ISSN 1650-674x
TRITA-ETS-2001-01

KUNGL TEKNISKA HÖGSKOLAN

ON DEVELOPMENT PLANNING OF ELECTRICITY DISTRIBUTION NETWORKS

Viktoria Neimane

Doctoral Dissertation
Royal Institute of Technology
Department of Electrical Engineering
Electric Power Systems
Stockholm 2001

ISSN 1650-674x
TRITA-ETS-2001-01

KUNGL
TEKNISKA
HÖGSKOLAN

ON DEVELOPMENT PLANNING OF ELECTRICITY DISTRIBUTION NETWORKS

Viktoria Neimane

Doctoral Dissertation
Royal Institute of Technology
Department of Electrical Engineering
Electric Power Systems
Stockholm 2001

Akademisk avhandling som med tillstånd av Kungl Tekniska Högskolan framlägges till offentlig granskning för avläggande av teknisk doktorsexamen torsdagen den 29 november 2001 kl 10.00 i Kollegiesalen, Kungl Tekniska Högskolan, Valhallavägen 79, Stockholm.

Abstract

Future development of electric power systems must pursue a number of different goals. The power system should be economically efficient, it should provide reliable energy supply and should not damage the environment. At the same time, operation and development of the system is influenced by a variety of uncertain and random factors. The planner attempts to find the best strategy from a large number of possible alternatives. Thus, the complexity of the problems related to power systems planning is mainly caused by presence of multiple objectives, uncertain information and large number of variables. This dissertation is devoted to consideration of the methods for development planning of a certain subsystem, i.e. the distribution network.

The dissertation first tries to formulate the network planning problem in general form in terms of Bayesian Decision Theory. However, the difficulties associated with formulation of the utility functions make it almost impossible to apply the Bayesian approach directly. Moreover, when approaching the problem applying different methods it is important to consider the concave character of the utility function. This consideration directly leads to the multi-criteria formulation of the problem, since the decision is motivated not only by the expected value of revenues (or losses), but also by the associated risks. The conclusion is made that the difficulties caused by the tremendous complexity of the problem can be overcome either by introducing a number of simplifications, leading to the considerable loss in precision or applying methods based on modifications of Monte-Carlo or fuzzy arithmetic and Genetic Algorithms (GA), or Dynamic Programming (DP).

In presence of uncertainty the planner aims at finding robust and flexible plans to reduce the risk of considerable losses. Several measures of risk are discussed. It is shown that measuring risk by regret may lead to risky solutions, therefore an alternative measure – Expected Maximum Value – is suggested. The general future model, called fuzzy-probabilistic tree of futures, integrates all classes of uncertain parameters (probabilistic, fuzzy and truly uncertain).

The suggested network planning software incorporates three efficient applications of GA. The first algorithm searches simultaneously for the whole set of Pareto optimal solutions. The hybrid GA/DP approach benefits from the global optimization properties of GA and local search by DP resulting in original algorithm with improved convergence properties. Finally, the Stochastic GA can cope with noisy objective functions.

Finally, two real distribution network planning projects dealing with primary distribution network in the large city and secondary network in the rural area are studied.

Acknowledgments

This document constitutes my doctoral dissertation, which covers work I have carried out at the division of Electric Power Systems, Royal Institute of Technology since 1995.

First of all I want to express my gratitude and appreciation to Professor Göran Anderson for his guidance and support throughout this project.

My deepest respect and gratitude I want to express to Professor Emeritus Janis Bubenko for his incredible support, encouragement and care at the initial stages of this project.

Many thanks to the staff of Electric Power Systems for providing stimulating and friendly ambiance for study and research and help in different aspects. My special and warmest thanks to the former secretary of the division Mrs. Lillemor Hyllengren for her assistance and concern. I would like to express my gratitude to the present secretary Mrs. Margaretha Surjadi for her help. I am very grateful to Valery Knyazkin for his help with numerious practical aspects. Thanks to Lina Bertling and Ying He for being a nice company on several conferences.

I would like to thank Mr. Daniel Terranova and Mr. Kjell Gustafsson from Birka Nät AB for their cooperation during elaboration of network planning project and sharing the experience concerning planning and operation of distribution networks.

Sincere thanks to the staff of the IDEA group at INESC, Porto, Portugal, where I spent six unforgettable weeks in June 1995. My special thanks and appreciation to Miguel Proenca for our fruitful discussions concerning distribution planning and especially application of Evolutionary Algorithms and Fuzzy Sets to power systems.

I am indebted to Professor Zigurd Krishans from the Institute of Physical Energetics of Latvian Academy of Science for sharing his experience in the field of power system planning. I am deeply grateful to the head of Strategic Planning department at Latvenergo Mr. Arturs Kruskops for his help in different planning aspects.

Financial support from KTH Faculty Board Foundation and Center of Competence in Electric Power Engineering at KTH is gratefully acknowledged.

And at last, but not at least I owe the most sincere gratitude to my wonderful family for their incredible support and encouragement.

Viktoria Neimane

Stockholm - Riga

May 2001

Contents

List of Acronyms

AC	Alternating Current
CAIDI	Customer Average Interruption Duration Index
CIC	Customer Interruption Cost
CLD	Current Limiting Devices
COG	Center of Gravity
CS	Classifier Systems
DA	Distribution Automation
DC	Direct Current
DLC	Distribution line carrier
DMS	Distribution Management System
DP	Dynamic Programming
DSM	Demand Side Management
EA	Evolutionary Algorithms
ENS	Energy Not Supplied
EMV	Expected Maximum Value
EP	Evolutionary Programming
ES	Evolution Strategies
GIS	Geographical Information Systems
GA	Genetic Algorithms
GP	Genetic Programming
MCDM	Multi-Criteria Decision Making
MOGA	Multi-Objective Genetic Algorithm
MOM	Mean of Maxima
O&M	Operation and maintenance
PCA	Principal Component Analysis
PDF	Probability Density Function
PLF	Probabilistic Load Flow
SAIDI	System Average Interruption Duration Index
SAIFI	System Average Interruption Frequency Index
SCADA	System for Control And Data Acquisition
TDC	Total Distance Criterion
VaR	Value-at-Risk
VEGA	Vector Evaluated Genetic Algorithm
VRTs	Variance Reduction Techniques

List of Figures

List of Tables

1 Introduction

1.1 Background

Future *development* and present operation of electric power systems along with other *large systems* must pursue a number of *different goals*. Above all, the power system should be *economically efficient*, it should provide *reliable energy* supply and should not have any detrimental impact on the *environment*. In addition to these global goals there is a number of supplementary goals, objectives and criteria. At the same time, operation and development of the system is influenced by a variety of *uncertain and random* factors. As a result, the development strategy can be chosen from a large number of *possible alternatives*. Obviously, that among the set of possible alternatives the planner attempts to find the best, or in accordance with accepted term, the *optimal alternative*. Thus, the complexity of the problems related to power systems planning is mainly caused by presence of multiple objectives, uncertain information and large number of variables.

This dissertation is devoted to consideration of the methods for development planning of only one part of the electric power system, namely *distribution networks*. However, a lot of problems arising during the elaboration of methods for strategic planning of power system objects are common apart of features of the object (voltage level, size etc.). Therefore, methods and approaches treated in this work could be useful also for planning of some other subsystems of electric power system.

The history of the methods for network planning comes along with the history of electric power industry. As the significance of the electric power for the national economy was increasing, more and more efforts were put to find the optimal network development strategies. Recognized, that there is a number of methodologies applied in practice, which result in feasible and decent solutions. However, it is evident, that these methods can be improved. The development of more and more efficient methods for the planning of distribution networks is constantly significant. This can be explained by the high investments involved in reinforcement and operation of the networks. In industrialized world nearly half of the investments in power industry is spent in distribution networks.

Furthermore, in recent years there is a worldwide wave of considerable changes in power industries, including the operation of distribution networks. Deregulation, open market, alternative and local energy sources, new energy conservation and communication technologies, these are the major factors, which on one hand increase the uncertainty level and on the other hand provide the alternative solutions to the planning problem. New conditions persuade the search for new comprehensive methods for planning of power system objects, including distribution networks.

Then again, the powerful tools for solution of the tasks of the given type become available. Computational capacities increase exponentially, and the new mathematical methods and algorithms are developed.

1.2 The Purpose and Contributions of the Dissertation

The major purpose of this dissertation is to contribute to the development of reinforcement planning in distribution networks. To archive this purpose the following principal missions were stated:

- Choose the planning objectives reflecting the efficiency of the planning alternatives. In the most general case the efficiency criteria can be reflected by the numerical parameters of the probability distribution of the expected revenues (or losses), which can be calculated by integration of the multidimensional probability distribution function.
- Select the methods for assessment of the attributes corresponding to the planning objectives and organize the optimization procedure. To assess the attributes it is suggested to use the method of Monte-Carlo. The optimization procedures are based on Genetic Algorithms (GA). In order to improve the performance of the algorithms it is suggested to use the method of importance sampling, the method of common random numbers and the modification of the GA, which allows for the simultaneous search of the whole Pareto optimal set.
- Provide the algorithm for optimization of complex combinatorial dynamic problems. The novel algorithm based on combination of Genetic Algorithm with Dynamic Programming (DP) is suggested. The efficiency of the algorithm is demonstrated on several examples.
- Account for different types of information serving as an input for the network planning tasks. It is concluded that for the network planning in general case the following types of information should be modeled
  - ✓ Deterministic
  - ✓ Probabilistic
  - ✓ Fuzzy
  - ✓ Truly uncertain.

  The dissertation first provides the review of the methods for modeling of uncertain parameters, then suggests the model, which includes all four informational conditions.
- Manage risk. The criteria, characterizing risks of the alternative solutions are recommended, and necessity and rationale of these criteria are shown.

- Apply the suggested methods to the real network. The real tasks of distribution network planning are analyzed. Based on the analysis it can be stated that the suggested algorithms are realistic and can be applied for solution of practical tasks.
- Represent multiple attributes. In order to simplify the trade-off between the alternatives obtained as a result of optimization it is suggested to use the Principal Component Analysis (PCA).

1.3 Scope and Organization of the Dissertation

The dissertation is organized as following:

- *Chapter 2* assigns the distribution system as an important part of the electric power system – one of the most complicated systems created by the mankind. The chapter states the main planning objectives: minimize power losses, capital investments and maintenance costs and energy not supplied due to interruptions in the network. It is declared that the complexity of the stated task of distribution planning is caused by multiple objectives, large number of variables, uncertainty of initial information and dynamic nature of the problem. Development of new technologies provides the extended opportunities for improvement of network operation, but simultaneously complicates the planning process.
- A broad theoretical base for the network planning is given in *Chapter 3*. The network planning problem is formulated in general form in terms of Bayesian Decision Theory. However, the difficulties associated with formulation of the utility functions make it almost impossible to apply the Bayesian approach directly. It is stated that when approaching the problem applying different methods it is important to consider the concave character of the utility function. This consideration directly leads to the multi-criteria formulation of the problem, since the decision is motivated not only by the expected value of revenues (or losses), but also by risk of not having the expected result. The chapter contains the description of several means, which allow reducing the complexity of the problem in its strict formulation. The conclusion is made that the difficulties caused by tremendous complexity of the problem can be overcome either introducing a number of simplifications, leading to the considerable loss in precision or applying methods based on modifications of Monte-Carlo or fuzzy arithmetic and Genetic Algorithms or Dynamic Programming.
- *Chapter 4* contains the description of the suggested model and the algorithms for assessment of the selected planning attributes. The model is very fundamental and there is large potential for its improvement or extension. Furthermore, the chapter contains the suggested algorithm for

probabilistic load modeling based on analysis of measured data and choice of suitable empirical probability distribution.

- *Chapter 5* states that in presence of uncertainty the planner aims at finding the robust and flexible plans to reduce the risk of considerable losses. Several measures of risk are discussed. It is shown that measuring risk by regret may lead to the risky solutions, therefore an alternative measure – Expected Maximum Value – is suggested. The general future model called fuzzy-probabilistic tree of futures, which integrates all classes of uncertain parameters (probabilistic, fuzzy and truly uncertain) is described in this chapter.
- *Chapter 6* contains three original applications of GA to the network planning. The first algorithm searches simultaneously for the whole set of Pareto optimal solutions. The hybrid GA/DP approach benefits from the global optimization properties of GA and local search by DP resulting in original algorithm with improved convergence properties. Finally, the Stochastic GA able to cope with a noisy objective functions is described. The chapter also contains the recommended modus operandi for the network planning tasks.
- *Chapter 7* includes the description of two real distribution network planning projects, namely primary distribution network in the large city and secondary network in the rural area. The stages of the “Large Swedish City” project are studied in details.
- Finally, several conclusions and suggestions for the future work are given in *Chapter 8*.

1.4 List of Publications

L.M. Proenca, V. Bochkareva, V. Miranda, “Towards a Comprehensive Methodology for Distribution System Planning”, ELAB’96, Porto, 1996.

V. Bochkareva, “Transmission Network Development Planning. A Literature Review”, Internal Report at Royal Institute of Technology, Stockholm, A-EES-9508, 1995.

V. Bochkareva, *Development Planning of Electricity Distribution Networks*, Lic. Eng. Thesis, Royal Institute of Technology, Stockholm, pp. 97, TRITA-EES-9702, ISSN 1100-1607, 1997.

Z. Krishans, V. Bochkareva, G. Andersson, “Dynamic Model for Reinforcement Planning of Distribution System under Uncertainty”, *Universities Power Engineering Conference (UPEC’97),* Manchester, UK, September 1997, pp.1089-1092.

V. Neimane, Z. Krishans, “Multi-criteria analysis in electric power network planning”, *Latvian Journal of Physiscs and Technical Sciences*, 1997.

V. Neimane, G. Andersson, Z. Krishans, “Multi-Criteria Analysis Application in Distribution Network Planning”, *33rd Universities Power Engineering Conference (UPEC’98),* Edinburgh, UK, September 1998, pp.791-794.

Z. Krishans, V. Neimane, G. Andersson, “Dynamic model for planning of reinforcement investments in distribution networks”, *Power Systems Computation Conference (PSCC’99),* Trondheim, Norway, June 1999.

V. Neimane, G. Andersson, “Distribution Networks Reinforcement Planning: A Dynamic Multi-Criteria Approach”, *Budapest Power Tech’99*, Budapest, Hungary, August 1999.

V. Neimane, “Distribution Network Planning Based on Statistical Load Modeling Applying Genetic Algorithms and Monte-Carlo Simulations", *Porto Power Tech’2001*, Porto, Portugal, September 2001.

V. Neimane, “Towards Bayesian Solution for Network Planning Problem", *RIMAPS’2001*, Porto, Portugal, September 2001.

2 Network as an Object under Optimization

This chapter assigns the distribution system as an important part of the electric power system – one of the most complicated systems created by the mankind. It is declared that the complexity of the stated task of distribution planning is caused by multiple objectives, large number of variables, uncertainty of initial information and dynamic nature of the problem. The chapter states the main planning objectives and identifies the stages of the planning process. Development of new technologies provides the extended opportunities for improvement of network operation, but simultaneously complicates the planning process.

2.1 Planning of Reinforcements in Distribution Networks

2.1.1 Essence of Distribution Planning Problem

The electric power systems are among the most complex systems created by the mankind. These include hundreds of thousands components: generators, transformers, transmission lines, control and protection equipment, etc. Construction of power systems and their operation and maintenance require billiards of dollars. The functions are interdependent: the processes going on in one of the system's components influence functioning of the other elements.

Operation conditions are continuously varying in time – new customers and power system objects appear, prices grow and legislation changes. Additionally, constantly changing weather conditions, e.g. temperature and wind speed, influence significantly operation of the system. The costly objects and elements have a finite life of several decades. This motivates the need to estimate the conditions, which may arise in a rather distant future. Clearly, these conditions cannot be predicted exactly. Therefore, it is necessary to account for uncertainties and random factors. At the same time, construction of some objects requires both considerable investments and certain time. Thus, the planning mistakes leading to the wrong decisions cannot be corrected fast and may result in substantial financial losses.

In planning of electric power systems a number of goals must be achieved and correspondingly a number of objectives, which often are conflicting, must be optimized. The planning goals include minimization of power losses and required investments, enhancement of reliability, personal safety and power quality, and consideration of environmental factors. If there is a number of objectives and it is impossible to define their relative importance or to express the corresponding attributes in monetary terms, the planner has to deal with multi-criteria optimization tasks. Significance of electric power for the national economies, high investment costs and considerable possible losses in case of planning mistakes encourage the development of well-motivated methods for robust and flexible planning of power systems.

Strong interdependence of the power system elements imposes the need to consider the system as a whole. However, the optimization of large power system is the task of remarkable complexity. One of the most powerful means to reduce the complexity of the problem is decomposition, i.e. the task is divided into several simpler sub-problems. Thus, traditionally transmission, sub-transmission and distribution systems can be treated independently (e.g. [10]), furthermore, the local distribution networks can also be handled separately, taking into consideration relatively weak connection between them.

It should be noted that decomposition of the initial complicated problem into several sub-problems is not free of charge. The decisions of each sub-problem should be consistent, moreover often there is a need to solve each sub-problem for several outcomes of other sub-problems [6].

In comparison to generation and transmission or sub-transmission planning, planning of distribution network seems to be an easier task. The decision-making responsibility and consequences of wrong decisions seem to be not so bad. However, precisely in distribution networks identified the major power losses, their maintenance require considerable efforts and the total length of distribution networks is by an order higher than that of transmission networks. In Figure 2-1 data are given about length of power transmission and distribution lines in Sweden in 1992 (after Elverksföreningens statistics). The comparison of the line length gives the sense of plurality of planning tasks in distribution and therefore emphasizes their significance.

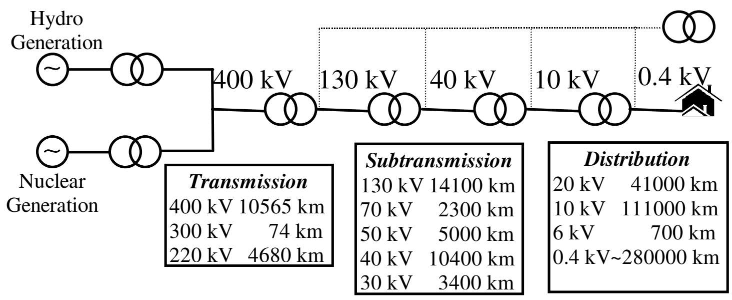


Figure 2-1 Lines length statistics for Swedish networks

Despite the possible simplifications the problem of distribution network planning remains an extremely complicated optimization and decision-making problem due to:

- Conflicting objectives
- Large number of variables
- Dynamic nature
- Uncertainties.

In recent years distribution network planning has become a subject of interest both for researchers and power utilities. There are several reasons for this. First of all there is the need from the industry to have such tools. The conditions for which the network was planned are changing: open market introduces new challenges and increases uncertainty, new technologies provide new possibilities for reinforcement, new communication channels and SCADA systems considerably extend the degree of information available for analysis, etc. All these changes encourage the efforts to improve the performance of the network and, therefore, the efficiency of the planning process. On the other hand increasing computation capacities and introduction of new powerful methods for the optimization problems decision provide a possibility to develop new tools for the network planning.

2.1.2 Process of Distribution Planning

A starting point of reinforcement planning is the existing network under the influence of external factors. Once it has been identified that network performance during the planning period is in any way inadequate, it is time to start the planning process.

Inadequacy of performance may be induced by internal or external changes, such as increase or decrease in existing loads or appearance of new loads, appearance of a new local generation source and obsolescence of equipment. Furthermore, new requirements to network performance criteria, such as improved reliability, decreased operation and maintenance cost, decreased losses may also require additional reinforcements. Information about inadequacy of the performance can be obtained from several sources, mainly from monitoring calculations, but also from customers' complaints, direct measurements and observations by the utility staff.

Reinforcement actions may include addition, upgrade or elimination of the network elements. Each problem may have several possible solutions. For example, monitoring calculations indicate that in 5 years voltage level will be too low at some parts of the network. Possible reinforcement actions may include for example building of new lines, selecting between overhead line and cable, providing alternative network configurations, installation of capacitors, change of transformers, enlargement of conductor cross-section or transition to the alternative voltage level. Moreover, appearance and development of new technologies may suggest alternative or additional options, which should also be considered in the planning process.

The planning process consists of several steps including identification of possible alternatives, their evaluation according to selected performance criteria and selection of the most suitable alternatives, which form the development strategy. For instance, in [138] the planning process is segmented into the following five stages:

Stage 1 Identify the problem: Explicitly define the range of application and its limits.

Stage 2 Determine the goals: What goals are to be achieved? What is to be minimized?

Stage 3 Identify the alternatives: What options are available?

Stage 4 Evaluate the alternatives: Evaluate all the options on a sound basis.

Stage 5 Select the best alternatives: Select the options that best satisfy the goals with respect to the problem.

However, it is important to add one more planning stage, which can be formulated as follows:

Stage 6 Make the final decision: Based on the results obtained on the previous stages select options, which will be carried out.

It should be noted, that in order to perform the first five planning steps, the companies may employ experts able to follow each stage of the process based on the detailed information and applying various modeling and optimization tools. In some cases this work may be even subject of outsourcing. But the sixth stage essentially should be accomplished by the management of the power company.

2.2 Objectives of the Network Reinforcement

Objectives of the network reinforcement may vary considerably from one utility to another and from one plan to another within the utility. However, it is possible to formulate the common objectives for the planning task in general in terms of planning attributes, which have to be minimized.

The approximate hierarchy of objectives for distribution network planning is presented in Figure 2-2. More or differently formulated objectives can be added, i.e. voltage quality or environmental impact.

Shaded rectangles in Figure 2-2 contain the attributes, which are common for the distribution planning problems, and are suggested for application in planning software presented in this dissertation. As a result, there are the following three general attributes to be minimized:

Attribute 1 Power Losses: Cost of power losses is calculated for the whole planning period. Different loading conditions may be modeled by duration of every mode

Attribute 2 Investments: Investment and operation and maintenance (O&M) costs are combined into the single attribute

Attribute 3 Reliability: Either energy not supplied or customer outage costs is used depending on the information available for the particular task.

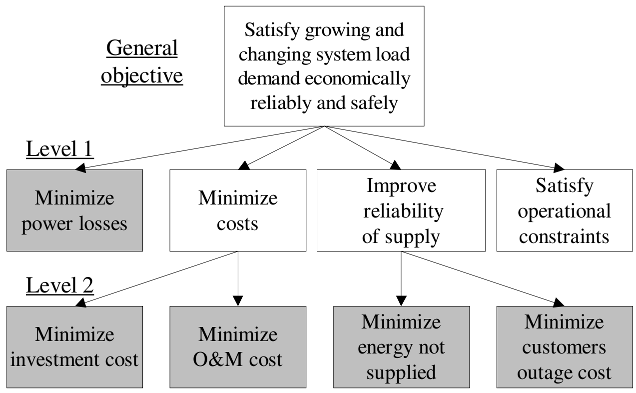


Figure 2-2 Hierarchy of objectives for distribution planning

The objectives, which are the subject of optimization, are open-ended. No matter how good the plan is, the planner is always challenged to do better. By contrast, the operational constraints must only be met, not exceeded [138].

There may be additional goals specific for particular project. For example, the project in section 7.1 considers reinforcement of the existing stations primarily in order to improve personal safety standards. Another goal in this project is to replace oil-paper cables, mainly for environmental reasons.

Part of the objectives may be formulated as attributes and taken into account during optimization, but some are considered as constraints. In any case, the objectives combine engineering and economics. The solution, satisfying all the objectives must be both technically feasible and economically efficient.

2.3 Problem Dynamics

The objective of distribution planning is to provide technically and economically efficient reinforcement plans to meet future electrical demand at the acceptable reliability level. The view of distribution planner on "future" is discussed in this sub-section.

Due to uncertainty in future circumstances, the planner may commit only to the planning alternatives, which have to be realized in the nearest future. However, the consequences of these alternatives must be estimated on a long-term basis. Thus, the planning process may be divided into two major time stages: short and long term. The purpose of short-term planning is to make certain that the system can continue to serve customer load while meeting all standards and criteria. The duration of the short-term period depends on the lead time for the particular level of power system. Then again, the purpose of the long-term plan is to assure that all short-term decisions have lasting value and contribute to a robust solution for network reinforcement.

The short and long-term planning periods shown in Table 2-1 are recommended in [138] for normal utility circumstances.

Table 2-1 Typical short and long-term planning periods for power system planning(after [138])

System level	Planning period – years ahead	
	Short-term	Long-term
Large generation (>150 MVA)	10	30
Small generation (<50 MVA)	7	20
Transmission	8	25
Sub-transmission	6	20
Distribution substations	6	20
Feeder system	6	20
Primary tree-phase feeders	4	12
Laterals and small feeder segments	1	4
Service transformers and secondary	0.5	2

Furthermore, the long-term planning must be considered from the dynamic perspective. Since power system is an object, which may change continuously taking into account load growth, economic indices, etc., estimation of the current decisions must allow for some potential additions when they needed at some moment in the future – for example in 10, 11, or 13 years. However, instead of modeling on the yearly basis, often it is sufficient and more convenient to divide the whole planning period into several time stages – for example 3 years each.

2.4 Uncertain Information

2.4.1 Classes of Initial Information

Uncertainties are introduced in any, especially in long-term, planning problem.

Then the planners face a dilemma. It is challenging to make a commitment of resources and facilities for the future that may develop and may not.

A wide scope of different information regarding

- present and future customers data and load profiles
- present network structure and trends of its future development
- characteristics of the network components
- technical and economic data of installed and potential equipment
- financial and labor resources of the company

is needed in order to solve the network planning problem.

Considerable part of this information is uncertain, i.e. it is vague, fuzzy, even ambiguous. Uncertainty or vagueness of the information is caused by errors in measurements as well as inevitable errors in estimation of future forecasts. Furthermore, since most of the data used for the planning tasks are not based on the direct measurements, the degree of information uncertainty may be quite high. Clearly, the data about the present situation in the network is much more accurate than the forecasted data about the network's future development.

From the descriptive viewpoint all the initial information may be categorized into the following several classes.

<u>Deterministic</u> information is the one defined explicitly, as for example nominal voltage levels in the network or present network configuration. Some data describing the process of the network development - such as possible sites for new substations or possible routes for new cables - can also be specified to be deterministic.

<u>Probabilistic</u> information based on available statistical data can be obtained in form of some known probability distribution law and its parameters. In some situations there are such data concerning existing loads, reliability data for the network components and power quality indices. However, more often there are situations when probabilistic information itself is uncertain - distribution law is unknown or its parameters are uncertain. The major part of the load data can be related to this class of information.

A large amount of valuable information can be obtained in linguistic form, e.g. "large", "average", "small", "many", "a few", "efficient", "less efficient", etc. This <u>fuzzy</u> information is often very subjective and is usually based on expert judgment, however it can be a huge aid during the decision-making process.

Finally, there is truly <u>uncertain</u> information, which do not have probabilistic nature. Usually all the information about the decisions to be taken in the future is uncertain, therefore uncertain is future network configuration. Furthermore, we may consider as truly uncertain the information, which is simply not available. In many cases truly uncertain information can be estimated based on some general considerations, the expertise of the planner and his intuition and modeled by *subjective* probabilities. If the assignment of subjective probabilities is too problematic and could not be done with satisfactory precision, then the corresponding uncertainties should be modeled using the scenario approach.

2.4.2 Methods to Reduce Information Uncertainty

The higher the degree of uncertainty the lower the quality of the decisions made by the planner. Therefore, the attempts to reduce uncertainty are well substantiated. However, even inconsiderable contribution to the reduction of information uncertainty usually requires considerable efforts and expenses. Therefore, the vital task prior to network optimization is to determine those parameters and characteristics, which influence most the optimization results (and therefore the final decisions) and to improve the quality of the corresponding models and the relevant information.

There are two basic approaches to reduce uncertainty in the initial data. The first one is based on enhancement of existing measuring, monitoring and forecasting methodologies. Applying this approach, the present network can be described with a very high degree of accuracy. To some extend this can be realized utilizing the functions of System for Control And Data Acquisition (SCADA).

The second method implies the principle of minimal period for decision-making. It is obvious, that the farther into the future we look the more uncertain information we have, and on the contrary, the information is more accurate for the less distant future. Therefore, the planner postpones making the final decision till the last reasonable moment, when the information is the most up-to-date. It should be noted, that it is possible to greater extend to reduce uncertainty concerning the present state of the network and to some extend to reduce uncertainty concerning conditions of the network development. Still, it is impossible to eliminate uncertainty entirely when it comes to the future forecasts.

Ideally, the developed plans must be able to confront any or at least the most likely eventualities. Alternatively, it should be possible to change the accepted plan easily when more information is available. These properties constitute the framework of *robustness* and *flexibility* in distribution planing.

2.5 Decision-Making

What is the best alternative? There is no simple answer to this question, this depends on the goals for the particular situation. Most of the real life tasks involve several goals, which often are contradictory. Similarly, the network planning involves a number of objectives and corresponding attributes (see Figure 2-2), which have to be minimized simultaneously. In presence of uncertainty and if it is impossible to express all the attributes in monetary terms, the problem results in a set of non-dominated solutions called Pareto optimal set instead of single optimum. Thus, even in deterministic formulation the planning problem rises above the optimization problem, and attains the class of decision-making problems.

Important ideas concerning decision-making will be discussed in Chapter 3. One of the main arguments is that there is one person (or a group of persons) who has the detailed information about the problem, states the task and performs calculations and another person who has the authority to make a decision. This fact is one of the main motivations why trade-off analysis and risk analysis, resulting in quite a large set of possible solutions should be preferred to the approaches reducing task to a single criterion optimization.

2.6 New Considerations and Challenges for Distribution Planing

2.6.1 Extension of the Traditional Distribution Network Planning

Taking into account the processes of deregulation going on in many countries and rapid development in technologies there may be a need to reconsider or to extend the traditional approach to the planning of electricity distribution networks. The following issues concerning modeling problems as well as application of distribution network planning by power companies are seen as important for consideration:

- Deregulation[*]: how does it affect network companies, what does it change in network planning, how should these changes be taken into account in network planning models?
- Appearance and development of new technologies: modeling as possible options for network reinforcement
- Fast development and wide spread of the SCADA: their utilization provides the possibility to collect the variety of the detailed statistical data
- Local generation: if present must be taken into account, which requires more complex analysis.

[*] The process of deregulation may be also called liberalization of the electricity market, which probably more precisely reflect its nature.

2.6.2 Deregulation

2.6.2.1 Transition from the Traditional System

The traditional system of electricity supply is characterized by regional monopolies and the exclusion of electricity to electricity competition. However, in many countries the electricity industries undergo the transformation from traditional vertically integrated utilities into a competitive business [52],[80],[135].

Deregulation of the Swedish electricity market was legally executed January 1, 1996 [126],[129]. The main objectives of the deregulation are to improve efficiency of the electric power industry and to reduce electricity prices. The focus is now put on the end-customers.

The reform completely separated electricity generation and trading from transmission and distribution. Generation and trading became a subject to competition, while network operations, continue as natural monopolies. Traditionally distribution utility comprises two separate activities: operating the distribution network and retailing. Retailing consists of trading electricity at the wholesale level and selling it to the end users. Distribution network business is a natural monopoly and must thus be regulated. In the Swedish model these two functions of the utility must be separated.

The electricity value chain as depicted in Figure 2-3 shows that competitive features do not exist on all levels of the chain.

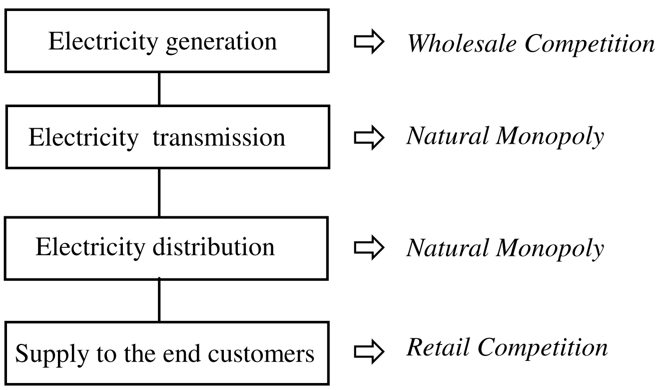


Figure 2-3 The electricity value chain

2.6.2.2 Deregulation in Sweden: The Law

According to the law [126],[129], holders of network concessions are obliged to connect customers' lines and installations on reasonable terms. All networks at all levels should be open. Terms for the transmission of electricity, irrespective of its supplier or purchaser, shall also be reasonable. Network tariffs shall consequently be unbiased and non-discriminatory. They should be based on objective calculations and provide a reasonable rate of return. The Network Authority (Nätmyndigheten) will examine network tariffs in this respect, not only as a result of complaints received from customers, but also on a spot-check basis. If a network company is found to have unreasonable tariffs, the Network Authority will serve it with an injunction to adjust its prices. A network company wishing to appeal against such a decision must go to a court of law.

The new legislation prescribes the point-of-connection principle tariff for the entire network system. The basic principle is that payment in one point, the point of connection, gives access to the whole network system, and thus the whole electricity market. This means, that consumer or producer connected to a local network pay network fees only to the owner of that network. Costs related to the external network owners are included when each network owner calculates the transmission tariffs. In this way the main grid tariffs are included into the regional tariffs, and the regional tariffs are included into the local tariffs.

2.6.2.3 Electricity Distribution in Open Market Conditions

On the distribution level there are two main activities to be provided:

- Energy transportation (distribution network): An adequate remuneration mechanism to the utility should be provided by the Network Authority. At the same time it should control power quality and the continuity of supply. Basic cost items for the network activities are new investments in the network installations and operation and maintenance costs.
- Energy supply (retail business): The Network Authority must establish objective rules for open access to the network in order to provide a non-discriminatory framework to the different competitive agents. The supply activity to non-franchised customers can be carried out under competition, while supply to franchised customers must be regulated: selling price and provided services should be controlled.

The distribution network utility remains responsible for the technical aspects of the power quality and the technical state of the network. Usually it is also responsible for the new network connections and installation of meters.

In the new competitive environment the distribution network has to cope with expectation from:

- the customer, who expects to get reliable supply, and since the prices for network operations become more transparent, can compare the tariffs and services with another networks prices; if the comparison is not satisfactory, the customer can complain;

- the owner of the distribution system, who expects a certain return on the capital employed in network operation;
- the Network Authority, who regulates the network tariff; which means that the companies have strong incentive to improve their efficiency as the profit will be the different between the price cap set by the Network Authority and the actual cost.

Thus, in order to satisfy the expectations from the customers, owners and the Network Authority the company has to be efficient mainly in order to keep the total cost low.

It seems, however, that under the new separation scheme of activities, the network operation and retailing, the distribution network business does not have enough natural incentives to decrease losses and to improve power quality. Furthermore, there is also lack of incentives to reduce investment costs and to avoid non-efficient installations. The reason for this is that according to the traditional remuneration scheme the capital investments are assured of a fixed percentage of return. All the utility costs are also included in the base for the tariff (Figure 2-4). Under this scheme the Network Authority must control the investments and operation costs as well as power supply quality.

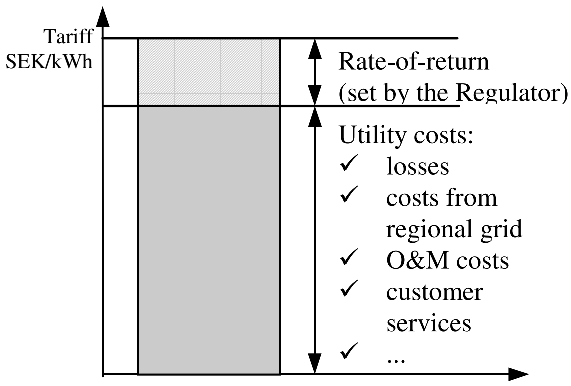


Figure 2-4 Traditional tariff structure

In Sweden there is a number of indirect incentives to improve network operation and to reduce the prices following from the benchmarking principle and distribution network ownership structure. Prior to deregulation there was about 260 local network companies (owned mostly by municipalities. Clearly, the politicians in these municipalities cannot allow unreasonable prices, which would lead to disappointment of their voters - distribution company customers. Eventually, potentials to maintain high quality standards, good quality service and low prices will require units larger than today's companies. Thus, by the same reason of ownership the company having lower tariffs has better chance to acquire or to merge with another company. On the other hand, due to separation of the accounting for network operations and retailing the utilities cost become more transparent and it is easier to compare and to conclude which company is efficient and to which degree.

In some countries there are real incentives to keep power quality high. In Norway, for example, the regulator has introduced a compensation for interruptions or outages in year 2001. Interruptions of more than 3 minutes are multiplied with a fee for households or a higher fee for industrial customers and deduced from the income for the next year. In that way the customer will get a reduced tariff and the network company will be punished if the quality is too low [128].

2.6.2.4 Changes in Distribution Network Planning

Deregulation will give a new focus to the traditional network planning. The most important changes are listed bellow:

- Distribution company is a business focused on making profit. Thus, in the majority of cases the company is tending to maximize utilization of existing assets and avoid redundant investments. On the other hand, a particular attention is paid on minimization of operation and maintenance costs.
- The utilization of methods, which incorporate uncertainty, is especially important in deregulated environment. Thus, risk-based and probabilistic planning will find more applications.
- Different customers value reliability differently and they are willing to pay certain tariff for a certain level of reliability. If until now reliability was just a measure of system performance, it is now becoming an explicit factor in the planning process.
- Network optimization and planning methodologies are serving as decision-making aids, which help to quickly find the best solutions. Now they are becoming also the aids to justify the need for certain investments and document the process for regulator and company investors.

2.6.3 New Technologies

2.6.3.1 Brief Review

There is a multitude of options for new technology presented in the literature, e.g. [22],[49],[59],[75],[135], proposing traditional, radical and sometimes conflicting solutions. Determining how to invest limited resources and capital the best to achieve the maximum possible advantage is a challenge for every utility. The challenge is to deliver even more reliable service at reduced cost and with fewer technical and support employees. This creates the driver towards further implementation of distribution automation and distribution network management tools.

Due to rapid development of technologies, especially in data communication and computing, there is a variety of distribution management functions which can aid the utilities to improve their operational efficiency. The main functional groups are presented in Figure 2-5 (after [22]).

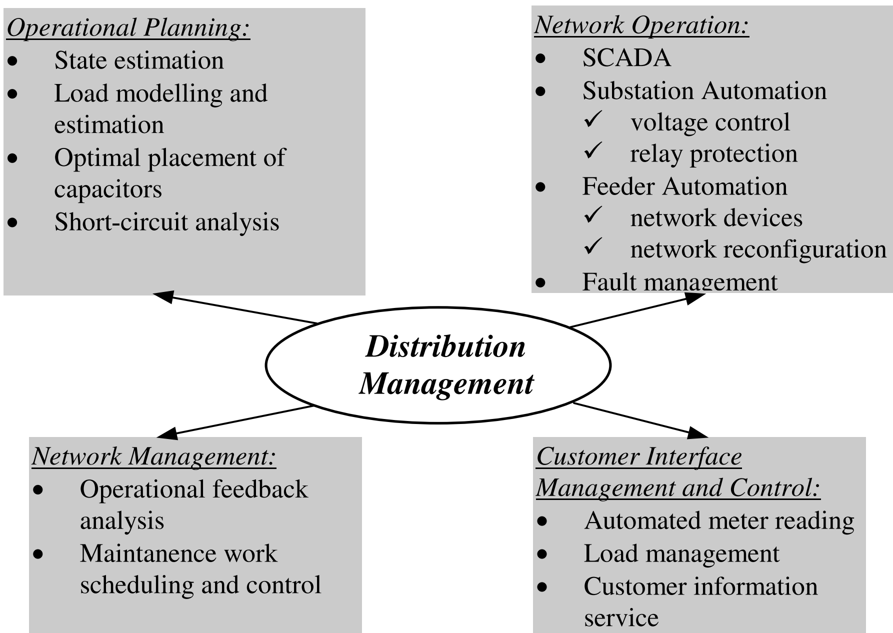


Figure 2-5 Main functional groups of Distribution Management Systems (DMS)

Most of these options are available for the utilities, however not all of them are applied, even fewer of them are applied in the integrated manner. There is always a gap between available technologies and their practical implementation. This is caused by several reasons:

- Cost of implementation
- Competing investments
- Rapid development of technologies
- Uncertainty in technology and market development trends.

The new technologies mentioned here include both the developments in automation devices and data communication and computing systems. Some of them are briefly described below.

