of futures*, not in the space of criteria. Maintaining the problem as multicriteria and multiscenario would make the analysis extremely complex (with regrets in both the criteria and the futures) and obscure the real purpose of this study. The weights were chosen within reasonable limits in order to make the analysis as interesting as possible. The software system developed allows the variation of these weights in order to see the sensibility of the solutions to the variation in weights for the criteria.

The genetic algorithm was able to find 49 non-dominated solutions (expansion plans). Note that the solutions are non-dominated in the space of criteria and not necessarily in the space of futures. The defuzzified weighted cost (including effects from investment and losses, power not supplied and system inadequacy) for each scenario is shown in table 1 for some selected solutions (cost in PTE.10^6). The ideal scenario is shaded - Solutions 32, 28 and 11 are the best for futures 1, 2 and 3 respectively.

Solution	SC1	SC2	SC3
Sol11	1029.38	1090.21	1291.77
Sol12	1061.26	1192.24	1553.40
Sol13	1008.27	1265.27	1724.74
Sol14	1015.86	1257.96	1682.45
Sol15	1014.60	1257.26	1687.42
Sol16	992.36	1224.30	1640.64
Sol17	992.27	1233.12	1674.14
Sol18	992.18	1233.65	1680.39
Sol19	1108.29	1319.32	1745.15
Sol20	1080.22	1245.72	1640.27
Sol21	1080.13	1247.11	1653.92
Sol22	1082.70	1278.96	1760.50
Sol23	1009.98	1075.85	1295.23
Sol24	1022.25	1092.10	1382.19
Sol25	1033.80	1107.24	1347.22
Sol26	1012.37	1082.88	1373.52
Sol27	1008.36	1074.94	1332.83
Sol28	1004.63	1070.62	1322.84
Sol29	1005.41	1079.53	1319.42
Sol30	1012.13	1085.74	1292.85
Sol31	1009.16	1074.61	1293.57
Sol32	977.19	1119.73	1424.01
Sol42	989.94	1118	1423.1

Table 17 Weighted costs for selected solutions (PTE . 10^6)

9.5.1.Comparing RA and PC results

Taking the planning problem as multicriteria over the three Scenarios, it was solved using: linear Probabilistic Choice as in expression 90 and Risk Analysis as in expression 94. The problem was also solved using the concept of Euclidean distance (expression 97 for i=2) and probabilistic choice associated to non-linear regrets (expression 95 combined with expression 91) The regret values were easily calculated from the results on table 1.

The interesting results become apparent when we represent the set of preferred solutions in a space of weights, by considering the w_k future weights as variable parameters in the whole range [0,1]. As we have

$$w_{SC1} + w_{SC2} + w_{SC3} = 1 \tag{99}$$

we can project the result of this exercise in a w_{SC1}/w_{SC2} plane, giving Figure 42 to Figure 45.

This projection is a triangle, as a result of equation 99. The vertices of this triangle correspond to the three scenarios, because at each of these points one of the weights has value 1 (at the origin, we have $w_{SC3} = 1$); a point inside this domain corresponds to a specific mix of (subjective) weights.

In this particular problem we can see that the PC model recommends solutions 11 and 32 (which are the ideals in SC3 and SC1) in a larger range of weights than the RA model - the PC paradigm tends to recommend extreme alternatives, more than the RA paradigm, while the RA points out earlier to compromise solutions.

However, the most striking observation is that the PC model suggests much less solutions than the RA model. The PC paradigm does not "see" some of the possible compromises - it does not consider alternatives 29, 33 or 42. Furthermore, solution 31 is recommended by RA in a much more reduced set of weights that by the PC model.

The fact that solution 28 is favored by the RA model is not in contradiction with the statement that the PC model is riskier. On the contrary, 28 is the ideal for SC2, *which is an intermediate scenario*, so in fact the RA model is favoring an intermediate solution.

The calculations with a pure Euclidean distance model give (not surprisingly) coherent results. These can be seen as intermediate between the PC and the RA models, uncovering more compromise solutions as one moves from the L_1 to the L_∞ metric.

Figure 45 shows the results for the non-linear regret PC model with a threshold value Tr=30, which means that regret values over this threshold value are squared. This model is considerably risk-adverse, since it strongly penalizes high regret values (in a sense, it tries to avoid catastrophic consequences, even for unlikely scenarios). Consequently, extreme solutions

(11, 32, 42, for example) have their influence area highly reduced, while compromise or intermediate solutions (28, 29) appear very strong. Reducing the threshold value leads to models that are more risk-adverse.

Superimposing PC and RA results, we can identify what we call the "stability areas" for the best decision: they correspond to the combinations of weights for which the recommended decision is constant, irrespective of the metric (or paradigm) chosen. For planning or decision making, these stability areas are quite interesting; they are depicted in Figure 47.

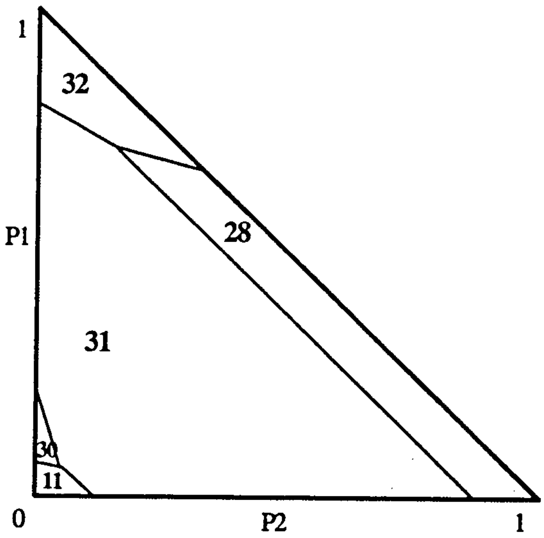


Figure 42 Best solutions for the linear PC model

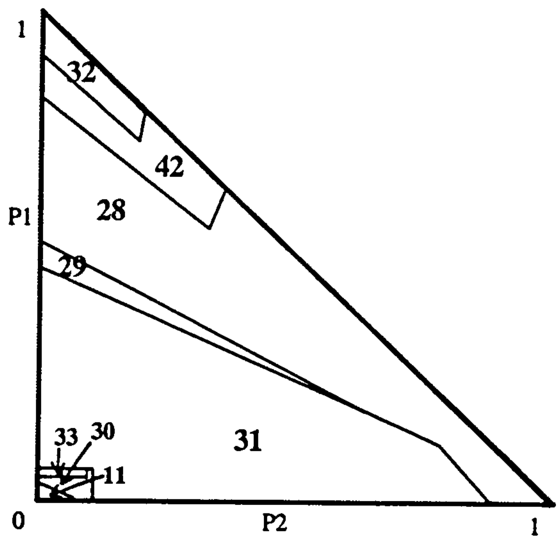


Figure 43 Best solutions for the RA model

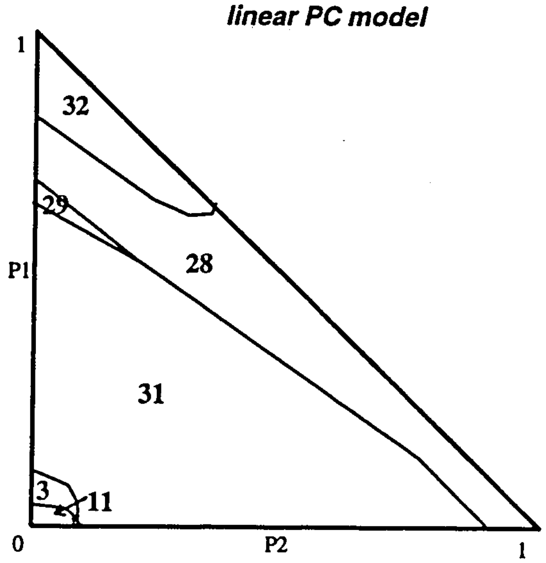


Figure 44 Best solutions for the Euclidean distance model

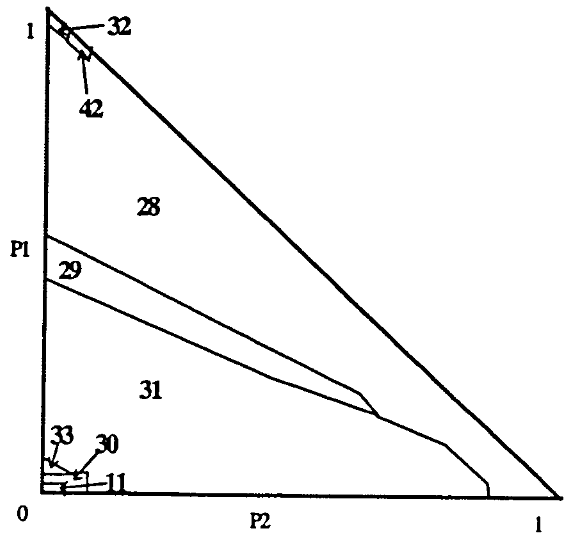


Figure 45 Best solutions for the non-linear regret PC model (threshold = 30)

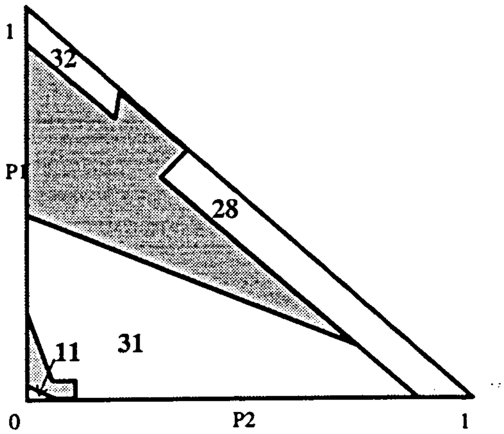


Figure 47 Stability areas

9.5.2. Decision by the fuzzy assessment of subjective probabilities

The interest of having this multi-criteria view of the problem can be appreciated in the very realistic case of having imprecise statements about the values of the future probabilities. These vague statements could have simply the form of declaration of intervals (for instance, "probability of future 1 will be between 0.3 and 0.5"), but we can accept more general declarations and translate them through a fuzzy representation (fuzzy probability, see chapter 4).

In the example we have been following, we defined the probability of each scenario through the declaration "the probability of future k is likely to be between 0.3 and 0.4, and is neither below 0.2 nor above 0.5". This may be represented by a trapezoidal fuzzy number, such as in Figure 47.

The imprecise future probability definitions open a "window" in the best solution diagrams; such a fuzzy window is superimposed both in Figure 48 and Figure 49 the inner triangle represents the points with membership 1 and the outer hexagon defines the limit where membership assumes value 0. Any point $P(w_1, w_2, w_3)$ receives a membership value

$$\mu(P) = \min \{\mu(w_k), k=1, 2, 3\} \tag{100}$$

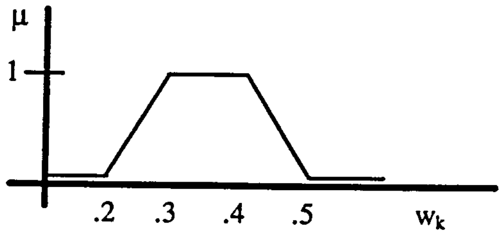


Figure 47 Trapezoidal fuzzy number describing an imprecise assessment of a future probability

The fuzzy window focuses our attention in a limited set of "best" solutions. If we were following the PC paradigm, the fuzzy window in Figure 48 just offers us Solution 31, regardless of the uncertainty in the probabilities. However, the image given by the fuzzy window in the RA results is more complex - 3 solutions are emerging as "best" within the range of uncertainties defined. How to select one?

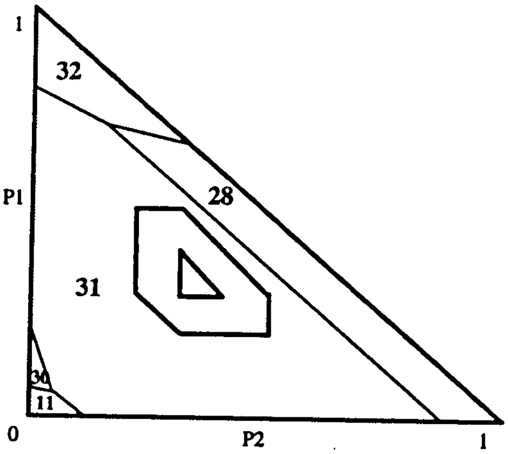


Figure 48 Fuzzy probability window superimposed on PC diagram

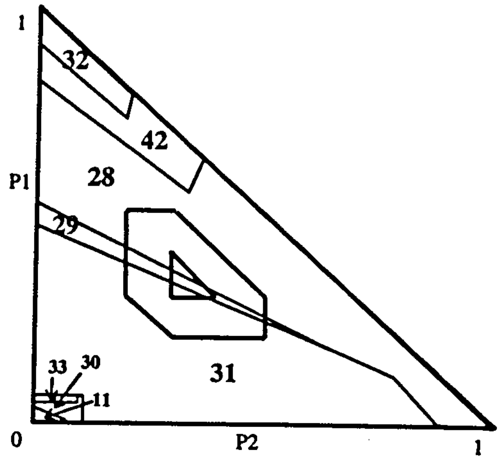


Figure 49 Fuzzy probability window superimposed on RA diagram

Matching the information on the regrets for solution and each point inside the window with the membership values distributed across the window, we can get a result depicted in Figure 50: a fuzzy evaluation of the regret associated with choosing each of the three solutions 28, 29 and 31 that can be found inside the window. For each solution, at a given α level, its fuzzy regret is given by an interval:

$$[\min \text{Regret}(\alpha), \max \text{Regret}(\alpha)] \quad (101)$$

Having calculated these intervals for several α levels and for the solutions in competition, we have the three fuzzy Regrets of Figure 50. These fuzzy values may be now ranked according to one of the usual criteria for defuzzification, such as the Center of Mass or the Removal. In this case, the result of the defuzzification using removal process gives:

$$29 < 28 < 31 \quad (102)$$

which means that solution 29 is preferred (and 28 preferred to 31), given the fuzzy assessment of the probability values for the three futures. It is interesting to notice, in comparison with a PC approach, that solution 31 is not even the second best.

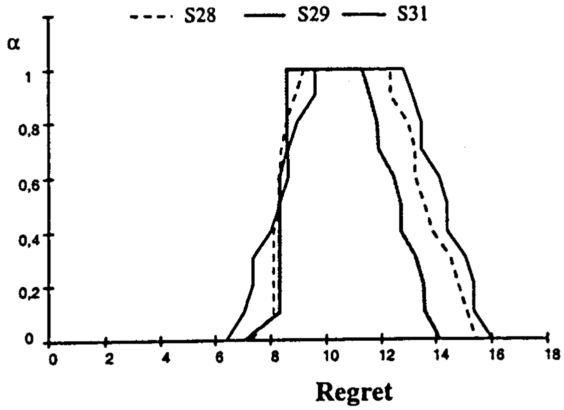


Figure 50 Fuzzy regrets in the fuzzy probability window

9.6. Conclusion

This chapter proposes a debate on the procedures and assumptions adopted in Power System planning. It demonstrates immediately that the discussion is not irrelevant, and that before applying one or the other of the two paradigms discussed, we should carefully examine the characteristics of the problem and of the data available. Having discussed the fundamentals of the Probabilistic Choice and the Risk Analysis paradigms, in system planning, and having analyzed the practical results obtained with the application of both models to a planning problem, it reached the following conclusions:

1. The assumptions behind the PC paradigm (that the repetition of events is achieved and that it has a large enough rate to give statistical assurance

that bad outcomes will be compensated by positive outcomes) are not verified in many cases - either because unwanted or catastrophic events may be foreseeable or because the planning time horizon is too short compared to the event repetition cycle.

2. The PC paradigm is blind to many compromise solutions, because it has a linear nature, in a space of scenarios.

3. The PC paradigm tends to recommend riskier solutions, namely giving extended preference to extreme alternatives and allowing the possibility for low probability futures with high regret values.

These conclusions substantiate the statement that *Risk Analysis* outperforms *Probabilistic Choice* as the effective decision making paradigm for Power System planning (in a general or typical case).

The conclusion must however be taken with the precaution of analyzing if the probabilistic model assumptions are fully verified in each case.

It was also demonstrated that taking in account non-linear regrets, within a PC environment, is not equivalent to a Risk Analysis approach. However, non-linear regrets may be seen as an intermediate or hybrid model between the PC and the RA paradigms. This may have some virtues in particular cases, which remain to be explored.