Fast switches and line switch remote control [49],[75]

If large number of remote controllable switches is incorporated into a distribution network, by varying open/closed state of the switch the operator could optimize the configuration of the network to minimize operation costs, including customer outage costs. The main effect would be reduction of average outage duration in the system. In addition to this savings would be achieved in labor costs of both fault clearance and routine connections.

Furthermore, addition of the fast switches to the system allows to operate the system as a more highly meshed, thus introducing more paths to feed any given load point. This also allows to reduce operation costs and, which is the most important, considerably to improve reliability of the system.

Power quality improvement devices [49]

As an example Current Limiting Devices (CLD) can be mentioned. The main effect of CLDs is to isolate parts of a system from the effects of faults in other parts of the system. The effect of CLD use is to cause voltage to remain within the acceptable region for more faults than it would be the case without CLD use. The problem is how to adjust fault detection and location functions in presence of these devices.

Automatic voltage control [75]

The load flows during peak demands can be reduced by improved voltage control and by reactive power compensation. Improved voltage control allows the utility to reduce active and reactive power flows during the peak demand period temporarily lowering the supply voltage. To implement this function a real-time network calculation system is needed.

Fault management systems [75],[80]

Fault management in electric power systems is becoming more and more sophisticated due to the development of the modern micro-processor based protection and fault detection equipment. Using the measurements before and during the fault, an estimate of the distance to the fault can be computed. Furthermore, there exist systems, which allow for allocating the fault taking into account network configuration, types of lines, weather, previous experience, etc.

Distribution line carrier (DLC) [22]

DLC offers utilities the possibility to create a cheaper communication infrastructure by using distribution networks as a communication medium. Implementation of DLC opens a number of possibilities to apply different Distribution Automation (DA) and DMS options.

This list can be almost infinitely continued and should be updated continuously.

It is expected in the nearest future, that the new technologies will be implemented by the utilities in a larger scale. Clearly, the priorities of the implementation differs from one utility to another, as well as the benefits from the various functions. Furthermore, it can be seen even from the description given above, that the functions are interrelated.

Distribution Automation includes equipment that allows for the real time global reconfiguration of the distribution system and its equipment in response to operating problems, system status and utility goals. Distribution automation provides the following two capabilities that permit distribution system to operate with less capacity margin compared to a traditional configuration:

- to monitor loading and electrical performance on a real-time basis
- to remotely control equipment – switches, regulators, switched capacitors.

Therefore, in the network planning process the DA, if present, must be considered in the model, since it would allow for system operation with less capacity margin. On the other hand, introduction of new DA functions may be considered as possible reinforcement option if the capacity limit is reached (or due to some other criteria).

SCADA, the new communication technologies and microprocessor-based registration devices, facilitate the creation of the comprehensive data bases, which can serve as a source of:

- The statistical data about the laws of the load distributions
- The reliability indices of the network elements as well as protection and control systems.

The Demand-Side Management (DSM) techniques are another alternative to control network inadequacy [136]. Instead of reinforcing the network to meet the future load demand, the load curves may be reshaped to correspond to the network capacities. New technologies make it easier to plan and perform DSM incentives. For instance, meters functioning on the hourly bases allow for introducing adaptive network tariffs.

It should be also mentioned that there is a number of technological advances available to the customers. Applied in a large scale these can considerably influence the level of demand. New technologies usually suggest more efficient solutions, for example for lighting, isolation of houses, refrigerators and other household products, therefore, in residential areas the demand may even decrease.

2.6.3.2 Changes in Distribution Network Planning

Therefore, the appearance and application of new technologies induce the following changes into the network planning process:

- The number of possible planning alternatives increase significantly

- The possibility to manage the load shape becomes accessible via introduction of the multi tariff system
- The detailed statistical data about operation of the network (loads, failures, etc.) become available
- The ability to monitor loading and electrical performance on a real-time basis permits operation very near to upper limit of acceptable loading.
- Failure duration times become shorter due to remotely controlled equipment.

2.6.4 Distributed Generation

2.6.4.1 Generation in Distribution Networks

There could be several types of small scale local generation sources such as wind power plants, small hydro power plants, small fuel cells or local combined heat and power plants connected to the network. For decades it has bee recognized [75] that introduction of distributed generation into distribution systems will significantly complicate distribution planning and operating practices and require substantially greater data collection and analysis efforts.

Normally, small power plants are connected to the distribution network at relatively low voltage level. Therefore, in distribution network calculations it is important to consider presence of the distributed generation in the network area and its influence at least to the losses, investments to the network and reliability of supply [122].

Clearly, distributed generation influences power losses in the network. Decrease or increase of losses depends on size of local power plant, its location and the power flows in the system. If consumers are located near the local plant and their size is of the same order, then the losses decrease, since energy does not have to be transported long distances and is generated locally. However, if local consumption is less than production, then losses can locally increase.

Introduction of the distributed generation source can also considerably affect investments into the network. Apparently, it is difficult to make a general conclusion about the influence of the local generation. There are different situations dependent on the correlation between local production and consumption as well as on configuration of the particular network and location of the distributed generation source.

Local producers are one of the actors in the open energy market. It is important to define how the actor which increases or on the contrary reduces expenses in the system, respectively pays or is paid for this. For example, if the local producer reduces losses in the distribution, regional and transmission networks.

The question is then how the local producer should be compensated?

According to the law [126] the owner of the small generation source should not pay any other part of network tariff, but only connection fee, fee for the metering equipment and its installation as well as annual fee for measuring of transmitted energy. The producer has rights to be compensated for reduction of losses it causes. The relationships between the producer and the network owners considering losses in the network can be summarized as follows:

- If local producer causes reduction of losses in local network it gets compensation;
- If losses in the local network increase, the local producer does not have to pay. The costs will be covered by the local network or more precisely by the consumers, since they will be included in network tariff;
- If losses in the higher level network decrease, production owner gets compensation according to that network tariff.

2.6.4.2 Changes in Distribution Planning

The influence of the distributed generation considering investments into the network is based on the following:

- The network owner's expenses of distributed generation source connection to the network are covered by the local generation owner;
- Normally it is illegal to charge local producer by any annual fee for connection to the network;
- If distributed generation in distribution network causes reduction in higher level network costs this reduction should be compensated to the generation owner.

On the other hand, the company that owns the network can also decide to install the distributed generation source in order to provide local peaking capacity for the local area it serves. The unit would be foreseen only for operating during times when, because of system-wide or local needs, the utility decides to activate the unit to provide power injection into the local distribution system.

2.7 Conclusions

- Distribution network is an important part of electric power system, which is one of the most complicated systems created by the mankind.
- Planning of the development of distribution networks pursues a number of conflicting objectives: minimization of power losses, capital investments, operation and maintenance costs and energy not supplied due to

interruptions in the network. The complexity of the stated task is caused by multiple objectives, large number of variables, uncertainty of initial information and dynamic nature of the problem.

- Planning is a multi-stage process. Primary calculations and analysis, and actually the decision-making traditionally is made by different group of persons.

- Information about the factors, which influence functioning of the network can be categorized into the following classes:
  - ✓ Deterministic
  - ✓ Probabilistic
  - ✓ Fuzzy
  - ✓ Truly uncertain.

- New tendencies and conditions in organization of electric power supply – deregulation, open market and appearance of local generation – increase level of uncertainty in planning tasks and inspire the search for the new methods.

- Development of new technologies provides the extended opportunities for improvement of network operation, but simultaneously complicates the planning process. Application of new informational technologies (SCADA) facilitates the collection of extensive information, which can be used to improve the efficiency of planning process.

3 Theoretical Basis for Planning

A broad theoretical base for the network planning is given in this chapter. The network planning problem is formulated in general form in terms of Bayesian Decision Theory. It is stated that when approaching the problem applying different methods it is important to consider the concave character of the utility function. This consideration directly leads to the multi-criteria formulation of the problem, since the decision is motivated not only by the expected value of revenues, but also by risk of not having the expected result. The conclusion is made that the difficulties caused by tremendous complexity of the problem can be overcome either introducing a number of simplifications, leading to the considerable loss in precision or applying methods based on modifications of Monte-Carlo or fuzzy arithmetic and appropriate optimization methods.

3.1 The Problem of Network Optimization from the Viewpoint of the Decision Theory

3.1.1 Bayesian Decision Theory

Statistical decision theory is well developed and extensively applied in different areas [36]. It comprises two fundamental approaches for the decision-making problems:

- Bayesian
- Orthodoxy (contains all known approaches, which are not based on Bayesian principles).

The main advantage of the Bayesian approach is the possibility to measure the efficiency of the alternative by a single attribute and therefore to formulate strict optimization task. The main disadvantage of this approach is that it is necessary to know a priory statistical distributions and their parameters of all the uncertain and random factors. Moreover, the formulation of the utility function is vital in preparation of the Bayesian mathematical model, since the optimization task is formulated as follows:

$$\max \int_0^T \overbrace{\int \ldots \int_\Omega}^{n} U(R)\,dF(R)\,, \qquad (1)$$

where the subject of maximization is a Stieltjes-Lebesgue integral, $U(R)$ is the utility function (see below) and $F(R)$ is the distribution function of revenues $R(r_1,...,r_n)$, T is the planning period and Ω is domain of existence of

$R(r_1,...,r_n)$. The revenues $R(r_1,...,r_n)$ depend on a number of random and uncertain factors $X\left(x_1(t),...,x_l(t)\right)$, which vary in time, as well as on the state of the network $S\left(s_1(t),...,s_k(t)\right)$, resulting in

$$R = R\left(X(t), S(t)\right). \tag{2}$$

Assuming that time can take only the discrete values $t_0, t_1, ...t_d$, (2) can be represented as γ-dimensional function:

$$R = R\left(x_{10},...,x_{l0},...,x_{1d},...,x_{ld}, s_{10},...,s_{kd}\right), \tag{3}$$

where $\gamma = (l+k)\cdot(d+1)$.

Taking into account (2) and (3) we can rewrite (1) as follows

$$\max_{s_{opt}\in S_f}\left(E\left[U,S\right] = \overbrace{\int...\int}^{l\cdot(d+1)}_{\Psi} U\left[R(X,S)\right] d\Phi(X) \right), \tag{4}$$

where $E\left[U,S\right]$ is the expected value of the utility function for the network state S, $\Phi(X)$ is the $l\cdot(d+1)$-dimensional function of probability distributions of parameters X, S_f is the set of all the feasible states of the network and Ψ is the domain of existence of the parameters X.

In the most common case the planning task is formulated as

$$U\left[R(X,S)\right] = R(X,S), \tag{5}$$

which implies that the utility can be entirely described by the revenues, for instance in monetary terms. In this case the task (1) is reduced to the revenues maximisation problem. However, such a formulation of the problem does not take into consideration the ambition of the decision-maker to reduce risk when planning high-priced objects under information uncertainty.

Therefore, in more general case

$$U\left[R(X,S)\right] \neq R(X,S). \tag{6}$$

There are methods (e.g. [6],[34]) which provide the possibility to build the utility function based on the particular task under consideration and considering the preferences of the decision-maker.

In order to reduce the planning task to the form (4) one should be able to obtain:

- The function of probability distributions $\Phi(X)$
- The relation between the revenues R and the random parameters X and feasible states S
- The utility function $U(R)$.

Furthermore, the procedures must be available for

- Calculation of the multiple integral in (4)
- Search of the parameters s_{opt}.

All the tasks listed above are complex. The sources of the complexity can be divided into the following two categories:

- Principal (informational): the difficulties in formulation of the functions $\Phi(X)$, $U(R)$ and $R(X,S)$. Formulation of the function $\Phi(X)$, which incorporates the information about the factors, which influence the solution, is a complicated, but realistic task, especially if one accepts the concept of the subjective probabilities and subjective distribution functions [36]. However, formulation of the utility function $U(R)$ for the network planning problems is a very difficult (in some cases almost impossible) task.
- Computational: Considerable difficulties are associated with calculation of the integral in (4) and arranging the corresponding optimization procedure.

A large number of publications (see for example [1], [8], [12], [13], [20], [44], [79], [85], [112], etc.) devoted to the methods of network planning illustrate the history of these methods as a search of a reasonable compromise between the degree of simplifications of the model and the reasonable precision in finding the solution approaching the "ideal" according to (4). This dissertation is not an exception and has a major goal to facilitate the search of solution approaching the Bayesian solution. For this purpose and due to the reasons, which will be explained later in this chapter, the dissertation contains the following:

- The attempts of formulation of the utility function and its direct utilization are refused. However, the importance of the utility function is emphasized and its indirect application is suggested (Chapter 3).
- The methods of evaluation of the dependencies $R(X,S)$ are considered. Instead of maximization of the revenues the task is to minimize a number of attributes, which are calculated based on known and well-developed models (Chapter 3).
- The methods and techniques of $\Phi(X)$ modeling are presented. Whenever it is possible, it is suggested to apply probabilistic models acquired either from the available statistical information or utilizing the subjective

probability distributions. The alternative modeling is recommended if probabilistic information is not available (Chapters 3 and 4).

- Direct and indirect methods for calculation of the integral (4) are analyzed (Chapter 3).
- Different optimization procedures are reviewed and the improved algorithms most suitable for the problem under consideration are suggested (respectively, Chapters 3 and 6).

3.1.2 The Utility Function

It can be assumed that some decision can result in some revenues R. It can also be assumed, that the distribution $F(R)$ is known. The function $U(R)$ is called a utility function if and only if the inequality

$$E\left[U(R)/F_1(R)\right] \leq E\left[U(R)/F_2(R)\right] \tag{7}$$

implies

$$F_1(R) \leq^* F_2(R),$$

where $\leq^*$ means that the distribution $F_2(R)$ is preferred to the distribution $F_1(R)$.

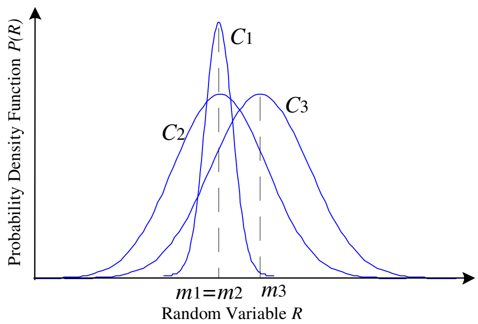


Figure 3-1 Comparison of distributions with different dispersions and means.

The utility function $U(R)$, which meets the condition (7), is usually non-linear and subjective, depending on the preferences of the person, who formulates the function. In case of considerable revenues (or losses) the function is concave, the decision-maker usually takes the risk-averse position [110]. This statement is strictly motivated in [36].

For the concave functions yields the following relation

$$E\left[U(R)\right] \le U\left[E(R)\right] \quad , \tag{8}$$

which follows from the Jensen's inequality [36]. In our case the inequality (8) can be interpreted as follows: deterministic revenue to be preferred to the expected revenue with the same mean value.

The subjective nature of the utility function can be illustrated on the well-known and widely used example of property insurance. Both the person, who owns the property, and the insurance company are ambitious to reach a deal about the insurance. Both parties have the objective to maximize their revenues. The reason why this deal is beneficial for both parties lies in different subjective utility functions they use. The person, who insures the property, prefers rather to have small negative income (insurance fee), than a very considerable loss, even at very low probability. The insurance company actions are opposite. The difference in the utility functions may be explained by considerable gap between the resources available for the parties.

The concave character of the utility function may be illustrated on the gambling example known as St. Petersburg paradox*. The expected prize in St. Petersburg game may be presented by the following infinite series:

$$\sum_n 2^n \cdot \left(\frac{1}{2}\right)^n . \tag{9}$$

Therefore, considering the possibility to win an expected prize according to (9) one should pay any amount of money to enter the game (paradox!?). In real life the reasonable payment is always limited and strictly depends on resources available to the players.

This observation, that the real-life utility function has a concave character, leads to the following important results. It is rather obvious, that the fixed income is more preferential than the expectation of the same income with some dispersion. Furthermore, in some cases the decision-maker will prefer the smaller value of fixed income rather than expected income of larger amount. This is illustrated in Figure 3-1. The solution C_1 is clearly more favorable than C_2 despite the fact that both solutions share the same mean value. The expected income is larger for C3, but the solution is more risky, therefore, C_1 still might be the preferred solution, especially if the large amounts are at stake.

---

* The St. Petersburg game is played by flipping a fair coin until it comes up tails, and the total number of flips, n, determines the prize, which equals $\$2^n$. The question is how much one should agree to pay to enter the game?

3.1.3 Why Refuse the Bayesian Approach?

Theoretically it is possible to define and accept both subjective probability distributions for uncertain parameters and for the utility functions, which would allow us to adapt pure Bayesian approach for distribution network planning. However, the conclusion is that unfortunately this approach is not applicable for the problem under consideration, taking into account the decision-making traditions and procedures in the utilities. Usually the corresponding procedure can be divided into two stages:

Stage 1 Technical and economic analysis of possible alternatives. The analysis is performed by the experts having the detailed information about the factors that influence the development strategies as well as about the possible reinforcement alternatives. At this stage it is viable to use the subjective probability distributions and define the problem in terms of monetary or some other revenues (or losses).

Stage 2 Making the final decision. The decision is made by the management of the power company, who may not be familiar with the details of the initial information. Instead, they have the experience and the authority to make a decision in tasks with considerable potential revenues or losses. Formulation of the utility function of the higher level management is a very problematic task at the initial stages of the planning process.

Therefore, the direct application of the Bayesian approach is not feasible. Furthermore, when preparing the analysis of the planning alternatives, it is insufficient to estimate the corresponding expected costs (revenues or costs). The additional information, which allows for estimation of risk (possibility) of getting the revenues that diverges from the expected mean, is needed. This information would facilitate the decision-making for the person in charge, who, may be subconsciously, will use his subjective utility function.

From the speculations stated above another important conclusion can be made. Even in cases when the planning task can be formulated as a problem of maximization of revenues (or minimization of costs), estimating reliability, power losses and other criteria by the monetary equivalents, the planner is dealing with multi-criteria optimization problem, which requires to achieve at least two global goals:

- maximize the expected value of the revenues (or minimize costs)
- minimize the risk of possible losses.

Furthermore, if the utility function is unknown, there is no clear and generally accepted definition of risk, as a result there is a number of possible approaches for measuring risk and risk management. However, taking into account the concave nature of the utility function it can be concluded that risk can be measured by the numerical value, which reflects the dispersion of the revenues around the mean value.

3.2 Means to Reduce the Complexity of the Problem

Taking into account the large number of variables in distribution planning, the complexity of the functions to be integrated and the domain of integration, equation (4) results in a formulation of the task of incredible complexity. The most important means to simplify the problem are listed in this section.

Deterministic model

In this case the difficulties associated with ambiguity of information "disappear". Consequently, there is no need to assign probability distribution functions, as well as there is no need for the utility function, therefore the optimization task can be reduced to:

$$\max_{s\in S_f} R\left(x_{10},\ldots,x_{l0},\ldots,x_{1d},\ldots,x_{ld},s_{10},\ldots,s_{kd}\right). \tag{10}$$

The problem is to organize the optimization procedure and to find the parameters of the state of the network s, which would maximize the vector of revenues R. Another problem is associated with making a decision in case if R is a vector.

It should be noted, that deterministic models are being widely used in present practice for planning of electricity networks. Moreover, the deterministic model serve as a core for the methods taking into account random and uncertain factors.

Static model

Even the problem (10) can be considerably simplified refusing the dynamics and accepting static model. Then the problem is reduced to:

$$\max_{s\in S_f} R\left(x_1,\ldots,x_l,\ldots,s_1,\ldots,s_k\right). \tag{11}$$

In (11) the number of variables has been reduced dramatically in comparison to (10). But still, the optimization problem can be sufficiently complicated.

Oversimplified model

In some cases instead of the optimization problem in (10) or (11) the planner may choose to perform an analysis of several planning alternatives. Based on experience the planner chooses a limited number of combinations among the set of possible options, checks their technical feasibility and makes a choice without the results of the extensive optimization.

Fuzzy arithmetic

One of the most powerful methods for complexity reduction is based on utilization of the Fuzzy Set Theory and Fuzzy Arithmetic. On one hand this method provides a mathematical theory to describe uncertain parameters unfeasible for a traditional modeling, on the other hand it withdraws the need for integration in (4). Thus, if the revenues R are functions of fuzzy variables $x_{F1},...,x_{FN}$, then the values in R can also be considered as fuzzy and their characteristics can be obtained using efficient and relatively simple arithmetic [143].

Methods of Monte-Carlo

If there is an ambition to solve the task (4) as strict as possible, the complicated high-dimensional integral must be solved. Practically the only method able to deal with such a problem is Monte-Carlo simulation.

Scenario approach

Complete rejection of uncertainty may lead to unfortunate decisions, therefore traditionally the network development is planned considering several possible scenarios. Then, the deterministic, fuzzy or probabilistic tasks are decided for every scenario resulting in a number of solutions. The final decision can be made based on one of the Game Theory criteria (see section 3.3).

Multi-Criteria optimization

If the utility function U is deficient and the revenues R cannot be described by a scalar function, there is a need to solve a complicated problem in multi-criteria formulation. The methods of Multi-Criteria Decision-Making are described in section 3.12.

Aggregation functions

In some cases the revenues from vector R can be aggregated into a scalar function. Those revenues, which undermine the aggregation, can be considered as constraints.

Optimization techniques

Most approaches for distribution network planning require application of the appropriate optimization procedure. A brief review of optimization algorithms applied to distribution network planning is given in section 3.3. Furthermore, Chapter 6 presents the novel applications of the optimization algorithms – multi-criteria optimization and estimation of the attributes applying Monte-Carlo simulation.

The methods listed above are dissimilar: different is the degree of complexity and the achieved accuracy. Choosing the method from this list is the typical example of multi-criteria task, while the choice represents the compromise solution. This choice can be made only based on particular task under consideration. However, the variety of the practical tasks is vary large, therefore none of the methods is absolutely the best, and on the contrary none of the methods can be declared as obsolete.

3.3 Scenarios Model

Until recently, the deterministic model was the only modeling option for distribution networks. Uncertainty in this case can be modeled by a number of scenarios, which lead to formulation of a number of deterministic tasks.

In the most general case the scenarios approach can serve as a base for solving planning tasks in presence of truly uncertain information in combination with fuzzy and probabilistic variables. Truly uncertain information is always modeled by a number of scenarios and the task is solved independently for every scenario using the corresponding methods presented below.

3.4 Decision-Making Criteria

In case of truly uncertain information modeled through scenario approach the decision can be made only by using some additional criterion. The most common decision-making criteria adopted from the Game Theory are listed bellow.

Expected-Cost Criterion (Bayes' Criterion)

Under the Expected-Cost criterion a probability or weight is associated with each scenario. The weighted average of costs of a strategy under the different scenario yields an expected cost for each strategy. If the cost associated with scenario j for the strategy i is f_{ij} and the probability of each scenario is w_j, then the selection is made as following:

$$\arg\left\{\min_i \sum_j w_j f_{ij}\right\}, \qquad (12)$$

where arg stands for "the argument of".

The advantage of this criterion is that each scenario is taken into account and the importance of the scenario is reflected trough its probability of occurrence. However, according to this criterion the solution is made without estimation of possible consequences after occurrence of a particular scenario, therefore it may lead to a risky decision.

Table 3-1 contains the results obtained after calculation of cost for each combination of scenario and strategy for some object under examination. Assuming that weights or probabilities of occurrence associated with each scenario are known the results summarized in Table 3-1 can be obtained.

Laplace's Criterion

In some cases, when the probability associated with the scenario is not available one can assume equal probabilities for each scenario. Then, the optimal solution is the one minimizing the arithmetical mean of costs over n different scenarios:

$$\arg\left\{\min_i\left(\frac{1}{n}\sum_j f_{ij}\right)\right\}. \tag{13}$$

Table 3-1 Example illustrating the Expected Cost and Laplace's criterion

Scenarios	Weights	Strategies				
		1	2	3	4	5
Scenario 1	*0.5*	55	62	60	40	54
Scenario 2	*0.25*	58	60	50	60	51
Scenario 3	*0.25*	55	43	55	65	70
Expected Cost: $\sum w_i f_{ij}$		55.75	56.75	56.25	***51.25***	57.25
Mean: $\frac{1}{n}\sum f_{ij}$		56	***55***	***55***	***55***	58.33
Minimum: $\max_i(f_{ij})$		***58***	62	60	65	70
Minimum: $\min_i(f_{ij})$		55	43	50	***40***	51

Table 3-1 contains the results obtained after calculation of cost for each combination of scenario and strategy for some object under examination. Assuming that weights or probabilities of occurrence associated with each scenario are known the results for the expected cost for each strategy can be obtained. Then, the decision according to the Expected Cost criterion corresponds to the strategy 4.

In this case the Laplace's criterion will give equal prospects to three decisions, namely strategies 2, 3 and 4 in Table 3-1.

Min-Max Criteria

a) The *Minimax* decision rule is to seek decision-maker's action, which minimize the maximum potential loss. Formally, the minimax decision is the strategy such that

$$\arg\left\{\min_i[\max_j f_{ij}]\right\}. \tag{14}$$

A decision-maker who uses the minimax criterion acts extremely conservatively. He seeks the actions that achieve the best outcome under the worst scenario. In other words, the emphasis is on the potential for extreme events.

b) The *Minimal Risk* criterion leads to the selection of the strategy, which implies the lowest extra cost under the most adverse scenario.

If the cost associated with scenario j for the strategy i is f_{ij}, while the optimal solution for scenario j is f_j^{opt}, the strategy selected according to the Minimal Risk criterion can be characterized as:

$$\arg\left\{\min_i\left[\max_j(f_{ij}-f_j^{opt})\right]\right\}. \tag{15}$$

To illustrate the application of this criterion we continue the example presented in Table 3-1. The corresponding regret matrix is presented in Table 3-2.

Table 3-2 Example illustrating Minimal Risk criterion

Scenarios	Strategies				
	1	2	3	4	5
Scenario 1	15	22	20	***0***	14
Scenario 2	8	10	***0***	10	1
Scenario 3	12	***0***	12	22	27
Maximal Risk: $\max_j(f_{ij}-f_j^{opt})$	***15***	22	20	22	27

It is important to note, that the strategy selected according to the Minimal Risk criterion in this example does not lead to the least cost solution in any one of the scenarios taken separately. The explanation is that the Minimal Risk criterion focuses on the most adverse scenario even if it has a very low probability.

Pessimism-Optimism Criterion (Hurwitz' Criterion)

The Hurwitz' criterion allows for representing the planers attitude towards risk. According to this criterion the best strategy is the one minimizing the linear combination minimal and maximal costs according to:

$$\arg\left\{\min_i\left[\lambda\cdot\max_i(f_{ij})+(1-\lambda)\cdot\min_i(f_{ij})\right]\right\}, \tag{16}$$

where $0\le\lambda<1$ is a parameter indicating planers attitude towards risk. The value $\lambda=1$ and reduces the Hurwitz' criterion to minimax criterion described above and corresponds to an extremely pessimistic decision-maker.

The value $\lambda = 0$ corresponds to an extreme optimist. The decision-maker seeks the strategy, which minimize the minimal potential loss according to:

$$\arg\left\{\min_i[\min_j f_{ij}]\right\}. \tag{17}$$

Application of Hurwitz' criterion for the case of extreme optimism is also illustrated in Table 3-1. The intermediate solutions are given by the straight line between two extremes as illustrated in Figure 3-2.

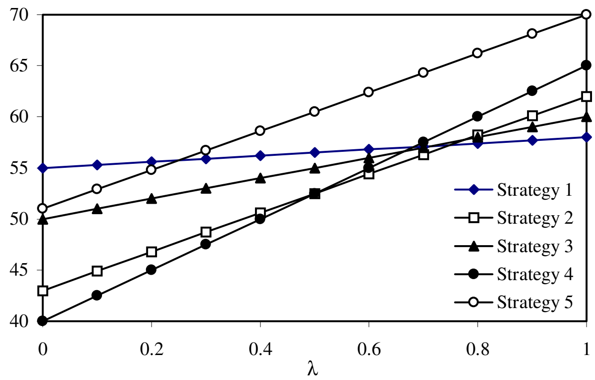


Figure 3-2 Hurwitz' criterion: strategies versus attitude towards risk

Analyzing the criteria listed above it can be stated that:

- Bayes', Laplace's and Hurwitz' criteria require data about the probabilities (weights) associated with every scenario. At the same time, if the corresponding weight would be assigned to the scenarios prior to optimization, it is possible to avoid utilization of scenarios and to apply fuzzy or probabilistic approach instead.
- Minimal Risk criterion is suitable for games with active intelligent opponent, who intentionally would choose worst for the second party conditions. However, this situation is not typical for the network planning tasks.

Therefore, the scenarios approach can be recommended only in such cases, when it is impossible to obtain information about the probabilities associated with particular scenarios.

3.5 Fuzzy Model

Fuzzy set theory, a generalization of the classical set theory was first introduced by Zadeh in [143]. Fuzziness describes sets that have no sharp transition from membership to non-membership. Fuzzy set theory provides a strict mathematical theory to describe uncertain and ambiguous structures and parameters unfeasible for a traditional modeling.

Classical Sets and Fuzzy Sets

Let X be a set with elements x, where A is called a fuzzy sub-set of X (or a fuzzy set). Membership of x in classical set A can be viewed as a characteristic function μ_A such that:

$$\mu_A(x) = \begin{cases} 1 & when \quad x \in A \\ 0 & when \quad x \notin A \end{cases}. \tag{18}$$

For a fuzzy set A of the set X the grade of membership of x in A is defined as $\mu_A(x) \in [0,1]$, where $\mu_A(x)$ is called the membership function. Fuzzy set elements are ordered pairs indicating the value of a set element and the grade of membership, i.e. $A = \{x, \mu_A(x)\}$ for $x \in X$.

Fuzzy Arithmetic

Fuzzy intervals are denoted by upper-case letters - A, B – and their α-cuts are denoted by the symbols ${}^{\alpha}A, {}^{\alpha}B$ respectively. The set of all fuzzy intervals is denoted by $\Re$.

The membership function of any $A \in \Re$ may conveniently be expressed for all $x \in \mathbb{R}$ in the canonical form:

$$A(x) = \begin{cases} f_A(x) & when & x \in [a,b) \\ 1 & when & x \in [b,c] \\ g_A(x) & when & x \in (c,d] \\ 0 & otherwise & \end{cases}, \tag{19}$$

where $a,b,c,d \in \mathbb{R}$ and $a \le b \le c \le d$, f_A is a real-valued function that is increasing and right-continuous and g_A is a real-valued function that is decreasing and left-continuous.

For any fuzzy interval expressed in the canonical form of (19), the α-cuts of A are expressed for all $\alpha \in (0,1]$ by the formula:

$$\alpha_A = \begin{cases} \left[f_A^{-1}(\alpha), g_A^{-1}(\alpha) \right] & when \quad \alpha \in (0,1) \\ [b,c] & when \quad \alpha = 1 \end{cases}, \tag{20}$$

where f_A^{-1} and g_A^{-1} are the inverse functions of f_A and g_A respectively.

The simplicity of presentation and manipulation make the trapezoidal fuzzy intervals (as well as triangular) desirable in most applications.

Any trapezoidal fuzzy interval *A* is fully characterized by quadruple $\langle a,b,c,d\rangle$ of real numbers in (19) via the special canonical form:

$$A(x)=\begin{cases}\dfrac{x-a}{b-a} & when \quad x\in[a,b) \\ 1 & when \quad x\in[b,c] \\ \dfrac{d-x}{d-c} & when \quad x\in(c,d] \\ 0 & otherwise\end{cases} \tag{21}$$

Let $A=\langle a,b,c,d\rangle$ be a shorthand symbol representing this special form. The set of all trapezoidal fuzzy intervals is denoted by $\mathfrak{R}_T$. It follows immediately from (20) and (21) that α-cuts of trapezoidal fuzzy intervals are expressed for all $\alpha\in(0,1]$ by the equation:

$${}^{\alpha}A=[a+(b-a)\cdot\alpha, d-(d-c)\cdot\alpha]. \tag{22}$$

Employing the α-cut representation, arithmetic operations on fuzzy intervals are defined in terms of the well-established arithmetic operations on closed intervals of real numbers. Given any pair of fuzzy intervals, *A* and *B*, the four basic arithmetic operations on the α-cuts of *A* and *B* are defined for all $\alpha\in(0,1]$ by the general formula:

$${}^{\alpha}(A*B)=\left\{a*b\,\middle|\,\langle a,b\rangle\in{}^{\alpha}A\times{}^{\alpha}B\right\}, \tag{23}$$

where $*$ denotes any of the four basic arithmetic operations; when the operation is division of *A* by *B*, it is required that $0\notin{}^{\alpha}B$ for any $\alpha\in(0,1]$.

Let the symbols ${}^{\alpha}[\underline{a},\overline{a}]$ and ${}^{\alpha}[\underline{b},\overline{b}]$ denote for each $\alpha\in(0,1]$ the α-cuts of fuzzy intervals *A* and *B* respectively. The endpoints $\underline{a},\overline{a},\underline{b},\overline{b}$ are functions of α defined for any fuzzy intervals by (20) and for trapezoidal fuzzy intervals by (22). The individual arithmetic operations on the α-cuts of A and B can be defined more specifically in terms of these endpoints by the following equations:

$${}^{\alpha}[\underline{a},\overline{a}]+{}^{\alpha}[\underline{b},\overline{b}]={}^{\alpha}[\underline{a}+\underline{b},\overline{a}+\overline{b}], \tag{24}$$

$${}^{\alpha}[\underline{a},\overline{a}]-{}^{\alpha}[\underline{b},\overline{b}]={}^{\alpha}[\underline{a}-\overline{b},\overline{a}-\underline{b}], \tag{25}$$

$$^{\alpha}[\underline{a},\overline{a}]\cdot{}^{\alpha}[\underline{b},\overline{b}]={}^{\alpha}[\min(\underline{a}\underline{b},\underline{a}\overline{b},\overline{a}\underline{b},\overline{a}\overline{b}),\max(\underline{a}\underline{b},\underline{a}\overline{b},\overline{a}\underline{b},\overline{a}\overline{b})],\tag{26}$$

$$^{\alpha}[\underline{a},\overline{a}]\big/{}^{\alpha}[\underline{b},\overline{b}]={}^{\alpha}[\underline{a},\overline{a}]\cdot{}^{\alpha}[1/\overline{b},1/\underline{b}]\quad if\quad 0\notin{}^{\alpha}[\underline{b},\overline{b}].\tag{27}$$

Employing the extension principle, the arithmetic operations on fuzzy intervals A and B are defined for all $c \in \mathbb{R}$ by:

$$(A*B)(c)=\sup_{c=a*b}\min\{A(a),B(b)\},\tag{28}$$

where $*$ denotes any of the four basic arithmetic operations.

Defuzzification

The comparison of fuzzy sets can be performed by means of defuzzification or by ranking. The most commonly used defuzzification methods are the center of gravity (COG) and the mean of maxima (MOM) [118].

The COG defuzzification index applied to a fuzzy set A defined over $\mathbb{R}$ and having a discrete membership function, that is $A=\{x_i,\mu_A(x_i)\}$ for $x\in\mathbb{R}$ and $i=1,2,...,N$ with N being a finite positive integer can be expressed as

$$G(A)=\frac{\sum_{i=1}^{N}x_i\mu_A(x_i)}{\sum_{i=1}^{N}\mu_A(x_i)}.\tag{29}$$

Or, when A is a fuzzy set with a continuous membership function $A=\{x,\mu_A(x)\}$ for $x\in\mathbb{R}$ then

$$G(A)=\frac{\int_{-\infty}^{\infty}x\mu_A(x)dx}{\int_{-\infty}^{\infty}\mu_A(x)dx}.\tag{30}$$

A reasonable and systematic solution for the ranking problem was proposed in [119] where the standard decision criteria for ranking crisp intervals were reformulated and generalized to become applicable to fuzzy sets over the real line. The optimistic minimax, the pessimistic maximin criteria and the Hurwitz' criterion for a parameter of ½ which are usually applied in decision-making under non-probabilistic uncertainty, i.e. when no probability distributions are available, were reformulated and generalized to fuzzy sets. Particularly, the Hurwitz' criterion for a parameter of ½ was reduced to the form of a ranking index and called the total distance criterion (TDC) which maps each fuzzy set to a crisp number. This is done by using integration along the membership axis of the arithmetic mean value of the α-cut of a fuzzy set considered. That is a convex and normal fuzzy set *A* over the real line is mapped to *F(A)* by:

$$F(A) = \int_0^1 M(\alpha_A) d\alpha \qquad (31)$$

where $M(\alpha_A)$ is the arithmetic mean of the α-cut of *A*.

Use of fuzzy numbers corresponds to performing sensitivity analysis on all uncertain parameters simultaneously.

3.6 Probabilistic Methods

Bayesian formulation of the optimization task is based on probabilistic positions. If the probability distributions of the functions in (4) are known, only the algorithm for calculation of the expected value of the utility function (or characteristics of revenues/losses if the utility function is unknown) is needed.

For this purpose three common methods can be identified:

- Methods of multiple integration [19].
- The *method of moments* based on the assumption that the density function of a random variable can be denoted as an infinite Taylor series with statistical moments. The method allows for reaching a high degree of accuracy, but is complicated and requires a function, which can be approximated by Taylor series.
- The *method of Monte-Carlo* implies repetition of deterministic calculations a large number of times. For each simulation the values of random variables are chosen via random number generator. The weakness of this method is a large computation efforts, since more accurate results require a large number of trials. At the same time often this is the only method, which gives the "real" estimation of the resulting probability distribution.

Recognized that the method of Monte-Carlo has significant advantages in comparison with the first two methods, therefore this was chosen for application in the planning algorithm considered in this dissertation.

If the model contains random parameters, the wanted quantities (power losses, nodal voltages and power flows) are also random. There is a variety of probabilistic load flow algorithms suggested in the literature. Thus, in [30] statistical load flow is applied in order to verify the compliance of the selected planning alternatives with technical constraints. The detailed load model is used: the year is represented by means of 14 standard days and 30 customer classes have been selected. The load curve of each typical day has been subdivided into 96 intervals, thus the mean value and the variance is determined for each quarter of an hour. Then the load is represented as a random variable. To represent unknown random variables, which are functions of load, a series expansion around the mean value has been performed.

It is shown in [122] how the energy losses can be approximated via statistical moments. From a statistical point of view the daily load curves are the time series of the chronological observations of power demand. They are specific for different kinds of customers and give the variation of the power supply throughout a 24-hour period. Daily load curves are treated as random variables, which allows for calculating their statistical moments. Knowing the moments a simple formula can be used to obtain the losses in any part of the network.

Probabilistic Load Flow (PLF) for transmission network expansion planning is described in [76]. The PLF is efficiently solved by Monte-Carlo simulation techniques and linearized load flow equations.

3.7 Methods of Monte-Carlo

3.7.1 Definitions

Several definitions of methods of Monte-Carlo can be found in the literature. Two of them are given bellow. Thus, the methods of Monte-Carlo comprise [19]:

- Modeling of random variables performed in order to calculate the characteristics of their probability distribution
- Approximate calculation of the integrals.

The method in its both definitions suits entirely for solution of the network optimization task given by (1).

Calculation of the probability distributions based on the modeling of random variables consists of the following tree stages:

Stage 1 Generate the random numbers according to the given distribution law

Stage 2 Model the functions of random variables

Stage 3 Estimate the characteristics of probability distributions.

3.7.2 Random Numbers Generators

The first stage of the Monte-Carlo algorithm can be built based on one of the many methods, which allow for simulation of the random values with almost arbitrary distribution law. The most frequently used algorithms are based on random numbers generators with uniform distribution, which are the standard function in the most high level programming languages, and the von Neumann algorithm [97], which allows for generation of the random numbers with any given probability density function.

In certain cases in order to obtain the variable distributed according to some particular distribution laws (normal, exponential, etc.) computationally efficient procedures can be created.

3.7.3 Modeling Probabilistic Variables

The problem formulated according to (1) involves modeling of revenues (or losses) depending on a number of factors. For this purpose well-developed methods and even commercial software packages, which perform deterministic estimation of the required functions (for instance, load flow and short circuit calculations, reliability assessment, etc.) can be used. The basics of these methods are presented in Chapter 4.