On another perspective, it was shown that, for a large set of probabilities allocated to scenarios, the two basic planning paradigms may in fact propose the same decisions as the best ones to be taken. The identification of these solution stability areas may be of great help in practical decision making, avoiding discussions about assumptions or paradigms and giving extra reassurance to the planners on how unquestionable their decisions may be.

Finally, It was shown that a vague or imprecise definition of the subjective probabilities allocated to each possible future is compatible with a decision making process. Working with fuzzy probability leads to the calculation of fuzzy regrets for the possible alternatives, and through a defuzzification process a ranking of alternatives can nevertheless be obtained.

10. CONCLUSIONS AND FUTURE WORK

The future is not what it used to be.

Paul Valery

10.1. Some reflections on the results

In a review of the paper (based on part of the work presented in this thesis) "Why Risk Analysis Outperforms Probabilistic Choice as the Effective Decision-Support Paradigm for Power System Planning" [Mir97b] the author came across this extraordinarily insightful excerpt by an anonymous referee:

> *"The authors' method, or any method, is a combination of technical compromises brought on by the limits of our numerical methods and the data we have available. More important though, is the conceptual basis for our decisions as planners. What are we really trying to accomplish? The authors make it clear that the avoidance of risk, in an organized manner, is at the heart of any uncertainty analysis."*

The message here is evident: no matter how fancy and complex our models and methods may be, the most important ingredient in planning, is the awareness of our goals, the perception of what we want to achieve. This is precisely what we mean by *the essence of planning.*

And that was exactly what this dissertation diagnosed as the basic cause for the manifest divorce between existing academic models and practical industrial applications: the fact that 30 years of models for distribution networks were not able to fully apprehend the basic nature of planning.

From this somehow provocative (and surely controversial) statement, the dissertation sets on to answer the following question: "*are we able, now, to develop a full comprehensive methodology for distribution system planning*"? This would mean developing a methodology that would embrace all the complex (both technical and instrumental) aspects of the problem and, at the same time, understand the real purposes of the people involved in the planning process.

The first and manifest problem is the need to deal with a tremendous amount of information. It was mentioned that, at the present time, this seems not to constitute a problem: we have the tools and we have the power. This power is based on vast computational resources, including parallel and distributed systems and modern geographical information systems for user interfaces. Finally, we have efficient and innovative algorithms that overcome some of the weaknesses of traditional algorithmic methodologies. Chapter 5 referred the use of Evolutionary Algorithms as robust search mechanisms, perfectly fitted to distribution system planning and well adapted to distributed implementations.

The thesis also makes the important recognition that Power System planning is a matter of *Decision-Making,* not *Optimization.* Therefore, it proposes the definitive abandon of *the Single Criterion/Optimal Solution* concept, replacing it by a fully multiobjective analysis. Once again, Evolutionary Algorithms and, in particular, genetic algorithms have been shown to be remarkably suited for the generation of multiple planning alternatives, enhancing the opportunity for multicriteria decision methods to be explicitly applied.

In addition, the dissertation offers a model for a broad representation of uncertainties where fuzzy sets, scenario trees and probabilistic models are employed according to the type of uncertainty being modeled. Thereafter, it presents the notions of hedging, robustness and exposure, simple but powerful concepts that turn up to be so familiar to the planner's experience and perspective.

The application of the proposed three-step methodology to a real-size distribution planning problem proved to be both suitable and resourceful, validating the relevance of the approach. It also established the importance of having powerful tools such as this in distribution planning. This is particularly correct when we are dealing with a large dimension problem, where a fast increase in load levels leads to a large number of possible investments and, consequently, to a large number of possible expansion strategies.

Finally, the thesis confirms the paradigm of Risk Analysis as the effective decision support paradigm for power system planning, in a rigorous comparison with the more traditional Probabilistic Choice paradigm, under a multiobjective approach to the scenario technique.

From all that was written, it seems that we have a positive answer to the question asked in the beginning of the dissertation: we do have the necessary power, tools and conceptual basis that will open the way for the development of planning aid tools for distribution design that represent reality and its constraints in a way closer to the point of view of decision makers and planners. Moreover, we have now the theoretic foundation to go one step beyond the simple portrayal of reality and model the decision maker himself and the way he reasons.

10.2. Future work

The subject of this dissertation was somehow extensive and certainly ambitious and there were a few aspects that were not covered with the thoroughness that could be expected. Still, this option was perfectly conscious and had the single objective of not obscuring the basic principles of the dissertation.

Among the aspects that could be further developed, the author would like to refer the use of Multicriteria Decision Methods (MCDM) in the decision process. It would be possible to include one or even two extra chapters on this subject. However, the author feels that the research on MCDM is well mature and that the inclusion of such chapters would, in a way, shift the focus and weaken the principle of the dissertation. This also holds for other questions that were not presented in depth in the dissertation pages.

On the other hand, this thesis opens the doors for some possible developments:

- Some research could be done in trying to apprehend the elementary concepts behind each type of problem in Power Systems (or even in other fields of engineering). This would involve typifying the different areas of planning according to the applicability of each decision making paradigm. Eventually, the results of these studies could be used in refining the models presently used and would unquestionably help to better understand the problems.

- Further investigate the virtues of Evolutionary Algorithms in Power System planning. This could be done by investigating on how to improve the algorithms, studying the possibility of using of hybrid systems (including Neural Networks or Classifier Systems) and finding new, efficient parallel or distributed implementations of Evolutionary Algorithms.

- Improve some technical aspects of the models, specifically regarding reliability and switching policies, dynamic security, capacitor bank planning, etc. The inclusion of subjective environmental and visual impact aspects in the models seems also to be relevant. The performance of some kind of sensitivity analysis (assessing the effects of small changes in the model on the final results) appears as fundamental in future developments.

- Perform some research on the planning procedures presently used by utilities in different regions and countries, in order to better understand the problems faced daily by planners and decision makers. The results of these studies would help improving our modeling of the decision making processes and even allow suggesting some possible improvements of these processes.

- Explore the exceptional possibilities opened by Geographical Information Systems in Power System spatial planning and the chance to integrate efficient search algorithms in their environment.

The previous list is far from being thorough. There is a great deal of other features that could be refined, a few additional paths to follow. Some of the referred aspects are already being studied or implemented both by utilities and research institutions, as there is still a lot to be achieved.

So, where do we go from now? It seems that the ultimate goal should be the integration of all these models and techniques in an inclusive tool aimed at having a wide acceptation in the industrial world.

But will we really have a universal tool for planning? The author believes that will not be the case since we'll have to integrate, not only the tools and the planning paradigms, but also the <u>cultural reality of each utility</u>. And this, above all, conditions the process of reasoning and behavior of decision makers.

Ultimately, any planning tool, sophisticated as it may be, will only have a decisive success when the planner feels that he has an efficient decision help instrument to deal with his practical problems, a tool that understands what planning is all about.

11. BIBLIOGRAPHY

[Ada73] R.N. Adams, M. A. Laughton, "A Dynamic Programming Network Flow Procedure for Distribution System Planning", Proceedings of the 8th PICA Conference, Minneapolis, MN, USA, June 1973.

[Ada74] R.N. Adams, M. A. Laughton, "Optimal Planning of Power Networks Using Mixed-Integer Programming Pt. 1 - Static and Time-Phased Network Synthesis", Proceedings IEEE 1974, 121, (2).

[Ada84] R.N. Adams, A. Afuso, A Rodriguez, V. Gerez, "A Methodology for Distribution System Planning", Proceedings of the 8th PSCC, Helsinki, Finland, Aug. 1984.

[Afu82] A. Afuso, V. Gerez, "An Integrated System for Distribution Planning", Proceedings of the 4th IEEE International Symposium on Large Engineering Systems Calgary, Alberta, Canada, Jun. 1982.

[Ajj91] V. Ajjarapu, Z. Albanna, "Application of Genetic Based Algorithms to Optimal Capacitor Placement", Proceedings First International Forum on Applications of Neural Networks to Power Systems, New York, NY, USA, IEEE 1991.

[All74] R. N. Allan, B. Borkowska, C. H. Grigg, "Probabilistic Analysis of Power Flows", Proceedings IEE, Vol. 123, No. 6, June 1976, pp. 531-536.

[All76] R. N. Allan, R. Billinton, M. F. Oliveira, "An Efficient Algorithm for Deducing the Minimal Cuts and the Reliability Indices of a General Network Configuration", IEEE Trans. on Reliability, R-25, 226-233, 1976.