The first two stages of the Monte-Carlo method result in a set of function values $r_1,...,r_N$, based on which the parameters of probability distribution can be estimated. For example, the expected (mean) value of the random variable can be estimated as:

$$E\left[r\right]\approx\frac{1}{N}\sum_{i=1}^{N}r_i \quad . \tag{32}$$

The error in estimation of (32) to a great extent depends on the number N. Assuming that the sum $s=\frac{1}{N}\sum_{i=1}^{N}r_i$ has a normal distribution, the error of can be estimated based on the Chebyshev inequality, which for the particular case can be reduced to the form [19]:

$$P\left\{\left|\frac{1}{N}\cdot\sum_{i=1}^{N}r_i-E\left[r\right]\right|<\frac{\sigma}{\sqrt{N\gamma}}\right\}>1-\gamma\,, \tag{33}$$

where $E\left[r\right]\approx\frac{1}{N}\sum_{i=1}^{N}r_i$ is an expected value of r and γ is a small probability.

It follows from (33) that the value $\frac{1}{N}\sum_{i=1}^{N}r_i$ differs from the mean value $E\left[r\right]$ no more than by $\sigma/\sqrt{N\gamma}$ with a probability $1-\gamma$. Therefore, for the given σ and γ the error declines as $\sqrt{N}$. The main disadvantage of Monte-Carlo methods is that the precision of the method depends significantly on number of trials and declines slowly as this number increases.

The theory of Monte-Carlo is to a large extend devoted to the methods of precision enhancement without increasing the number of trials N. Such methods of manipulating the simulation so as to improve the accuracy of the estimators are known as *Variance Reduction Techniques* (VRTs) [77]. The following two techniques are used in this dissertation:

- Method of importance sampling [19]
- Method of common random numbers [77]

The idea of the importance sampling consists of concentrating the distribution of sample points in the parts of the integration domain that are of most "important" instead of spreading them randomly. The idea of the method can be extended, which results in decomposition of the problem into two parts one of which is solved applying Monte-Carlo and another one can be solved by more simple methods. Projection of this idea on distribution network planning result in the algorithm, which consists of the following two stages:

Stage 1 Use the simplified methods (possibly deterministic and static) to reject a large number of unfeasible and obviously inferior alternatives

Stage 2 Solve the problem (4) in the domain, which was considerably limited on the first stage.

More detailed description of this algorithm can be found in Chapter 6.

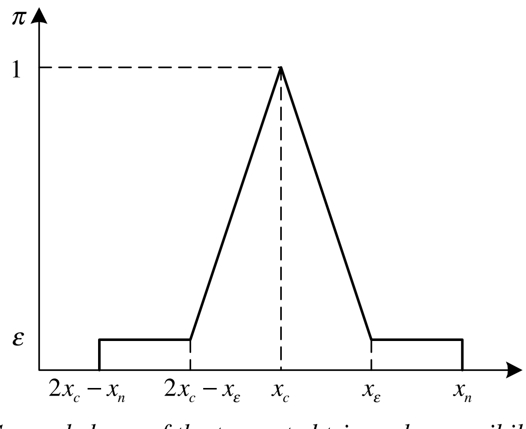


Figure 3-3 General shape of the truncated triangular possibility distribution

The common random numbers method is used when estimating the difference in the expected performance measures of more than one system. The rationale is that the same underlying random numbers are used for all the system under consideration, which means that the noise or experimental conditions will be the same for all the systems. The observed differences will be only due to the differences between the systems and not to the fact that one has been more lucky than the other in picking its random numbers. The idea of common random numbers is used in organization of the search procedure in noisy GA (section 6.4).

3.8 Probability - Possibility Transformations

One of the most controversial issues in uncertainty modeling and information sciences is the relationship between probability theory and fuzzy sets. The literature pertaining to this debate is surveyed in [38]. In [39] the same authors propose an optimal transformation – the transformation which looses as little information as possible.

Let X be a universe of discourse*, P the probability measure associated to the probability distribution p and Π the possibility measure associated to the possibility distribution π. The transformation gives the possibility distribution, which has the smallest area under the following conditions:

- the consistency principle

$$P(A) \leq \Pi(A), \qquad \forall A \subset X \tag{34}$$

- the preference principle

$$\begin{aligned} &p(a) < p(b) \Rightarrow \pi(a) < \pi(b) \\ &p(a) = p(b) \Rightarrow \pi(a) = \pi(b) \qquad \forall a, b \in X \end{aligned} . \tag{35}$$

The transformation for the continuous case is defined below.

There is one solution for a unimodal continuous probability distribution p with support $[a, b]$ such that p is increasing on $[a, x_0]$ and decreasing on $[x_0, b]$, and x_0 is the modal value of p. Let a function $f\colon [a, x_0] \to [x_0, b]$ be defined by $f(x) = \max\{y / p(y) \geq p(x)\}$. Then the optimal solution transforming p into a possibility distribution π is given by the following equation:

$$\pi_{opt}(x) = \pi_{opt}(f(x)) = \int_{-\infty}^{x} p(y)dy + \int_{f(x)}^{\infty} p(y)dy . \tag{36}$$

The value $\pi_{opt}(x)$ can be interpreted as the confidence degree associated with the confidence interval $[x, f(x)]$ when the "one point" estimation is x_0, the modal value of $p(x)$. Then the cut of level α of the resulting fuzzy subset is the confidence interval of confidence level $1\text{-}\alpha$. Therefore, the considered possibility distribution is a collection of the confidence intervals with a center

* Universe of discourse is the subject of the database or a model: a part of the "word" under discussion.

at the modal value. For example, for the real value to be inside the confidence interval [x_1, $f(x_1)$] with $\pi(x_1)=\pi(f(x_1))=0.01$ the probability is given by the confidence degree $1-\pi(x_1)=0.99$ (or 99%).

In this dissertation, the simple transformation suggested in [74] has been adopted. The transformation is called truncated triangular transformation and is an approximation of the optimal transformation. The general shape of the transformation applied to symmetric probability functions is depicted in Figure 3-3.

The corresponding curve is defined as:

$$\pi(x)=\begin{cases} 1-\dfrac{1-\varepsilon}{x_\varepsilon - x_c}\left|x-x_c\right| & if \quad \left|x-x_c\right| \le (x_\varepsilon - x_c) \\ \varepsilon & if \quad (x_\varepsilon - x_c) \le \left|x-x_c\right| \le (x_n - x_c) \\ 0 & if \quad \left|x-x_c\right| \ge (x_n - x_c) \end{cases} \tag{37}$$

Table 3-3 Parameters of the truncated triangular possibility distribution for four probability distributions and of the generalized distribution

	x_c	x_n	x_ε	ε
Gausian distribution $p(x)=\dfrac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-x_m)^2}{2\sigma^2}}$	x_m	$x_m+2.58\sigma$	$x_m+1.54\sigma$	0.12
Exponent distribution $p(x)=\dfrac{1}{\sigma\sqrt{2}}e^{-\frac{\sqrt{2}\left|x-x_m\right|}{\sigma}}$	x_m	$x_m+3.2\sigma$	$x_m+1.46\sigma$	0.13
Triangular distribution $p(x)=\begin{cases}\dfrac{1}{\sigma\sqrt{6}}-\dfrac{\left|x-x_m\right|}{6\sigma^2} & if \quad \left|x-x_m\right|<\sigma\sqrt{6} \\ 0 & otherwise\end{cases}$	x_m	$x_m+2.45\sigma$	$x_m+1.63\sigma$	0.11
Uniform distribution $p(x)=\begin{cases}\dfrac{1}{2\sigma\sqrt{3}} & if \quad \left|x-x_m\right|<\sigma\sqrt{3} \\ 0 & otherwise\end{cases}$	x_m	$x_m+1.73\sigma$	$x_m+1.73\sigma$	0
Generalized distribution	x_m	$x_m+3.2\sigma$	$x_m+1.73\sigma$	0.086

In [74] the truncated triangular possibility distribution was applied on four symmetric probability density functions, namely normal, exponential, triangular and uniform distribution. In this case the resulting possibility distribution can be described by four parameters x_c, x_n, x_ε and ε. The choice of these parameters is based on:

- x_c coincides with the mean value x_m
- x_n contributes to $P[2x_m - x_n] = 0.99$, where *P* is the probability measure associated with the probability density function *p*; thus the fuzzy subset is bounded by 99% confidence interval
- x_ε corresponds to the minimum area, while the condition (34) is satisfied for the possibility distribution
- ε is deduced from x_ε as follows

$$\varepsilon = \pi_{opt}(x_\varepsilon) = 2\int_{x_\varepsilon}^{\infty} p(t)dt\ . \tag{38}$$

It can be shown mathematically that the parameter ε does not depend on σ, therefore the value ε is the same for any σ of a given law. The results of the transformation for four distribution laws are summarized in Table 3-3.

The parameters of the generalized distribution can be applied to any of the four probability laws.

3.9 Fuzzy-Probabilistic Model

The situation when information is uncertain due to several uncertainty sources is very common for planning tasks. If these sources are of different nature it can in some cases be an irrational approximation to choose one model over another to represent the uncertainty.

One of the most prevalent examples of such situations is the following. It is known that the variable is stochastic and follows some probability density function, however, the knowledge about the parameters of this function is also vague. This vagueness can be modeled, for instance, via representation of PDF parameters by fuzzy numbers [6]. The example of merging one type of uncertainty into another resulting in fuzzy-probabilistic model is given below.

Consider a normally distributed random variable X with μ and σ as its mean and standard deviation, respectively. The probability density function $g(x;\mu,\sigma)$ is

$$g(x;\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]. \tag{39}$$

The next assumption is that the mean of random variable X is uncertain and can be described by a triangular fuzzy number as shown in Figure 3-4. The relationship between α and μ can be obtained as follows:

$$\alpha(\mu)=\begin{cases} h\cdot\dfrac{\mu-a}{b-a} & when \quad \mu\in[a,b) \\ h\cdot\dfrac{c-\mu}{c-b} & when \quad \mu\in(b,c] \end{cases}, \tag{40}$$

where h is the height of the triangle and $\alpha(\mu)$ is the weight function of the mean value μ. When $h=1$ $\alpha(\mu)$ is the membership function of μ. The function $\alpha(\mu)$ can also be seen as the PDF for the random variable μ; therefore the area of the triangle must be equal to 1. Employing this condition the height of the triangle can be determined as

$$\int_{-\infty}^{\infty}\alpha(\mu)d\mu=\int_{a}^{c}\alpha(\mu)d\mu=1=\frac{h(c-a)}{2}\Rightarrow h=\frac{2}{c-a}. \tag{41}$$

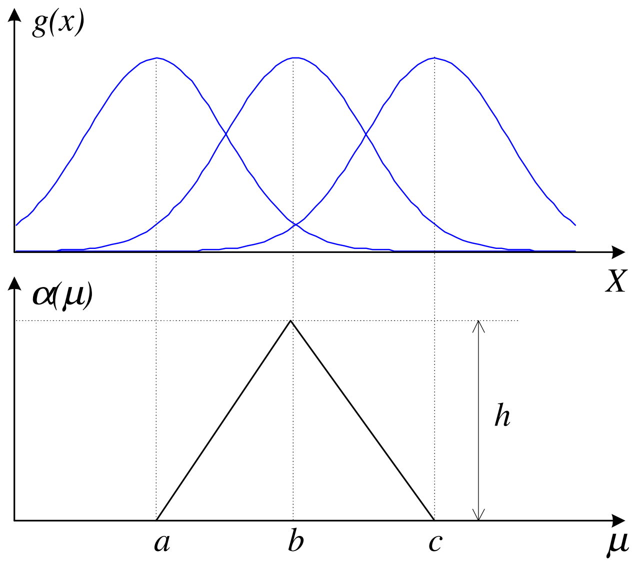


Figure 3-4 Random variable X with fuzzy mean value $\mu=\langle a,b,c\rangle$

Substituting h into (40) produces the normalized membership function

$$\alpha(\mu)=\begin{cases}\dfrac{2}{c-a}\cdot\dfrac{\mu-a}{b-a} & when \quad \mu\in[a,b)\\ \dfrac{2}{c-a}\cdot\dfrac{c-\mu}{c-b} & when \quad \mu\in(b,c]\end{cases}. \tag{42}$$

It can be shown that the triangular distribution in (42) has the following mean value $\hat{\mu}$ and variance $\hat{\sigma}^2$ respectively:

$$\hat{\mu}=E[\mu]=\frac{a+b+c}{3} \tag{43}$$

and

$$\hat{\sigma}^2=Var[\mu]=\frac{a^2+b^2+c^2-ab-bc-ac}{18}. \tag{44}$$

A joint probability density function $f(x;\mu,\sigma)$ can be obtained from multiplication of the $g(x;\mu,\sigma)$ and the $\alpha(\mu)$ as

$$f(x;\mu,\sigma)=\alpha(\mu)\cdot g(x;\mu,\sigma)=$$
$$=\begin{cases}\dfrac{2}{c-a}\cdot\dfrac{\mu-a}{b-a}\cdot\dfrac{1}{\sigma\sqrt{2\pi}}\exp\left[-\dfrac{(x-\mu)^2}{2\sigma^2}\right] & when \quad \mu\in[a,b)\\ \dfrac{2}{c-a}\cdot\dfrac{c-\mu}{c-b}\cdot\dfrac{1}{\sigma\sqrt{2\pi}}\exp\left[-\dfrac{(x-\mu)^2}{2\sigma^2}\right] & when \quad \mu\in(b,c]\end{cases} \tag{45}$$

The probability density function can be obtained integrating (45) over μ:

$$f_X(x)=\int_{-\infty}^{\infty}f(x;\mu,\sigma)d\mu=\int_a^c f(x;\mu,\sigma)d\mu=\int_a^c\alpha(\mu)\cdot g(x;\mu,\sigma)d\mu. \tag{46}$$

Finally, by definition the expectation and variance of X based on $f_X(x)$ can be shown to be

$$E[X]=\int_{-\infty}^{\infty}x\cdot f_X(x)dx=\int_{-\infty}^{\infty}x\left(\int_a^c f(x;\mu,\sigma)d\mu\right)dx=\frac{a+b+c}{3} \tag{47}$$

and

$$Var[X]=\int_{-\infty}^{\infty}(x-\overline{x})^2\cdot f_X(x)dx=\sigma^2+\frac{a^2+b^2+c^2-ab-bc-ac}{18}. \tag{48}$$

It can be observed that the expectation of a fuzzy-probabilistic variable is equal to the mean value of its fuzzy parameter and the variance of the variable is equal to the sum of variance for random variable and the variance of its fuzzy mean.

Similarly, even more complicated uncertainties – for example, if standard deviation of the random variable is fuzzy - can be merged into one model.

3.10 Fuzzy Arithmetic versus Method of Monte-Carlo

Existence of the probability-possibility transformation presented in section 3.8 provides the opportunity to choose one of the following two approaches to the problem of network planning:

- Application of Fuzzy variables and Fuzzy arithmetic
- Method of Monte-Carlo.

Both methods allow for consideration of uncertain and random factors.

Method of Monte-Carlo has only one disadvantage in comparison with fuzzy approach: it requires considerable computational efforts. For a long time this disadvantage was decisive. Only a few cases applied the method of Monte-Carlo in optimization of real electric power networks [76]. However, development of computational capacities provides the possibility for practical application of Monte-Carlo methods in complicated optimization problems. This statement is based on two tendencies:

- Rapid development in computational speed and capacities
- Development of systems suitable for organization of parallel computations.

It should be noted that method of Monte-Carlo fits ideally for parallel computations. In fact, to provide the required accuracy a large number of samples N is needed. For this calculations m computers can be engaged. Then, the computational time will be reduced approximately by a factor of m, since the feature of the method is that there is no need for interaction and information exchange between the processors working in parallel.

The choice of Monte-Carlo method does not mean that utilization of fuzzy information and fuzzy variables is rejected. These variables can be and must be used if the statistical information is not available. However, here it is suggested to transform the membership function into probability distribution and apply the last for organization of the statistical samples.

3.11 Review of Optimization Algorithms Applied to Distribution Network Planning

3.11.1 Historical Review

During the last decades several methods were proposed for distribution system planning [8], [9], [17], [22], [26], [44], [49], [58], [59], [62], [63], [72], [75], [82], [85], [104], [105], [108], [114], [127]. These techniques have evolved together with the development of scientific knowledge and benefiting from the increase in computational capacities. The history of these methods is the history of the conflict between the precision of the model (number and type of simplifications) and the computation efficiency of the solution method for the model.

The first computer-aided distribution network planning tools were presented in the seventies [8],[75]. Different optimization techniques were first applied only to simplified models. During the last two decades a lot of research efforts have been made to include more details in the models. In chapter 4 the formulation of the power network planning task will be discussed and the problem identified as inherently non-linear – first of all due to variable costs, which are squarely dependent on loads. Linearization of objective function was applied by several researchers [44],[49],[127],[132]. In [132] the problem is formulated as a transportation problem and linear programming is applied as an optimization tool. In [44] and [49] the authors apply mixed-integer programming to a linearized model. However, besides resulting in less accurate solution, disadvantages of this simplification are that it has restricted application and is not suitable for network reinforcement problems. As a result, most of the present network planning models utilize nonlinear cost functions [9],[22],[26],[58],[105], therefore applying either more complex optimization techniques, namely nonlinear programming methods [105] or different kinds of heuristic algorithms [9],[22],[26],[58].

In some models first applied to the network planning, the problem was decomposed into two subsystems: substations subsystem and feeders subsystem [8],[75]. In these models it is assumed, that the problem of optimal substation allocation and sizing to be solved, based on which the optimal feeder routing can be provided. This approximation was avoided in later models.

The importance of time consideration is obvious in planning tasks. However, dynamic problem formulation results in dramatic increase of computational efforts. Until recently, most of the models considered the study period as a single stage, providing so called “horizon year” planning [8],[44],[105],[132]. To expand the study period over several time-stages a number of either dynamic models [34],[71],[84],[104] or pseudo-dynamic models

[9],[72],[114],[127] were introduced. In [34],[71],[104] dynamic programming is applied as an optimization tool, while in [49],[62] mixed-integer programming, which in the last reference is combined with Bender's decomposition. Some kind of heuristic algorithm is utilized in most of pseudo-dynamic models. State-of-the-art is that timing issue is present in most of the recent studies, e.g. [20], [71],[93],[112].

Traditionally only deterministic methods have been applied to the distribution network planning. However, uncertainty is a natural attribute of long-term planning task and since the eighties there are some attempts to take it into account. Starting from some rather simple intuitive studies and sensitivity analysis, the model of network planning under uncertainty becomes much more complex in later investigations [63],[71],[85]. Thus, in [63] scenario representation of future forecasts is applied with subsequent estimation according to the criterion of minimal risk. Minimal risk analysis is performed also in [84],[85], however uncertain parameters are modeled by fuzzy numbers.

In the nineties the necessity was recognized to deal with multi-criteria nature of the problem [85]. This fact together with exceptional complexity of the model requires new optimization tools. Furthermore, it changes the whole philosophy of network planning applied so far: instead of search for a single optimum the search in multi-objective domain should be provided resulting in a set of competitive solutions, and none of which is optimal. Firstly, came evolutionary algorithms, which seems to be feasible for such problems [47],[87],[99].

This work does not aim at a detailed description of the whole variety of models and algorithms that have been applied to the problem. Nevertheless, the main techniques for distribution planning and their advantages and limitations are presented and discussed.

3.11.2 Traditional Mathematical Optimization Methods

Numerical optimization may be considered the traditional approach for optimization. Depending on problem formulation, it can involve either use of algorithms tailored to discrete or continuous analysis. Regardless, numerical optimization applies computed numerical formula and procedure to search for the optimal solution.

General matrix formulation of mixed-integer model can be represented as follows:

$$\min c_1 x + c_2 y$$

$$\begin{array}{lcl} subject \quad to: & & \\ A_1 x + A_2 y & \geq & b \\ A_3 y & \geq & d, \end{array} \tag{49}$$

where x is the vector of continuous variables containing power flows, power supplies and voltage drops and, y is the vector of integer decision variables. Cost coefficients c_1 and c_2 reflect fixed and variable costs associated with both integer and continuous variables. Matrices A_1, A_2, A_3 as well as right hand side constraints vectors b, d depend on constraints of the problem and can be derived from the problem formulation.

The weak point of this formulation is that linearization of quadratic terms is required.

The task in form of (49) for multi-stage distribution network planning was presented in [49],[62],[114], where mixed-integer programming was used as an optimization tool. Furthermore, in order to spare computer time and to speed up calculations Bender's decomposition was applied to the mixed-integer model in [62]. To separate continuous and integer variables, (49) can be rewritten as:

$$\min \left\{ c_2 y + \min \left[c_1 x, subject \;\; to{:} A_1 x \geq b - A_2 y \right], subject \;\; to{:} A_3 y \geq d \right\} \tag{50}$$

with subsequent formulation of dual problem to the inner problem:

$$\begin{array}{l} \max \; u(b - A_2 y), \\ subject \quad to{:} \\ uA_1 \leq c_1, \; u \geq 0. \end{array} \tag{51}$$

As a result, the Master problem, the outer problem, contains only integer variables and the inner problem has to deal only with continuous variables.

Decomposition approach simplifies the optimization task. However, as any simplification, it has disadvantages from the point of view of accuracy of the calculation. Another weak point of decomposition approach is the complexity dealing with multiple criteria.

The main advantage of numerical optimization approaches is the convergence, at least in theory, to the optimal solution and not just to a "good" solution. However, methods based on this type of optimization can, in practice, hardly be applied to real dimension cases. This is due to extreme mathematical and computational complexity introduced by the discrete and non-linear nature of the problems to be considered.

3.11.3 Heuristic Methods

3.11.3.1 Branch Exchange Methods

A large class of methods, which are widely applied both to the transmission and distribution network planning can be related to the class of heuristic methods. The major part of them is based on the implicit enumeration.

Mathematically, the general problem formulation can be represented by (49), where the task is to define a vector of state variables x, and one of the decision variables y minimizing the objective function.

The idea behind the algorithms is that decision and state variables can be separated. Then for every network configuration defined by decision variables the state variables can be calculated. A search algorithm is applied to find the optimal configuration.

According to the literature [26] these heuristic approaches can be classified according to the search space exploration method: constructive and destructive (greedy search) or branch exchange approaches. The common principle used can be described as follows: starting from some reasonable initial plan, the current configuration is replaced with one of its neighbors, which is obtained by applying an elementary modification to the current configuration.

The different kinds of algorithms preferring branch exchange approach to other heuristic techniques are probably the most prevalent for distribution network planning [8],[72]. The search starts from some feasible configuration (radial) and by opening and closing branches (one at the time), it accepts configurations if the objective function is improved and rejects otherwise. An example of branch exchange procedure is illustrated by Figure 3-5.

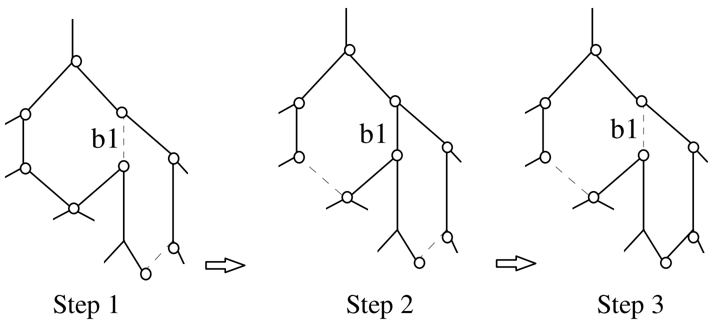


Figure 3-5 An example of Branch Exchange approach

As it was stated, besides siting and sizing in network planning problem we have to cope also with the problem of timing which is much more challenging. A number of heuristic approaches suggest different combination of forward/backward procedures. The problem is decomposed into several single-year sub-problems and each sub-problem is solved independently which gives pseudo-dynamic solutions [9],[72]. In [72] the advantages of multi-year heuristic optimization approach in comparison with single-year planing approach are shown.

The main advantage of heuristic algorithms is that a good solution can be found for the real-size (large) network with comparably small computational effort. But the global optimum cannot be guaranteed, especially for time-variant tasks due to their pseudo-dynamic nature.

3.11.3.2 Simulated Annealing

Some researchers relate to the Heuristic Methods the optimization technique known as Simulated Annealing. It is based on the analogy between the simulation of annealing in solids and the problem of solving large combinatorial optimization problems. The objective function is referred to as the energy function. The system to be optimized starts at a high temperature and is cooled down until the system freezes and reaches the global optimum.

The algorithm can be illustrated by the following three steps:

Step 1 Generation of candidate solutions by perturbation of current solution according to probabilistic distribution function;

Step 2 Acceptance test of solutions. A new solution is accepted as current when its cost is lower than that of the current solution. If cost is higher, a new solution is accepted with a probability of acceptance:

$$P_r(\Delta F) = 1/(1+e^{\frac{\Delta F}{t}}), \tag{52}$$

where ΔF is the increment of cost of the new solution compared to the current solution and t is the temperature level.

Step 3 Iterative procedure. The last accepted candidate solution becomes the initial solution for the next iteration. The temperature of the next iteration is reduced according to the cooling schedule:

$$t_k = r^{(k-1)} t_0, \tag{53}$$

in which t_k is the temperature at the k^{th} iteration, t_0 is the initial temperature and r is the temperature reduction rate ($0 < r < 1$).

The iterative process is terminated when there is no significant improvement in the solution or the maximum allowable number of iterations is reached.

Application of Simulated Annealing approach to distribution network planning is presented in [59] and [108]. Another application of Simulated Annealing to combinatorial planning problem is illustrated in [27],[28], where the optimal capacitor placement problem is addressed.

A basic characteristic of Simulated Annealing is that the quality of the final solution does not depend on the initial configuration. It can be shown mathematically that the algorithm converges asymptotically to the global optimal solution with probability one. Although this may turn out to be computational expensive, it is a valuable feature of the approach. Normally, in practice, a faster solution could be obtained with faster cooling schemes, which may yield optimal solution. Another important feature of Simulated Annealing as well as another heuristic approaches is that there are no special requirements of the model; the problem can be modeled as non-linear, non-differential and constrained.

3.11.3.3 Tabu Search

The basic concept of Tabu Search as described in [47] is a meta-heuristic superimposed on another heuristic. The overall approach is to avoid entrapment in cycles by forbidding or penalizing moves which take the solution, in the next iteration, to points in the solution space previously visited (hence “tabu”). The method is still actively researched, and is continuing to evolve and improve. The Tabu method was partly motivated by the observation that human behavior appears to operate with a random element that leads to inconsistent behavior given similar circumstances. Thus, the resulting tendency to deviate from a charted course might be regretted as a source of error but can also prove to be a source of gain. The Tabu method operates in this way with the exception that new courses are not chosen randomly. Instead the Tabu search proceeds according to the supposition that there is no point in accepting a new (poor) solution unless it is to avoid a path already investigated. This insures new regions of a problem’s solution space will be investigated in with the goal of avoiding local minima and ultimately finding the desired solution.

The Tabu search begins by marching to local minima. To avoid retracing the steps used, the method records recent moves in one or more Tabu lists. The original intent of the list was not to prevent a previous move from being repeated, but rather to insure it was not reversed. The Tabu lists are historical in nature and form the Tabu search memory. The role of the memory can change as the algorithm proceeds. At initialization the goal is to make a coarse examination of the solution space, known as “diversification”, but as candidate locations are identified the search is more focused to produce local optimal solutions in a process of “intensification”. In many cases the differences between the various implementations of the Tabu method have to do with the size, variability, and adaptability of the Tabu memory to a particular problem domain.

The following steps can illustrate the basic algorithm:

Step 1 Initialization:

- ✓ Select an initial solution $x_{now} \in X$;
- ✓ Initialize the best with the initial solution $x_{better} = x_{new}$;
- ✓ Initialize the tabu list H with x_{now} .

Step 2 Search:

- ✓ Determine a neighborhood of $x_{now} \in N(x_{now})$;
- ✓ Select a subset $Candidates\_N(x_{now}) \subset N(x_{now})$;
- ✓ Evaluate each solution $x_{new} \in Candidates\_N(x_{now})$ and choose the best according to the objective function $F(H, x_{new})$;
- ✓ Store the best solution $x_{now} = x_{new}$;
- ✓ If x_{now} is better than x_{better} , then assign $x_{better} = x_{now}$;
- ✓ Update the history of the search H with x_{now} .

Step 3 Termination of the process.

- ✓ Stop the process if the termination criterion is verified, otherwise return to the *Step 2*.

Application of Tabu Search to distribution network optimization is presented in [92] and [114].

3.11.4 Dynamic Programming (DP)

The methods based on dynamic programming seem to be very attractive since they naturally allow for representing the dynamic nature of the development process. Another advantage of the method is that there is no need for linearization of the objective function used in the optimization process. This also means that the objective function can contain present values of costs, which reduces the influence of the investments made far in the future. Thus, the decisions made for the nearest future will be correct, but the decisions for the distant future can be corrected when more accurate forecasts are available.

The only challenge, which makes DP not applicable to the real-size network planning problem is the so-called “curse of dimensionality”; the method demands very big computational effort for large dimension problems. On the other hand, when talking about network planning, in many cases it means reinforcement of existing network. These are the types of tasks where dynamic programming could be applied efficiently.

The planning task could be represented as a graph where the nodes represent particular states of the network and the branches represent certain investments made to reinforce the network (realized actions) when moving from one state to another. Each column depicts a certain time stage and each horizontal line one possible action (Figure 3-6). For a particular task some of the graph branches can be absent corresponding to the logical (or others) constraints. On the other hand, some investments can be made simultaneously, in which case the graphical problem representation would not be so obvious.

The idea behind DP is that the decision at the t^{th} stage is obtained from the decision made at stage $(t-1)$ minimizing the transfer cost of moving from the starting point to this stage, which mathematically can be expressed as follows:

$$F(t,e) = \min_{\{G(t,e)\}} \left[g(0,e(0)) + g(1,e(1)) + \ldots + g(t,(e(t))\right], \quad (54)$$

where $\{G(t,e)\}$ is the set of acceptable strategies during the time t and until state e is reached and, $g(t,(e(t))$ is the component of the objective function at t^{th} stage for the state $e(t)$.

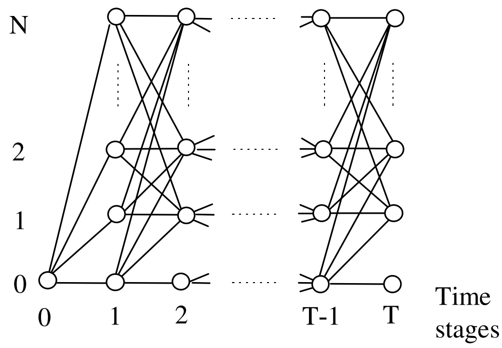


Figure 3-6 Dynamic Programming graph for network reinforcement problem

Furthermore, it can be shown [34],[70] that (54) can be reduced to the following recursive equation of DP:

$$F(t,e) = g(t,e) + \min_{\{e(t-1)\subseteq e\}} F\big(t-1, e(t-1)\big), \quad (55)$$

where $\{e(t-1) \subseteq e\}$ stands for the set of states $e(t-1)$ from which the transition to state e is feasible.

Then the optimization process can be accomplished by decision of some set of equations according to (55) minimizing the objective function for the period from the initial to the final stage.

In order to overcome the difficulties connected with high dimensions, there are attempts to reduce computational capacities needed for realization of the dynamic programming method. For instance, the modified method of dynamic programming called the Optimal Initial States method [34],[70],[71] which actually is a heuristic time-variant optimization algorithm based on dynamic programming. The idea behind this algorithm is that as dynamic optimization proceeds at each stage, only some states could lead to the optimal solution. Only these states, called Optimal Initial States, should be kept for further consideration. It gives the great savings in computer time and memory. It is proven that Optimal Initial States for the particular task of power system planning could be found by applying technical economic characteristics of the power object which are regular.

A number of researchers have applied DP to the distribution network reinforcement problem [34],[71] and [104] attracted by its advantageous features. However, the problem of the "green-field" network planning (planning of a new network) is likely to be addressed by some other optimization techniques.

3.11.5 Evolutionary Algorithms

3.11.5.1 An Overview

The complexity introduced by planning concepts such as uncertainties, multiple objectives, etc., associated with the combinatorial complexity of the problem, lead to the perception of the limitation of the traditional methods referred to in the previous points.

The technique known as Evolutionary Algorithms (EA) includes several algorithms, which share the same conceptual base of simulating the evolution of individual structures by processes of selection, mutation and recombination. These processes depend on the perceived performance of each one of these structures in a certain environment.

Interest in EA-related research [98] as well as in EA application in Power Systems [17],[82],[84],[86],[87] and [113] increased rapidly in recent years. The main advantages of these methodologies can be summarized as follows:

- ✓ search is performed starting from several points and is based on probabilistic transition rules; consequently, there is less chance for convergence to a local optimum
- ✓ EA do not require "well behaved" objective functions, discontinuities can be tolerated
- ✓ EA are very good for multi-criteria optimization.

These features have caused increased interest in EA application as an optimization tool to the tasks (non-differentiable, discrete, and non-convergent) where it is difficult to apply any other optimization method.

Different variants of EA exist, the most popular of which are the following:

- ✓ Genetic Algorithms (GA)
- ✓ Evolution Strategies (ES)
- ✓ Evolutionary Programming (EP)
- ✓ Genetic Programming (GP)
- ✓ Classifier Systems (CS).

In application to Power Systems, the most attention among EA has been received by GA [17],[82],[84],[86] and [113].

3.11.5.2 Genetic Algorithms

The standard GA operates on a population of binary strings, referred to as "chromosomes", which consist of bit-genes, referred to as "individuals". Each individual represents a solution coding all the decision parameters. A population of individuals is replaced during each generation cycle. Individuals for reproduction are selected according to their "fitness", which reflects the quality of the particular solution, thus biased towards the best. Then the recombination of selected strings takes place through crossover according to some high probability. An example of Single Point crossover is shown in Figure 3-7.

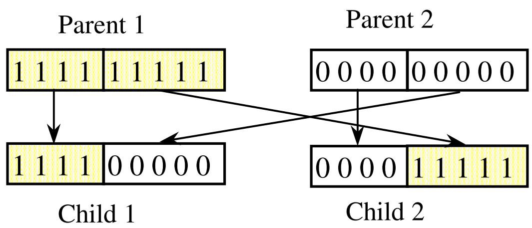


Figure 3-7 An example of Single Point crossover

The resulting offspring may undergo mutation according to some mutation probability, which usually is very low. Utilization of this operator ensures, that the probability of searching a particular subspace of the problem space is never zero, thereby tending to inhibit the possibility of ending the search at a local, rather than global optimum.

The whole process continues until a pre-specified termination criterion holds. The basic form of GA is represented by the pseudo-code in Figure 3-8.

```
begin
   Initialize(pop)           //Create initial population
   for g=1 to maxgen do
   begin
      Evaluate(pop)          //Evaluate fitness of all individuals
      newpop=Select(pop)     //Select the new population
      Crossover(newpop)      //Apply Crossover
      Mutate(newpop)         //Apply Mutation
   end
end
```

Figure 3-8 Basic GA in pseudo-code

There are many variations of this basic algorithm. There are a number of different schemes for Selection (Stochastic Tournament, Boltzman Tournament, Roulette, Integral, Uniform etc.), Crossover (Single Point, Uniform or Multiple Crossovers) as well as there can be different schemes for Replacement and Mutation.

The technique for encoding the solution to the problem on strings also may vary from problem to problem and within GA. The following general guidelines are given in [47]:

- ✓ A coding should be selected so that short, low-order schemata* are relevant to the underlying problem and relatively unrelated to schemata over other fixed positions.
- ✓ The smallest alphabet that permits a natural expression of the problem should be selected.

Most optimization tasks require consideration of constraints. There are the following two ways to represent constraints for optimization by GA:

- ✓ the most effective way is to embed constraints in the coding
- ✓ alternative way is to apply penalty function method.

3.11.5.3 Other Evolutionary Algorithms

Evolutionary Programming (EP) starts from the assumption, that evolution optimizes behavior (phenotype level) and not the encoding genetics. EP, therefore, has no restriction on problem representation (coding is not essential). This is a beneficial feature of EP in comparison with GA. Mutation is the only source of variations in the algorithm. EP typically does not use recombination

* In low order schemata the number of fixed string positions is small.

or other genetic operators.

The basic EP method starts from some initial population of trial solutions created randomly. Then the algorithm proceeds with the following two steps until a termination criterion holds:

- ✓ Off-springs are created from parent solutions by duplicating them. In basic EP mutation is implemented as adding normally distributed random variables with zero mean and dynamically adjusted variance to the components of all new trial solutions. Mutation variance is derived from the parent's fitness score.
- ✓ Each offspring solution is estimated according to its fitness. Some form of tournament between individuals leads to selection of a new population of a pre-specified size.

<u>Genetic Programming</u> [45] is a variant of GA with a different problem representation. The main difference from GA is that Genetic Programming operates by computer programs, which are candidate solutions, instead of strings that encode possible solutions. Similarly, each program is evaluated in terms of fitness by running it on a number of test problems and averaging the result. Usual Genetic Operators are used except Mutation, which usually is not applied.

<u>Evolution Strategies (ES)</u> are similar to GA with some notable differences [45]. The real-valued vector of the objective variables is processed instead of binary strings. Mutation is the dominating operator. It adds normally distributed random variables with zero mean and dynamically adjusted standard deviation to all components of each solution in the population. An additional feature of some ES is the self-adaptation of mutation variances and covariances. ES method is very similar to EP, although independently developed.

<u>Classifier Systems (CS)</u> are rule-based machine-learning systems capable of learning by examples [45]. It takes a set of inputs and produces a set of outputs, which indicate some classification of the inputs. There is functional similarity of CS to neural networks.

3.12 Multi-Criteria Decision-Making (MCDM) Methods

During the last two decades a wide variety of multi-criteria methods has been developed for identifying and resolving the conflict in decision situations where a large number of objectives is taken into consideration [67], [120].

Multi-criteria decision-making methods have two major purposes:

- To describe trade-offs among different objectives;
- To help the decision-maker or a number of participants in the planning process to define and to articulate their values, apply them rationally and consistently and document the results.

Almost all methods for solving multi-criteria problems involve two general sub-processes: articulation of the decision-maker's preference structure of the multiple criteria and optimization over the preference structure. One of the ways to categorize the MCDM methods is according to the timing of these sub-processes relative to one another [41],[88].

Thus, the prior articulation of preferences is in which the preference structure is obtained prior to the optimization. Techniques based on the prior articulation include goal programming, and use of a multi-attribute value or utility function [55]. In these approaches the decision-maker's preferences are obtained prior to the start of the optimization process. The major disadvantage of this approach is the difficulty the decision-maker faces in providing the required information. The optimization process, however, is usually relatively simple since the problem has typically been reduced to a single criterion.

Progressive articulation of preferences, in which the elicitation of information about the preference structure is interspersed with the optimization, involves an interaction between the decision-maker and the computer throughout the process. There are many different types of information, which could be required in these methods, for instance ranking of points in the outcome space, readjustment of aspiration levels from one iteration to the next, marginal rates of substitution between the various criteria.

Algorithms that rely on the posterior articulation of preferences seek to first find all or almost all non-dominated solutions. Then these solutions are presented to the decision-maker to select the best through some arbitrary process. The major disadvantage of these approaches is the large number of non-dominated solutions that must be analyzed.

Several methods that span the entire range of MCDM techniques are presented and compared in [55]. Several of these techniques are briefly described below.

The ***Pareto optimality*** concept defines the total set of non-inferior points. The techniques based on this principle search the set of Pareto optimal points.

The concept of Pareto optimum was formulated by Vilfredo Pareto in the XIX century [100] and constitutes by itself the origin of research in multi-objective optimization. The point $\bar{x}^*$ is *Pareto optimal* if there exists no feasible vector $\bar{x}$, which would decrease some criterion without causing a simultaneous increase in at least one other criterion.

The major criticism of multi-criteria decision-making based on Pareto optimality is that the technique is limited to a small number of criteria. If the number of criteria is more than two the visualization is not apparent enough. A solution to this problem can be the projection of the k-dimensional criterion space containing Pareto optimal set into a two dimensional space. Through factorization, using principal component analysis of the matrix containing the set of Pareto optimal points, followed by projection, it is possible to visually inspect the total criteria space.

A second point of criticism is the rigidity of the method. A solution to this problem could be creation of tolerance region around the Pareto optimal set. This also makes it easier to make a more robust decisions, because the point that is inferior to a Pareto set may sometimes is preferred for its robustness, i.e. its insensitivity to small variations in the independent variables.

Response surfaces can be depicted by three-dimensional plots or contour plots when there are no more than two independent variables. Thus, the *overlay plots* are the projection of several contour plots of criteria surfaces onto one figure. Minimum and/or maximum boundaries that go with acceptable criterion values need to be set and marked in the contour plots. Projection of these marked areas in one figure makes it possible to perform a visual selection of variable settings that results in acceptable values for all the criteria.