[All76a] R. N. Allan, M. R. G. Al-Sharkarchi, "Probabilistic A/C. Load Flow Analysis", Proceedings IEE, vol. 124, No. 2, Feb 1976, pp. 154-160.

[All76b] R. N. Allan et al. "Probabilistic Power Flow Techniques Extended and Applied to Operational Decision Making", Proceedings IEE, Vol. 123, No. 12, Feb 1976, pp. 154-160.

[Aok90] K.Aoki, K. Nara, T. Satoh, M. Kitagawa, K. Yamanaka, "New Approximate Optimization Method for Distribution System Planning", IEEE Transactions on PWRS, vol. 5, no.1, Feb. 1990.

[Bac79] Y. Baklund, J. A. Bubenko, "Computer-Aided Distribution System Planning", Electrical Power & Energy Systems, vol.1, no.1, Apr. 1979.

[Bäc91] T. Bäck, F. Hoffmeister, H.-P. Schwefel, "A Survey of Evolution Strategies", Proceedings of the 4th International conference on Genetic Algorithms, 1991.

[Back96] Jürgen Backes, Hans-Jürgen Koglin, Lutwin Klein, "A Flexible Tool for Planning Transmission and Distribution Networks with Special Regard to Uncertain Reliability Criteria", 12th Power Systems Computation Conference, Dresden, Germany, Aug. 1996.

[Bak85] J. E. Baker, "Adaptive Selection Methods for Genetic Algorithms", Proceedings 1st Int. Conference on Genetic Algorithms, 1985.

[Bez81] J. C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms", Plenum press, New York, 1981.

[Bia95] S. Biato, G. C. Oliveira, "A Heuristic Procedure to Cope with Multi-Year Transmission Expansion Planning", SPT 1995.

[Bil84] R.Billinton, R.N.Allan, "Reliability Evaluation of Power Systems", Pitman Publishers, 1984.

[Bla96] M. Blanchard et al., "Experience with Optimization Software for Distribution System Planning", IEEE Winter Meeting, Jan 1996.

[Bla96] Marie Blanchard, Louis Delorme, Christian Simard, Yvon Nadeau, "Experience with Optimization Software for Distribution System Planning" IEEE Transaction in Power Systems, vol. 11, No 4, November 1996.

[Boc97] V. Bochkareva, "Development Planning of Electricity Distribution Network", Licentiate thesis, Royal Institute of Technology, Stockholm, Sweden, 1997.

[Bor74] B. Borkowska, "Probabilistic Load Flow", IEEE Trans. On PAS, Vol. Pas-93, No 3, May/June 1974, pp. 752-759.

[Bur85] T. Burkhardt, H.-J. Koglin, K. Werth, "Decision Making Including Forecast Uncertainties and Optimal Routing in Distribution Networks", Proceedings of CIRED'85, s.6, Brighton, U.K., Apr. 1985.

[Cai91] Cai Kai-Yuan et al., "Fuzzy Variables As a Basis for a Theory of Fuzzy Reliability in the Possibility Context", Fuzzy Sets and Systems 42 (1991) 145-172.

[Che96] R. Cherkaoui, A. Bart, A. J. Germond. "Optimal Configuration of Electrical Networks Using Heuristic Methods".

[Cle91] "Making Hard Decisions - an Introduction to Decision Analysis", PWS-Kent Publishing Company, Boston, 1991.

[Dal90] V. Dale, Z. Krisham, O. Pagle, "Dynamic Optimization of Power Networks", Riga, Zinatne, 1990.

[Del96] B. Delfino et al., "Black-Start and Restoration of a Part of the Italian HV Network: Modeling and Simulation of a Field Test", IEEE Transactions in Power systems, Vol. 11, August 1996, n. 3, pp. 1371-1379.

[Dja96] Djalma M. Falcão, "High Performance Computing in Power Systems Applications", Jornadas de Informática, 1996, Porto, Portugal.

[Dub87] D. Dubois, H. Prade, "Theorie Des Possibilités - Application à la Representation des Connaissances en Informatique", Edition 2, Masson, Paris, 1987.

[Ebe90] W. Ebeling, "Applications of Evolutionary Strategies", System Analysis Modeling Simulation, vol. 7, Issue 1, pp. 3-16, 1990.

[Faw83] T.H. Fawzi, K. F. Ali, S. M. El-Sobki, "A New Planning Model for Distribution Systems", IEEE Trans, 1983, PAS-102, (9).

[Fer96] Pedro Carvalho, L. Ferreira, Fernando Lobo, Luis Barruncho "Distribution Expansion Planning Under Uncertainty: a Hedging Algorithm in an Evolutionary Approach".

[Fog65] B. Fogel, "Evolving Artificial Intelligence", Doctoral Dissertation, University of California at S. Diego, 1965.

[Fog66] B. Fogel, A. J. Owens, M. J. Walsh, "Artificial Intelligence Through Simulated Evolution", New York, John Wiley & Sons, 1966.

[Fon93] C. M. Fonseca, Peter J. Fleming, "Genetic Algorithms for Multiobjective Optimization : Formulation, Discussion and Generalization", Proceedings of the 5th International conference on Genetic Algorithms, San Mateo, CA, USA, 1993.

[Fre89] H. Freund, G. Lotter, U. Kammerer, L. Klein, A. Brozio, Wellssow, "Utilization of a Graphic-Based Information System for Computer Aided Planning of Urban Medium Voltage Networks", CIRED 89, Brighton, England, April 1989.

[Gel89] C. W. Gellings, W. M. Smith, "Integrating Demand-side Management Into Utility Planning", Proceedings IEEE, vol. 77, n. 6, June 1989, pp. 908-918.

[Glo89] F. Glover, "TABU Search – Part I", ORSA J. computing, vol. 1, pp. 190-206, 1989.

[Glo90] F. Glover, "TABU Search – Part II", ORSA J. computing, vol. 2, pp. 4-32, 1990

[Gol89] D. E. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning", 1989, Addison-Wesley.

[Gol91a] D. E Goldberg, "A Note on Boltzman Tournament Selection for Genetic Algorithms and Population-Oriented Simulated Annealing", 1991, Complex Systems 3.

[Gol91b] D. E Goldberg, K. Deb, "A Comparative Analysis of Selection Schemes Used in Genetic Algorithms", Genetic Algorithms in Search, Optimization and Machine Learning Summer School, 1991, Stanford.

[Gon81] T. Gönen, B.L. Foote, "Distribution System Planning Using Mixed-Integer Programming", IEEE Proceedings, Generation Transmission & Distribution, 1981, 128, (2).

[Gon82] T. Gönen, B.L. Foote, "Mathematical Dynamic Optimization Model for Electrical Distribution System Planning", Electric Power & Energy Systems, 1982, 4, (2)

[Gon86] T. Gönen, I.J. Ramirez-Rosado, "Review of Distribution Planning Models: A Model for Optimal Multistage Planning", IEEE Proceedings, Vol. 133, Pt. C, No. 7, Nov. 1986.

[Gor93] B.G. Gorenstin, N.M. Campodónico, J.P. Costa, M.V.F. Pereira, 1993, "Power System Expansion Planning Under Uncertainty", 11th PSCC, Avignon, France, Set 1993.

[Gorg92] M. Gorges-Schleuter, "Comparison of Local Mating Strategies in Massively Parallel Genetic Algorithms", Parallel problem solving from nature, 2, North-Holland, 1992.

[Gre91] J. Grefenstette, "Conditions for Implicit Parallelism", Navy Center for Applied Research in Artificial Intelligence, Internal Report, 1991, Washington.

[Hin77] K.S. Hindi, A. Brameller, "Design of low-voltage distribution networks", Proceedings IEEE, 1977, 124, (1)

[Hlt76] S. B. Holt, D. M. Crawford "Distribution Substation Planning Using Optimization Methods", IEEE Tutorial Course, 1976.

[Hof91] Hoffmeister, T Bäck, "Genetic Algorithms and Evolution Strategies: Similarities and Differences", Parallel Problem solving from Nature, Proceedings of the first workshop PPSN I, Dortmund, October 1991, Berlin: Springer. 455-469.

[Hol75] J. H. Holland, " Adaptation in Natural and Artificial Systems", the University of Michigan Press, 1975 Ann Arbor.