There is no simple procedure for combining the magnitude of several properties into a single quantitative measure. There is a number of *aggregation methods,* the simplest of which defines the objective function as the sum of weighted criteria values. There is a large variety of methods for assignment of weights in the aggregated function [67], which may be considered as a main disadvantage of the method since introduces arbitrariness prior to the optimization process. There is a number of approaches using other methods of criteria aggregation into a single qualitative property, one of which the so-called method of displaced ideal [144] will be described in section 6.2.2.

A completely different category of multi-criteria decision-making techniques is formed by so-called *outranking methods*. The object of these methods is to explicitly rank the complete collection of alternative settings of the independent variables. The outranking methods, such as for example PROMETHEE [15], were originally developed for use in economics. A clear problem with PROMETHEE, as well as with certain other methods, is mentioned to be the attachment of weights to the criteria.

The comparison of these techniques according to some important performance criteria is given in Table 3-4.

The performance with many criteria is best for PROMETHEE and worst with overlay plots. Information on robustness is another important criterion of the method performance. Robustness measures the sensitivity of a response or criteria value against small variations of the independent variables. Overlay plots give the best information. No information on robustness is available when employing Pareto optimality or PROMETHEE. Information on robustness of the utility function is available only when it is used as a simultaneous technique with prior articulation of preferences.

Table 3-4 Performance evaluation of multi-criteria decision-making techniques ('+'-good, '-'-poor), after [55]

	Overlay plots	Pareto optimality	Utility functions	PROMETHEE
Articulation	Posterior	Posterior	Prior or progressive	Prior and posterior
Many criteria	+/-	+/-	+	++
Information on robustness	+	--	+/-	--
Foreknowledge*	+	++	-	--
Calculation complexity	+	+/-	+/-	-
Comprehensi-bility	+	++	+/-	-

*Minimum required foreknowledge is represented by '+'

Another point to consider when choosing which MCDM method to use is the amount and complexity of the performed calculations. Comprehensibility and surveyability of a certain method is also an important criterion.

3.13 Conclusions

- The network planning problem cannot be formulated in terms of Bayesian Decision Theory due to the difficulties associated with formulation of the utility functions. However, when approaching the problem applying different methods it is important to consider the concave character of the utility function. This consideration directly leads to the multi-criteria formulation of the problem, since the decision is motivated not only by the expected value of revenues (or losses), but also by risk of not having the expected result.

- In the most general case the planning model must account for uncertain and random parameters. These considerations, as well as extensive number of decision variables in the model, make it difficult to organize the procedure to search for the best alternatives.

- Several means to reduce the complexity of the problem are known and widely applied in network planning practice. Unfortunately, the simplification causes loss of precision and insufficient consideration of uncertain information.

- In presence of truly uncertain information one of the Decision-Making criteria should be applied. The particular choice is subjective and may lead to considerable losses.

- Uncertain information can be modeled by fuzzy variables and the corresponding attributes calculated by means of fuzzy arithmetic.
- Existence of probability-possibility transformation provides the opportunity to choose application of either fuzzy arithmetic or method of Monte-Carlo for calculation of the planning attributes. Method of Monte-Carlo requires considerable computational efforts, however it is more universal and with sufficient number of samples provides better accuracy.
- The difficulties caused by tremendous complexity of the problem can be overcome either introducing a number of simplifications, leading to the considerable loss in precision or applying methods based on modifications of Monte-Carlo or fuzzy arithmetic and Genetic Algorithms or Dynamic Programming.

4 Assessment of Planning Attributes

This chapter contains the description of the suggested model and the algorithms for assessment of the selected planning attributes. The model reflects the basic technical and financial aspects of the network elements. The set of chosen attributes is calculated for the planning period as a sum of the annual and discounted terms considering economic principles. The chapter also contains the suggested algorithm for probabilistic load modeling based on analysis of measured data and choice of suitable empirical probability distribution.

4.1 The Planning Objectives

The relative goodness of each alternative can be measured in terms of chosen set of attributes. In order to estimate the attributes the corresponding model is needed. Modeling is an approximate reflection of the reality. The good model must essentially consider the most important features of the real system and neglect excessive details. The mission of the model is to gather numerous data about the problem under a single framework, and to process this data in such a way that the planning objectives can be expressed numerically in terms of attributes.

The model proposed here is dynamic and multi-criteria, i.e. the identified optimization objectives are treated separately. The corresponding attributes are calculated for the planning period as a sum of the annual and discounted terms.

Annual losses are obtained from the load flow calculations. For the primary networks the different loading conditions can be modeled by duration of every mode, such as winter maximum, summer minimum, day, night, weekend, etc.

The reliability criterion in the general case combines network utility unavailability data with customers' view on unavailability of supply. If the interruption costs are not available the Energy Not Supplied (ENS) can be used.

Investments for the particular state of the network are calculated as a sum of annuities of each investment in reinforcement action realized during the planning period.

During the whole planning period the network must satisfy a number of security and configuration constraints. Distribution networks are usually operated radially. Radiality of each particular configuration is retrieved from load flow calculations and disconnection at corresponding points.

Furthermore, there is an option to apply a number of logical constraints on the order and compatibility of reinforcement actions realization. These logical constraints together with the assumption that once the action is realized at some stage, it cannot simply be canceled on the later stages leading to the certain policy of transitions between the states. Not all the transitions are feasible or admissible.

The model treated here is very fundamental and there is a large potential for its improvement or extension. First and foremost, the set of attributes can be extended by for example power quality or environmental impact criteria. Furthermore, new technologies, DSM and local generation encourage further development of the models. The corresponding advances are widely discussed in the literature (e.g. [136]).

Previously it was discussed that in presence of uncertainty and whenever it is impossible to assign a utility function, the risk associated with acceptance of one or another alternative must be estimated in order to insure the robustness of chosen alternatives. Therefore under uncertainty the number of criteria will increase. This issue will be addressed in the next chapters.

4.2 Economic Considerations

The planning process essentially involves estimation of economic consequences of each alternative plan. The planner is usually restricted by the operational constraints, standards and guidelines, but within these frames the goal is to minimize cost. Every alternative plan implies certain costs: equipment, installation labor, operating and maintenance, and many others. The total costs are important, but also when the costs are incurred – how much must be spent now and how much later.

The fact that different expenses or incomes might not coincide in time when comparing the costs of alternative solutions, especially for the existing system reinforcement, is a difficulty to deal with. It means, that a method of economic assessment to take into account the timing of cash flows is needed.

Present Value (or Worth) analysis is a method of measuring and comparing costs and savings that occur at different times on a consistent and equitable basis for decision making.

To convert the single payment at some year t in the future into equivalent amount at present and vice versa the Present Value method can be described as:

$$PV = \frac{1}{(1+i)^t} \cdot FV \text{ or} \tag{56}$$

$$FV = (1+i)^t \cdot PV, \tag{57}$$

where FV is a value of future amount in year t, PV is the value of the same amount at time zero and i is the interest rate.

If there are uniform series of the annual payments from today through T years the present worth of these payments can be found by using the Annuity method:

$$PV = \left[\frac{(1+i)^T - 1}{i(1+i)^T} \right] \cdot A, \tag{58}$$

where A stands for value of annual payments, which is considered constant and T corresponds to the planning period.

In the network planning tasks different alternatives are usually analyzed over a longer period of time corresponding to the lifetime of the equipment. However, the lifetimes of different units of the equipment may differ considerably. One solution to the problem of dynamic allocation of assets is to use one of the depreciation accounting methods. Depreciation may be defined as the lessening in value of a physical asset with the passage of time. Thus, the alternative investments, which do not coincide in time, can be compared based on the Present Value of the investments and the salvage value. Another, conceivably more general approach is to reduce a single investment to a series of annualized costs:

$$A = \left[\frac{i(1+i)^T}{(1+i)^T - 1} \right] \cdot PV \ . \tag{59}$$

If one defines the lifetime of the particular unit of the equipment as depreciation time and assign the planning period, the following cases may need to be compared with each other:

- Planning period is shorter than the unit depreciation time and the investment is made at present time. The planner is only interested in payments to be made during the planning period. A series of annualized costs can be found from the following equation:

  $$A_{Depr} = \left[\frac{i(1+i)^{T_{Depr}}}{(1+i)^{T_{Depr}} - 1} \right] \cdot PV \tag{60}$$

  The present value of the investments during the planning period may be found applying the equation (58).
- Planning period may be shorter than (or equal to) the unit depreciation time, but the investment is postponed by a number of years more than $T_{Depr} - T_{Pl}$.

  In this case a series of annualized costs to be found from (60), and used to find the future investment value as follows:

  $$FV_{Pl} = \left[\frac{(1+i)^{T_{Pl}-T_0} - 1}{i(1+i)^{T_{Pl}-T_0}} \right] \cdot A_{Depr}, \tag{61}$$

where T_0 is the time of delay of the investment in comparison to the present time. The present value of the investment can be obtained either from (61) applying (56) or directly from the physical value of investment according to:

$$PV_{Pl} = (1+i)^{-T_0} \cdot \left[\frac{1-(1+i)^{-(T_{Pl}-T_0)}}{1-(1+i)^{-T_{Depr}}} \right] \cdot PV \ . \qquad (62)$$

Equation (62) was obtained from equations (56), (60) and (61).

- Unit depreciation time is shorter than (or equal to) planning period. In this case the present value of the investment is equal to its physical value, but annuity can be calculated using (60).

4.3 Modeling Network Elements

In modern distribution planning software systems the same basic circuit model provides the basis for all performance simulators of the distribution system, including load flow, short circuit, reliability and economic analysis. It contains representation of lines, loads, and equipment along with connectivity information. The most basic decision in building a circuit model for distribution planning relates to how much detail is used. For the purposes of long-term planning usually single-phase circuit model is an adequate answer.

The transformers, regulators, capacitors, line drop compensators and other elements of the distribution system can be modeled in varying levels of details from simplistic to greatly detailed [138]. Generally, distribution planning requires less detail in the modeling of equipment behavior and control than distribution engineering. The key aspects are technical – capacity and basic electric behavior, and financial – costs.

The relatively short length of <u>distribution lines</u> enables simple modeling. It is usually sufficiently accurate to ignore the capacitance of a distribution circuit and represent it by a series impedance.

<u>Transformers</u> can be represented by shunt and series impedances. The smaller distribution transformers have a larger series resistance than reactance, while the larger power transformers have negligible resistance compared to reactance.

<u>Capacitors</u> can be modeled as an impedance. <u>Tap-changing</u> devices can be approximated for planning purposes as step-less devices which maintain voltage at a control bus at a constant level regardless of upstream voltage level.

Static model of the electrical <u>loads</u> normally falls into one of the following three categories: constant current, constant power or constant impedance.

4.4 Representation and Modeling of Loads

4.4.1 Customer Demand

The forecasted load is the most sensitive parameter affecting the solution of the problem of the network planning.

Since the pattern of electrical demand of each customer cannot be determined precisely, it is usually necessary to calculate system loading on a statistical basis, whether considering existing loads or forecasted values. Plans to meet future demands are generally based on the assumption that load patterns will not change significantly, unless there is positive evidence of the contrary.

The modeling of loads is a complex task. The complexity is motivated by the existence of daily and annual load cycles, which is different for different customers.

4.4.2 Actual Measurements

Full loading information for the particular nodes of the network can be collected from automatic load metering equipment. Then the correspondingly processed actual data together with the adequate load forecasts represent a good input for planning process. The example of processing of the measured data is discussed in section 4.5.

4.4.3 Velander's Formula

If only the energy demand is measured, the load forecasts are also presented as annual energies. Therefore, a transformation method is needed. The method known as Velander's formula [131] is widely used in Scandinavia for transformation of annual energies into peak power:

$$P_{\max} = k_1 E + k_2 \sqrt{E} \tag{63}$$

where k_1, k_2 are the empirical coefficients, E is an annual energy and $P_{\max}$ is a peak power demand.

Some examples of Velander's coefficients (if E in (63) is given in kWh and $P_{\max}$ is obtained in kW) are given in Table 4-1.

Table 4-1 Some values of Velander's coefficients

Customer group	k_1, h^{-1}	k_2, h^{-1}
Domestic:		
- without electrical heating	0.00033	0.050
- cottage with electrical heating	0.00030	0.025
- large house with electrical heating	0.00028	0.025

The formula is based on the assumption that the load studied is the sum of several individual loads, which have very similar behavior. Therefore, the results of Velander's formula may be reasonable if the number of customers is not too small and if the load is fairly homogeneous.

4.4.4 Coincident Load Behavior

If the possible maximum demand is known, it is still necessary to apply corrections to individual loads in order to carry out load-flow studies in medium-voltage (MV) and low-voltage (LV) networks. This is because the sum of the maximum values of all the loads will result in too high a value for the total current flow and therefore the overall voltage drop, if the loads do not peak at the same time.

Consider n loads supplied trough the same line. The instantaneous power will be $\sum_{i=1}^{n} P_i$, but since the different loads can have different variation in time, peak power demand will be $P_{\max} \leq \sum_{i=1}^{n} P_{\max i}$. In order to find the maximum demand of a group of loads a *coincidence factor* is used. This factor depends on the types of the load and their numbers and is defined as follows:

$$k = \frac{P_{\max}}{\sum_{i=1}^{n} P_{\max i}} \leq 1\,. \tag{64}$$

4.4.5 Typical Load Curves

Traditionally the consumers with different types of activities have been sorted into different classes each of which had a characteristic load curves [31,32]. Load curves depict seasonal and daily consumption pattern for different types of customers and their variation is approximated by mean value and standard deviations of Gaussian distribution. Load curves can then be used to simulate the load of a customer belonging to a particular group by scaling the annual units of that customer to the average units of the group. A load curve for a group of customers can be constructed by summing the individual customer load curves statistically. It means, that if there is no correlation between the loads then mean values for each hour can be summed directly and the total standard deviation can be found as $\delta = \sqrt{\sum \delta_i^2}$.

If there is positive correlation between customer demands the following relationship is valid for the values of mean probability and a given excess probability level:

$$P_p = P_m + k_p\delta , \tag{65}$$

where P_p is a power having an excess probability of p %, P_m is mean power, k_p is a coefficient related to p and δ is standard deviation [32].

Normally, both Velander's formula and load curves give almost the same results when there are many customers from the same customer group. However, if there is a need to sum the demands of the customers from different groups, which do not peak at the same time, the Velander's formula can result in too high values.

4.5 Probabilistic Load Modeling

4.5.1 Load Variations: Annual, Seasonal, Weekly, Daily, and Hourly

Load is the most uncertain constantly varying parameter. During one year the load may vary greatly from season to season, from day to day and from hour to hour.

Typically, the value of most interest to the planner is the annual peak demand during the year since it serves as an indication of capacity requirements for equipment. On the other hand, for the economic comparison it is necessary to calculate the recurring annual system losses.

Among the fundamental drivers of electricity demand can be mentioned:

- Economic trends
- Time factor
- Weather
- Random effects.

The driver associated with economic environment affects the annual peak load on the long-term basis. It includes such factors as demographics of the area and level of industrial activity, which define the system load increase/decline trends. These factors must be taken into account in long-term load forecasting, which is one of the starting points for the network planning studies.

The other uncertainty source is related to time/weather factors. The main time factors can be identified as seasonal variations, weekly-daily cycle and holidays. Temperature is the most relevant weather variable affecting the load. The effect of the weather varies from country to country - for Northern countries the most important is the effect winter low temperature may cause. Furthermore, the weather affects in a different way different types of customers – in Sweden residential loads are probably the most sensitive group.

In this section the attempt is made to establish a probabilistic model of short-term load uncertainties caused by time/weather factors. The hourly data from two substations 110 kV and 33 kV measured during three years were subjected to a detailed analysis according to the algorithm described in the next sub-sections. The maximal and mean values of the load for the two nodes are depicted in Figure 4-1. Statistics for three year period do not allow to make a well founded conclusion about the load growth trends, but the given data together with the additional information allow for a qualified guess that the load remains at approximately the same level. The seasonal load variations as well as the difference between working day and weekend peaks can be clearly seen on the graph. Prior to the analysis and modeling the data should be pre-processed and cleaned. The data containing discontinuities and abnormally large disturbances must be excluded from consideration.

The examples of the daily load curves for typical winter and summer days are presented in Figure 4-2.

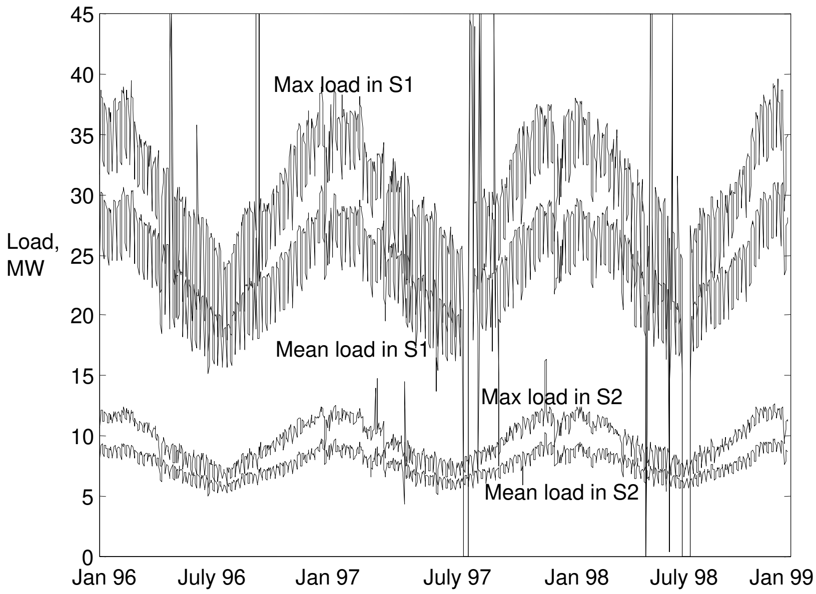


Figure 4-1 Max and mean load measurements for two nodes during 3 years period

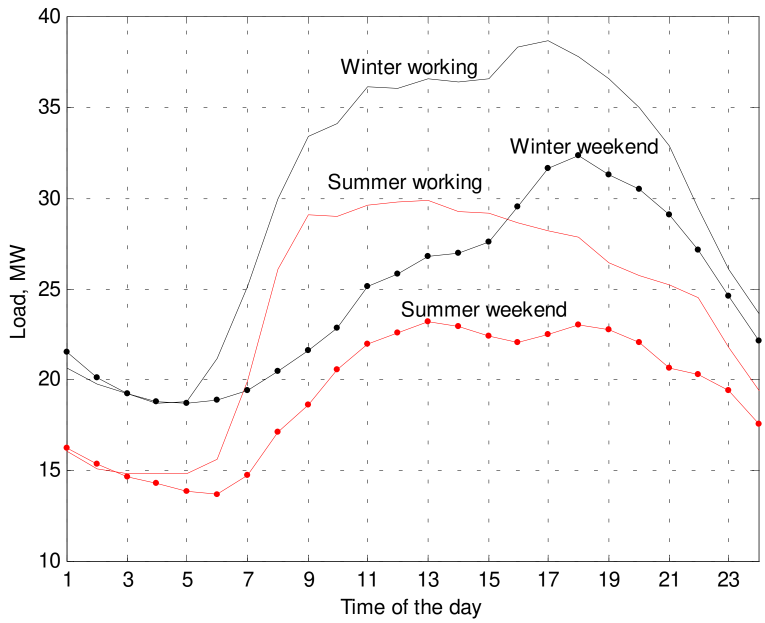


Figure 4-2 Example of daily load curves: winter and summer loads in S1

4.5.2 Processing the Measured Load Data

To take into account the time factor, which influences the load variations, the load data can be sorted according to several modes. One example of such a model for data processing is presented in Figure 4-3. Three seasonal modes, each of which consists of two modes depending on the weekday, result in six load modes.

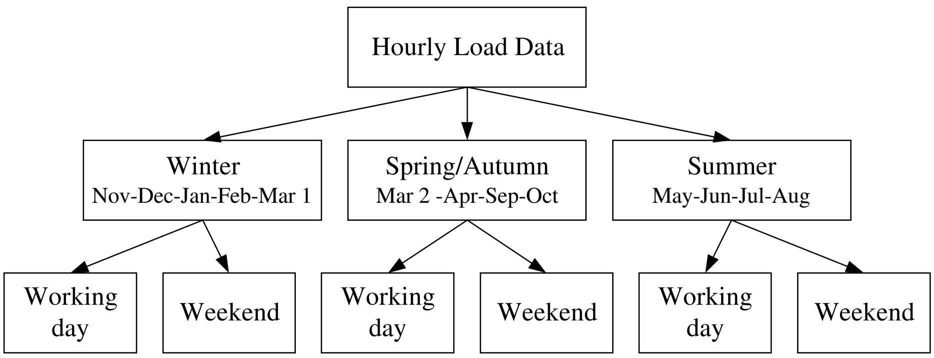


Figure 4-3 Processing measured load data for the time factor

The last time factor is the daily load variations, which are considerable for most types of customers. Depending on the purpose of the studies several models following the model in Figure 4-3 may be suggested (Figure 4-4).

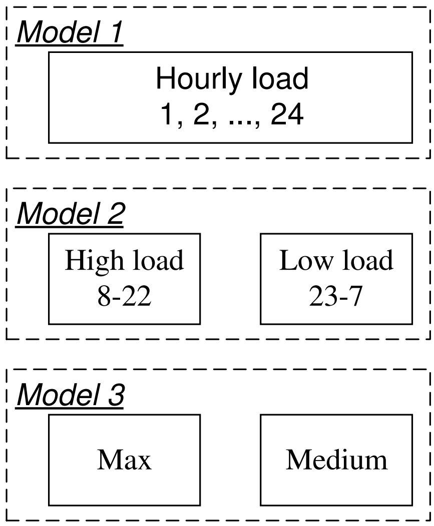


Figure 4-4 Three models for daily load variation classification

For the detailed studies the first model can be adopted (Figure 4-4). It assumes that the load is modeled on the hourly basis. The second model simulates the daily load to be consisting of two stages high load (during the day) and low load (during the night). The model is well-founded even from the point of view of economic calculations – there are often different electricity tariffs for customers at the time of high and low demand. Finally, for the planning tasks the most important information concerning load is the information about the peak demand. Thus, for the planning studies the third model evaluating load maximum and average value of the load for each mode may be adopted.

4.5.3 Algorithm for Assessment of Statistical Model from Measured Load Data

4.5.3.1 Measures of Distribution

The behavior of a random variable is completely described by its *Probability Density Function* (PDF).

Let $f(x_i)$ be the PDF of a discrete random value. The measures characterizing distribution function of the variable are given bellow.

The expected (mean) value is defined by

$$E[X] = \sum_i f_i x_i \ . \tag{66}$$

In addition to the central value, the usual measures of distributions are spread, symmetry and peakedness. These characteristics may be summarized by the moments of the distribution.

If M_1 denotes the expected value $E[X]$ of the random variable X. The k^{th} moment about the mean or central moment is defined as

$$E[X - M_1]^k = \sum_i (x_i - M_1)^k f(x_i) \ . \tag{67}$$

The second moment about the mean is a measure of dispersion. It is known as variance, and is defined by

$$\mu_2 = Var[X] = \sigma^2 = E[X - M_1]^2 \tag{68}$$

The standard deviation is defined by

$$\sigma = \sqrt{E[X - M_1]^2} \ . \tag{69}$$

The interaction of two random variables *X* and *Y* can be summarized by joint moments of their distributions. The covariance of *X* and *Y* is defined as the central moment:

$$\mathrm{cov}[X, Y] = E[XY] - E[X] \cdot E[Y] \ . \tag{70}$$

The correlation coefficient of *X* and *Y* is defined by

$$\rho = \frac{\mathrm{cov}[X, Y]}{\sqrt{\sigma(X)\sigma(Y)}} \tag{71}$$

where $-1 \le \rho \le 1$.

A negative correlation coefficient indicates opposite movements, whereas a positive correlation coefficient indicates coordinated movement.

The third moment about the mean is related to the asymmetry or skewness of a distribution and is defined as

$$\mu_3 = E[X - M_1]^3 \ . \tag{72}$$

The quantity

$$\sqrt{\beta_1} = \frac{\mu_3}{(\mu_2)^{3/2}} \tag{73}$$

measures the skewness of the distribution relative to its degree of spread. This standardization allows us to compare the symmetry of two distributions whose scales of measurement differ.

The fourth moment about the mean is related to the peakedness – also called kurtosis – of the distribution and is defined as

$$\mu_4 = E[X - M_1]^4 . \tag{74}$$

The quantity

$$\beta_2 = \frac{\mu_4}{{\mu_2}^2} \tag{75}$$

is a relative measure of kurtosis.

A distribution is completely specified once all its moments are known. Many distributions can be adequately described by the first four moment, which play an important role in fitting empirical distributions and in approximating the distribution of a random variable.

The probability distributions can be derived from historical data or estimated subjectively. However, in order to apply probabilistic model the uncertainty modeled must be of random nature and its probability distribution should be estimated with reasonable accuracy.

4.5.3.2 The Suggested Algorithm

If the statistical measured data for the load is available it is easy to apply an algorithm for finding the appropriate distribution of data variation and fitting its parameters. The procedure presented below is based on estimates of the moments and on utilization of the Pearson's chart.

1. Sort out the data with unreasonably large disturbances.
2. Arrange the load data according to Figure 4-3 and one of the models presented in Figure 4-4 and proceed with this algorithm for each data set.
3. Calculate the estimates for the central moments.
4. Compute the estimates of $\sqrt{\beta_1}$ and β_2.
5. Plot the resulting points on Pearson's chart (Figure 4-5) and allocate the appropriate distribution.
6. Estimate parameters of the distribution chosen in the previous position.

4.5.3.3 Estimation of Central Moment from the Empirical Data

If the form and parameters of the probability density function are unknown, the estimated central moments denoted by m_k can be calculated by replacing M_1 by $\frac{1}{n}\sum_{i=1}^{n} x_i^k$ in the expressions (67), where x_i, $i = 1, 2, ..., n$, are the values of n given observations. Thus

$$m_2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 , \tag{76}$$

$$m_3 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^3 \text{ and} \tag{77}$$

$$m_4 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^4 , \tag{78}$$

where $\overline{x}$ is the mean value.

The estimates of $\sqrt{\beta_1}$ and β_2 can be calculated using equations (73) and (75), respectively, with corresponding estimates of the moments.

4.5.3.4 Approximation by Empirical Distribution

The fitting of distributions to data has a long history and many different procedures have been advocated. The most common of these is the use of a normal distribution. Taking into account the Central Limit Theorem one can expect that normal distribution would provide a reasonable representation in many cases.

The Pearson's chart in Figure 4-5 shows the regions in the (β_1, β_2) plane for different normal, beta (special case - uniform), gamma (special case – exponential), Student *t*-distribution and log-normal distributions. β_1 and β_2 are respectively the square of the standardized measure of skewness and the standardized measure of peakedness.

For all normal distributions $\beta_1 = 0$ and $\beta_2 = 3$. Therefore this distribution is represented in Figure 4-5 by a single point, as are also the exponential and uniform distributions. The gamma, log-normal and Student t distributions involve curves. Thus gamma distributions can be fitted for all values of β_1 and β_2 that fall on the curve shown near the center of the chart. The beta distribution, which has two shape parameters, occupies a region in Figure 4-5 and provides greater generality than any of the other distributions.

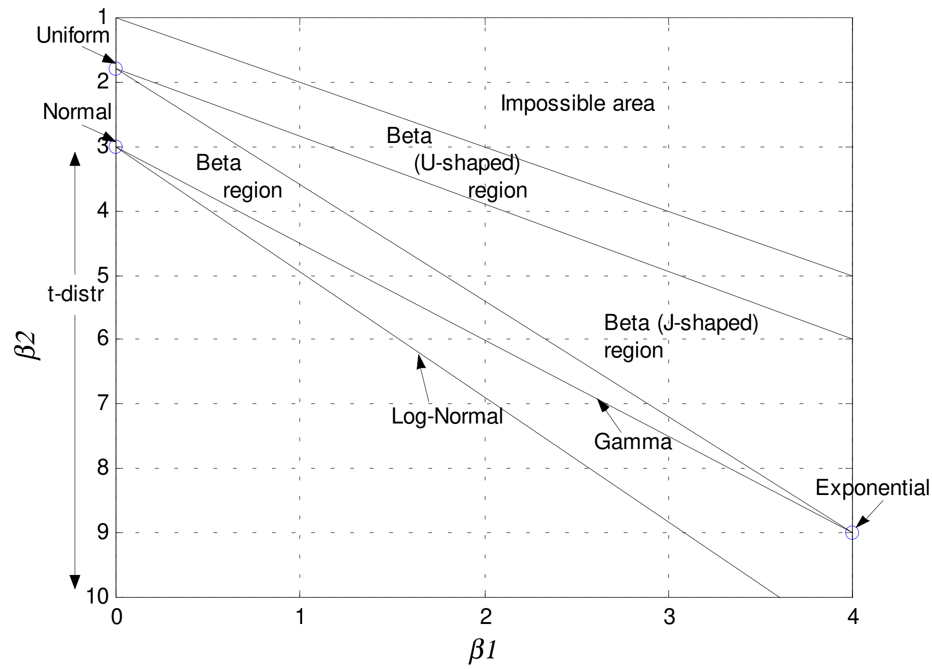


Figure 4-5 Pearson's chart: regions in (β_1, β_2) plane for various distributions [49]

4.5.3.5 Estimation of Distribution Parameters

The chart on Figure 4-5 may be used to provide an indication of whether or not given data can be appropriately represented by one of the distributions shown. This is done by obtaining the estimates of $\sqrt{\beta_1}$ and β_2 using equations (73) and (75) with corresponding estimates of the moments, and plotting this point on Figure 4-5. If the plotted point is reasonably close to a point, curve, or region corresponding to one of the models, this distribution can be used to represent the data. One can proceed to obtain estimates for the distribution parameters using the appropriate formulas.

Normal Distribution Parameters

The probability density function of the normal distribution is

$$f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \tag{79}$$

where μ and σ are location and scale parameters respectively and $-\infty < x < \infty$, $-\infty < \mu < \infty$, $\sigma > 0$.

In many problems μ and σ must and can be estimated from the available data. The estimates of the parameters are denoted by hat. Estimate of μ can be calculated as

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}, \tag{80}$$

where x_i, $i = 1, 2, ..., n$ are the values of the n data points.

The standard deviation can be obtained as

$$\hat{\sigma} = \sqrt{m_2} = \left[\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n} \right]^{\frac{1}{2}} = \left[\frac{\sum_{i=1}^{n} x_i^2}{n} - \frac{\left(\sum_{i=1}^{n} x_i \right)^2}{n^2} \right]^{\frac{1}{2}}. \tag{81}$$

<u>Beta Distribution Parameters</u>

The beta probability density function defined over the interval (0,1) is

$$f(x; \gamma, \eta) = \begin{cases} \dfrac{\Gamma(\gamma + \eta)}{\Gamma(\gamma)\Gamma(\eta)} x^{\gamma-1}(1-x)^{\eta-1} & 0 \le x \le 1, 0 < \gamma, 0 < \eta \\ 0 & elsewhere \end{cases}, \tag{82}$$

where $\Gamma(\chi)$ is gamma function defined by $\Gamma(\gamma) = \int_0^\infty x^{\gamma-1} e^{-x} dx$.

The beta distribution can be generalized to cover the interval (μ_0, μ_1). This leads to the probability density function

$$f(x; \gamma, \eta, \mu_0, \mu_1) = \begin{cases} \dfrac{1}{(\mu_1 - \mu_0)} \dfrac{\Gamma(\gamma + \eta)}{\Gamma(\gamma)\Gamma(\eta)} \left(\dfrac{x - \mu_0}{\mu_1 - \mu_0} \right)^{\gamma-1} \left(1 - \dfrac{x - \mu_0}{\mu_1 - \mu_0} \right)^{\eta-1} \\ \qquad\qquad \mu_0 \le x \le \mu_1, 0 < \gamma, 0 < \eta \\ 0 \qquad\qquad elsewhere \end{cases} \tag{83}$$

When $\gamma > 1$ and $\eta > 1$ the distribution is single peaked, for other values of parameters the distribution can be U shaped, J shaped or reverse J shaped.

The distribution can be generalized to cover the interval (μ_0, μ_1).

The following equations can be used to obtain estimates of beta distribution parameters:

$$\hat{\eta} = \frac{(1-\bar{x})}{s^2}\left[\bar{x}(1-\bar{x}) - s^2\right] \tag{84}$$

$$\hat{\gamma} = \frac{\bar{x}\hat{\eta}}{1-\bar{x}} \tag{85}$$

where $\bar{x}$ can be obtained from (80) and

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}. \tag{86}$$

Log-Normal Distribution Parameters

The log-normal distribution is the model for a random variable whose logarithm follows the normal distribution with parameters μ and σ. Thus the probability density function for x is

$$f(x;\mu,\sigma) = \begin{cases} \dfrac{1}{\sigma x\sqrt{2\pi}}\exp\left[-\dfrac{1}{2\sigma^2}(\ln x - \mu)^2\right] & x > 0, \sigma > 0 \\ & -\infty < \mu < \infty \\ 0 \quad elsewhere & \end{cases} \tag{87}$$

The log-normal distribution can be generalized to cover an interval other than $(0,\infty)$ by introducing the location parameter ε. Thus

$$f(x;\mu,\sigma,\varepsilon) = \frac{1}{\sigma(x-\varepsilon)\sqrt{2\pi}}\exp\left[-\frac{1}{2\sigma^2}(\ln(x-\varepsilon) - \mu)^2\right], \tag{88}$$

where $x \geq \varepsilon, \sigma > 0, -\infty < \mu < \infty, -\infty < \varepsilon < \infty$.

It is relatively easy to estimate the remaining parameters of log-normal distribution if the location parameter ε is known. However, if the parameters should be estimated from experimental data, the value of this parameter is usually unknown. In this case the approximate algorithm which can be applied to obtain the estimates is given below. Three equations are obtained of the form

$$z_\alpha = \gamma^* + \eta\ln(x_\alpha - \varepsilon), \tag{89}$$

where z_α is the $\alpha\,100^{th}$ percentile for a standard normal variate and x_α is the corresponding percentile calculated from the data. For example, if the 5^{th}, 50^{th}

and 95[th] percentiles are used for this purpose the three equations are

$$\begin{aligned} -1.645 &= \gamma^* + \eta \ln(x_{0.05} - \varepsilon) \\ 0 &= \gamma^* + \eta \ln(x_{0.5} - \varepsilon) \quad , \\ 1.645 &= \gamma^* + \eta \ln(x_{0.95} - \varepsilon) \end{aligned} \tag{90}$$

since $z_{0.05} = -1.645, z_{0.5} = 0, z_{0.95} = 1.645$.

Solution of these equations yields

$$\hat{\eta} = 1.645 \left[\ln \left(\frac{x_{0.95} - x_{0.5}}{x_{0.5} - x_{0.05}} \right) \right]^{-1}, \tag{91}$$

$$\hat{\gamma}^* = \hat{\eta} \ln \left(\frac{1 - e^{-1.645/\hat{\eta}}}{x_{0.5} - x_{0.05}} \right) \text{ and} \tag{92}$$

$$\hat{\varepsilon} = x_{0.5} - e^{-\hat{\gamma}^* / \hat{\eta}}. \tag{93}$$

Finally, the estimates for $\hat{\mu}$ and $\hat{\sigma}$ can be found as

$$\hat{\sigma} = \frac{1}{\hat{\eta}} \text{ and} \tag{94}$$

$$\hat{\mu} = -\hat{\sigma} \cdot \hat{\gamma}^*. \tag{95}$$

4.5.3.6 Example: Assessment of Statistical Model from Measured Load Data

The load data recorded at two load points 110 kV and 33 kV (see section 4.5.1) during three years were subjected to a detailed analysis according to the suggested algorithm. The load data were arranged according to Figure 4-3 and two the models presented in Figure 4-4, namely, *Model 2* for high/low daily loads and *Model 3* for the daily peak load. Each data set was processed - some of the results are presented in this section.

The estimates for the central moments give us the estimates of $\sqrt{\beta_1}$ and β_2, which in its turn allows for allocation of appropriate distribution function.

For this illustrative example the load variations were modeled by three probability density functions, namely normal, log-normal and beta distribution. The parameters for each PDF were chosen according to the methodology given in section 4.5.3.5.

Pearson's chart with beta estimates for daily peak loads is presented in Figure 4-6. The following notation is used: "LW" corresponds to the load during the winter mode, "LSp" to spring/autumn mode and "LSo" to summer mode. Number 1 or 2 after the seasonal mode denotes working day or weekend, respectively. The chart reveals that most of the points lie in beta-distribution area. However, according to the chart some of the modes (summer) can be modeled even by normal distribution.

To demonstrate the association between the Pearson's chart and actual PDFs Figure 4-7 and Figure 4-8 depict probability distributions modeled for winter and summer respectively (both working day and weekend) in comparison with histograms of the corresponding loads. Solid line corresponds to the normal distribution, log-normal distribution is depicted by circles and, finally, crosses represent beta probability distribution.

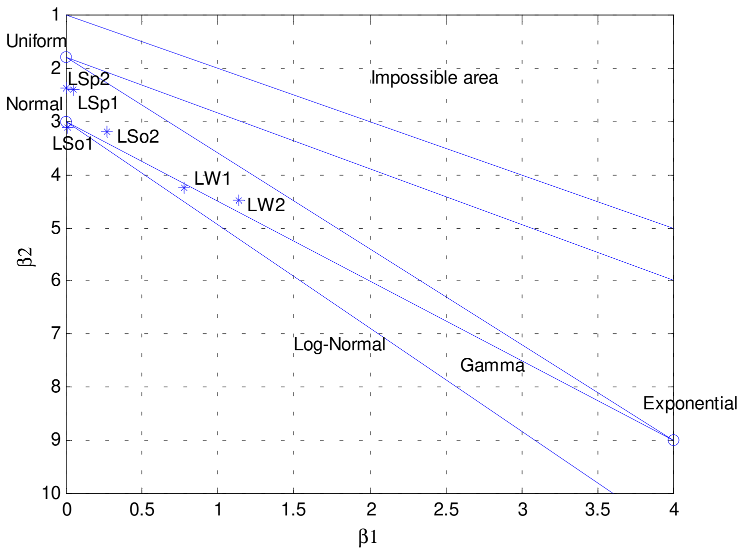


Figure 4-6 Pearson's chart with S2 peak load data: "LW" corresponds to the load during the winter mode, "LSp" to spring/autumn mode and "LSo" to summer mode. Number 1 or 2 after the seasonal mode denotes working day or weekend, respectively.

From the comparison (visual or using Chi square test of statistical significance) of statistical data and empirical probability distribution it can be concluded that in symmetric cases all three distributions provide reasonably good representation of the load variations. However, if variations of modeled parameter are non-symmetrical log-normal or beta distribution would give a better approximation.

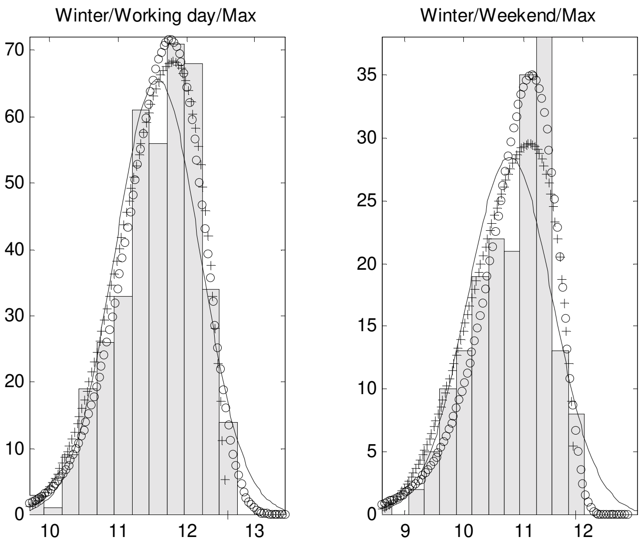


Figure 4-7 Histograms of winter peak loads in S2 (both working day and weekend) in comparison with modeled probability distributions: solid line corresponds to the normal distribution, circles to the log-normal and crosses to beta probability distribution.

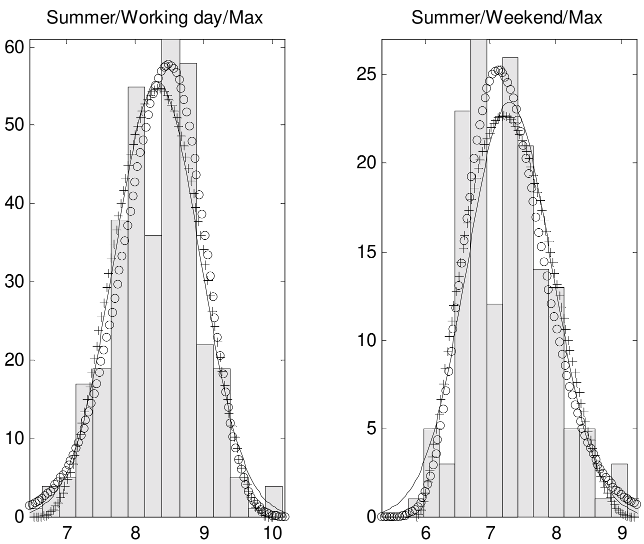


Figure 4-8 Histograms of summer peak loads in S2 (both working day and weekend) in comparison with modeled probability distributions: solid line corresponds to the normal distribution, circles to the log-normal and crosses to beta probability distribution.

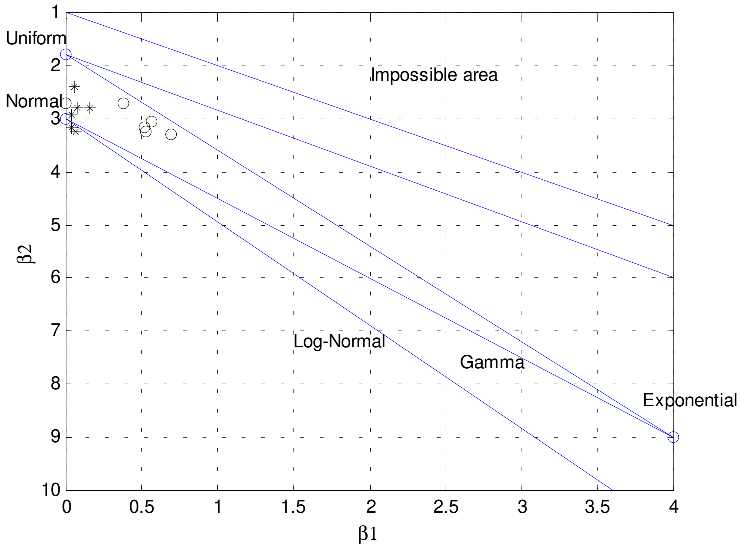


Figure 4-9 Pearson's chart with S2 high/low load data: the points corresponding to the high loads are depicted by asterisks, the points corresponding to the low loads are depicted by circles.

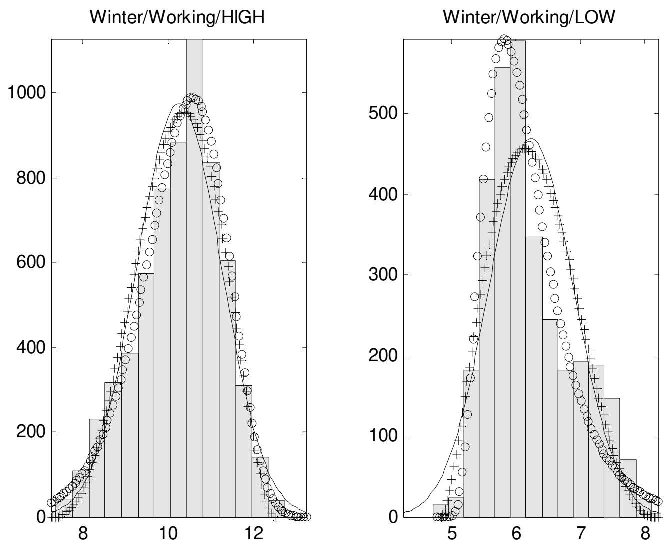


Figure 4-10 Histograms of winter high/low loads in S2 and modeled probability distributions: solid line corresponds to the normal distribution, circles to the log-normal and crosses to beta probability distribution.

Similar observations can be made from Figure 4-9 and Figure 4-10, which depict respectively Pearson's diagram and probability distributions modeled for low and high load modes (Model 2 in Figure 4-4). In Figure 4-9 the points corresponding to the high load mode (Winter/Working/High, Winter/Weekend/High, Summer/Working/High, etc) depicted by asterisks are grouped near the normal distribution. The points corresponding to the low loads depicted by circles are placed somewhat away in the area of beta-distribution. Figure 4-10 illustrates histograms and modeled probability distributions of low and high loads during winter working day. In correlation with Pearson's diagram high mode load variations have higher degree of symmetry, and can be quite realistically approximated by normal distribution.

4.6 Load Flows and Power Losses

4.6.1 AC load flow

Load flow calculations in network planning tasks are needed for two purposes: to check whether the network meets operational constraints and to find power losses for the particular state of the network. The main feature of load flow calculation for the network planning tasks is that it would be done many times and therefore it must be very fast.

In general case the network can be described by the its π equivalent. Then, omitting shunt capacitance of the distribution lines the transmitted can in be expressed as:

$$P_{ik} = V_i \left[V_i g_{ik} - V_k \left(g_{ik} \cos\delta_{ik} + b_{ik} \sin\delta_{ik} \right) \right], \tag{96}$$

$$Q_{ik} = V_i \left[V_i (-b_{ik}) - V_k \left(g_{ik} \sin\delta_{ik} - b_{ik} \cos\delta_{ik} \right) \right], \tag{97}$$

where P_{ik} and Q_{ik} is respectively transmitted real and reactive power from node i to node k, V_i is the voltage of the node i, δ_{ik} is the angle between two voltage vectors $\delta_{ik} = \delta_i - \delta_k$, b_{ij} and g_{ij} are respectively the line suseptance and admittance.

The state of the system is defined if the voltages and angles at all the nodes are known. These can be obtained solving the following system of power balance equations:

$$P_i = \sum_{k=1}^{n} P_{ik} = V_i \sum_{k=1}^{n} V_k \left[G_{ik} \cos\delta_{ij} + B_{ik} \sin\delta_{ij} \right] \tag{98}$$

$$Q_i = \sum_{k=1}^{n} Q_{ik} = V_i \sum_{k=1}^{n} V_k \left[G_{ik} \sin\delta_{ij} - B_{ik} \cos\delta_{ij} \right], \tag{99}$$

where n is a number of buses, P_i and Q_i are respectively real and reactive net power production at node i, B_{ik} and G_{ik} are imaginary and real parts of the of the admittance matrix.

The classical approach applied to solve this system of equations is Newton-Raphson method [89].

However, the distribution networks have certain features in comparison with other power system objects. The main differences can be listed as:

- Radial or weakly meshed structure
- High R/X ratio
- Unbalanced loads
- Dispersed generation.

Therefore, distribution networks fall into category of ill-conditioned power systems for generic Newton-Raphson and fast decoupled load flow methods.

Single-phase Alternating Current (AC) representation is the most popular analysis method for distribution. Numerous algorithms developed specially for calculation of AC load flow in distribution networks are available [124]. A computationally efficient solution scheme based on Newton-Raphson method is proposed in [42].

A large group of methods exploits the radial configuration of distribution networks [9],[123]. These algorithms consist of two basic steps: backward sweep and forward sweep. The backward sweep is a current or power flow summation with possible voltage updates. The forward sweep is a voltage drop calculation with possible current or power flow updates.

Implicit Z-bus method is presented in [25]. The method is based upon the principle of superposition applied to system bus voltages. The voltage of each bus is considered to arise from two different contributions: the specified source voltage and equivalent current injection.

4.6.2 DC load flow

The simplest representation is a so-called Direct Current (DC) model in which the electrical behavior is not analyzed in a complex manner, but represented simply as scalar quantities. This approximate model can be used where speed is more important than absolute accuracy, which is exactly the case in the preliminary planning studies.

During normal load-flow conditions the angle between the receiving and sending end voltages is only a few degrees, therefore for most practical cases the approximation $\delta_i - \delta_j = 0$ is acceptable. Therefore, equation (98) can be reduced to:

$$P_i = V_i^2 g_{ii} - V_i V_k \sum_{\substack{j=1 \\ k \neq i}}^{n} g_{ik} \ , \qquad (100)$$

where $g_{ij} \approx \dfrac{1}{R_{ij} + X_{ij} \tan\varphi}$, where $\cos\varphi$ is the load factor.

The equation (100) results in a system of linear equations, which can be written in matrix form as

$$P = Y\Delta \ , \qquad (101)$$

where $Y_{ij} = \dfrac{V_i}{R_{ij} + X_{ij} \tan\varphi}$, and Δ is the vector of nodal voltage drops.

The system of equations can be solved by any either direct or iterative method. The model treated in this dissertation performs DC load flow for simplified preliminary analysis solved by conjugate gradient method. AC load flow is performed for the final estimation of the alternative solution and based on Newton-Raphson algorithm.

4.6.3 Planning for the Peak Demand

Until recently most planning models were based on peak demand conditions. Then, the losses during the given period of time were approximated using losses utilization time. This approach is briefly described in this sub-section.

Two types of losses are distinguished: fixed losses (e.g. no load losses in transformers and reactors) and variable losses (caused by current flows through network components). It is assumed that utilization time of fixed losses is 8760 hours per year, while variable losses depend on peak load utilization time. This dependence can be approximated for example by the following equation [16]:

$$\tau_l = 8760\left[0.13\frac{\tau_p}{8760} + 0.87\left(\frac{\tau_p}{8760}\right)^2\right], \qquad (102)$$

where τ_l is losses utilization time and τ_p is peak load utilization time.

The annual energy losses can be found by integrating the load duration curve:

$$E = \int_0^T P(t)dt \,. \tag{103}$$

Alternatively, if peak load and therefore losses utilization times are known, the annual energy losses can be obtained from simple multiplication:

$$E = P_{L\max}\tau_l \,, \tag{104}$$

where $P_{L\max}$ stands for losses during the maximal demand.

The annual cost of losses can be approximated knowing loss utilization time τ and cost of energy C_E expressed in monetary unit per kWh. No-load losses P_{NLL} are assumed to take place during utilization time *T*.

$$C_{Losses} = \left[(P_L + P_{LL}) \cdot \tau + P_{NLL} \cdot T\right] \cdot C_E \cdot K \,, \tag{105}$$

where P_L is the total power losses in the network elements (excluding transformers) and P_{LL} is variable load dependent losses in transformers. The costs, which take place in more distant future, are converted to the present value using principles described section 4.2 with $K = (1+i)^{-t}$.

4.6.4 Consideration of Different Loading Conditions

Losses utilization time can be determined based on the collected past data about the demand and projected into the future. In case of entirely regulated power industry this approach in many cases gave a good approximation of total losses and their costs could be approximated via fixed energy tariffs. However, under deregulation when the power market determines the prices, this approach can give too approximate results. Therefore, the new condition encourage utilization of more detailed models for estimation of costs of power losses.

Thus, each year can be divided into N characteristic modes (see Figure 4-3 and Figure 4-4), which are modeled by their duration T_i, fixed load and fixed energy price C_{Ei} during each mode. Then, annual cost of losses can be approximated as a sum:

$$C_{Losses} = \sum_{i=1}^{N} \left(P_{Li} + P_{LLi} + P_{NLLi}\right) \cdot T_i \cdot C_{Ei} \cdot K \,. \tag{106}$$

The load flow calculation procedure is performed for each mode using the corresponding load forecasts and the forecasts for the energy prices.

4.7 Reliability Assessment for Network Planning Problems

4.7.1 Feasibility of Reliability Estimation in Planning Tasks

Quantitative reliability estimation is being recognized as necessary and is becoming feasible in the planning of electric power distribution systems [10]. The improvement in the network reliability level, or the decrease in interruption cost, usually leads to increase in investment cost.

Reliability evaluation of a complete electric power system including generation, transmission and distribution is normally not conducted due to the enormity of the problem. Instead, reliability evaluation of generating facilities, of composite generation and transmission systems and of distribution system segments are conducted separately [10]. Reliability assessment of real-size systems is a complicated task. Accurate methods require considerable computational efforts and are based on Monte-Carlo simulation techniques [3],[11]. Fortunately, there exist well-developed methods for approximate reliability assessment for distribution networks [10], which are suitable for planning purposes, since they allow for compelling reliability estimation for each state of the network.

4.7.2 Basic Reliability Indices

At distribution level, basic power supply reliability is defined by two sets of indices [10], namely, the load-point indices and the system performance indices. The primary reliability indices at a customer point are:

- expected frequency of failures, λ;
- the average duration of a failure, r;
- the average annual outage time (unavailability), U.

These indices depend on many factors such as the reliability of individual items of equipment, circuit length and loading, network configuration, load profile and available transfer capacity.

In radial distribution systems the calculation of reliability indices involves a system consisting of series components from source to load. Supposing there are n components in series the system failure rate λ_s will be:

$$\lambda_s = \lambda_1 + \lambda_2 + \ldots + \lambda_n \tag{107}$$

and the system failure duration r_s will be:

$$r_s = \frac{\lambda_1 \cdot r_1 + \lambda_2 \cdot r_2 + \ldots + \lambda_n \cdot r_n}{\lambda_1 + \lambda_2 + \ldots + \lambda_n}. \tag{108}$$

The system interruption time U_s will be:

$$U_s = \lambda_s r_s = \lambda_1 r_1 + \lambda_2 r_2 + \ldots + \lambda_n r_n . \tag{109}$$

Equipment failure rates and failure durations are the data obtained from statistics and their values vary in certain ranges. Even for the same equipment there are many types and sizes. These values depend also on age of the particular piece of equipment.

4.7.3 System Performance Indices

4.7.3.1 Customer-Orientated Indices

There is a wide range of possible system performance indices [10] of which the ones most commonly used and the most suitable for the problem under consideration are the System Average Interruption Frequency Index (SAIFI), the System Average Interruption Duration Index (SAIDI), the Customer Average Interruption Duration Index (CAIDI) and Energy Not Supplied index (ENS).

System average interruption frequency index:

$$SAIFI = \frac{\text{total number of customer interruptions}}{\text{total number of customer served}} = \frac{\sum \lambda_i N_i}{\sum N_i} \tag{110}$$

Customer average interruption frequency index:

$$CAIFI = \frac{\text{total number of customer interruptions}}{\text{total number of customers affected}} \tag{111}$$

System average interruption duration index:

$$SAIDI = \frac{\text{sum of customer interruption durations}}{\text{total number of customers}} = \frac{\sum U_i N_i}{\sum N_i} \tag{112}$$

Customer average interruption duration index:

$$CAIDI = \frac{\text{sum of customer interruption durations}}{\text{total number of customer interruptions}} = \frac{\sum U_i N_i}{\sum \lambda_i N_i} \tag{113}$$

Average service availability index:

$$ASAI = \frac{\text{customer hours of available service}}{\text{customer hours demanded}} = \frac{\sum N_i \cdot 8760 - \sum U_i N_i}{\sum N_i \cdot 8760} \tag{114}$$

4.7.3.2 Load and Energy Orientated Indices

Energy not supplied index:

$$ENS = \text{total enegy not supplied by the system} = \sum \mathrm{P_i U_i} \tag{115}$$

Average energy not supplied:

$$AENS = \frac{\text{total energy not supplied}}{\text{total number of customers served}} = \frac{\sum P_i U_i}{\sum N_i} \tag{116}$$

4.7.4 Customer Interruption Cost

Reliability estimation has been recognized as an important part of the system planning task. But it is also important to take into account the market value of the particular customer [2],[24],[132]. It could be done through a Customer Interruption Cost (CIC), which is defined as a measure of the monetary losses for customers due to an interruption of electric service. Customer interruption costs reflect the service value provided by a utility to the customers and the inconvenience or damage experienced by its consumers if a power failure occurs.

For many types of customers the issue of service reliability is simply a question of whether the supply is available or not. Other customers have quality requirements more stringent. Therefore, in the nearest future the utilities will face the problem of providing differentiated levels of reliability for different customers.

4.7.5 Reliability as Planning Attribute

Thus the value, which combines network utility unavailability data with customers' view on unavailability of supply can be used as reliability criterion in planning tasks [24]. The corresponding attribute is calculated according to the following equation:

$$C_{\mathrm{Re}liab} = \sum_{i=1}^{n} IC_i = \sum_{i=1}^{n} \left(\sum_{j \in m(j)} \lambda_j \cdot r_j \cdot P_i \cdot CIC_i(r_i) \right), \tag{117}$$

where the reliability criterion $C_{\mathrm{Re}liab}$ is calculated as a sum of load node interruption costs IC_i. The interruption cost is calculated for each node in the network as a sum of interruption cost due to possible failure of each upstream element $m(j)$ from the node to the feeding point. Finally, λ_j is a failure rate of the element j, r_j is its average outage time, P_i is average load at the load point and $CIC_i(r_i)$ is customer interruption cost due to failure of duration r_i.

If the information about the customer interruption costs is not available, the Energy Not Supplied itself can serve as reliability criterion. The expanded equation (115) takes the form:

$$ENS = \sum_{i=1}^{n} \left(\sum_{j \in m(j)} \lambda_j \cdot r_j \cdot P_i \right), \qquad (118)$$

where unavailability is calculated for each node in the network as a sum of interruption cost due to possible failure of each upstream element $m(j)$ from the node to the feeding point, λ_j is the failure rate of element j, r_j is its average outage time, P_i is the load at load point i.

The reliability attribute must be calculated for each load mode. Moreover, the economic principles must be taken into consideration even in case of (118), despite the fact that the ENS is an energy value. Thus the annual value of ENS must be multiplied by $K = (1+i)^{-t}$.

4.8 Investments and Other Costs

A major attribute of planning is reduction of cost. An assessment of various planning alternatives may be based on the capital investment cost alone if the additional network capacity provided by each option is comparable and if system maintenance costs are effectively the same. If they are not the same, the supplement in the costs must be taken into account. The change in maintenance costs usually associated with the addition of new objects (for example a new substation).

Besides, there are costs of equipment, land, labor, etc. These values can be estimated by the utility for each possible planning option.

The investment criterion can be calculated as a sum of annuities of the investments and the corresponding supplements to the maintenance costs.

4.9 Environmental Concerns

Environmental concerns in cable distribution networks are caused mainly by oil leakages from pressurized oil-filled insulated cables. However, cables of this type are being more and more replaced by XLPE insulated cables. Therefore, for example the total length of oil-filled cables in the network can be suggested as environmental criterion.

The visual impact and the land usage in some cases may become the major factors in planning of the overhead lines.

4.10 Power Quality

A traditional approach to distribution planning typically does not consider explicitly the customer costs associated with poor power quality.

Table 4-2 General categories of power quality: impact, indices and countermeasures

Category	Impact	Indices [16],[41]	Countermeasures
Short-duration rms variations - voltage sags, swells or interruptions, which last less than one minute	Equipment malfunction	SARFI SIARFI SMARFI STARFI	Reduce the incidence rate, amount of variation or the duration of rms variations
Sustained interruptions – loss of power for more than one minute	Work or production stoppage for the customer	SASIFI SATIFI SASIDI ASIDI*	Limit the incidence rate and duration of sustained interruption
Voltage regulation – steady state problems with voltage magnitude, including unbalance	Excessive heating of equipment and damage of insulation	SAEVUR SAENSR SAEVDR	Improve voltage regulation and balance
Harmonics – components in the supply voltage with frequencies at an integral multiple of the fundamental frequency	Excessive equipment heating and control malfunction	$\mathrm{SAENR_hI}$	Use active or passive filters to reduce harmonic voltage distortion
Transients – impulsive and oscillatory overvoltages mainly due to lightning and utility capacitor switching	Equipment damage or control malfunction	SATMORI	Reduce the magnitude and incidence of switching transients and lightning surges entering the customer facility

The potential economic impact to the customers from each of these categories differs according to the sensitivity and thermal ratings of customer devices. In

* The indices are analogous to the 5-minute sustained interruption indices SAIFI, SAIDI and CAIDI defined in [9] and used for quantification of interruptions longer than 5 minutes.

general, sustained interruptions and short-duration rms variations result in the most significant impact over time. In [40] the methodology, which allows for considering power quality indices of these categories in distribution planning process is suggested.

Traditionally, the sustained interruption indices have been calculated as measures of service quality. However there is a number of indices [16],[41], which provides a qualification of service quality with regard to short duration rms variations, harmonic distortion and other power quality phenomena. Five general power quality impact categories [40] are summarized in Table 4-2. In addition, the table contains the primary indices used to quantify the effect of each category and possible countermeasures.

4.11 Constraints

During the whole planning period the network must satisfy the following constraints:

- Kirchhoff equations and supply-demand balance
- electrical performance
- radiality of configuration
- reliability indices
- standard equipment types
- logical conditions.

The first set of constraints is obvious and it can be omitted in the optimization task if load flow calculations are provided externally.

Thus, security constraints are related to standards for acceptable electrical performance both under normal and abnormal operation. They involve a number of factors, such as lines and transformers transmission capacities, voltage drops, power losses, thermal limits, short circuit currents etc. Some of them (transmission capacities) may always be critical, while others may be either critical or not depending on particular network component or the circumstances. In [73] the circumstances are summarized, when some of the security constraints are critical.

Table 4-3 The circumstances when the particular constraints become critical

Constraints	Components and Circumstances
Voltage Drop	long lines in medium and low voltage networks
Losses	medium voltage overhead main lines
Thermal Limit	medium and low voltage underground cables
Short Circuit Current	• lines supplying the small load near primary substations (too high) • long low voltage lines with small load (too low)

Since distribution networks at both medium voltage and low voltage levels are normally operated radially, radiality constraints, which force only radial solutions are introduced. However, in some cases, radiality constraints may not be considered explicitly - then it is a separate task to find optimal open points for radial operation.

Reliability of supply may be considered both as a planning criterion, when unavailability of supply is minimized, and as a constraint. In the last case the critical reliability indices must not be exceeded.

Each utility may have additional constraints on layout and design of their systems. Thus, the planner should keep in mind, that he only may use the equipment of approved types and sizes. The approved list of equipment and sizes generally provides a range of capabilities to meet most needs, but uses only a small portion of equipment that is available in the industry.

And finally, there might be given logical constraints, which on one hand strictly depend on particular task and imply the restriction on the number of new possible network components; on the other hand they prevent locating several components at location foreseen for one.

4.12 Conclusions

- Fundamental set of planning attributes essentially includes three objectives: minimization of losses, energy not supplied and capital investments.
- Different expenses do not coincide in time, therefore in order to compare the costs of different alternatives the timing of the cash flows must be taken into account.
- Electric model of the network can be built with different levels of details. However, the model essentially must reflect the capacity requirements and basic electrical behavior of the network elements.
- Customer demand is a parameter, which changes considerably during the day, week and year. This is the cause of difficulties in estimation of power losses. The methods with different degrees of approximation are available.
- Reliability assessment for real-sized systems is a complicated task. Fortunately, approximate reliability assessment methods are suitable for planning purposes.
- Probabilistic load modeling and calculation of power losses may be based either on:
  - ✓ Direct utilization of the results of measurements
  - ✓ Choice of suitable empirical probability distribution (normal, log-normal, beta, etc.)

5 Risk Management

In presence of uncertainty the planner aims at finding the robust and flexible plans to reduce the risk of considerable losses. Several measures of risk are discussed in this chapter. It is shown that measuring risk by regret may lead to the risky solutions, therefore an alternative measure – Expected Maximum Value – is suggested. The general future model called fuzzy-probabilistic tree of futures, which integrates all classes of uncertain parameters (probabilistic, fuzzy and truly uncertain), is described in this chapter.

5.1 Managing Risk

Uncertainty and risk management is becoming an essential and integrated part of the planning process in the electric power industry. For the past decades the uncertainties in load growth, capital costs and regulatory standards have challenged the planners. Uncertainty imposes risk and explicit risk management strategies are being developed.

General strategies for dealing with risk include the following [4]:

- Investing for flexibility so that the changes can be easily and inexpensively made
- Investing in projects that are robust, i.e. perform well across a variety of futures
- Hedging against uncertainty
- Ignoring risk.

Flexibility allows for easy and inexpensive changes to be made if the adaptation to the future conditions is needed. This can be achieved by building the system from small, modular components or by specifying investments with a relatively short lifetime. On the other hand projects, which can easily be adapted to changing future circumstances, are also preferred. Additional information typically improves the decision-making process, therefore if the degree of uncertainty is high, it may be advisable to postpone the investments and wait for additional information.

The planner can also choose the alternatives that perform well across a variety of possible futures – *robust* solutions. Robust investments may reduce the variance in possible outcomes by reducing the use of an uncertain input.

Rather than reducing the variance in performance of a single investment by making it robust, one could reduce variance in the performance of the overall portfolio of the firm's investments. Such a *hedging* strategy would couple investments with complementary vulnerabilities.

Finally, in some cases it is rational to ignore the effects of uncertainty – on small projects, short-term projects or whenever one feels confident in predicting the course of future events.

5.2 Approaches for Flexible and Robust Planning

5.2.1 Scenario Analysis – General Approach

The scenario technique is presently the most widely used method for representing uncertainties in planning tasks [29]. Furthermore, in [138] multi-scenario planning is appointed as "the only completely valid way to handle uncertainty in transmission and distribution forecasting and requirement planning".

The planner faces a difficult task of identifying uncertainties that could be of importance and may seriously influence the final solution, and those, which do not. The most useful and the easiest approach is one that is termed "thematic". This approach starts with themes (such as "rapid load growth" or "low load growth"). Important variables are then identified and values for these variables are chosen that would correspond to each of these different themes. Some guidelines for building scenarios are given in [45]: scenarios should challenge assumptions, discount extrapolations and question historical trends and eventually take into account possible technological wonders. It is advised not to assign probabilities to events or trends, since this task is very difficult and there is no real benefit. Some of these general recommendations may be adopted to network planning, while some may be argued.

The scenario technique can be described by the following four steps:

Stage 1 Selection of the alternatives to be examined.

Stage 2 Construction of scenarios by assigning plausible values to uncertain parameters.

Stage 3 Calculation of attributes for every scenario combining each future with each alternative plan.

Stage 4 Selection of a strategy according to a given decision criterion.

A number of criteria can be used in order to select the final strategy. The most commonly used and the most suitable criteria for the problem under consideration are presented in section 5.7.

5.2.2 Scenario Approach with Internal Optimization

Deterministic optimization algorithms may be adapted to better acknowledge uncertainty by using scenario analysis. Unlike sensitivity analysis that changes only one variable at the time relative to a base case, scenario analysis constructs several different futures and identifies optimal and near-optimal plans for each of them. Robust elements are those included in most of the optimal plans generated for the range of scenarios. Probabilities may be assigned to the various scenarios to allow expected value comparison of alternative plans.

Reference [37] presents a powerful decision-making tool for transmission planing, which allows the planner to qualify and hedge risk and lead to identification of robust plans. Although, this type of scenario analysis can be used to facilitate robust planning, it is not well suited for evaluating flexibility.

5.2.3 Scenario Approach with External Optimization

In order to assess flexible options and optimize on multiple criteria a modeling framework can be built around some planning package, which is used as a simulator to evaluate the plans generated externally. The framework generates many scenarios, the simulator evaluates them and either decision analysis [67] or trade-off analysis techniques identify a preferred plan [33].

5.2.4 Stochastic Optimization

Uncertainty can be brought inside of the optimization process either relying on internal comparison of scenarios or performing Monte-Carlo simulation. Stochastic optimization has not been extensively used in distribution planning, but it found a common application in resource and transmission planning [50].

5.2.5 Decision Trees

The decision trees technique is an important modeling tool for integrating a number of component of a decision-making problem into a formal layout [67]. Decision trees provide a graphic tool for modeling and analyzing probabilistic multi-stage problems. The decision-maker can represent the states of the system, its multiple stages and corresponding decisions and the potential outcomes with associated probabilities. Like a real tree the decision tree contains a number of limbs which continue to branches as one proceeds along the tree. Decision trees have two types of nodes – decision nodes and chance nodes. A decision node represents a juncture where the choice of which branch to follow depends on the decision-maker, who can choose from any of the available alternatives at that node. A chance node represents a situation where the decision-maker has no control over which branch is selected. The path taken though the tree determines the payoff.

Decision trees provide an excellent tool for modeling decisions that involve a sequence of choices over time. The traditional approach of solving a decision tree is to start from the terminal branches and to calculate the expected utility proceeding downstream, choosing at each decision node the branch with minimal expected costs. The procedure is repeated until the base of the tree is reached.

In [130] the scenario technique is compared with decision trees analysis illustrating some advantages of the last on elementary example. In [105] it is suggested to use decision trees to enhance transmission and sub-transmission planning, but decision trees are used to identify a set of rules which combine criteria based on heuristics and sensitivities. In [109] a multi-objective algorithm for determining a restoration plan for radial distribution network supply is based on decision tree, efficiently selecting a specified number of "high quality" variants of alternate supply.

The decision tree analysis however is only suitable when the number of options is not too big, otherwise the task becomes huge. Another limitation of the decision tree analysis is that probabilities must be associated with each combination of decisions and events.

5.3 Measures of Risk

5.3.1 Robustness, Exposure and Regret

Robustness, the most fundamental measure of risk, is the likelihood that a particular decision will not be regrettable [106]. If the plan chosen by the decision-maker is optimal for all the scenarios, then the choice is 100% robust. More often the choice is optimal only for a subset of the possible future. Supposing the probability of 80% that the future from this subset will occur, the robustness of the chosen plan can be assigned to 80%.

Exposure is a measure of loss if an adverse materialization of uncertainties occurs for a particular choice. *Regret*, which is the difference in exposure between some choice and the best choice for a particular realization of the uncertainties is the measure most commonly used to quantify risk in power system planning and a number of other applications. Consider that the value of the objective function for the particular plan and future is f_{ij}, and the optimal plan for the future j is f_j^{opt}, then the regret is:

$$R_{ij} = f_{ij} - f_j^{opt} . \qquad (119)$$

Regret is zero for an optimal plan for a particular future. If the regret is zero for the same plan and for all the futures, then the plan is robust.

5.3.2 Standard Deviation

When dealing with uncertain information of probabilistic nature, or the subjective probabilities can be obtained the alternative measure of risk can be used. In this case it might be convenient to estimate deviation of the parameter from its mean value. These can be estimated for example as a standard deviation according to:

$$\sigma = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(f_{ij} - E\left[f_i\right]\right)^2}, \qquad (120)$$

where n is a number of possible futures, f_{ij} is the value of the objective function for the particular plan i and future j, and $E\left[f_i\right]$ is the expected value of function for the plan i.

However, direct application of measure of variation according to (120) would be insufficient to quantify risk. In [33] the inconsistency of such an approach is illustrated on a simple example (see Table 5-1 in section 5.3.3).

5.3.3 Value-at-Risk and the Expected Maximum Value

The Expected Maximum Value (EMV) of the attribute at the given level of confidence can be suggested as an alternative measure of risk. This measure is similar to the Value-at-Risk (VaR) extensively used in financial management [5].

Under assumption that the considered uncertain parameter is continuous and follows some probability distribution, the Expected Maximum Value for each alternative can be calculated as:

$$EMV_i = M\left[f_i\right] + K \cdot \sigma_i, \qquad (121)$$

where K is the coefficient, which depends on particular distribution and level of confidence. For example, for a normal distribution and 95% one-tail confidence limit K is equal 1.65.

If the uncertain parameter takes only the discrete values, the Expected Maximum value can be calculated as a maximum value of the attribute for each alternative and for all the plans, which may occur with a probability higher than a given level of confidence.

If the level of confidence is not given and unknown to the planer (but is the subject of future assessment by the decision-maker) then the whole probability distribution function is needed instead of single value. This is less convenient for two reasons:

- larger amounts of information must be analyzed in order to compare the plans,
- more Monte-Carlo trials are needed to obtain the function of probability distribution.

Knowing that in the general case the uncertainty in the network planning tasks results from the sum of the random parameters in many network elements, and applying the Central Limit Theorem [54], normal distribution appears to be a reasonable assumption. Therefore, if the level of confidence is not given, the mean and standard deviation should be maintained, but the EMV can be calculated later assuming normal probability distribution.

To compare three measures of risk consider the example adapted from [33]. Suppose that four plans are measured in terms of two attributes with tree futures (Table 5-1). Suppose the task is to minimize both attributes and for simplicity assume that all three futures may occur with equal probability.

Table 5-1 Comparison of Minimax Regret Criterion, standard deviation and Expected Maximum Value as a measure of risk

	Plan A		*Plan* B		*Plan* C		*Plan* D	
	Att 1	*Att* 2	*Att* 1	*Att* 2	*Att* 1	*Att* 2	*Att* 1	*Att* 2
Future 1	3	13	1.6	15	1.2	30	2.3	23
Future 2	3.1	14	1.1	19	1.1	31	2.3	23
Future 3	3.2	11	0.8	13	0.9	29	2.3	23
$E[f]$	3.10	12.65	1.16	15.64	1.07	29.99	2.30	230
σ	0.14	2.16	0.57	4.32	0.22	1.41	0	0
$\max(f_{ij})$	3.2	14	1.6	19	1.2	30	2.3	23
Regret 1	1.8	0	0.4	2	0	17	1.1	10
Regret 2	2	0	0	5	0	17	1.2	9
Regret 3	2.4	0	0	2	0.1	18	1.5	12
Max(*R*)	2.4	0	0.4	5	0.1	18	1.5	12

Minimization of the expected values of the attributes results in Pareto optimal set, which consists of three plans: Plan A, Plan B and Plan C.

Consider the variance was taken as a measure of risk. Plan D minimizes variance of both attributes. It is very illustrative, that Plan B, which is the best compromise choice, maximizes the variance of both attributes.

In Figure 5-1 four plans are compared in terms of regrets. Plan B, which appears to be the best for all three futures, is also the best trade-off for both regrets.

Similar conclusion can be drawn comparing the solutions in terms of the Expected Maximum Value for both attitudes (Figure 5-2).

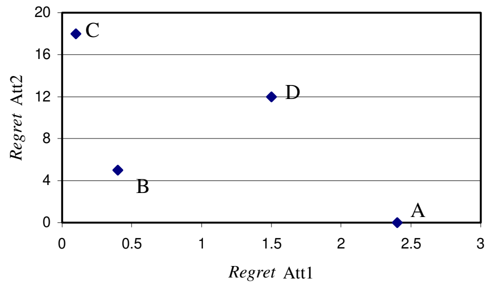


Figure 5-1 Comparison of the plans measured in terms of regret

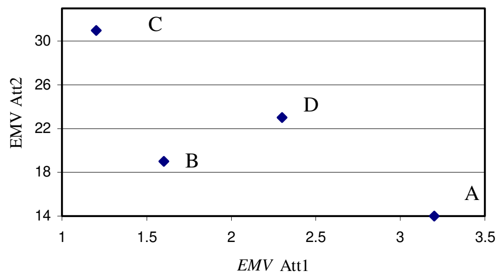


Figure 5-2 The comparison of the plans measured in terms of the Expected Maximum Value

Now consider the following rudimentary example published in [106]. Two options have to be evaluated under two possible scenarios. The cost of option A for scenarios 1 and 2 are respectively 0 and 60. The costs of B for the same scenarios are 40 and 30. Assuming that the scenarios have the same probability, the expected costs are 30 and 35, respectively (see Table 5-2), therefore under the least expected cost criterion Plan A would be selected. We observe, that A presents extreme results – very good in Future 1, but exposed in Future 2. Therefore, intuitively most people would prefer option B as less risky.

Table 5-2 Comparison of Minimax Regret Criterion and the Expected Maximum Value

	Plan A	*Plan* B
Future 1 *Future* 2	0 60	40 30
$E[f]$	30	35
EMV	60	**40**
Regret 1 *Regret* 2	0 30	40 0
Max(R)	**30**	40

Next, let us try to evaluate risk numerically in terms of regret and the Expected Maximum Value. The results are summarized in Table 5-2. In this case the methods give the opposite answers to the question which solution should be preferred. But intuitively it seems that the answer given by the Expected Maximum Value is more adequate.

Furthermore, it can be added, that it is appropriate to measure risk by regret when playing with the opposing competitor, who chooses his moves depending on our selection. On the contrary, network planning task can be considered as a game against the nature – the natures actions are obviously independent on our moves.

5.4 Classical Trade-Off/Risk Analysis

Power network planning is a decision-making problem with a number of planning objectives. Often it is not possible to identify a single plan, which satisfies all objectives – a trade-off between the objectives is required. Furthermore, in presence of uncertainty risk evaluation becomes the essential part of analysis.

The trade-off/risk method for multi-criteria planning with uncertainty has been developed in the beginning of eighties to support the identification of robust plans [18],[81]. It represented a new theoretical advance in decision-making to support generation expansion, demand planning, maintenance scheduling and transmission planning.

The method has the following tree main steps:

- Formulate the problem and compute attributes for very many scenarios
- Use trade-off concepts to identify the "decision set" – the set of plans left after all inferior plans have been rejected
- Analyze the plans in the decision set to eliminate more plans and support the development of a final strategy.

In [18] the Conditional Decision Set, containing the candidates for the final decision is defined as a union of a trade off curve and a knee set (see Figure 5-3), defined as:

- The trade-off curve is a set of all plans that are not strictly dominated by any other plan on a particular future
- The knee set is a set of all plans that are not significantly dominated by any other plan on a particular future.

The significant and strict dominance are defined bellow. Let $a_i(P_1)$ and $a_i(P_2)$ be the values of attribute i for two plans P_1 and P_2. Plan P_2 Strictly dominates plan P_1 if $a_i(P_2)$ is better than $a_i(P_1)$ for all attributes. Plan P_2 significantly dominates plan P_1 if at least one attribute $a_i(P_1)$ is "much worse" that $a_i(P_2)$ and if no attribute is no "significantly better" than $a_i(P_2)$.

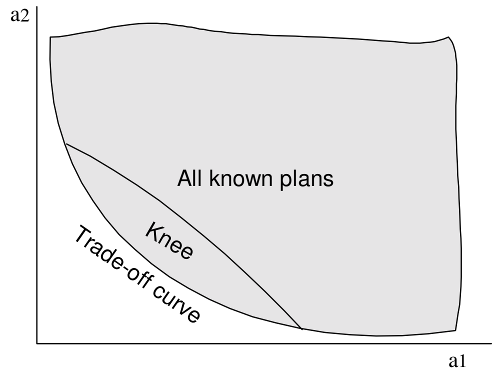


Figure 5-3 Example of conditional decision set

The global decision set contains the plans from all the conditional decision sets (for all the futures). The final decision can be made after analyzing the global decision set and identifying the most robust plans.

In [33] the most robust plans are identified as those present in all conditional decision sets. If there is no such plan, then the robustness of each particular plan can be measured according to the frequency of occurrence in the conditional decision sets.

5.5 On Choice of the Decision-Making Criteria

The decision made having qualitative estimates of risk depends on the decision-makers risk-aversion attitude. Undoubtedly, minimization of risk will lead to robust solution. However, in many cases it may be considered too conservative for a utility to plan for the worst possible future, which might occur with a very low probability.

Consider again the example of property insurance. Both parties – the company and the person buying the insurance – have the objective to maximize their revenues. The deal is beneficial for both, since they have different risk aversion profiles. The person, who insures the property, uses the Minimum Risk criterion. He prefers rather to have small negative income (insurance fee), than a very considerable loss, even at very low probability. The insurance company will use the Expected Value criterion. The different attitude can be explained by different level of resources available to both parties – the person have very limited resources, while the company has a high financial capability. Furthermore, the company has a large number of clients, therefore in average it will gain, even if in some cases it will have to pay a premium.

Similar speculation can be extended to the network utilities. If the project under consideration is only one of many, and possible losses are small compared to the budget of the company, then the Minimum Risk criterion is far too conservative.

In some situations the preferable solution may be found performing a multi-criteria decision analysis and try to find the trade-off between the expected costs and risks. According to the accepted standpoint, here the EMV is used as a measure of risk. Consider the situation illustrated in Table 5-3. Three possible plans are analyzed and both the expected costs and possible EMVs are calculated. All the scenarios have the same probability.