[Hua96], Yann-Chang Huang, Hong-Tzer Yang, "Solving the Capacitor Placement Problem in a Radial Distribution System using TABU Search approach", IEEE transactions on Power Systems, vol. 11, No 4, November 1996.

[Iba93] K. Iba, "Reactive Power Optimization by Genetic Algorithm", PICA'93, Phoenix, AR, USA, May 1993.

[Iba93] K. Iba, "Reactive Power Optimization by Genetic Algorithm", PICA'93, Phoenix, AR, USA, May 1993.

[ICL97] http://www.icl.co.uk/business/utils/backup/utdinis.html, ICL information on DINIS.

[Jon96] S. Jonnavithula, R. Billinton, "Minimum Cost Analysis of Feeder Routing in Distribution System Planning", IEEE/PES Winter Meeting, Jan 1996.

[Jor85] Jorma Mörski, Jorma Seppälä, "Integrated Control and Planning of Several Voltage Levels with Special Reference to Electricity Supply Quality", Proceedings of CIRED'85, s.6, Brighton, U.K., Apr. 1985.

[Kad84] M.A el-Kady, "Computer Aided Planning of Distribution Substation and Primary Feeders", IEEE Transactions on PAS vol. 103, Jun. 1984.

[Kag90] N. Kagan, R.N. Adams, "Application of Benders Decomposition technique to the distribution planning problem", Proceedings of the 10th. PSCC, Graz, Austria, Aug. 1990, ed. Butterworths, London.

[Kag93] N. Kagan, R.N. Adams, "Electrical Power Distribution Systems Planning Using Fuzzy Mathematical Programming", Proceedings of the 11th PSCC, Avignon, France, 1993.

[Kau88] A. Kaufmann, M.M. Gupta, "Fuzzy Mathematical Models in Engineering and Management Sciences", North Holland ed., 1988.

[Kau93] K. Kauhaniemi, "Fuzzy Models and Techniques for the Calculation of Radial Distribution Networks", Proceedings of IEEE/NTUA Athens Power Tech Conference, Athens, Greece, September 1993, pp. 423-428.

[Kir93] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, "Optimization by Simulated Annealing" 1983 Science, Vol 220, number 4598.

[Kuw96] H. Kuwabara, K. Nara, "Multi-year and Multi-state Distribution System Expansion Planning by Multi-stage Branch Exchange", IEEE/PES Summer Meeting, Jul, 96.

[Lea95] Mª Teresa Leão, "Distribution Network Planning with Independent Production", Ph.D. thesis, FEUP, 1995.

[Lei90] A. M. Leite da Silva, V. L. Arienti, "Probabilistic Load Flow by a Multilinear Simulation Algorithm", proceedings IEE, vol. 137, Pt. C, No. 4, July 1990, pp. 276-282.

[Lop95] J. A. Peças Lopes, J. V. Ranito, J. Neto, N. Hatziargyriou, "Derivation of Classification Structures for Fast Evaluation of Dynamic Security Assessment in Power Systems Using Genetic Algorithms", Proceedings of IEEE/KTH Stockholm Power Tech Conference, Stockholm, Sweden, June 18-22, 1995, pp. 641-646.

[Mas74] E. Masud, "An Interactive Procedure for Sizing and Timing Distribution Substations Using Optimization Techniques", IEEE Transactions Vol. PAS-93, September 1974, pp. 230-235.

[Mas78] E. Masud, "Distribution Planning: State-of-the-art and extensions to substation sizing", Electrical Power systems Review, 1978, 1, pp. 203-212.

[Mat88] M. Matos, "Multicriteria Decision Aid - New Contributions", (in Portuguese), Ph.D. thesis, FEUP, Porto, 1988.

[Mat92] M. Matos, L. M. Proença, "Successive Amplification Method", (in Portuguese), Investigação Operacional, vol. 12., June, 1992.

[Mer90] H. Merril, A.J. Wood, "Risk and Uncertainty in Power System Planning", 10th Power System Computation Conference - PSCC, Graz, Austria, Aug 90.

[Mil56] G. A. Miller, "The magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity For Processing Information", Psychological Rev. 63 n. 2, 1956.

[Mir83] V. Miranda, F. M. Barbosa, "Three Phase Load Flow for Radial Networks", Proceedings of MELECON'83, report D4.07, Athens, Greece, 1983.

[Mir89] V. Miranda, M. A. Matos, "Distribution System Planning with Fuzzy Models and Techniques", Proceedings of CIRED'89 (IEE/AIM), Brighton, UK, April 1989.

[Mir90a] V. Miranda, "Reliability Calculations in Power Systems with Fuzzy Indices" (in portuguese), 1. Jornadas Hispano-Lusas de Ingeneria Electrica, Vigo, Spain, July 1990.

[Mir90b] V. Miranda, M.A. Matos, J.T. Saraiva, "Fuzzy Load Flow - New Algorithms Incorporating Uncertain Generation and Load Representation", 10th PSCC, Graz, August 1990; in Proceedings of the 10th PSCC, Butterworths, London, 1990.

[Mir91a] V. Miranda, "Using Fuzzy Reliability Indices in a Decision Aid Environment for Establishing An Interconnection and Switching Location Policy in Distribution Networks", Proceedings of CIRED'91 (AIM/IEE), Liège, Belgium, April 1991.

[Mir91b] V. Miranda, L.M.V.G. Pinto, "A Model to Integrate Uncertainties in Power System Operation" (in Portuguese), Proceedings of SNPTEE, Rio de Janeiro, Brazil, 1991.

[Mir91c] V. Miranda, "Using Fuzzy Reliability Indices in a Decision Aid Environment for Establishing Interconnection and Switching Location Policies", Proceedings CIRED'91, s.6, Liege, Belgium, Apr. 1991.

[Mir92a] V. Miranda, M. A. Matos, J. T. Saraiva, "Fuzzy Flows in Linear Networks with Correction for Some Dependencies in Nodal Uncertainties", APDIO'92 Congress, Évora, Portugal, May 1992.

[Mir92b] V. Miranda, J. T. Saraiva, "Fuzzy Modeling of Power System Optimal Load Flow", PICA'91, in IEEE Transactions on Power Systems, May 92.

[Mir94a] V. Miranda, "Power System Planning and Fuzzy Sets: Towards a Comprehensive Model Including all Types of Uncertainties", Proceedings of PMAPS'94, Rio de Janeiro, Brazil, September 1994.

[Mir94b] V. Miranda, L. M. Proença, J. V. Ranito, "Genetic Algorithms and Fuzzy Models - An Application to Gas and Electricity Distribution Planning Under Uncertainty", 3rd Int. Workshop on Rough Sets and Soft Computing (RSSC'94), San Jose, Cal, USA, Nov. 1994.

[Mir94c] V. Miranda, J. V. Ranito, L. M. Proença, "Genetic Algorithms in Optimal Multistage Distribution Network Planning", IEEE Transactions on Power Systems, Vol. 9, No. 4, November 1994, pp. 1927-1933.

[Mir95] V. Miranda, L. M. Proença, "A General Methodology for Distribution Planning Under Uncertainty, Including Genetic Algorithms and Fuzzy Models in a Multi-criteria Environment", Proceedings of Stockholm Power Tech, SPT'95, Stockholm, Sweden, June 18-22, 1995, pp. 832-837.

[Mir96a] V. Miranda, D. Srinivasan, L. M. Proença, "Evolutionary Computation in Power Systems", Proceedings of PSCC'96, August 1996, Dresden, Germany.

[Mir96b] V. Miranda, "Fuzzy Reliability Assessment in Power Systems" PSCC'96, Dresden, Germany.

[Mir96c] V. Miranda, L. M. Proença, "Dynamic Planning of Distribution Networks Including Dispersed Generation", Proceedings of CIRED'96, 2-5 Dec. 1996, Buenos Aires, Argentina.

[Mir97a] V. Miranda, L. Proença "Probabilistic Choice Vs. Risk Analysis - Conflicts and Synthesis in Power System Planning", Proceedings of PICA'97, May 1997, Columbus, Ohio, USA.

[Mir97b] V. Miranda, L. Proença "Why Risk Analysis Outperforms Probabilistic Choice As the Effective Decision-Support Paradigm for Power System Planning", Proceedings of SPM'97, July 1997, Berlin, Germany, to be published in IEEE transactions on Power Systems.

[Mom95] J. A. Momoh, X. W. Ma, K. Tomsovic, "Overview and Literature Survey of Fuzzy Set Theory in Power Systems", IEEE Trans. Power Systems, Aug. 95.