Table 5-3 Expected Costs and Regrets: situation "no conflict"

	Plan 1	Plan 2	Plan 3
Sc1	50	52	70
Sc2	56	54	60
Sc3	59	60	50
Expected Cost	**55**	55.4	60
EMV	**59**	60	70

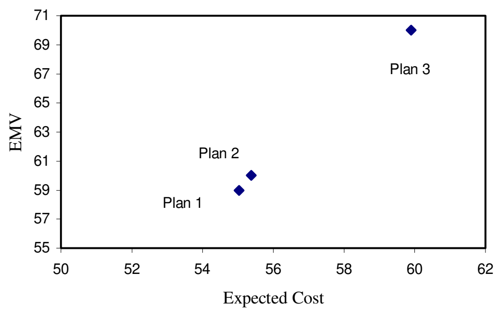


Figure 5-4 The trade-off between the Expected Cost and EMV: no conflict situation

In this example both Expected Cost criterion and minimization of EMV give the same answer, there is no conflict. As a matter of fact such situations are quite common in real-life applications. The "no-conflict" situation is illustrated in Figure 5-4, where the EMV for each plan is plotted versus the corresponding expected cost.

However, consider a different situation (Table 5-4). Minimization of Expected Cost and EMV give two different answers. If the planner chooses Plan 1, which minimizes risk, the solution can be expensive if future 1 occurs. However, Plan 3, which has minimal expected cost and very cheap for scenarios 1 and 2, is very exposed in scenario 1. The situation is illustrated in Figure 5-5. Plan 2 represents a good trade-off between the expected cost and possible risk.

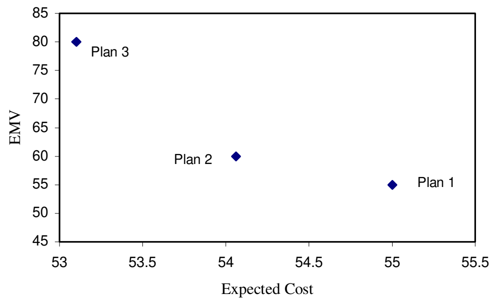


Figure 5-5 The trade-off between the Expected Cost and EMV criteria: conflict situation

In this case the solution minimizing the regret for tree alternatives coincides with the choice made according to minimal EMV. Both criteria recommend Plan 1. Assume, however, that there is the fourth plan under consideration. In Table 5-4 this alternative is denoted as “bad” plan, since it is apparently the wrong choice – the alternative is too exposed in case of the future 1. The appearance of this alternative will have no influence on the expected value and EMV for the initial three plans. However, the regrets must be recalculated. The new values are given in shaded cells.

Table 5-4 Expected Costs and Regrets: ”conflict” situation

	Plan 1		Plan 2		Plan 3		“Bad” Plan
Sc1	55		48		80		130
Sc2	55		54		50		20
Sc3	55		60		30		30
Regret 1	7	7	0	0	32	32	82
Regret 2	5	35	4	34	0	30	0
Regret 3	25	25	30	30	0	0	0
Max Regret	**25**	35	30	34	32	**32**	82
Expected Cost	55		54		**53**		60
EMV	**55**		60		80		130

It is illustrated, that in presence of the fourth alternative the Plan 3 is declared to be the least risky according to the regrets. This example puts emphasis on the drawback of the regret minimization approach in planning tasks: the alternatives, which may be good for some futures, but very exposed for the others, and therefore which could not be valid as candidates, influence the relative position of the alternatives, which are robust and can be selected as candidates for the final decision.

5.6 Network Development Model

The degree of uncertainty increases in time - the further into the future we look, the more uncertain this future is. The future may be very well predictable for the nearest few years. Therefore, a corresponding model of network development, which captures the dynamics of uncertainty, is suggested. The model [34], [71] divides the planning horizon into several periods defined below.

Planning period corresponds to the economic life cycle of most electric power equipment. The planning period is divided into several *development stages* (Figure 5-6). The duration of different stages can vary. It is assumed that action (if any) is realized during the first year of the stage.

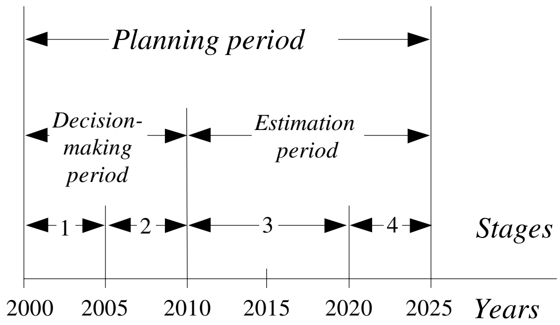


Figure 5-6 Illustration of the planning period consisting of decision-making and estimation period

Realization of the most essential action, the economic efficiency of which should be estimated immediately, is foreseen during the *decision-making period*.

Planning period and decision-making period have a common starting point. The time interval after the decision-making period until the end of the planning period is termed *estimation period*.

The dynamic approach implies that in order to estimate the economic efficiency of every alternative the whole economic life cycle (about 25-30 years) should be considered. However, the decisions can be made only for the nearest future (3-5 years). Furthermore, it is reasonable to assume that uncertain parameters (such as loads, prices, and economic indices) are known during the decision-making period, but can vary from one scenario to another during the estimation period.

Thus, the planner is interested in finding reinforcement actions to realize during the decision-making period. However, the consequences of these actions must be estimated for one or several futures during the whole planning period.

Therefore, it is consistent to assume the following definition of the *equality of* *the plans*: two plans are said to be equal if the actions to be realized during the first stage are equal. The remaining actions may be *adapted* to the particular future in order to provide a proper comparison of the attributes for these plans.

To illustrate the suggested modeling approach consider the following real-life example. A 20 kV electricity distribution network (Figure 5-7) is situated 30 km from the feeding substation 110/20 kV. A new customer - a sawmill - with initial load 3 MW, which will probably double during the next 20 years, has to be connected to the network.

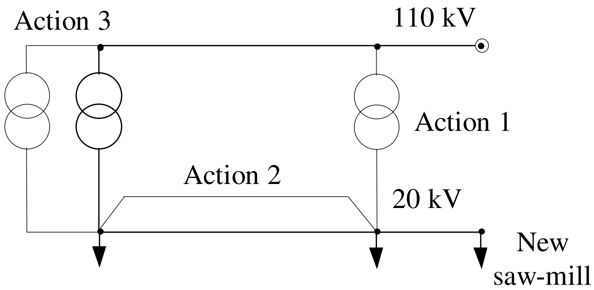


Figure 5-7 Alternative actions to provide electricity supply of a new customer

The task is to find a scheme, which would provide a reliable, qualitative and economic sawmill electricity supply during a 25-year planning period.

For simplicity the planning period 2000 - 2025 is divided into two time stages: decision-making period (5 years) and estimation period. Load flow calculation for the existing network shows that already at the first stage the voltage drops threshold is exceeded, at the next stage the transformer is overloaded, and by the end of the planning period, the line, which connects the sawmill with substation, is overloaded. Therefore, it is necessary to perform some reinforcement actions in the network starting from the first time stage. Possible actions suggested to reinforce the existing network are presented in Table 5-5 and depicted in Figure 5-7 by dotted lines.

Table 5-5 Possible reinforcement actions for the sawmill example

Action NR	Sort of action	Cost, USD·10^3
A_1	Construction of a new substation	800
A_2	Construction of a new line	300
A_3	Reinforcement of existing substation – change of transformer. May be realized only starting from the second stage.	80

Table 5-6 contains all the possible development scenarios of the network reinforcement for two futures – Future 1 corresponds to the rapid load growth and Future 2 to the intermediate load growth. The objective function consists of two attributes – investment cost and hypothetical cost of losses. On the first stage, there are only two possible options: action A_1 or action A_2. For each option there are four possible continuations on the second stage, but despite that there are only two plans, the optimal continuation of the plans have to be chosen for comparison.

Thus, for the Future 1 and Plan 1 the action A_1 has to be supplemented by action A_3 at the second stage. These results are comparable to the results for the Future 2, where Plan 1 contains only action A_1. Similarly for the Plan 2: for the Future 1 the action A_2 has to be supplemented by action A_3 at the second stage and results are comparable for the Future 2, where Plan 2 contains only action A_2.

Table 5-6 All possible scenarios for the sawmill example

Actions realized during the whole planning period	Decision-making period		Estimation period		
	C_{Inv} USD·10^3	C_{Loss} USD·10^3	C_{Inv} USD·10^3	C_{Loss} USD·10^3	C_{Total} USD·10^3
	Future 1 - Plan 1				
A_1	800	1500	0.0	1900	4200.0
A_1+A_3	800	1500	54.4	1800	4154.4
A_1+A_2	800	1500	204.2	1750	4254.2
A_1+A_2+A_3	800	1500	258.6	1700	4258.6
	Future 2 - Plan 1				
A_1	800	1500	0.0	1630	3930.0
A_1+A_3	800	1500	54.4	1600	3954.4
A_1+A_2	800	1500	204.2	1550	4054.2
A_1+A_2+A_3	800	1500	258.6	1500	4058.6
	Future 1 - Plan 2				
A_2	300	1650	0.0	2400	4350.0
A_2+A_3	300	1650	54.4	2200	4204.4
A_1+A_2	300	1650	544.5	1750	4244.5
A_1+A_2+A_3	300	1650	598.9	1700	4248.9
	Future 2 - Plan 2				
A_2	300	1650	0.0	1950	3900.0
A_2+A_3	300	1650	54.4	1920	3924.4
A_1+A_2	300	1650	544.5	1550	4044.5
A_1+A_2+A_3	300	1650	598.9	1500	4048.9

The resulting values for the sawmill example are summarized in Table 5-7. It is assumed that probabilities of occurrence of both futures are comparable. From these values the planner may assign a final decision according to the chosen decision-making criterion.

If EMV is applied as a measure of risk then Plan 1 contributes the most robust solution.

Table 5-7 Resulting values for the sawmill example

	Plan 1	Plan 2
	C_{Total} USD·10^3	C_{Total} USD·10^3
Future 1	4154.4	4204.4
Future 2	3930	3900
EMV	4154.4	4204.4

5.7 Suggested Model for Uncertainties

5.7.1 Classes of Uncertainties Considered in the Model

In Chapter 3 the types and sources of uncertainties, as well as the main modeling techniques were described in an explicit manner. With the purpose of modeling long-term uncertainties, the information can be categorized only into four principal classes:

- Deterministic information.
- Uncertain information, either continuous or discrete, with known either probability distributions or probabilities of events, respectively. The same class includes the information, if the statistical characteristics are not available, but the subjective probabilities can be assigned.
- Fuzzy – often linguistic information about the possible future trends and their range of variation.
- Truly uncertain information with unknown probability distributions.

In reality the boundary between all the classes is very thin and categorization of information into one or another class can depend on statement of the particular problem and preferences of the planner. However, the approaches to the problem decision may vary depending if the probabilities (or subjective probabilities) are assigned. If the probabilities (or subjective probabilities) are available the expected values for the attributes can be found. As stated previously, the expected value is not a sufficient criterion for comparison of the alternatives – it does not capture the volatility of the results across different scenarios – therefore even the probabilistic model requires additional risk analysis to facilitate the well-motivated decision-making. In this case the EMV can serve as a measure of risk.

5.7.2 Suggested Model: Fuzzy-Probabilistic Tree of Futures

If relevant information belongs to the class of truly uncertain information, the only possible modeling approach is a scenario technique. It is convenient to represent the scenarios in a tree like structure (Figure 5-8) [49]. The tree structure reflects the reality in a sense, that the degree of uncertainty increases with the time horizon the planer considers. Then the attributes for each alternative in every future must be compared and the decision can be made according to one of the criteria presented in section 3.3.

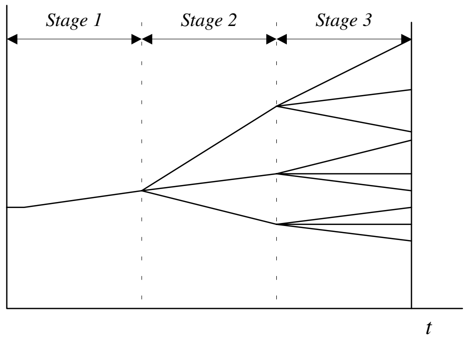


Figure 5-8 Tree of futures

The occurrence or absence of some events – for instance the regulatory decisions – may be modeled in terms of parameters for distribution planning. Such events as for example possible appearance of the big customers directly lead to the discrete scenarios of network development, with or without assigned probabilities.

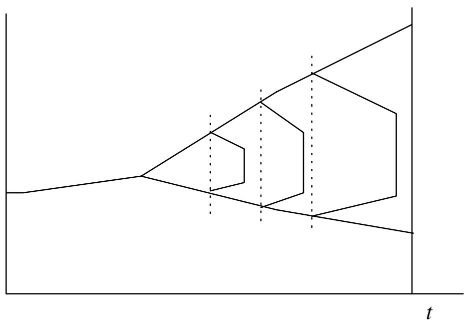


Figure 5-9 Fuzzy future

However, normally the scenarios can be estimated only approximately. Typically they are described in linguistic form, for example “in 10 years the load will be *about* 50 MW”. Thus, the range of variation of the uncertain parameter should be taken into account, which in general case can be done introducing fuzzy futures (Figure 5-9).

Furthermore, there are informational situations when the random parameters can be described by probability distribution, but the parameters of the distribution can be set only approximately. The example of such a parameter is load. In this case the fuzzy-probabilistic model described gives the most general representation of the available data.

The suggested general future model integrates all three classes of uncertain parameters – probabilistic, fuzzy and truly uncertain. The cross-section of the suggested model with two fuzzy futures is depicted in Figure 5-10.

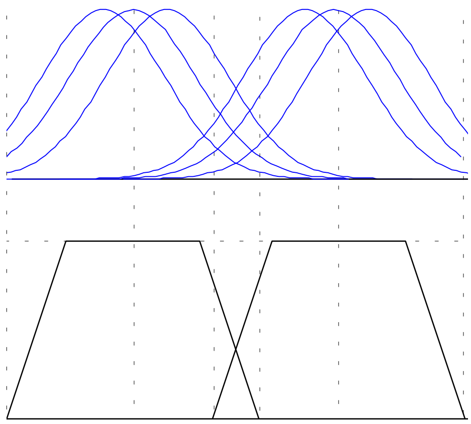


Figure 5-10 Fuzzy-probabilistic futures

The degree of uncertainty increases with time. Clearly, the further into the future we look, the more uncertain this future is. On the contrary, in many cases the future may be very well predictable for the nearest few years. Therefore, it is logical to build a corresponding model of network development, which would capture the dynamics of uncertainty.

Consider we want to estimate the value of some uncertain parameter looking into the future. Suppose, the period of time of interest can be divided into several stages each of which begins at $t_0, t_1, ..., t_l$. Variation of the parameter due to uncertainty can be decomposed into two components. For example, one of them can be discrete and another one continuous. Then, the expected value of parameter Π at time t_1 looking from the present time t_0 can be estimated as a sum:

$$E\left[\Pi_1, t_0\right] = \Pi_0 + \sum_{k=1}^{m} \Delta\Pi_{k1} , \tag{122}$$

where $\Delta\Pi_{k1}$ are the components of variation in random parameter at time t_1 looking from the present time and m is the number of components, which can be 1, 2 or more.

Similarly, the expected value of Π at t_2 looking from the time t_1 can be estimated as

$$E[\Pi_2, t_1] = \Pi_1 + \sum_{k=1}^{m} \Delta\Pi_{k2} . \tag{123}$$

However, if looking two stages ahead, the uncertainty will increase leading to:

$$E[\Pi_2, t_0] = \Pi_0 + \sum_{k=1}^{m} \Delta\Pi_{k1} + \sum_{k=1}^{m} \Delta\Pi_{k2} . \tag{124}$$

Therefore, the expected value of parameter Π at time t_n looking from the time t_j can be expressed by the following general model:

$$E\left[\Pi_n, t_j\right] = \Pi_j + \sum_{i=j}^{n} \sum_{k=1}^{m} \Delta\Pi_{ki} . \tag{125}$$

Similar speculations may be extended in order to model the dispersion.

It is remarkable, that for the loads it is convenient to assume presence of two components of load variation and giving them physical meaning. Two types of uncertainties affecting loads can be distinguished [32]. The first one is long-term, associated with economic environment of the service area. This type may be characterized by certain trends depending on the areas demographics and levels of industrial activities.

The second type can be considered as a short-term, and is related to time/weather factors. This one contributes randomness of the information, however in many cases the statistical data are available. In many systems temperature is the most relevant weather factor affecting the load. On the other hand, there are three main time factors, namely seasonal effects, weekly-daily cycle and holidays.

The majority of the network planning problems in the developed countries have to deal with planning having well developed existing network as a base. The existing loads are often steady. The uncertainties may be associated with regulatory changes leading to different load levels. Another possible source of uncertainty is the realization of some large projects – a new industrial customer or a local generation source.

Then again, the short-term uncertainties caused by time-weather factors are present in any planning task. Furthermore, their influence increases as the voltage level decreases. In addition, in distribution networks there might be strong correlation between these random load related variables.

5.7.3 Future Modeling Using Monte-Carlo Sampling

Applying the method of Monte-Carlo the values of random variables are chosen by a random number generator according to the given probability distribution. Referring to equation (125) the random parameter can be defined at each time stage based on the results of sampling on the previous stages as follows:

$$\Pi(t_n) = \Pi_j + \sum_{i=j}^{n} \sum_{k=1}^{m} \Delta\Pi_{ki} , \tag{126}$$

where $\Pi(t_n)$ is the sample of the random parameter at the time stage t_n, the $\Delta\Pi_{ki}$ is the component of variation in random parameter. The sampling is performed in such a way, that uncertainty decreases with every performed step. This is illustrated in Figure 5-11. Looking from the first stage the planner faces the uncertainty in a range $c_3 - c_4$ by the end of the planning period. However, after the sample at the second stage was taken (point b), the possible range of variation of the random parameter is reduced to $c_1 - c_2$.

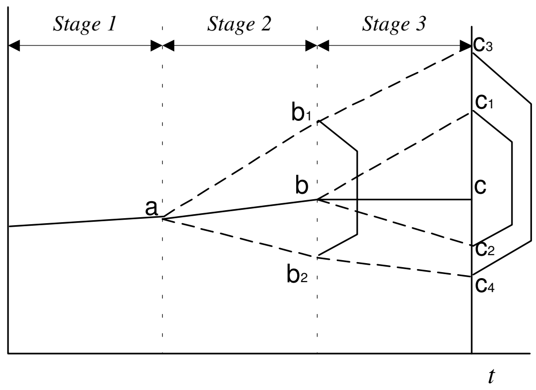


Figure 5-11 Illustration of future sampling scheme

A similar model is valid for the discrete component of the parameter variation. In this case the sampling is performed across several given future outcomes according to the probabilities associated with every future.

5.8 Conclusions

- The planner's intention is to find the robust and flexible solution to reduce the negative influence of the uncertain and random factors in dynamic network planing tasks.

- Regret is a widely used measure of risk. However, it may lead to risky decisions depending on the subjective factors.
- It is recommended to use the Expected Maximum Value as a measure of risk. This criterion concurs with the Bayesian formulation of the problem and promotes the most risk-averse choice.
- In order to achieve a flexible strategy it is rational to divide the planning period into two smaller periods: the decision-making period and the estimation period. The planner is interested in finding reinforcement actions to realize during the decision-making period. However, the consequences of these actions must be estimated for one or several futures during the whole planning period.
- The suggested general future model called fuzzy-probabilistic tree of futures integrates all classes of uncertain parameters – probabilistic, fuzzy and truly uncertain.

6 Suggested Applications of Genetic Algorithms and the Corresponding Software

The chapter starts with brief description of the previous application of GA to power systems. Three original applications of GA to the network planning are suggested. The first algorithm searches simultaneously for the whole set of Pareto optimal solutions. The hybrid GA/DP approach benefits from the global optimization properties of GA and local search by DP resulting in original algorithm with improved convergence properties. Finally, the Stochastic GA able to cope with a noisy objective functions is described. The chapter also contains the recommended modus operandi for the network planning tasks.

6.1 Previous Applications of GA to Power System Planning

Although evolutionary algorithms are relatively new techniques, their application to different aspects of power system planning has been performed by a number of researchers. First application of GA in power systems area can be found in [93], where the authors suggest solving distribution systems reconfiguration problem by means of GA.

Since then, GA has been applied to almost all kinds of planning problems in power systems area. GA are especially attractive for application to combinatorial optimization problems such as:

- distribution systems planning and operation [82],[89],[113],[142];
- generation and transmission network planning [101];
- reactive power planning [14],[68];
- generator scheduling and unit commitment [140] ;
- economic dispatch [22];
- load forecasting [139].

The problem of loss minimum re-configuration of distribution networks was presented in [93]. Aiming at finding the location for the tie switches so that the total power losses are minimized, the problem is formulated as a complex mixed integer programming problem, which is difficult to solve by traditional mathematical programming approaches. In [93] a simple GA is applied to the problem.

A least cost generation expansion problem is to find a set of optimal decision vectors over a planning horizon that minimizes investment costs, expected fuel, and O&M costs subject to a number of operational and reliability constraints. In [101] an improved GA is applied to solve this highly constrained nonlinear dynamic optimization problem. An improved GA incorporates a stochastic crossover technique and an artificial initial population scheme in order to provide a faster search mechanism and the robustness of performance.

In the optimal capacitor placement problem [14], the location, types, sizes and control schemes for the capacitors were studied. The distribution system considered consists of n_c possible locations for capacitors and n_t different loading conditions. The objective function, consisting of two terms: the cost of capacitor placement and the total cost of energy losses is discontinuous and non-differentiable. Therefore, traditional non-linear optimization approaches cannot be used.

In the course of pursuing the lowest possible cost over a time period of a day or a week, the unit commitment problem schedules the generating units to satisfy the power demand and the diverse constraints of the system and the individual units. The problem often comprises thousands integer as well as continuous decision variables and a wide spectrum of equality and inequality constraints. Therefore, mathematically it is also defined as a non-linear, large-scale, mixed-integer combinatorial optimization problem. In [140] a parallel GA approach is suggested for solving thermal unit commitment problem.

The objective of economic dispatch is to minimize the total generation cost of a power system over some appropriate period taking into account the transmission losses and satisfying various constraints. A genetic approach for solving the economic dispatch problem in large-scale systems is illustrated in [22].

A short-term load forecasting algorithm based on application of EP is presented in [139]. Typically, the surface of forecasting error function possesses multiple local minimum points. Therefore, the solution of the traditional gradient search methods may stop at the local optimum, which lead to inadequacy in the optimization model.

Summarizing the applications mentioned above, the following two factors can be mentioned as a motivation of the great interest to Evolutionary Computation:

- ✓ favorable features of evolutionary algorithms: they work when any other method would fail;
- ✓ complexity of the problem: it is difficult to find the traditional method for problem decision unless considerable approximations are done.

Probably the most explicit application of GA to multistage distribution network planning was presented in [86] and further developed in [84],[85] and [112].

The model presented in [86] is aimed at solving the problems of the optimal sizing, timing and location of distribution substations and feeders. The objective is to provide expansion plans that minimize the total cost including both new facility installation costs and network operation costs, as well as achieving an acceptable level of reliability. Furthermore, fuzzy models represent uncertainties in future events.

The model has the following characteristics: fuzzy variables and constraints, both continuous and integer variables, non-linear expressions, multiple criteria, large dimension, dynamic properties. As it is stated in [86] traditional optimization methods fail in dealing with the model of such a degree of complexity.

In [86] the variables represent the usage of a particular investment in a particular stage. Variables can directly be coded into binary strings (1 if the specific element is used at some stage, 0 otherwise). However, this approach results in a huge number of unfeasible solutions, since there are requirements on nodes connections and only radial configurations. A new coding technique was devised. The technique is based on the connections of a node to the feeders. A load node must be connected to at least one feeder. If there are e.g. four possible ways to feed the node, two bits are required to represent all possible connections.

There is a number of other applications of GA to distribution planning. For example, in [113] the problem of optimal design of large distribution system is formulated as a mixed-integer optimization problem. The authors present application of properly configured GA for solution of this problem. Application of GA addressing the stochastic problem of planning of large distribution networks under generation uncertainty is suggested in [20].

6.2 Multi-Criteria Optimization

6.2.1 Optimization and Decision-Making

Previously, the problem of distribution network reinforcement planning was formulated as a multi-criteria problem. It means that the planning problem must reasonably combine the concept of decision-making with the concept of optimization in order to find the best trade-off between the criteria. The subjective factors such as the expertise of the planner, the policy of the utility, the previous planning experience, available resources along with the others play an important role.

Most of the real world optimization problems are multi-criteria meaning that several objectives have to be optimized simultaneously. In many cases it is not possible to find the solution, which is optimal with respect to all objectives, since the objectives are often competing and there are many trade-off solutions. In this case the solution is represented by a set of solutions called Pareto optimal set, set of non-dominated solutions or non-inferior set. It is the task of the human decision-maker to consider all the factors and to make the final decision about the best solution.

6.2.2 Method of the "Displaced Ideal"

In order to describe the relative preference of each candidate strategy the method of "Displaced Ideal" [144] is accepted. This method is based on the concept of the ideal point and makes an appraisal of the alternatives overall performance by comparing them to extreme (best and/or worst) reference solutions. These solutions are often defined with respect to the set of the examined alternatives. Let $a = (a_1, a_2, ..., a_i)$ denote the ideal point, $b = (b_1, b_2, ..., b_i)$ the anti-ideal point, and $c_j = (c_{1j}, c_{2j}, ..., c_{ij})$ the vector of criteria values associated the alternative j. The ideal point is a non-feasible solution presenting in every criterion the best among the archived performances. In a similar way the anti-ideal point reflects the specific decision frame-work and may change with the addition or extraction of an alternative to the initial set of actions. The first step of the analytical procedure is to calculate for every criterion i a "degree of closeness" - the dimension-free relative distance from the ideal point defined by:

$$d_{ij} = \frac{c_{ij} - a_i}{b_i - a_i}. \tag{127}$$

The value d_{ij} maps the performances of the examined alternatives to the interval $[0,1]$, where 0 is assigned to the alternatives with the best performance and 1 to the alternatives with the worst performance under the i^{th} criterion.

The obtained partial ratings d_{ij} are then aggregated to provide a new composite function. This function represents the degree of closeness of the j^{th} alternative to the ideal point and can be described by the following relation:

$$D = \left(\sum_i w_i^p d_{ij}^p \right)^{1/p}, \tag{128}$$

where w_i are the relative-weight measures, which normally are set equal for all the criteria and p refers to the Lp-metrics and $p \geq 1$. For $p = 1$ deviations from the ideal values of the examined criteria are simply summed up, while for higher values of p a large deviation from the ideal value in a single criterion becomes increasingly important and the corresponding alternative is placed in the worse position in the overall ranking. In present work the Euclidean norm was used, i.e. $p = 2$.

This aggregation method was chosen since it has the following features:

- The method uses a transparent and easily comprehensible analytical model;
- The results are given as numerical values assigned to each alternative and reflecting their overall score;
- The method reflects the common human attitude in the comparative assessment of different alternatives by using points of reference for constructing relative preference orderings among them.

The task in this formulation would be a perfect candidate for application of the GA/DP described in section 6.2.4. However, the outcome from the optimization applying relative measure strictly depends on definition of the ideal and anti-ideal points. Therefore, prior to optimization it is essential to define both points for each time stage. For this purpose the overview of the possible consequences from each decision, i.e. each strategy and sub-strategy would be a great aid for the planner. This kind of information may be obtained applying Multi-Criteria GA presented in section 6.2.3.

6.2.3 Multi-Criteria Genetic Algorithm

The usual approach to deal with multiple criteria is to weight their composition into one objective function. The set of non-dominated (or Pareto optimal) solutions is obtained by multiple task decision changing the combinations of weights. These approaches have two limitations: increase in computational expenses and theoretically there is a need to consider an infinite number of weight combinations.

Due to their parallel search properties, the evolutionary algorithms are especially suitable for multi-criteria optimization. Their application for simultaneous search for the set of non-dominated solutions has a 15 years history starting from the Vector Evaluated Genetic Algorithm (VEGA) [121]. In VEGA subpopulations of each generation are formed from the existing population by using proportional selection according to each of the objectives. Therefore the offspring created by parents from different subpopulations is expected to perform well in both objectives.

The idea for the alternative kind of multi-criteria genetic algorithm using a ranking procedure was contributed in [48]. According to this algorithm all individuals are assigned a rank related to the number of individuals dominating them. The fitness is assigned based on the rank of the individual. The similar ranking method is used in [46], where Multi-Objective Genetic Algorithm (MOGA) is described. A nished Pareto GA for multi-criteria optimization was first presented in [57]. A comprehensive overview of evolutionary approaches to multi-objective optimization can be found in [31]

In this work the method, which allows allocating simultaneously the whole Pareto optimal set is suggested to apply for distribution planning. The appropriately modified Genetic Algorithm is used as an optimization method.

The effective approach to search for Pareto optimal set similar to the one described in [46] may be derived from the conventional GA modified in order to search for the set of non-dominated solutions instead of single optimum. Modification implies addition of the supplementary operator at the evaluation stage. The operator ranks the population according to the criteria values and assigns the fitness corresponding to the rank. Accordingly, non-dominated individuals are assigned the lowest fitness value and highly dominated individuals - the highest. Thus, the set of Pareto optimal solutions may be found after one run of GA.

The algorithm implemented based on the GAlib package involves the following steps:

<u>*Step 1*</u>. Initialization of the initial population. Pre-set number of individuals is randomly generated.

<u>*Step 2*</u>. Evaluation. Evaluation procedure is population based and involves several subroutines:

- Estimation of the criteria for each individual;
- Ranking individuals by interpolating from the best (rank 1) to the worst (rank $n \le N$) (see Figure 6-1)
- Assignment of the fitness value corresponding to the rank to each individual.

<u>*Step 3.*</u> Creation of a new offspring from the modified population obtained at the previous step using basic GA operators - reproduction, crossover and mutation.

<u>*Step 4.*</u> Repetition of the step 2 and step 3 until the pre-set stopping criterion is satisfied

However, for larger tasks it would be incorrect to claim that the Pareto optimal set found by this method contains the complete set of non-dominated solutions. Furthermore, since the optimization is proceeding simultaneously in many directions corresponding to different sets of weights, each solution hardly represents the global optimum in its own direction – it is rather somewhere near the optimal solution.

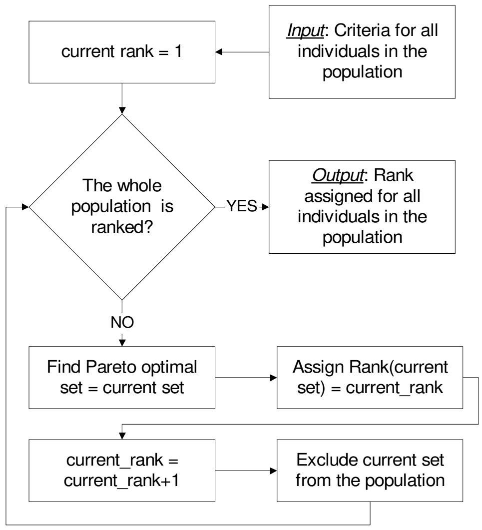
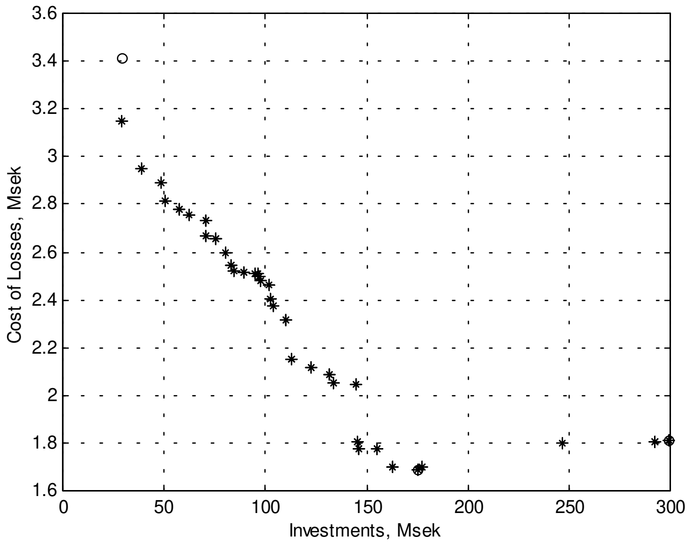


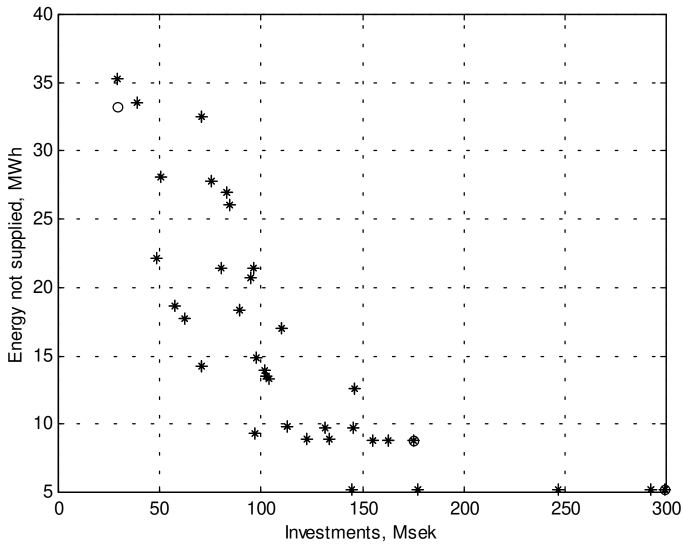
Figure 6-1 Population ranking algorithm

The characteristic performance of the algorithm can be illustrated by the real network application example. Consider the network planning task, which will be described in Chapter 7 in details (Large city network). The task involves 27 variables. Pareto optimal set obtained as a result of multi-criteria GA optimization is depicted in Figure 6-2 by asterisks. The efficiency of the algorithm can be judged comparing the Pareto set presented in this Figure with the Pareto set which would be obtained from the weighted composition of criteria in one objective function changing the combinations of weights. The results for three marginal combinations of weights (each time only one criterion was minimized) are depicted in Figure 6-2 by zeros. The results verify the initial statement that multi-criteria GA results in near optimal solution for the Pareto optimal set.

Clearly, the parameters of the GA, such as population size and number of generations have a great effect on the performance of the algorithm. In order to make the comparison somewhat conclusive, in this example the parameters of both GA optimizing the single objective and the multi-criteria GA were chosen exactly the same, 100 individuals in the population and 200 generations (which takes about 1.5 min CPU time). The multi-criteria GA proves to come up with at least reasonable results for Pareto optimal set even in this case. Consequently, the better results can be expected increasing the number of generations and the size of population.



(a)



(b)

Figure 6-2 Pareto optimal set obtained by multi-criteria GA
(a) Cost of losses and (b) Energy Not Supplied versus Investments

The outcome from multi-criteria GA gives the planner the perfect indication of the possible planning directions at each time stage during the planning period. Applying this information together with the planner's expertise and other related data, the ideal and anti-ideal points for more precise optimization may be easily obtained (section 6.2.1).

Finally, for the faster convergence when minimizing the distance to the ideal point the part of the final population from the multi-criteria GA may be used as an initial population in GA/DP (section 6.2.3).

6.2.4 Principal Component Analysis

If the number of attributes in multi-criteria optimization is more than three, it may be difficult to interpret the obtained results. Principal component analysis may serve as one of the possible solutions to these problems.

Principal Component Analysis (PCA) reduces the dimensionality of the problem by forming a new set of variables, which are a linear combination of the original measured variables and which explain the maximal amount of variability of the data. PCA seeks for a few linear combinations, which can be used to summarize the data with a minimal loss of information.

Let $X = x_1, x_2, ..., x_m$ be an m-dimensional data set describing the information. The first principal component is that linear combination of the columns of X, i.e. the variables, which describe the greatest amount of variability in X, $t_1 = X \cdot p_1$ subject to $|p_1| = 1$. In the m-dimensional space, p_1 defines the direction of greatest variability and t_1 represents the projection of each object onto p_1. The second principal component is the linear combination defined by $t_2 = E_1 \cdot p_2$ which has the next greatest variance subject to the condition that it is orthogonal to the first principal component t_1, therefore

$$E_1 = \left(X - t_1 \cdot p_1^T\right). \tag{129}$$

This procedure is essentially repeated until m principal components are calculated. In effect, PCA decomposes the observation vector X as:

$$X = TP^T = \sum_{i=1}^{m} t_i p_i^T , \tag{130}$$

where p_i is an eigenvector of the covariance matrix of X. P is defined as the principal component loading matrix and T is defined to be the matrix of principal component scores. The loading provide the information as to which variables contribute the most to individual principal components.

One of the features of PCA is that the less important components often describe the noise in the data. If the process variables are co-linear, k principal components will explain the variability in the data. Consequently it is desirable to exclude these components:

$$X = TP^T + E = \sum_{i=1}^{k} t_i p_i^T + E . \quad (131)$$

In practice two (or three) principal components provide an adequate description of the data.

To illustrate the performance of the approach, consider again the elementary example from Table 5-1. Four plans are compared according to four criteria – two attributes and two Expected Maximum Values as the measures of risk. The results of the PCA are presented in Figure 6-3.

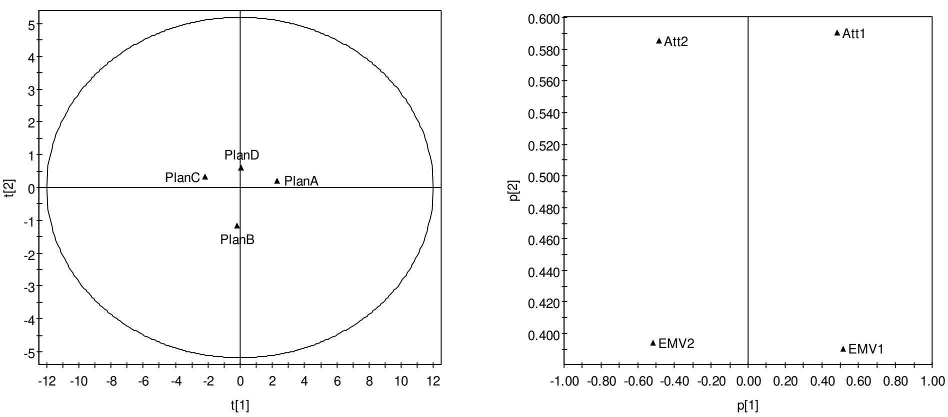


Figure 6-3 Scores plot (left) and loadings (right) for the example from Table 5-1

The loading plot shows which variables describe the similarity and dissimilarity between the plans. However, it also contains the information about the directions of improvement of each particular criterion. Since the task is to minimize all four objectives, the direction to the south (down) on the scores plot corresponds to minimization of all the objectives. Similarly, minimization of the Attribute 1 and the corresponding EMV1 result in down-left direction, but minimization of the Attribute 2 and the corresponding EMV2 result in down-right direction. Thus, Plan A has minimal Attribute 1 and EMV1, Plan C has minimal Attribute 2 and EMV2, while Plan B is the best trade-off between all four objectives. The results are fully consistent with those obtained previously and summarized in Table 5-1 and depicted in Figure 5-2.

6.3 Hybrid Genetic Algorithm/Dynamic Programming (GA/DP) Approach

6.3.1 The GA/DP Approach

GA is the efficient global optimization technique, however slower convergence near the optimum can be mentioned among its major drawbacks. A number of methods have been suggested in the literature resulting in hybrid approaches combining the global search applying GA with local search by some other method. A hybrid approach can be applied whenever it explores better trade-off between computational cost and global optimality of the solution found. For instance in [141] it is suggested to apply the simplex method as a local search engine, a number of researchers apply simulated annealing in combination with GA (see for example [35]), numerous heuristic engines and gradient search methods have also been applied.

In this section it is suggested to combine GA and DP into one algorithm in order to utilize their advantages. The idea behind the suggested approach is to incorporate the DP function performing local search at each generation during the course of genetic algorithm. Simultaneously, the whole (or some parts of) population should be accordingly modified in order to carry the correct genetic information to the next generations. The only comparable application of DP in hybrid with GA was found in [102], where tunnel-based DP is applied to search for the local optima, however unlike the method described below the local solutions were not given back to the GA, but stored as the useful information.

The suggested procedure can be described as follows:

<u>*Step 1*</u>. Initialization of the initial population. Pre-set number of individuals is randomly generated according to the encoding pattern described in section 6.5.3.

<u>*Step 2*</u>. Evaluation. Evaluation procedure is population based and involves several subroutines:

- Estimation of the state quality criterion $g(t, e(t))$ for each individual at each time stage. For the network planning tasks the criterion can be calculated aggregating the planning attributes according to the equation (128).
- Application of the forward search by DP, calculation of the objective function $F(x, \mathrm{P}(T, x))$ for each strategy found by DP, maintaining the paths.
- Application of the DP backtracking procedure, simultaneously adjusting the timing components of the individuals corresponding to the information retrieved from the path $\mathrm{P}(T, x)$.

- Assignment of the fitness value $F(x, \mathrm{P}(T, x))$ to each individual.

<u>Step 3.</u> Creation of a new offspring from the modified population obtained at the previous step using basic GA operators - reproduction, crossover and mutation.

<u>Step 4.</u> Repetition of the step 2 and step 3 until the pre-set stopping criterion is satisfied.

Evidently, the GA/DP differs from the pure GA only by the evaluation step. However, this difference is substantial, since the natural genetic optimization process is externally influenced. As a result, two methods in one algorithm compliment each other: the GA explores the global search space, while DP explores local subspaces found by GA; the GA performs search for the best state of the variables, while DP searches locally for the best realization times.

To illustrate the essence of the suggested method, consider some function of four binary state variables x. The function contains a dynamic parameter, which changes from one time stage to another.

Assume the GA has randomly generated the initial population (Table 6-1).

Table 6-1 DP operations on the initial population

Individual	*t=1*		*t=2*		*t=3*	
	g(1,x)	*f(1,x)*	*g(2,x)*	*f(2,x)*	*g(3,x)*	*f(3,x)*
Individual 1						
0 0 1 0 *** * 1 ***	**0 0 1 0**		**0 0 1 0**		**0 0 1 0**	
	6.6	6.6	10.49	17.09	15.49	32.58
Individual 2						
1 0 0 1 **2 * * 1**	**0 0 0 1**		**1 0 0 1**		**1 0 0 1**	
	6.54	6.54	9.47	16.01	13.22	29.23
Individual 3						
0 1 1 0 *** 3 2 ***	**0 0 0 0**		**0 0 1 0**		**0 1 1 0**	
	9	9	10.49	17.09	13.1	30.19
Individual 4						
1 1 0 0 **1 2 * ***	**1 0 0 0**		**1 1 0 0**		**1 1 0 0**	
	7	7	9.74	16.74	13.93	30.67
Individual 5						
1 1 0 1 **1 3 * 3**	**1 0 0 0**		**1 0 0 0**		**1 1 0 1**	
	7	7	11.67	18.67	12.3	28.31

Generated individuals are shown in the first column and for illustrative purposes extended for each particular time stage. If the variable is zero the timing information is not used – these time variables are shown by asterisks. For each state and stage the quality criterion is calculated. At that point the DP forward routine starts to search for the best feasible transitions among the available states. In the example presented in Table 6-1 only two transitions were both feasible and improving the strategy generated by the GA (between Individuals 1 and 3 transition to the second stage, and between Individuals 2 and 5 to the third stage). According to the transitions, the objective function is obtained by summation of the state quality components at each stage. Note that for the example in Table 6-1, the dynamic procedure allows us to minimize the function locally - Individual 5 corresponds to the local optimum and would not be found at this stage of optimization without application of the additional search.

The next step of the algorithm is backtracking. In our example only two individuals 3 and 5 should be modified. Modification is very simple, since the path contains transitions only between two states in each case. The modified individuals are shown in Table 6-2.

Table 6-2 Result of Performed Modification

Before	After
Individual 3	
0 1 1 0 *** 3 2 ***	**0 1 1 0** *** 3 1 ***
Individual 5	
1 1 0 1 **1 3 * 3**	**1 1 0 1** **2 3 * 1**

The suggested algorithm has the following features:

- Minor increase in computational expenses – no additional calculations of the criteria are required;
- Minor interference in the genetic optimization process – the structure of the state variables remains unchanged, only their timing is tuned.

Furthermore, as it will be illustrated in the next section the algorithm clearly explores better trade-off between computational expenses and global optimality of the solution.

6.3.2 Illustration of the Performance

6.3.2.1 Test Function

It is convenient to test the performance of the algorithms on a test function, which simulates the behavior of the network planning objective function.

Consider the state quality criterion calculated as:

$$g(t)=\sum_{i=1}^{N} a_i \cdot x_i + \frac{K(t)^2}{\sum_{i=1}^{N} b_i + 1}, \tag{132}$$

with

$$b_i = (a_0 + \alpha \cdot a_i) x_i \ . \tag{133}$$

The binary variable x_i represents the reinforcement actions in the planning task, a_i is the corresponding investment, while b_i stands for the positive effect from this particular investment; $K(t)$ corresponds to the system load during the time stage t, while a_0 and α are the function parameters. Thus, the first term in function (132) corresponds to the investment cost, while the second to the losses in the hypothetical network.

Consider a function of 40 variables that has to be optimized for 8 time stages. In Figure 6-4 the convergence characteristics of two random runs of both GA/DP (filled markers) and pure GA (empty markers) are compared. Both algorithms were run with the same parameters, namely population of 40 individuals and 200 generations.

From the figure we can see that in case of GA/DP the performance is enhanced considerably.

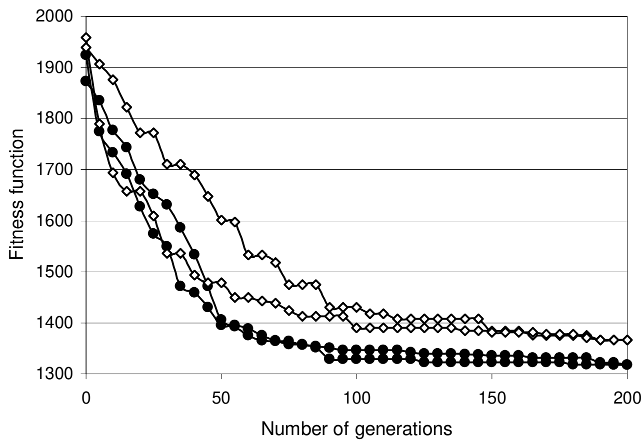


Figure 6-4 Convergence characteristics of GA/DP (filled markers) in comparison with conventional GA (empty markers) – function application example

The important issue, which must be discussed when assessing the performance of the algorithm, is computational time. With given parameters the average

CPU time for GA/DP algorithm was 20 seconds, while for conventional GA only 0.5 second. Then again, in this case it is particularly easy to calculate the function value itself, which takes the most computational time and efforts in real-life cases. This allows us to conclude that for the given parameters only about 20 seconds of the CPU time is consumed by DP operations.

The comparison was made on a PC with Pentium III 450 MHz processor.

6.3.2.2 Network Optimization Example

In Figure 6-5 the secondary distribution network in urban area of a small town is presented. The network consists of 57 lines and 55 nodes. Consumption is mainly residential with the annual load growth of about 1%. Presently the network is supplied by one substation. Calculations indicate that even in the present state voltage drops in the network exceed permissible limits (voltage drop threshold is 5%). Therefore a reinforcement strategy is required.

A number of actions for network reinforcement (totally 40 possible actions) is suggested to be analyzed:

- Introduction of a new substation TP Rigas;
- Construction of new lines (doted lines in Figure 6-5);
- Replacement of conductors in several lines, mainly along the main feeder.

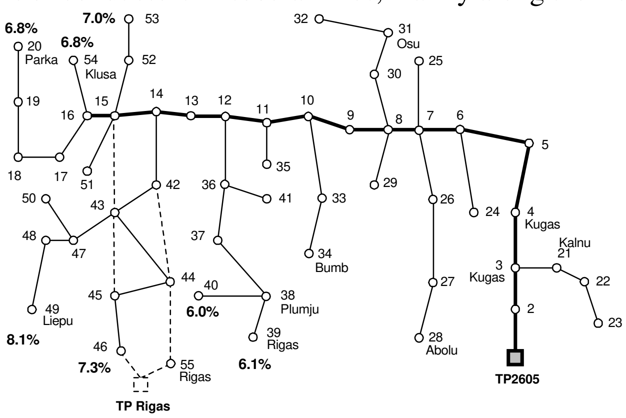


Figure 6-5 Secondary distribution network in Jelgava

The planning period is 24 years forward divided into 8 time stages. The interest rate is 10%.

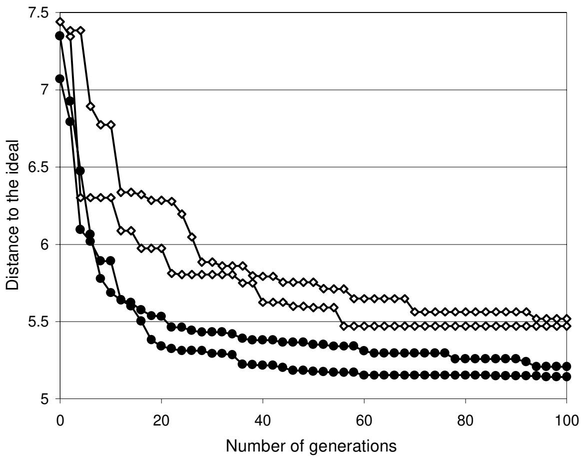


Figure 6-6 Convergence characteristics of GA/DP(filled markers) in comparison with conventional GA (empty markers) – network application example

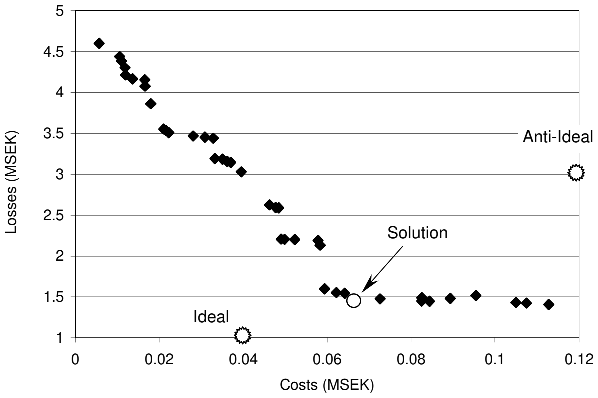
At the initial optimization phase the set of Pareto optimal solutions was obtained applying multi-criteria GA. The final set over the whole planning period is presented in Figure 6-7. Information about each development strategy is stored, thus the similar plots representing the development of each non-dominated strategy from the initial to the final time stages can be easily obtained.

This information allows the planner to identify the ideal and anti-ideal points for each time stage. In this example, the upper left corner of the final set from the Figure 6-7 corresponds to the strategies without the new substation. At the later stages these solutions become unfeasible due to violation of the voltage drop constraint. The lower part contains the strategies where the substation is introduced from the first time stage. Finally, the middle part of the plots contains the set of strategies, which imply introduction of the substation at the intermediate planning stages.

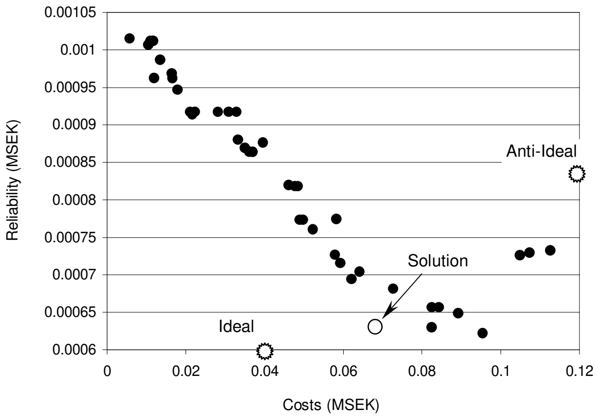
After careful consideration of possible planning directions the planer is able to identify the ideal and anti-ideal points for the final set of strategies (Figure 6-7), as well as the corresponding points for each time stage.

In order to find the best compromise strategy within the selected range, the GA/DP procedure is applied. In Figure 6-6 the convergence characteristics for two random runs of GA/DP (filled markers) are compared with for two random runs by conventional GA (empty markers). The best strategy found by GA/DP is depicted in Figure 6-7.

Suggested combination of GA with DP indicates very good convergence properties and reduces the drawback of GA – slow convergence near the optimum.



(a)



(b)

Figure 6-7 Final set of Pareto optimal solutions a) Cost of Losses and (b) Energy not Supplied versus Investments criterion

6.4 Genetic Algorithm in Noisy Environment

6.4.1 Optimization in the Presence of Noise

Any optimization task involves estimation of the quality of solution. Usually it is assumed that the quality can be determined by calculating the deterministic objective function. However, for many real-world applications this assumption is not adequate enough. Practical problems often involve noisy evaluations. The real system to be optimized might be too complex to fit into a well-defined deterministic mathematical model. In this case a stochastic model can replace the deterministic model.

The method of Monte-Carlo is the most powerful (and the most general) method of problem solution in stochastic formulation. However, it has a major drawback – the precision of the method to the highest degree depends on the number of trials during the simulation. Therefore for the realistic precision the function will become noisy – even identical decisions will not lead to the identical results. Therefore, the chosen optimization method must be robust enough to cope with noisy objective function. Sensitivity of the optimization methods (for instance gradient methods) to the noise was the main limitation on utilization of Monte-Carlo simulations in optimization.

There is a great deal of research devoted to optimization of noisy functions by means of evolutionary and algorithms based on other heuristics. An extensive background of this research is given in [99].

In [116] the effects of additive Gaussian noise on both local search and genetic search were investigated on a suite of several test functions. The authors concluded that adding noise could have a soft annealing effect in some cases but might also have the negative effect of adding false optima to the search space. They found that genetic search had very stable performance with and without noise on their test functions and that re-sampling of noisy data points mainly improved the performance of local search methods.

In [99] the authors compare two population-based optimization techniques, a GA and a ES with two point-based heuristic methods. The obvious conclusion was that increasing the amount of noise generally deteriorates heuristic performance while increasing the sample size per individual solution generally improves performance by reducing the amount of uncertainty in the evaluation. It was revealed that the point-based methods have convergence difficulties even with moderate levels of noise, although good results were achieved in deterministic case. And the most important conclusion is that the population-based GA and ES show a remarkable robustness at all noise levels.

It can be pointed out that population based optimization is by design less dependent on the quality of individual solutions. It moves from one set of solutions to the next and is, consequently, not so much affected when a mediocre solution receives a particularly good evaluation through stochastic influence.

6.4.2 Nonlinear Function of Losses

Analysis shows that mathematical expectation of power losses $\overline{P}_{Los}(x)$, and similarly of the criteria which are functions of losses, differs from the value of losses found at the average or mean value of random parameters $P_{Los}(\overline{x})$. Furthermore, it disagrees also with the value of losses calculated as the average value of losses at marginal values of random parameters

$\frac{P_{Los}(x_{\max}) + P_{Los}(x_{\min})}{2}$. Moreover, the correlation between these values can be described by the following relationship:

$$P_{Los}(\overline{x}) \le \overline{P}_{Los}(x) < \frac{P_{Los}(x_{\max}) + P_{Los}(x_{\min})}{2}. \tag{134}$$

The higher is the level of uncertainty, the more distinct is the shift described by (134). If the information is considered to be deterministic, all three values coincide. The relationship (134) is the result of nonlinear dependence of loss function from random parameters, in particular load values. It is well-known that for the most tasks, the quadratic relation between the functions of losses and the corresponding loads is valid.

As an illustration of this statement the histograms of losses for two different levels of load uncertainty are presented in Figure 6-8. The left chart represents distribution of losses found at the actual short-term uncertainty level of loads assuming their normal distribution. The right chart depicts similar distribution, but the uncertainty level was increased four times. The losses on both graphs can be compared with the value for the deterministic case $Ploss$ =3.14 MSEK. It can be observed from the charts, that the expected value of losses is larger than the value calculated for the deterministic case. Furthermore, logically the higher level of uncertainty leads to the larger dispersion, but remarkably it also leads to the considerable shift of the expected value.

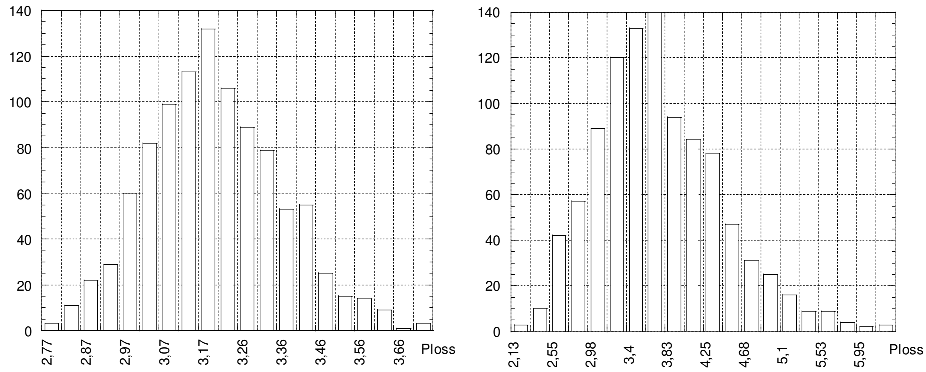


Figure 6-8 Histograms of losses for different levels of load uncertainty

6.4.3 Stochastic GA

From the previous subsections two main conclusions can be drawn:

- ✓ population based optimization algorithms, including GA, perform efficiently in the presence of noise;

- ✓ stochastic loads shift the corresponding mean value of losses to the right in comparison with equivalent deterministic computations.

The second conclusion immediately lead to the speculation that in some cases deterministic load modeling may lead to the false solutions. And on the contrary, stochastic modeling of loads may improve the quality of the results.

Therefore, the algorithm, which would consider the statistical variation of parameters, would enhance the performance of the network planning software. The algorithm suggested here is based on the basic GA, where the usual evaluation block is replaced by the procedure of random parameters simulation following by evaluation of the fitness function. This procedure is repeated several times corresponding to the chosen number of trials. It should be noted that for better performance the Variance Reduction Technique known as common random numbers [77] is used, or in other words the same disturbance is applied to the whole population. The suggested Stochastic GA can be described by the following algorithm:

Step 1. Initialization of the initial population. Pre-set number of individuals is randomly generated according.

Step 2. Evaluation. Evaluation procedure is population based and involves several subroutines:

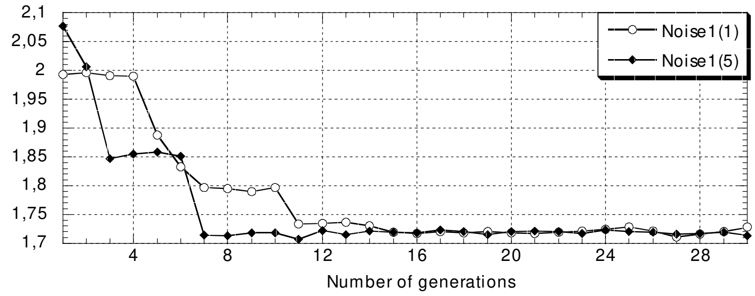
- Simulation of random parameters and exposure of stochastic parameters to perturbation;
- Estimation of the criteria for each individual;
- Assignment of the corresponding fitness value to each individual.

Step 3. Creation of a new offspring from the modified population obtained at the previous step using basic GA operators - reproduction, crossover and mutation.

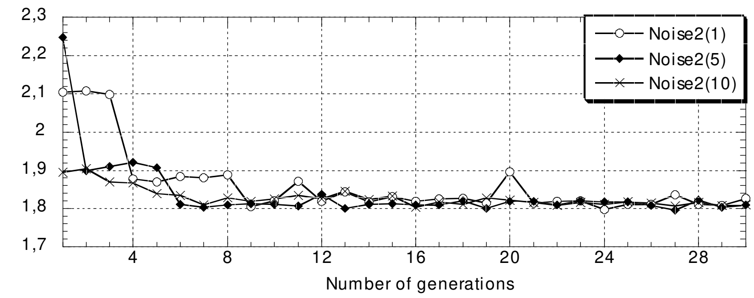
Step 4. Repetition of the step 2 and step 3 until the pre-set stopping criterion is satisfied.

Generally, the larger is the number of trials the better performance from the GA can be expected, due to reduced level of noise. However, evaluation of each population involves rather heavy computations and, as a result, increased number of trials leads to longer computational times.

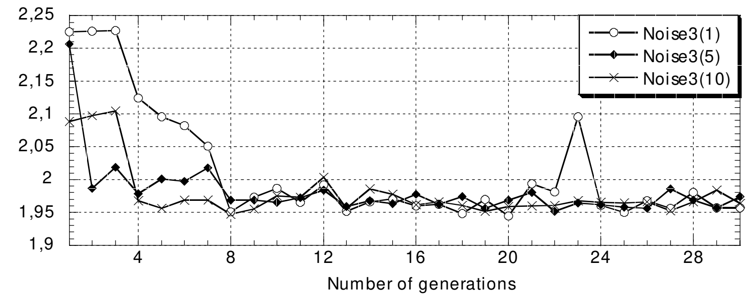
It can be shown, however, that for the moderate levels of noise the GA demonstrates robust performance even with a minimal number of trials. In [99] similar conclusions have been illustrated on four test functions, here the GA is tested on real network optimization task described in section 7.1. To simplify the understanding and interpretation of the results the single criterion, namely the cost of power losses is minimized.



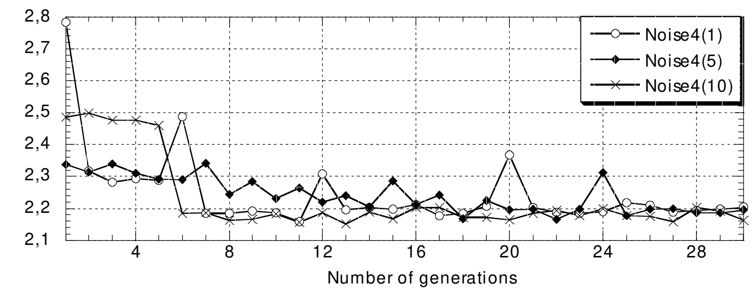
(a)



(b)



(c)



(d)

Figure 6-9 Convergence of Stochastic GA for different levels of noise

The level of noise in optimization task depends on stochastic parameters of random variables. These parameters can be estimated if the corresponding statistical data is available. The algorithm, which can be used for assessment of the stochastic load model from the measured data is described in section 4.5.3. The normal distribution parameters obtained for the example in section 4.5.3.6 are considered to represent the realistic and reasonable level of variation of load and denoted below as Noise 1 (the standard deviation is $0.055 \cdot P_{Load}$). Correspondingly, the notation Noise 2, Noise 3 and Noise 4 represent respectively 2,3 and 4 times higher levels of variation of load.

To test the convergence properties of Stochastic GA the task was optimized several times with different number of trials for each level of load variation. The results of this investigation are presented in Figure 6-9, where each chart represents one level of load variation and each curve corresponds to the given number of trials (number in parenthesis). Minor variations in fitness function can be explained by its stochastic nature. However, more considerable peaks (best seen on the Figure 6-9 (*c*) and (*d*) reveal deficient stability in performance of GA with the given number of trials. Accordingly, the stability improves with the number of trials increased. Thus, it can be observed that for the real level of load variation (Noise 1) the GA achieves robust convergence even with a single trial while calculating the stochastic fitness function (Figure 6-9 (*a*)). For load variation three times higher than the real five trials would be recommended (Figure 6-9 (*c*)), but for load variation four times higher - ten trials would give the adequate results (Figure 6-9 (*d*)). The actual level of noise should be estimated prior to the solution of the particular task, and the algorithm could be correspondingly adapted.

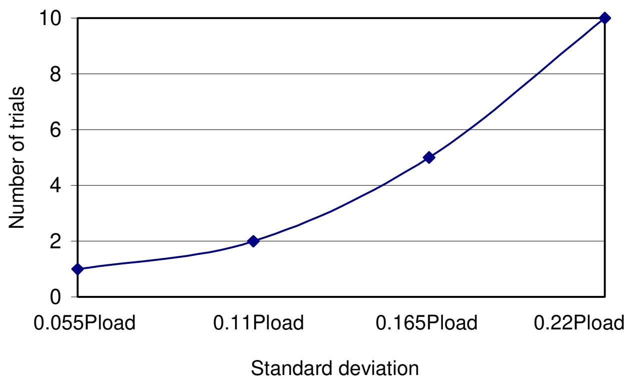


Figure 6-10 Suggested number of trials depending on the level of noise

Another important observation made while investigating the convergence characteristics of Stochastic GA is that even within the stable range of performance convergence properties of the algorithm can be improved increasing the number of trials. However, the higher is the number of trials the higher are computational efforts. Therefore, the optimal number of trials would be the compromise solution. An example of suggested number of trials for different levels of noise based both on stability of GA performance and optimal convergence rate is presented in Figure 6-10.

The last presented here observation from the Figure 6-9 is that the stochastic optimum is moving to the right increasing level of load variation. Furthermore, for deterministic case and the first two load variation levels there is one optimal solution, while for the next two sets of stochastic data the optimization algorithm finds a different solution. The mean values of power losses for different levels of load variation for both solutions (after 10000 trials) are summarized in Table 6-3.

Table 6-3 Mean of power losses for different levels of load variation (after 10000 evaluations)

	Solution 1	Solution 2
Deterministic	1.6896	1.7023
Noise 1	1.7198	1.7300
Noise 2	1.8090	1.8161
Noise 3	1.9583	1.9554
Noise 4	2.1726	2.1659

The example above illustrates the importance of considering short-term uncertainties in the planning model. Apparently, there is the possibility for the informational situation when stochastic solution would differ from the deterministic one. In these cases inadequacy of deterministic load representation is evident.

6.5 The Software

6.5.1 Overview of the GAlib Package

The software for this work used the GAlib genetic algorithm package, written by Matthew Wall at Massachusetts Institute of Technology [134]. The principal part of the library consists of two classes: a genome and a genetic algorithm. Each genome instance represents a single solution to the problem. The genetic algorithm object defines how the evolution should take place. The genetic algorithm uses an objective function to determine how “fit” each genome is for survival. It uses the genome operators and selection/replacement strategies to generate new individuals.

The *genetic algorithm object* determines which of the individuals that should survive, reproduce, and die. The library contains the following four flavors of genetic algorithms:

- ✓ 'Simple genetic algorithm' described in [48] uses non-overlapping populations and optional elitism;
- ✓ 'Steady-state genetic' algorithm uses overlapping populations with optional overlapping degree;
- ✓ 'Incremental genetic algorithm', where each generation consists of only one or two children;
- ✓ 'Deme genetic algorithm' evolves multiple populations in parallel using a steady-state algorithm.

The genetic algorithm object contains the statistics, replacement strategy and parameters for running the algorithm.

The following three things must be defined in order to solve the problem: representation, genetic operators and the objective function.

The data structure appropriate for the particular problem should be used. A *representation* that is minimal but completely expressive would be the best choice. According to the general guidelines a chosen representation:

- ✓ should be able to represent any solution to the problem, but if at all possible all the solutions represented must be feasible (otherwise the objective function must be designed to give partial credit to unfeasible solutions)
- ✓ should not contain information beyond that needed to represent a solution to the problem, otherwise the size of the search space may be unreasonably increased, thus, hindering the performance of the genetic algorithm.

The actual number of possible representations is endless. The GAlib package is flexible enough to allow for purely numerical representation, such as an array of real numbers or an order-based representation – either list or array. Finally, the package allows for representation of the solution explicitly as trees and for performing genetic operators on the trees directly. For the more complicated cases, where a various types of information should be represented in one structure, a corresponding structure must be created.

Each *genome* has three primary operators: initialization, mutation and crossover. GAlib pre-defines these operators for each type of genome, but the user may easily modify each of them.

The initialization operator determines how the genome is initialized. This operator does not actually create new genomes, rather it "stuffs" the genomes with the primordial genetic material from which all solutions will evolve.

The mutation operator defines the procedure for mutating each genome. Mutation means different things for different data types. For example, a typical mutator for a binary string genome flips the bits in the string with a given probability.

The crossover operator defines the procedure for generating a child from two parent genomes. Like the mutation operator, crossover is specific to the data types. Unlike mutation, however, crossover involves multiple genomes.

The *population* object is a container for genomes. Each population object has its own initializer (the default simply calls the initializer for each individual in the population) and evaluator (the default simply calls the evaluator for each individual in the population). It also keeps track of the best, average, deviation, and other statistics for the population.

As it was mentioned above the main advantage of genetic algorithms is that they do not require complicated differential equations or a smooth search space as for example the gradient methods. The genetic algorithm needs only a single measure of how good a single individual is compared to the other individuals. The *objective function* provides this measure.

It is important to note the distinction between fitness and objective scores. The objective score is the value returned by the objective function; it is the raw performance evaluation of a genome. The *fitness score*, on the other hand, is a possibly transformed rating used by the genetic algorithm to determine the fitness of individuals for mating. The fitness score is typically obtained by a linear scaling of the raw objective scores.

The individuals in the population can be evaluated either using an individual-based evaluation function, or alternatively a population-based evaluator. The last option was extensively used in the planning software.

6.5.2 Realization in C++

A typical optimization program has the following form:

```
float Objective(GAGenome&);
main(){
  GA2DBinaryStringGenome genome(width,height,Objective)        //create a genome
  GASimpleGA ga(genome);                                       // create the GA
  ga.evolve();                                                 // evolve the GA
  cout << ga.statistics() << endl;                             // print out the results
}
float Objective(GAGenome&) {
 // the objective function goes here
}
```

Setting various parameters can change the behavior of the genetic algorithm. Some of the more common ones are set as

```
ga.populationSize(popsize);
ga.nGenerations(ngen);
ga.pMutation(pmut);
ga.pCrossover(pcross);
GASigmaTruncationScaling sigmaTruncation;
ga.scaling(sigmaTruncation);
```

A typical (albeit simple) objective function looks like this (this one gives a higher score to a binary string genome that contains all 1s):

```
float
Objective(GAGenome & g)
{
  GA1DBinaryStringGenome & genome = (GA1DBinaryStringGenome &)g;
  float score=0.0;
  for(int i=0; i<genome.length(); i++)
    score += genome.gene(i);
  return score;
}
```

The objective function can be defined either as a static member of a derived class, or defined as a function and used with the existing GAlib genome classes. Originally, each objective function returns a single value that represents the objective score of the genome that was passed to the objective function.

6.5.3 The Objective Function

6.5.3.1 Optimization Criteria

The model applied for calculations is multi-criteria, i.e. the identified optimization objectives are treated separately. All the criteria are calculated for the planning period as a sum of the annual and discounted terms. For the details on the criteria calculation see Chapter 4. Three attributes are presently used in the model, they reflect the objectives to minimize power losses, investment costs and improve reliability through minimization of the ENS. The attributes are not calculated for the alternatives, which are technically unfeasible. Feasibility of the solution implies, that all the nodes must be connected and the voltage drops and power flows through the network elements must be within acceptable limits.

6.5.3.2 Static and Dynamic Representation of the Criteria

In terms of considered details in the planning process in time it is suggested to use two types of representations, which can be called static and dynamic.

Static optimization makes several assumptions about the planning process. First of all it is assumed that the state of the network does not change during the planning period. All the investments foreseen for the planning period are made during the first year. The values of the criteria are calculated only once and summed up for the planning period using the annuity principle. These simplifications may lead to simplified and not totally adequate results. However, as a preliminary analysis tool static optimization allows us to identify interesting network configurations and corresponding reinforcement actions prior to analyzing times of actions realization.

Dynamic optimization suggests not only the best actions to be realized, but also the optimal sequence of the events and the best timing for realization of each action. The values of the criteria are calculated for each time stage, therefore the network configuration may be modified at each subsequent time stage. Thus the most essential actions are promoted during the optimization and the actions, which are less critical, may be postponed. Thus, the solutions become more flexible in comparison with results of static analysis. Therefore, in the dynamic case, the investments can be distributed during the planning period and correspondingly the network configuration can differ from stage to stage.

6.5.4 Encoding of Variables

In application of GA it is important to find appropriate encoding of state variables into a binary string. Since in reinforcement planning tasks the number of the reinforcement actions is usually much less than the number of the existing network elements it is possible to apply direct binary codes to represent the variables. The percentage of unfeasible configurations is very low. The direct approach implies that the variable is 1 if the action is realized and 0 otherwise. For "green field" planning applications some alternative encoding must be used.

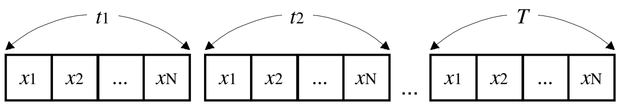


Figure 6-11 Direct dynamic actions encoding

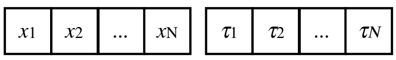


Figure 6-12 Alternative dynamic code

Straightforward extension of the binary direct codes to several time stages results in the string structure depicted in Figure 6-11. However, in this case the string may contain the unfeasible transitions, when individual either must be declared unfeasible or decoding depending on the previous history must be implemented.

Alternative encoding is shown in Figure 6-12. In the suggested string structure the information about the state of the variables appears only once and corresponds to the state of the actions realized by the end of the planning period. The rest of the string contains the information about action realization times $\tau_1, \tau_2, ..., \tau_N$. This information can either be integer or represented by binary code. In the present implementation with GAlib the binary codes are used. Each variable is represented by tree bits (or genes on evolutionary terminology). The alternative encoding overcomes the disadvantages of directly coded strings and simultaneously provides savings in the size of the codified string. To encode N variables during T stages into the binary string the direct straightforward approach results in $T \cdot N$ long string, but for the alternative dynamic code the string length is only $(1 + \log_2 T) \cdot N$.

6.5.5 Structure of the Program

6.5.5.1 Initializer

An initializer is invoked when the population is initialized. Two options are realized in the software, both of them are population-based. The first one invokes the initializer for each genome in the population, which in C++ looks as following:

```
void
PopInitializer (GAPopulation &p){
for (int i=0; i<p.size(); i++){
p.individual(i).initialize();
}return;}
```

The second type initializer allows to re-apply the population from the previous run of GA.

6.5.5.2 Evaluator

One of the vital parts of the program is the evaluator. The evaluator is population-based and it is used to set the score to each genome, rather than invoking an evaluator for each genome.

In Figure 6-13 the structure of the evaluator with multiple choices of the algorithms and criteria representations is depicted. The selections along the left column comprise the least details in computations and can be used in preliminary evaluation studies. The right-column model with either Multi-Criteria GA or Distance to the Ideal minimization routine can be used for the detailed studies.

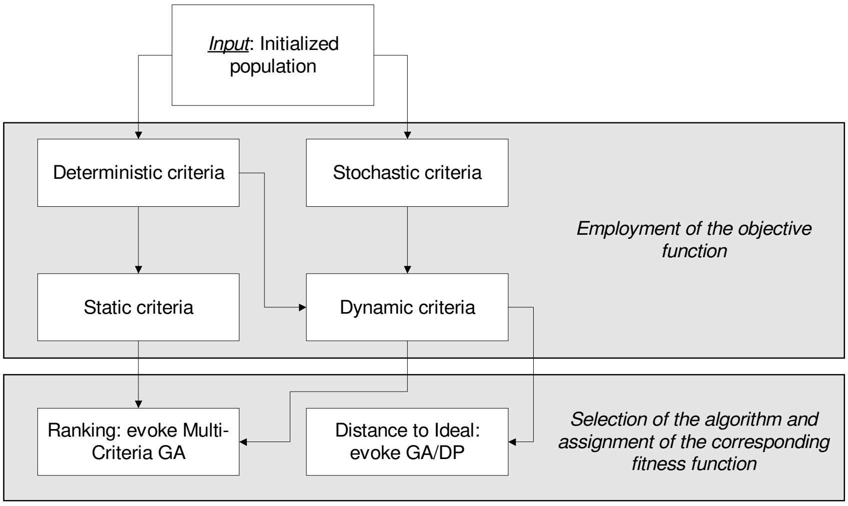


Figure 6-13 The structure of the evaluator with multiple choices of the algorithms and criteria calculation

6.6 The Recommended Modus Operandi

6.6.1 Modus Operandi: Single Future

The planning process consists of several large stages starting form the collection of the necessary data and identification of possible alternatives. For certain planning tasks it may take several years to come up with a final decision. For example, the following multi-stage procedure can be recommended for the projects concerning reinforcement planning in distribution networks:

Stage 1 Identification of possible reinforcement actions and network development strategies. Among the factors, identified as "drivers" of reinforcement process, can be mentioned operation cost, and environmental concerns.

Stage 2 Collection of the relevant network data. The data include information about cables, transformers, reliability indices, costs, etc. The exemplary structure of the required data can be found in Appendix A.

Stage 3 Preliminary static analysis and network optimization. The results include total values for the selected attributes (cost of losses, energy not supplied, and investments).

Stage 4 Presentation of the dynamic analysis and optimization. Dynamic analysis allows distributing the investments and the corresponding reinforcement actions during the planning period, which provides a more realistic and flexible model. "Broad search" searches simultaneously for the set of Pareto optimal solutions using the Multi-Criteria GA, while "deep search" by GA/DP minimizes distance to ideal and attempts to find the best trade-off among the candidate alternatives.

Stage 5 Decision-making after uncertainty analysis for the candidate alternatives obtained at the previous stages.

The software and the algorithms presented in this chapter may aid the decision-maker in the analysis during the intermediate stages of the planning process. The primary goal of these analyses is to identify a number of candidate solutions for further analysis at the decision-making stage.

After the stages of collection of the relevant network data and identification of possible reinforcement actions are completed the planer can use the software described here first for preliminary basic analysis. The preferences of the evaluator described in the previous section and depicted in Figure 6-13 must be correspondingly set. The resulting structure is depicted in Figure 6-14. The first sketchy analysis involves the static deterministic criteria and results in the first approximate Pareto optimal set. Despite the sketchy character of the results they may reveal important information, such as range of criteria variations, and allow identifying preferable planning directions meaning that in the future the planner can choose to concentrate only on some specific range in the criteria domain. Moreover, the results directly can help to detect several alternatives – candidates for the final solution – to be analyzed in more details.

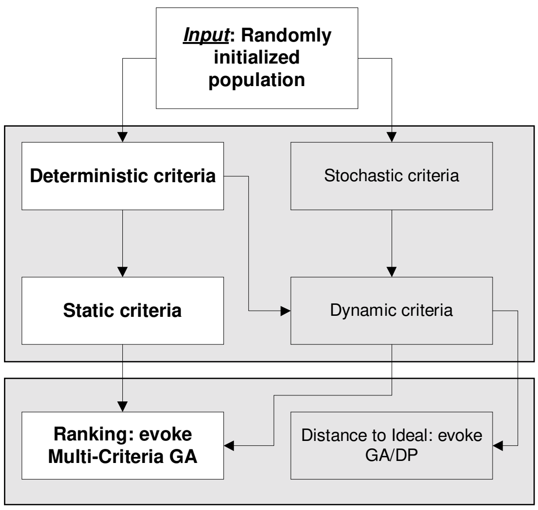


Figure 6-14 The structure I of the evaluator

The planner may continue the analysis of the problem by running more precise Dynamic optimization. There is an option to choose the population from the previous run as an input. Furthermore, the search space may be restricted taking into account the preferences of the planner. Example of the evaluator structure for continuation of analysis is depicted in Figure 6-15. The outcome from this analysis is the Pareto optimal set in the criteria range most interesting for the planner. Each strategy contains the information about the actions to be realized, the optimal sequence of events, as well as the best timing for realization of each alternative.

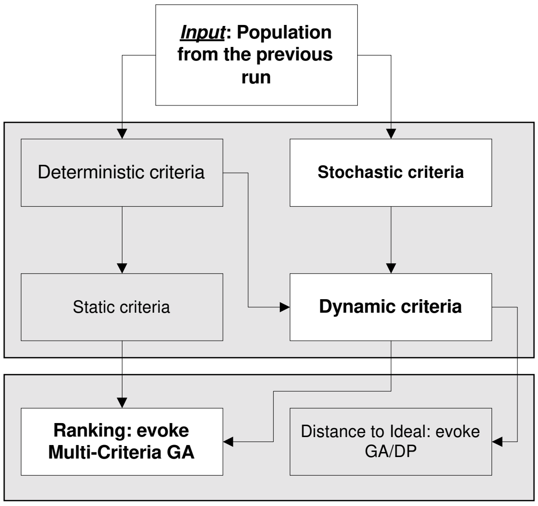


Figure 6-15 The structure II of the evaluator

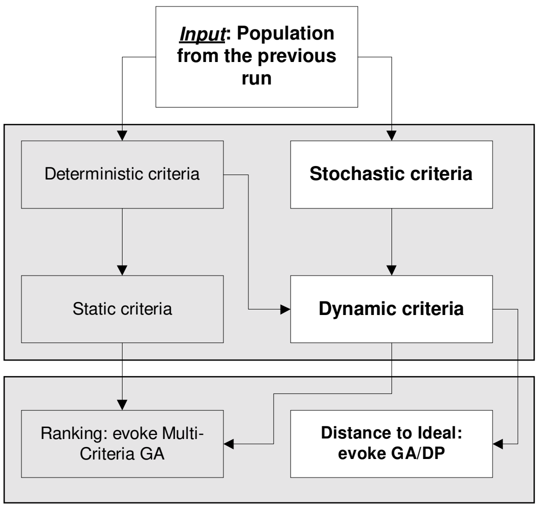


Figure 6-16 The structure III of the evaluator

The next level of precision in the model allows for consideration of the stochastic parameters during the optimization process. The planner may also be interested in finding the best trade-off between the considered planning criteria with help of GA/DP algorithm minimizing Distance to the Ideal (see section 6.2.1). The information obtained at the previous planning stages is required to define Ideal and Anti-Ideal point. The structure of the evaluator in this case may look for example as depicted in Figure 6-16.