[Mon96a] C. Monteiro, "PROJECT SOLARGIS - Methodology and Application in Cape Verde" (in Portuguese), Proceedings of ELAB'96, Porto, Portugal.

[Mon96b] C. Monteiro, "Integration of Renewable Forms of Energy in the Decentralized Production of Electricity Using GIS" (in Portuguese), MSc thesis, October 1996, Faculdade de Engenharia da Universidade do Porto, Portugal.

[Nar91] K.Nara, T. Satoh, K. Aoki, M. Kitagawa, "Multi-Year Expansion Planning for Distribution Systems", IEEE Transactions on PWRS, vol 6, no.3, Aug. 1991.

[Oli79] M.F. Oliveira, V. Miranda, "Etudes D'optimisation Dans Les Réseaux De Distribution Comprenant Des Calculs De Fiabilité", Proceedings of CIRED'79, s.6, Liège, Belgium, Apr. 1979.

[Pam95] N. Pamudji, R. J. Kaye, H. R. Outhred, "Network Effects in a Competitive Electricity Industry: Non Linear and Linear Nodal Auction Models", Proceedings of Stockholm Power Tech, SPT'95, Stockholm, Sweden, June 18-22, 1995, pp. 698-673.

[Par90] J. Partanen "A Modified Dynamic Programming Algorithm for Sizing, Locating and Timing of Feeder Reinforcements", IEEE Transactions on Power Delivery, Vol 5. No 1, January 1990.

[Pro93] Luís Miguel Proença, "Genetic algorithms in Distribution Expansion Planning", (in Portuguese) MSc thesis, Faculdade de Engenharia da Universidade do Porto, October, 1993.

[Pro97] Luís Miguel Proença, Vladimiro Miranda, "Coordenação Hidrotérmica Mediante a Utilização De Programação Evolucionária", 5as jornadas Luso-Espanholas, Salamanca, Espanha, Julho 1997.

[Ran93] J. V. Ranito, "Algoritmos Genéticos na Optimização de Perguntas Relacionais", Tese de mestrado, Faculdade de Engenharia da Universidade do Porto, Outubro 1993.

[Saa96] J. J. Saade, "A Unifying Approach to Deffuzification and Comparison of the Outputs of Fuzzy Controllers", IEEE Transactions on Power Systems, vol. 4, No 3, August 1996.

[Sar91a] J. T. Saraiva, V. Miranda, M. A. Matos, "Generation and Load Uncertainties Incorporated in Load Flow Studies", MELECON' 91, Ljubljana, Slovenia, May 1991.

[Sar91b] J. T. Saraiva, V. Miranda, "Analysis of the Results of a Fuzzy Power Flow Model" (in Portuguese), 2nd Portuguese-Spanish Meeting in Electrical Engineering, Coimbra, Portugal, July 1991.

[Sar92] J. T. Saraiva, "Application of Fuzzy Sets in Power System Modeling and Operation" (in Portuguese), Ph.D. Thesis, Porto, Portugal, Mar 1992.

[Sar94] J. T. Saraiva, V. Miranda, L. M. V. G. Pinto, "Impact On Some Planning Decisions From a Fuzzy Modeling of Power Systems", PICA May 93, in IEEE Trans. on Power Systems, May 1994.

[Sar95] J.T.Saraiva, V.Miranda, L.M.V.G.Pinto, "Generation-Transmission Power System Reliability Evaluation By Monte Carlo Simulation Assuming a Fuzzy Load Description", PICA May 95, to be published in IEEE Trans. on Power Systems.

[Sar96] J.T.Saraiva, "Some Reflections On Electrical Energy System Planning and Operation in a Deregulated Environment", (in Portuguese), proceedings of ELAB'96, Porto, Portugal.

[Sei94] T. Seitz, "Reliability Analysis of Electricity Distribution Systems with Decentralized Generation with Contributions From Fuzzy Set Theory" (in German), PhD. thesis, RW Tech. Hochschule Aachen, Germany, Jan 94.

[Spe91] W. M. Spears, "An Analysis of Multi-Point Crossover", Foundations of genetic algorithms, G. J. E. Rawlins (Ed.), San Mateo, CA : Morgan Kaufmann.

[Sul77] R. L. Sullivan, "Power System Planning", McGraw-Hill Book Co. New york, 1977.

[Sun82] D. I. Sun, D. R. Farris, P. J. Cote, R. R. Shoults, M. S Chen, "Optimal Distribution Substation and Primary Feeder Planning Via the Fixed Charge Network Formulation", IEEE Transactions on PAS-101 (3) 1982

[Swa96] K. S. Swarup, M. Yoshimi, S. Shimano, Y. Izui, "Optimization Methods Using Genetic Algorithms for Reactive Power Planning in Power Systems", Proceedings of PSCC'96, August 1996, Dresden, Germany.

[Tom92] K. Tomsovic, "A Fuzzy Linear Programming Approach to the Reactive Power/Voltage Control Problem", IEEE Trans. PWRS, vol. PWRS-7, no. 1, January 1992.

[Wan96] Lou Chin Wang, L. M. Proença, Vladimiro Miranda, "Demonstrating an Efficient Capacitor Location and Sizing Method for Distribution Systems - Application to the Macau Network", Proceedings of ELAB'96, Porto, Portugal.

[Whi91] D. Whitle, T. Starkweather, D. Shaner, "The Traveling Salesman and Sequence Scheduling : Quality Solutions Using Genetic Edge Recombination", Handbook of genetic algorithms, van Nostrand Rheinhold, NY, 1991.

[Wil96] H.L. Willis, "Spatial Electric Load Forecasting" (book), ed. Marcel Dekker Inc., NY, 1996.

[Wil97] H.L. Willis, "Power Distribution Planning Reference Book", ed. Marcel Dekker Inc., NY, 1997.

[Yau90] T. S. Yau, R. G. Huff, H. L. Willis, W. M. Smith, L. J. Vogt, "Demand-side Management Impact on the Transmission and Distribution System," IEEE Transactions on Power Systems, May 1990, p. 506.

[Zel82] M. Zeleny, "Multiple Criteria Decision Making", McGraw Hill ed., 1982.

[Zim85] H.-J. Zimmermann, "Fuzzy Set Theory and its Applications", Kluwer-Nijhoff Publishing, Boston, 1985.

APPENDIX A - Distribution system representation

The graphical representation of a distribution system is based on a network of nodes and branches. Nodes represent locations for primary substations or load nodes. Branches represent distribution lines, that may be of two different types:

- Existing lines
- Lines in Project

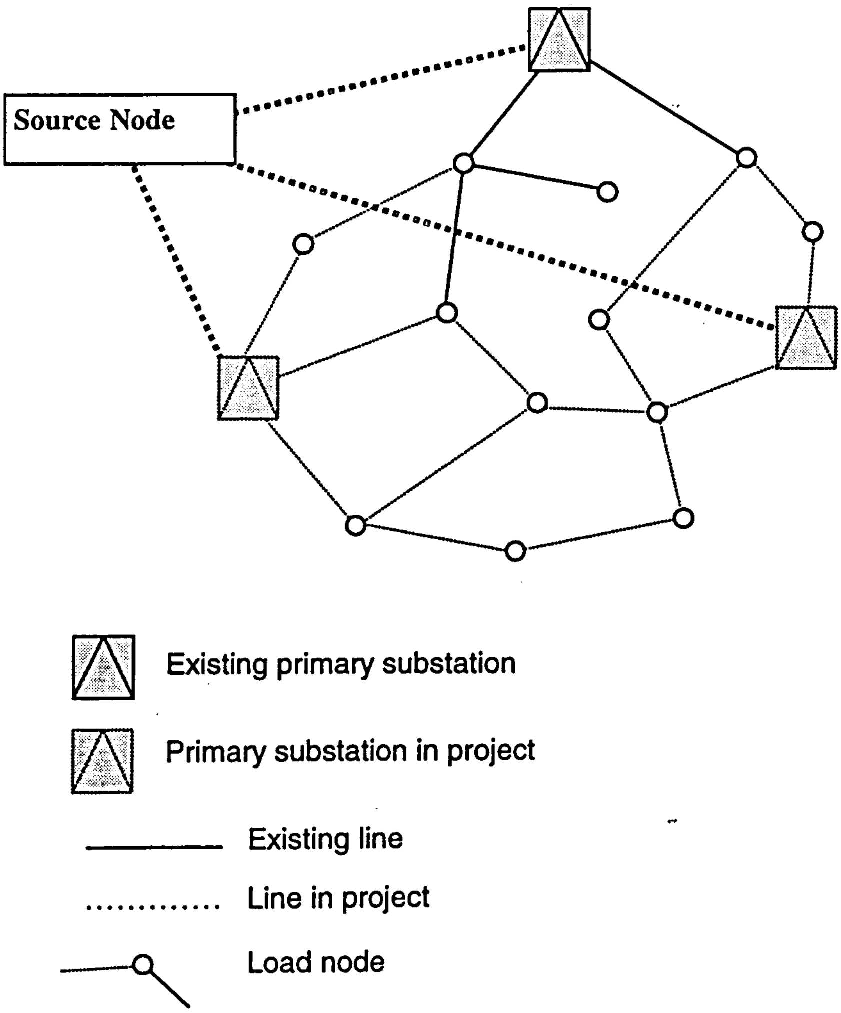


Figure 51 Graphical representation of a distribution system

Power flow in branches is permitted in any sense except between the source node and the primary substations. Therefore, as an oriented graph, lines are considered to be two arches in opposite directions.

Primary substations are viewed as branches between the node where they are located and the source node. The nodes where a substation is located may also be considered as load nodes.

This representation is illustrated by Figure 51.

The representation of the fictitious source node will be ignored from now on. The following points refer some aspects related to the representation of the distribution system.

Decision Variables

$F_{is} = 1$	If project line **i** is used in stage **s**
$F_{is} = 0$	Otherwise
$S_{is} = 1$	If substation **i** is used in stage **s**
$S_{is} = 0$	Otherwise
$E_{is} = 1$	If substation reinforcement **i** is used in stage **s.**
$E_{is} = 0$	Otherwise

Some important considerations:

- These variables define the network for the whole planning horizon.
- The algorithm considers that a line is built for the first stage it is used, unless it was built together with another item in order to reduce costs (for example, putting two lines in the same trench).
- A line may exist without being used. Radiality conditions are imposed to lines that are used in the operation of the system on a given stage.

Primary substations

Substations are considered to be special lines between the node they are located on and the source node. The cost for these lines is the investment cost for the substation referred to the stage where it is first used. The capacity of the equivalent line is the same as the capacity of the substation (including reinforcements). The same holds for reliability parameters.

Lines in Project

Lines in project may be used in a given stage. These lines correspond to a given project. If different capacities for the lines can be considered (a not common situation in distribution networks), then we would have to consider **k** decision variables for that line, being **k** the number of possible capacities for the line.

Lines in the same trench

The developed model may also consider the economical benefits of the simultaneous construction of two items, such as installing two lines in the same trench in a given stage.

Generally, a condition such as this one is very hard to include in a traditional mathematical programming model. However, the use of Genetic algorithms makes this issue rather easy to model.

Existing lines

Existing lines (lines in the initial system) may be dealt with in two different manners, accordingly to the general philosophy defined by the user.

a) The planner may accept that lines in the initial system will necessarily be used in the subsequent stages. In this case, we will not have decision variables related to these lines. The solution for all the stages in the planning horizon will always include these lines.

b) The planner may accept the declassifying (and eventual reuse) of lines in the initial system. In this case these lines will be viewed exactly as lines in project with null costs. This possibility increases the complexity of the problem, by increasing the number of decision variables.

The formulation of the model allows any of these two possibilities to be used. However, in some cases, one will necessarily have to use the possibility of declassifying existing lines in order to reach a feasible solution, as one can see in the following example:

Example:

In this example, the substation S2 has to be built (Figure 52). However, one of the existing lines **{L1,L2,L3,L4}** has to be declassified, in order to guarantee radiality. For example, the declassification of line L3 would lead to the radial system shown in Figure 53.

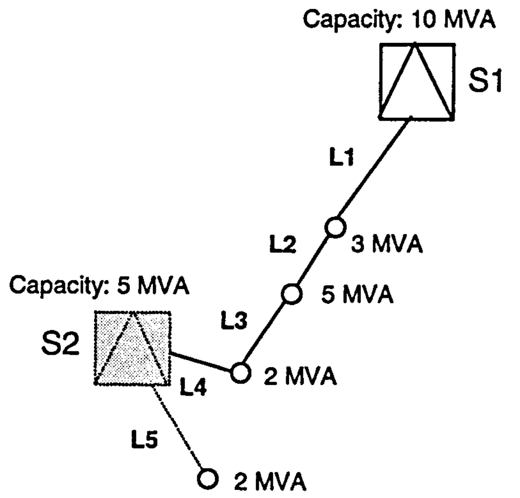


Figure 52 Substation S2 has to be put into service. One of the existing lines will have to be declassified

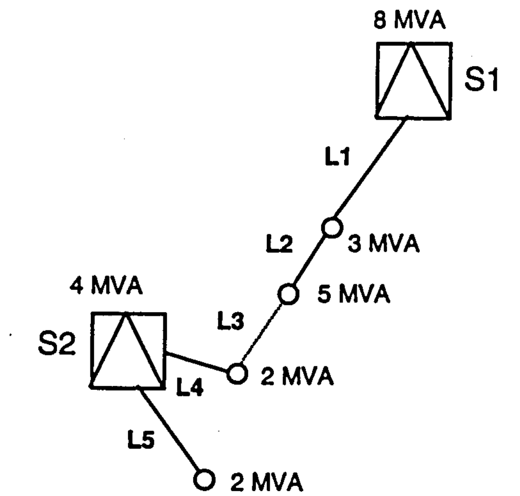


Figure 53 Radial network after declassifying L3

Reinforcements in substations

The model allows the consideration of projects of reinforcements in the capacity of the substations. A decision variable E_{is} will correspond to each one of these projects. If, for a particular substation we consider **f** different projects, then the model should include the corresponding **f** variables E_{is}.

APPENDIX B - Chromosome coding

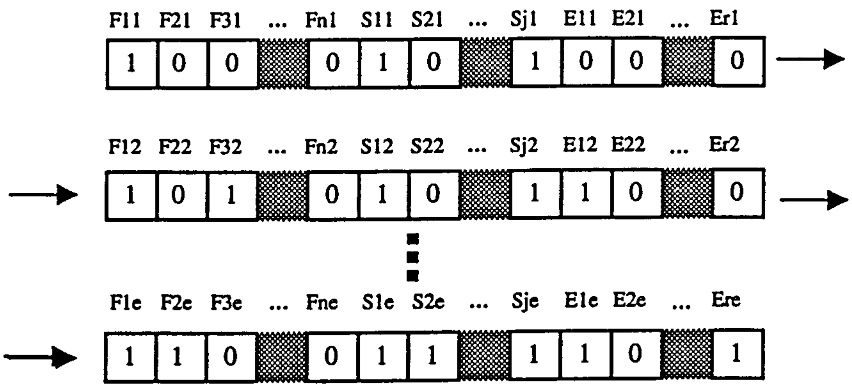
The use of genetic algorithms for the search of expansion plans for the distribution system imposes the use of a coding process for the distribution network for each of the stages within the planning horizon. Thus, each network expansion plan will have to be codified in a chromosome.

<u>Direct codification</u>

The simplest process is the direct codifying of the several decision variables:

F_{is}, S_{is}, E_{is}

For each stage, a certain bit of the chromosome would represent one of the binary decision variables. The following figure shows an example of direct codification.



n number of lines

j number of substations

r number of possible reinforcements in substation capacity

e number of stages

Figure 54 Example of direct codification in the case of binary decision variables

Different capacities for lines would need extra bits in the representation.

This type of coding has the following advantages:

- The representation is direct and therefore there is no need for decoding, leading to lower computation times for the fitness function.

- There is a direct "physical" relation between the topology of the network and the chromosome. Using an adequate mapping of the network into the chromosome, may result in that a certain block of the chromosome would correspond to a given region in the network.

However, this type of representation will lead to an exaggerated percentage of topologically unfeasible solutions. In a real-size network, there would be too many solutions with not connected load nodes, and other solutions would have an excessive number of lines. This factor renders this type of codification rather inefficient.

Improved coding process

The improved coding process appeared from the need of having a technique with the following basic characteristics:

- It should lead to a minimum number of topologically unfeasible solutions.
- The decoding process should be easy and fast.
- It should not lead to exaggeratedly large chromosomes.
- It should allow that all the feasible solutions have a corresponding representation in a chromosome. The coding should also not lead to an excessive bias towards a single solution or group of solutions.

The coding process adopted has all the characteristics mentioned above and is based on some specific properties of the problem:

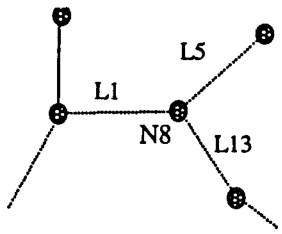
- Each node with an attributed load in a given stage has, necessarily, to be connected to a branch.
- A node with no load attributed in a given stage, may not have connections or, alternatively, be connected to two or more lines in order to allow the flow of load towards other nodes in the network.

- A radial network with n nodes has a number of lines **l < n** .
- A substation may be considered as a line between the source node and the node where the substation is placed. Form the coding point of view, a substation may be considered exactly as a line.

From these factors, a coding process was developed, and it may be summarized in the following steps:

From an initial system, for each node not connected to an existing line, we establish a list of project lines that are connected to it.

<u>Example</u>



For N8 we establish the following list:

N8(L1,L5,L13)

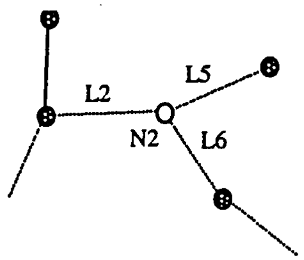
Obviously, at least one of this lines should exist, since node N8 has load attributed.

---

2. For a node with no load attributed and not connected to an existing line, the process for obtaining the list is identical. However, we add to the list the possibility of not being connected (***ISOLATED***).

This technique allows that, in the solution, a node with no load can be connected to two or more lines, if necessary.

<u>Example</u>



For N2 the algorithm establishes the list

N2(L2,L5,L6,ISOLATED).

In this case, N2 may be connected to one of the lines in the list or simply disconnected from the network.

---

4. In the end of the process, we will have a set of lists, corresponding to each one of the nodes not connected to the existing system. This set is obtained before the GA starts and it is kept in the computer memory in order to facilitate the decoding process in the fitness function.

The coding will determine for each one of this nodes a single entry in the corresponding list. For that, being t the number of entries in the list, we may calculate the number of necessary bits to codify an entry in each one of the lists:

$$b = \log_2(t)$$

(rounded to the nearest integer higher than b)

For example, if the number of entries in a given list is t=3 or t=4, the number of bits necessary to codify this list is b=2.

<u>Example</u>

Let us suppose the following list:

N3(L2,L4,L5,ISOLATED)

The number of entries in the list is **t=4**, and therefore, the number of bits necessary to codify the list is:

$$b = \log2(4) = 2$$

Each group of bits would correspond to the following situation for N3:

(0,0) Connected to L2
(0,1) Connected to L4
(1,0) Connected to L5
(1,1) ISOLATED

Let us see another situation:

N8(L3,L6,L8)

(b = 2)

The possible situations for N8 are:

(0,0) Connected to L3

(0,1) Connected to L6

(1,0) Connected to L8

(1,1) UNDEFINED

The case **(1,1)** is, therefore, undefined. In order to avoid that a case such as this would be considered unfeasible, undefined situations will correspond to defined situations, by the order initially established:

(1,1) - Connected to L3

One could think that there would be a bias toward to the construction of L3. However this bias is negligible in terms of the GA, since a week solution will always have a small fitness value. Conversely, if this combination of bits corresponds to a good solution, it will thrive in the GA.

This process guarantees that all the topologically feasible configurations can be codified in a subchromosome and, at the same time, assures that the total number of lines will never exceed the number of nodes. On the other hand, allowing a variable number of lines, the coding algorithm contemplates the possibility that a node with no load attributed is connected (or not) to a line. This is an important characteristic of this model.

5. Finally, the variables corresponding to reinforcements in substation capacity (E) are codified directly in the chromosome.

The following example shows the full decoding process for a single stage.

<u>Example:</u>

The following figure represents a small distribution network and, on the right, its equivalent from the coding point of view, for a given stage. Line L8 was eliminated in the preprocessing stage, since N6 has no load attributed in this stage. S1 and S2 have load attributed. At last, there is a reinforcement project for substation S1 corresponding to variable E_{s1}.

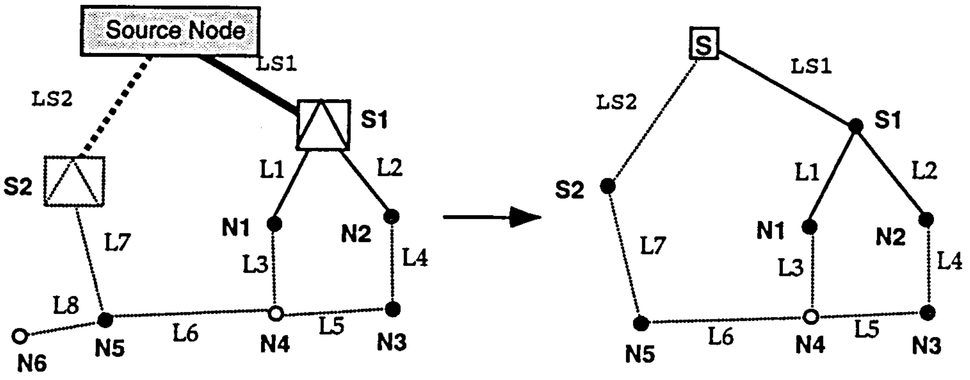


Figure 55 Example of a distribution network and its coding equivalent

The first step is to establish the list of lines for each node in the network not connected to the existing system, and the total number of bits needed to codify the network in a single stage.

N3	**(L4,L5)**	**- 1 bit**
N4	**(L3,L5,L6,ISOLATED)**	**- 2 bits**
N5	**(L6,L7)**	**- 1 bit**
S2	**(L7,LS2)**	**- 1 bit**
E_{s1}		**- 1 bit**
Total		**- 6 bits**

Therefore, for a multitemporal problem with **e** stages, we would need **6.e** bits. Figure 56 represents a possible subchromosome for a given stage.

N3	N4		N5	S2	ES1
0	0	0	0	1	0

Figure 56 Example of a 6 bit chromosome

This chromosome corresponds to the construction of the following lines:

L4 (from N3) ;

L3 (from N4) ;

L6 (from N5);

LS2 (from S2)

The network corresponding to this chromosome is shown in Figure 57.

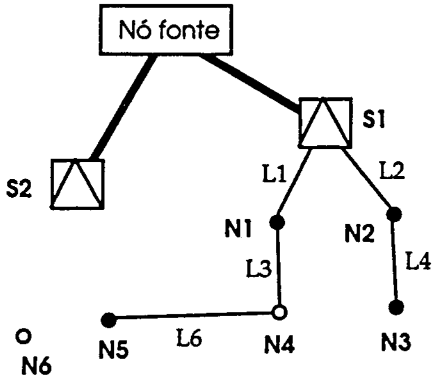


Figure 57 Network corresponding to the chromosome in Figure 56.

As it is possible to verify, this chromosome corresponds to a topologically feasible solution. It is also noticeable that node N4, even if it has no load attributed, is connected to two lines in order to allow power flow to N5.

However, the chromosome in Figure 58 does not correspond to a feasible solution as it possible to see in Figure 59.

N3 N4 N5 S2 ES1

1	1	0	1	1	0

Figure 58 Another possible chromosome

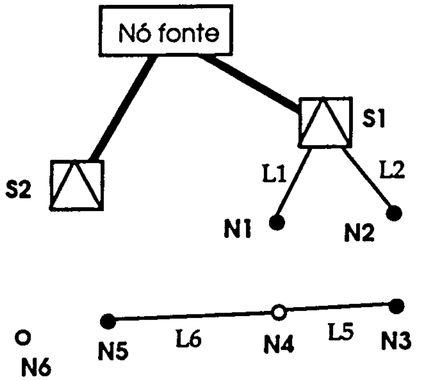


Figure 59 Unfeasible network

The final chromosome will be obtained by the concatenation of the subchromosomes corresponding to each stage inside the planning horizon. Therefore, this chromosome will codify the whole expansion plan. The coding adopted in this model has several positive aspects in it. However, there has been some research on more efficient coding algorithms.

FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO
BIBLIOTECA