6.6.2 Modus Operandi under Long-Term Uncertainty

6.6.2.1 The Ideal Algorithm

It would not be difficult to extend the algorithm presented in the previous section to deal with long-term uncertainties. Three major changes can be pointed out:

- The Monte-Carlo procedure to simulate the futures according to the given parameters of probability distribution (or via random selection of paths in the tree of futures) must be introduced.
- For each individual and each future the adaptation procedure must be performed in order to calculate the corresponding attributes. The options to be realized during the decision-making stage are fixed, while the estimation period options may be adapted to the particular future.
- The number of criteria will double, since the risk measure corresponding to each attribute must be calculated.

However, due to both Monte-Carlo and the adaptation procedures the algorithm becomes very computationally expensive. In many cases such an explicit exploration of the search space is not necessary and even may be considered redundant.

In accordance with the method of the importance sampling [19] and without significant loss in precision it is possible to apply more efficient algorithms gaining a considerable savings of computational capacities.

6.6.2.2 Decomposition into Deterministic Scenarios

The problem may be decomposed into several sub-problems, which are deterministic in terms of long-term uncertainties. Decomposition into deterministic sub-problems, but allowing for the margin of tolerance, leads to significant limitation of the search space. Experience shows [34], that in reality the plans that are robust would lie near the trade-off surface. The planner should provide the tolerance margin for each particular problem. For most tasks the margin 10-20% can be considered as reasonable (see Figure 6-17 and Figure 6-18).

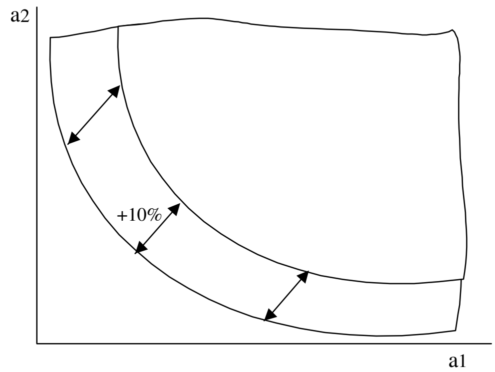


Figure 6-17 Tolerance margin for multi-criteria optimization

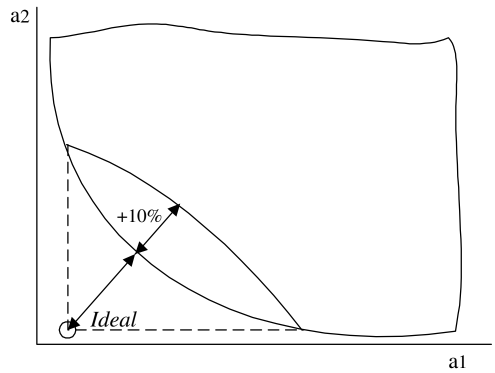


Figure 6-18 Tolerance margin for "Distance to Ideal" optimization

In this case the suggested procedure consists of the following stages:

Stage 1 Identify the tree of futures, probability distributions and the membership functions for the uncertain and random parameters.

Stage 2 Run deterministic optimization according to Stage 4 of the algorithm presented in section 6.6 for the basic future in order to identify a conditional decision set containing both the Pareto optimal set and the solutions minimizing distance to the ideal with a given margin of tolerance.

Stage 3 Repeat the procedure with the given margin of tolerance (Figure 6-17 and Figure 6-18) for the chosen set of futures; identify a conditional decision set for each future.

Stage 4 Find the global decision set via the union of the conditional decision sets. In addition the plans can be supplied manually - taking into account the planner's experience. If the intersection of the conditional decision sets is not empty, the alternatives, which are the members of the resulting set, are 100% robust.

Stage 5 Perform the risk analysis and choose the set of possible candidates for the final solution.

Alternatively, for some tasks the planner may decide to skip the Stage 3, but to form the global decision set from the Stage 2 results and based on the planner's judgment.

6.6.2.3 Risk Analysis

It is important to keep in mind that the plans are considered to be equal if the decision-making period options are equal. Depending on the number of plans included in the global decision set, the risk analysis (Stage 5 in the previous sub-section) may either involve the adaptation of each plan to the given set of futures, or, for a large global decision set, it may involve another run of multi-criteria optimization.

To perform risk analysis on the global decision-set the Multi-Criteria GA does the following:

- Checks if the individual is a member of the global decision set.
- Calculates the expected values of the attributes and the corresponding Expected Maximum Values for a given set of futures. The procedure involves adaptation.

Then the algorithm proceeds according to the process described in section 6.2.3, but the number of criteria is double with respect to the number of planning attributes.

6.6.2.4 Adaptation: Multi-Criteria Case

In order to compare the plans obtained for different futures the procedure of adaptation – an additional optimization in correspondence with dynamic model discussed in section 5.6, where the principle of adaptation was illustrated on one-attribute function – must be involved. The adaptation procedure can be expanded for the multi-criteria case. Consider again the sawmill example, which was discussed in section 5.6. Instead of the total monetary objective function two attributes, namely investments and cost of losses, can be treated separately. In this case, the method of the Displaced Ideal (see section 6.2.1) can be used to compare the performance of different continuations of every plan. The results are presented in Table 6-4, where only the resulting values of two attributes are shown (the intermediate values can be found in Table 5-6). The table contains the outcome from Distance to the Ideal calculations according to the equations (127) and (128). The corresponding values for the Ideal and Anti-Ideal points are given in Table 6-5.

Table 6-4 The scenarios for the sawmill example - multi-criteria approach

All actions during the whole planning period	C_{Inv} USD·10^3	C_{Loss} USD·10^3	d_1	d_2	D
	Future 1 - Plan 1				
A1	800	3400	0.67	0.20	0.70
A1+A3	854.4	3300	0.73	0.10	0.73
A1+A2	1004.2	3250	0.89	0.05	0.89
A1+A2+A3	1058.6	3200	0.95	0.00	0.95
	Future 2 - Plan 1				
A1	800	3130	0.67	0.22	0.70
A1+A3	854.4	3100	0.73	0.17	0.75
A1+A2	1004.2	3050	0.89	0.08	0.90
A1+A2+A3	1058.6	3000	0.95	0.00	0.95
	Future 1 - Plan 2				
A2	300	4050	0.11	0.85	0.86
A2+A3	354.4	3850	0.17	0.65	0.67
A1+A2	844.5	3400	0.72	0.20	0.74
A1+A2+A3	898.9	3350	0.78	0.15	0.79
	Future 2 - Plan 2				
A2	300	3600	0.11	1.00	1.01
A2+A3	354.4	3570	0.17	0.95	0.97
A1+A2	844.5	3200	0.72	0.33	0.79
A1+A2+A3	898.9	3150	0.78	0.25	0.82

Table 6-5 Ideal and Anti-Ideal points for the sawmill example

	Future 1		Future 2	
	C_{Inv} USD·10^3	C_{Loss} USD·10^3	C_{Inv} USD·10^3	C_{Loss} USD·10^3
Ideal	200	3200	200	3000
Anti-Ideal	1100	4200	1100	3600

Table 6-6 Attributes summarized for the sawmill example

	Plan 1		Plan 2	
	C_{Inv} USD·10^3	C_{Loss} USD·10^3	C_{Inv} USD·10^3	C_{Loss} USD·10^3
Future1	800	3400	354.4	3850
Future2	800	3130	844.5	3200
Regret–Future 1	445.6	0	0	450
Regret–Future 2	0	0	44.5	70
Max regret	445.6	0	44.5	450
Expected value	800	3265	599.5	3525
EMV	800	3400	844.5	3850

The highlighted rows for each scenario correspond to the optimal continuations of the first-stage (the decision-making period) plans for the planning period. The corresponding attributes are summarized in Table 6-6. Furthermore, the table contains the values of regret for all the plans, attributes and futures, and the maximal values of the regret from both futures considered in this example. The next row contains the expected values of the attributes assuming equal probability of both futures.

In this example neither Minimal Risk criterion nor the Expected Cost criterion give an adequate answer about which plan to be preferred, but according to the Expected Minimum Value Plan 1 is more robust. The equal probability of occurrence is assumed for both futures, but the table can be easily adapted for different future probabilities. Furthermore, the row containing EMV will change only when the probability of one of the futures will become very low (under the confidence level).

6.7 Conclusions

- GA and other Evolutionary Computation techniques have been widely applied to power systems. The great interests to these methods can be explained by two major reasons: the EA have a number of favorable features in comparison with other methods, and the problems under consideration are often too complex for the traditional optimization methods.
- The planner may quickly obtain the possible planning directions from the outcome of the suggested Multi-Criteria GA, which searches simultaneously for the whole set of Pareto optimal solutions.
- It is convenient to use PCA for graphical representation of the optimization results if the number of criteria is more than three.
- GA is the efficient global optimization technique, which has a slow convergence near the optimum. The suggested hybrid GA/DP algorithm overcomes this disadvantage of GA. With only minor increase in computational expenses and minor interference in the genetic evolution process the convergence properties of the suggested algorithm are considerably improved in comparison with conventional GA.
- Application of the method of Monte-Carlo for calculation of the planning attributes requires a robust optimization method able to cope with noisy objective function. It is shown that the suggested Stochastic GA performs well even with a minor number of trials in Monte-Carlo simulation.
- It is claimed that without significant loss in precision it is possible to decompose the problem into several deterministic sub-problems. Thus, the whole search space is first explored by the approximate methods. Then, the detailed analysis is applied to the limited number of alternatives.

7 Case Studies

This chapter contains the description of two real distribution network planning projects, namely primary distribution network in the large city and secondary network in the rural area. The stages of the "Large Swedish City" project are studied in details.

7.1 Network Planning Project in Large Swedish City

7.1.1 Present Situation in 220-33 kV Distribution Network and General Description of the Problem

A real-life distribution network planning project and the corresponding results are described in this section.

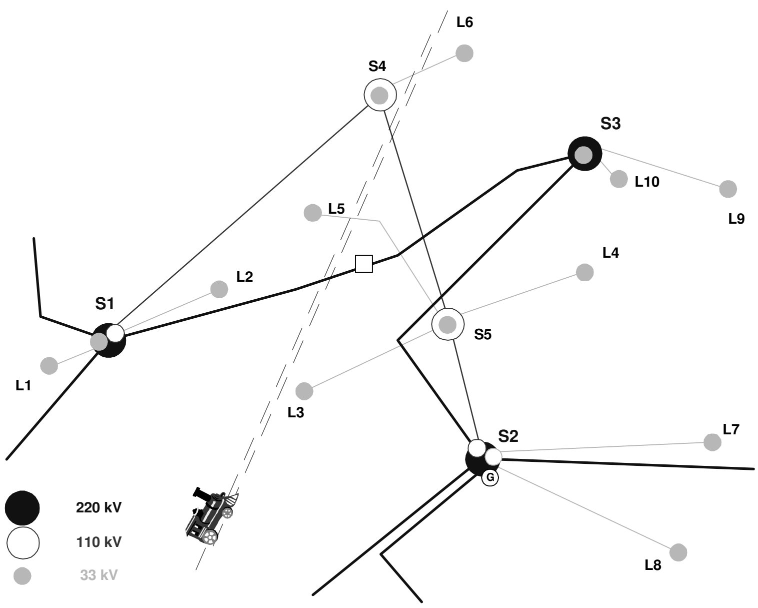


Figure 7-1 Present configuration of the network

The project deals with long-term planning of 220-33 kV distribution network in large city. The part of network under investigation is shown in Figure 7-1. The overhead lines connect two primary 220 kV substations to the national grid. In total there are five substations in the network, including three substation with primary voltage 220 kV and two with 110 kV. Energy is distributed via 110 kV, 33 kV and 11 kV cable systems. The notation is as follows: the 220/110/33 kV and 110/33 kV stations are denoted by "S" followed by number, while load nodes (33/11 kV substations) are denoted by “L” followed by number. When the station S4 or S5 is transformed into the load node, the notation changes correspondingly to L11 and L12.

Present situation in the network is not entirely adequate in several ways, therefore the corresponding reinforcement actions have to be taken. Inadequacy in the present network performance includes the following:

- Environmental impact from leaking oil-filled cables;
- High transformation losses (too many transformation stages – 220/110/33/11 kV);
- Obsolete equipment in substations in substations S1, S2 and S3, in addition:
  - ✓ reinforcement of S1 is desirable from operation point of view;
  - ✓ deficit of capacity in S1;
  - ✓ substation S5 overreached its capacity limit;
  - ✓ short circuit powers are too high.
- Not sufficient personal safety regarding switching operations and works on electrical equipment in older substations - S1, S2 and S5.

7.1.2 The Main Stages of the Planning Project

This main stages of the problem can be summarized as follows:

Stage 1 Identification of possible reinforcement actions and network development strategies. Among the factors, identified as “drivers” of reinforcement process, can be mentioned operation cost, environmental concerns, restrictions and possibilities provided by deregulation of the electricity market in Sweden.

Stage 2 Collection of the relevant network data. The data include information about cables, transformers, reliability indices, costs, etc.

Stage 3 Identification of several network development strategies for further analysis. The selection is based on reasonable acceptable investment level as well as on operational experience.

Stage 4 Preliminary static analysis and network optimization. The results include total values for three selected criteria (cost of losses, energy not supplied, and investments). An important information for the decision-making can be obtained from the joint charts of the Pareto optimal set together with pre-selected solutions (see *Stage 3*) - losses versus investments, and Energy Not Supplied versus investments. The charts serve as the powerful tool for visualization and comparison of different alternatives.

Stage 5 Presentation of the dynamic analysis and optimization as a logical continuation of the static stage. Dynamic analysis allows for distributing the investments and the corresponding reinforcement actions during the planning period, which provides a more realistic and flexible model. The comparison of the results of both static and dynamic optimizations allow us to identify several additional alternative solutions, as well as to restrict the solution domain by maximal value of the investments.

Stage 6 Dynamic simulation studies on several candidate alternatives (from the *Stage 6*). The goal is to investigate the influence of actions realization time on criteria values and, to suggest the best sequence for actions realization.

Stage 7 Uncertainty analysis for the candidate alternatives.

7.1.3 Fundamental Information about the Existing Stations

Substation S1 has reached its capacity limit. It means that in order to connect any new load and to satisfy a future load growth the substation must be reinforced. The situation can be improved by adding the third 110 kV XLPE cable S1-S4, which would increase capacity on 33 kV level. The third cable (or 3x33 kV cables) is an essential condition for the elimination of 110 kV oil-filled cables S2-S5-S4.

In the present situation the capacity of S2 is adequate. However, the equipment is old and, most important the configuration of substation switchgear does not meet the needs of operation and personnel safety. Therefore, reinforcement of S2 is one of the highest preferences. In addition, new transformers and extra connection units are needed to connect more large customers.

The capacity of S3 is rather large. It is estimated that the capacity of S3 is enough to supply its present connections even when considering load growth during the planning period. However, new connections require reinforcement: new switchgear units for the connections within the capacity limit, and installation of the third transformer to feed loads, which exceed the capacity limit.

S5 is another old substation in the network, which needs to be reinforced in the nearest future due to aging. New 33 kV switchgear has to be build if S5 continues to feed the loads presently supplied via this substation (e.g. substations L3, L4 and L5, see Figure 7-1).

7.1.4 Network Reinforcement Alternatives

7.1.4.1 List of suggested reinforcement alternatives

The following network reinforcement strategies and corresponding alternatives were suggested:

- step-by-step elimination of 110 kV cables
- a new 220/33 kV substation in S18/S20
- two new 220/11 kV substations (S18 and S19), development of 11 kV network.

Furthermore, the alternatives listed above can be considered together with the following supplementary alternatives:

✓ meshed network configuration
✓ introduction of local generation.

This section contains a detailed description of each suggested strategy.

7.1.4.2 Step-by-step elimination of 110 kV cables

The first step would include reconstruction of substation in S2. S2 feeds loads L3, L4 and L12 as well as L7 and L8. Then the 110 kV cables S1-S5 and S5-S4 can be removed*.

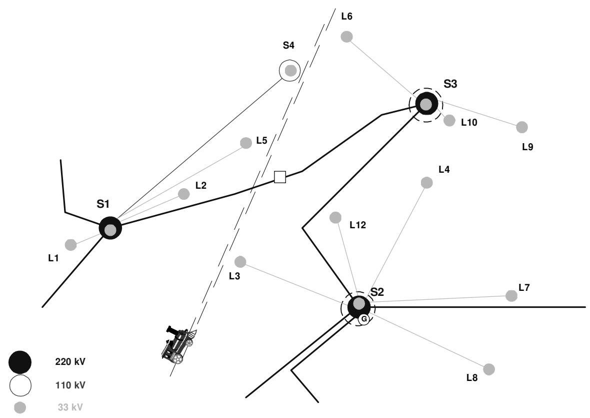


Figure 7-2 Step-by-step elimination of oil-filled 110 kV cables (a)

As further development of this alternative, two possible scenarios of the network reinforcement may be considered:

- Reinforcement of substation S3 – addition of the third transformer 220/33 kV, node 13 is fed from S3 (Figure 7-2);

---

* Later analysis showed that elimination of cables S2-S5-S4 is feasible only after addition of the third XLPE cable S1-S4.

- Addition of the third 110 kV XLPE cable between S1 and S4, node 13 is fed from S4 (Figure 7-3).

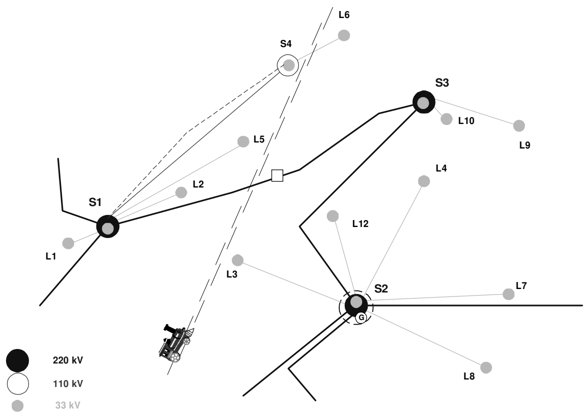


Figure 7-3 Step-by-step elimination of oil-filled 110 kV cables (b)

In more distant future it is possible to consider total elimination of 110 kV voltage level in the network. Cables from S1 to node 4 can be used at 33 kV.

Advantages:

- the main goals, namely reinforcement of S2 and elimination of oil-filled 110 kV S2-S5-S4 cables are accomplished;
- flexible and comparably inexpensive alternative;
- alternative, which provides a long technical life with personnel safety for the existing substations.

Disadvantage:

- requires considerable investments.

7.1.4.3 A new 220/33 kV substation S18 or S20

The major investment in this alternative is investment in new 220/33 kV substation S18. However there are two possible sites for the substation separated by the existing railway. Therefore, two possibilities of the network reinforcement may be considered:

- building a new 220/33 kV substation S18, which feeds nodes L5, L11 and L6; but nodes L3, L4, L12 as well as L7 and L8 are fed from S2 (Figure 7-4);
- building a new 220/33 kV substation S20, which feeds L5, L12 and possibly partly the loads of L3 and L4 (Figure 7-5).

Both sub-alternatives assume also:

- reinforcement of the substation S2;
- elimination of the 110 kV oil-filled cables S2-S5-S4.

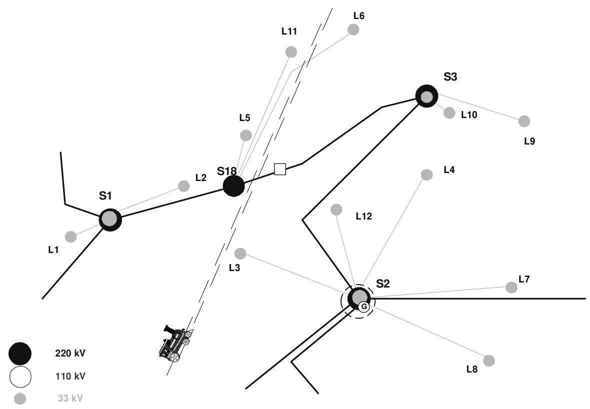


Figure 7-4 New substation in S18

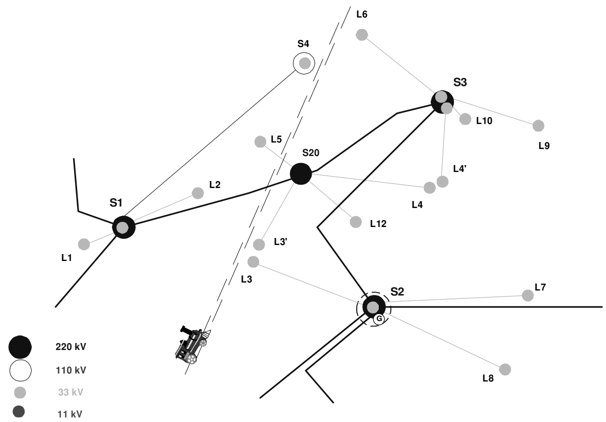


Figure 7-5 New substation in S20

Advantages:

- the main goals, namely reinforcement of S2 and elimination of oil-filled 110 kV cables are accomplished;
- considerable reinforcement of the network capacities.

Disadvantages:

- considerable investments.

7.1.4.4 Two new substations, development of 11 kV network

The strategy suggests starting with a reinforcement of the 220 kV switchgear in S2. Substation S5 is going to disappear in this strategy too. When this happens, a new 33 kV switchgear and new 220/33 kV transformers have to be installed in S2.

Two new substations 220/11 kV can be built in this area (providing the reinforcements on both sides of the railway) (Figure 7-6) – S18 and S19.

Under this strategy the following changes in the present network configuration are expected:

- 33/11 kV substation L4 disappears, the loads are distributed between S3 and S19;
- 33/11 kV substation L3 disappears, the loads are distributed between S18 and S19;
- 33/11 kV substation L5 disappears, the load is supplied from S18;
- 33/11 kV substation L12 disappears, the 11 kV load is supplied from S19;
- a large 33 kV customers are converted to 11 kV.

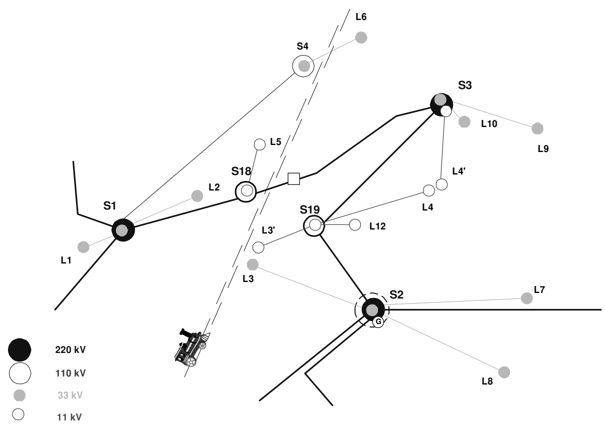


Figure 7-6 Two new substations, development of 11 kV network

Advantages:

- the main goals, namely reinforcement of S2 and elimination of leaking oil-filled 110 kV cables are accomplished;
- considerable reinforcement of the network capacities.

Disadvantages:

- extremely high investment costs;
- traditional 33 kV customers must be converted to 11 kV;
- a large number of new 11 kV cables must be installed (the usual three 33 kV cables from 220/33 kV substation to 33/11 kV station must be replaced by about twenty to thirty 11 kV cables – depending on the power capacity of the substation).

7.1.4.5 Supplementary alternatives

<u>Mesh network</u>

In addition to the traditional radial configuration, mesh or ring network structure on 33 kV level may be considered.

Advantages:

- the meshed network provides the possibility for back-up supply in case of contingency.

Disadvantages:

- additional switchgear units must be reserved;
- higher investments.

Distributed generation

Attracting the investments in new local generation sources and providing corresponding changes in existing substations – a new possibility provided by deregulation. This alternative may be realized, when some deficit of supply may occur.

Advantages:

- lower investments;
- higher available capacities in the network.

Disadvantages:

- not very realistic for practical realization.

7.1.5 A Model for Computerized Calculations and Optimization

The extensive examination of the strategic alternatives presented in section 7.1.4 allows for introducing a model for computerized calculations and optimization. The model, which generalizes all reinforcement actions suggested for all the corresponding alternatives is depicted in Figure 7-7.

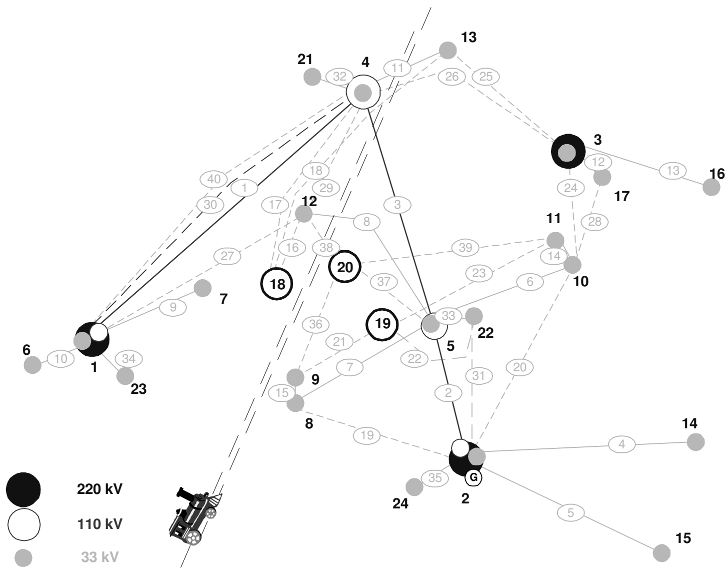


Figure 7-7 Network model for computerized calculations

The existing connections are depicted by solid lines, the dashed lines represent the possible routes for the new cables. Present loads of L4 and L3 are divided into two equal parts. It allows us to model the possibility to supply these loads from different substations – at least theoretically. For example L4' can be supplied from S3, while L4 from S20.

The nodes 21, 22, 23 and 24 are artificial and where created for modeling purposes. They correspond to the load associated directly with respective substations. For example node 21 carries the load of S4 and node 24 the load of S2.

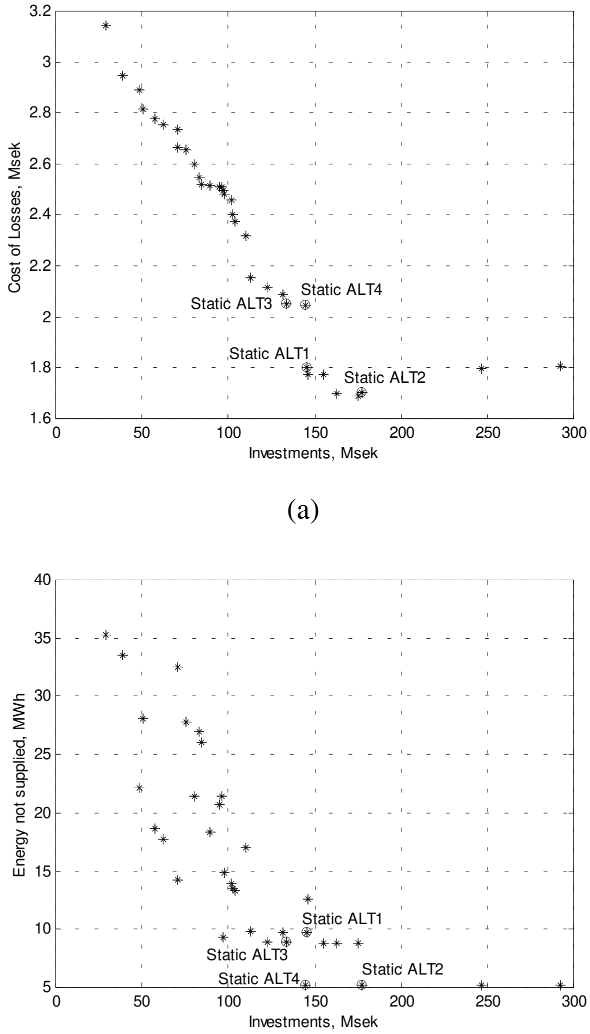
7.1.6 Optimization Results

7.1.6.1 Static optimization

Static optimization makes several assumptions about the planning process. First of all it is assumed that the state of the network does not change during the planning period. All the investments foreseen for the planning period are made during the first year. The values of the criteria are calculated only once and summed up for the planning period using the annuity principle. These simplifications may lead to simplified and not totally adequate results. However, as a preliminary analysis tool static optimization allows us to identify interesting network configurations and corresponding reinforcement actions prior to analyzing times of actions realization.

The results of static optimization are depicted by asterisks in Figure 7-8. Some alternatives within the most interesting range of investments, which were obtained as a result of static optimization, are depicted by circles. The corresponding final network configurations are shown in Table 7-1.

One preliminary conclusion from the results of static optimization is that for the given network the alternatives with new substations are not among the most advantageous. Instead, it is more profitable to use and reinforce the present assets. Another conclusion, which can be made, is that elimination of the 110 kV voltage stage will result in decreased losses.

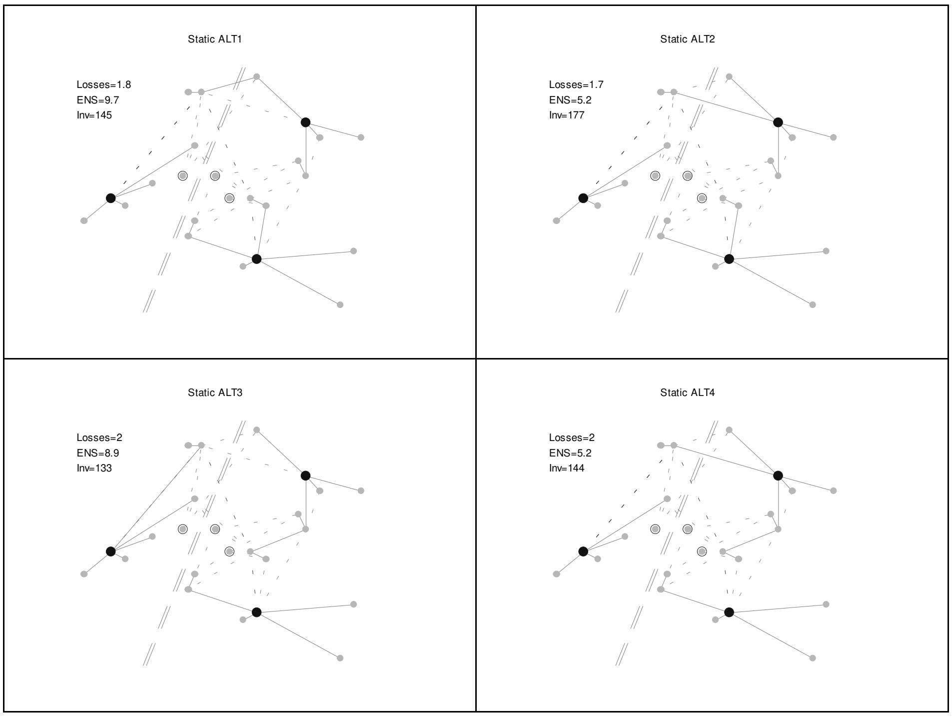


(a)

(b)

Figure 7-8 Pareto optimal set obtained as a result of static optimization (a) Cost of losses and (b) Energy Not Supplied versus Investments

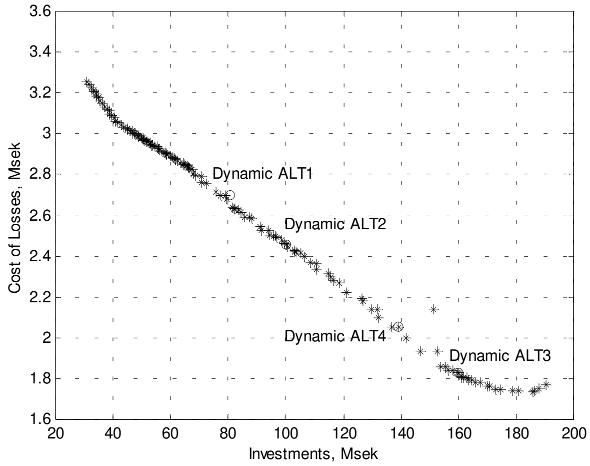
Table 7-1 Final configurations obtained from the results of static optimization



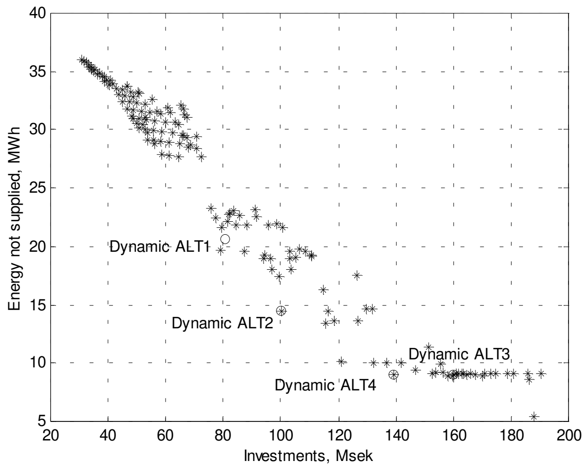
7.1.6.2 Dynamic optimization

Dynamic optimization suggests not only the best actions to be realized, but also the optimal sequence of the events and the best timing for realization of each action. The values of the criteria are calculated for each time stage, therefore the network configuration may be modified at each subsequent time stage. Thus the most essential actions are promoted during the optimization and the actions, which are less critical, may be postponed. Thus, the solutions become more flexible in comparison with results of static analysis. Therefore, in dynamic case, the investments can be distributed during the planning period and correspondingly the network configuration can differ from stage to stage.

The results of dynamic optimization are presented in Figure 7-9. Some alternatives (depicted by circles in the figure) in a wide investment range have been chosen for more detailed analysis. Configurations, corresponding to the last time stage of the strategies described by these alternatives are summarized in Table 7-2.



(a)

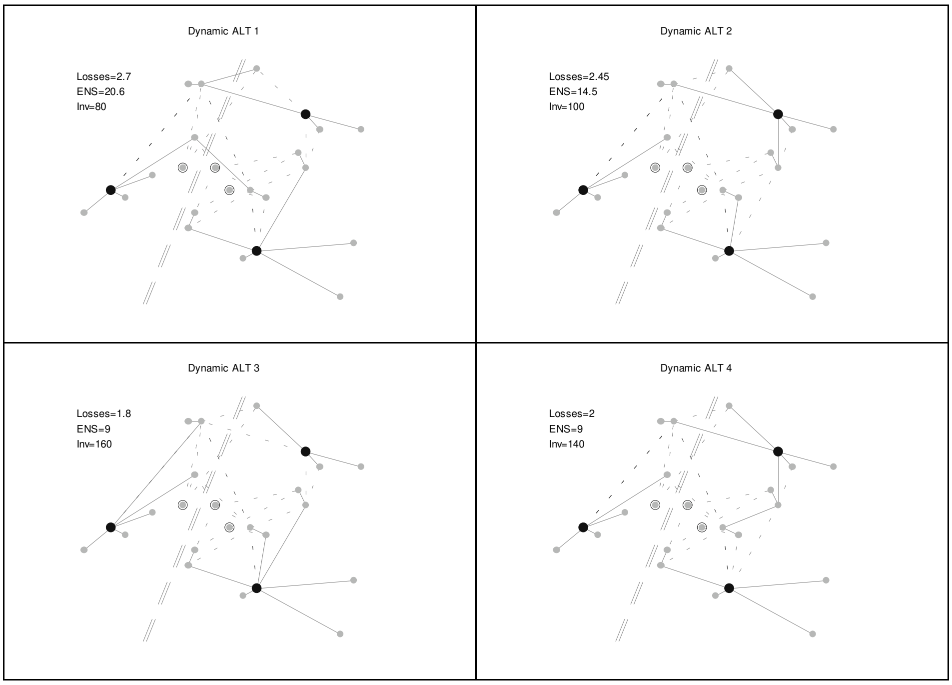


(b)

Figure 7-9 Pareto optimal set obtained as a result of dynamic optimization (a) Cost of losses and (b) Energy Not Supplied versus Investments

Preliminary conclusions from the results of dynamic optimization coincide with those made previously in case of static analysis. Furthermore, we can see that the dynamic last stage configurations from Table 7-2 are very similar to static optimization results given in Table 7-1. Thus, for example the configuration for Static ALT2 is exactly the same as for Dynamic ALT2, however the criteria values differ considerably. This can be explained by the fact that the network may undergo several changes during the planning period in dynamic case and remains unchangeable in static. In the first case the reinforcement actions may be distributed during the planning period, while in the second – all actions are realized at the first year. In this particular case, all the investments for the alternative Static ALT2 are made during the first year, while for Dynamic ALT2 the investments are postponed partly to the second and partly to the fourth planning stage – this results in considerable difference in total sum of investments.

Table 7-2 Final configurations obtained from the results of dynamic optimization

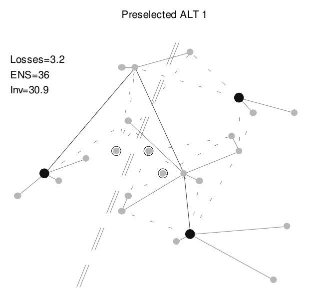
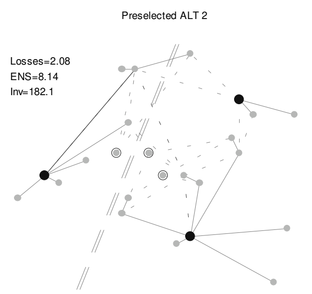
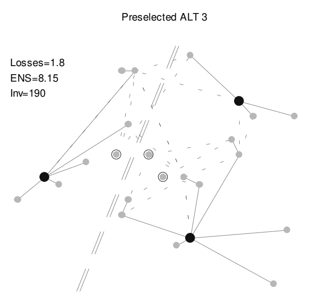


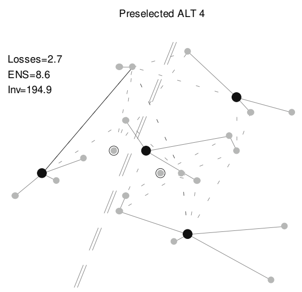
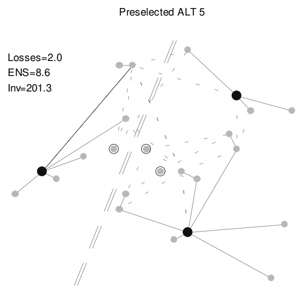
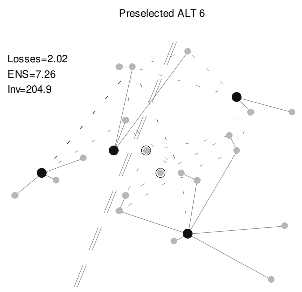
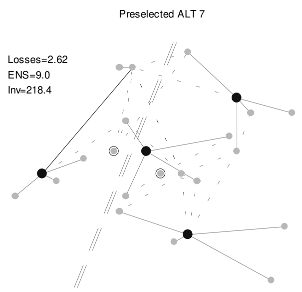
7.1.6.3 Analysis of pre-selected alternatives in comparison with the results of optimization

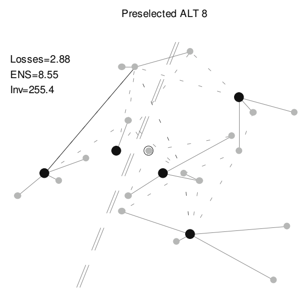
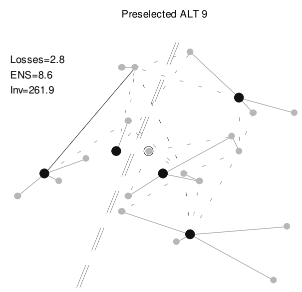
The choice of pre-selected alternatives, which are analyzed in this section, is based on different alternatives described in chapter 7.1.4. However, some alternatives were modified to more realistic form. Thus, the idea of development of the 11 kV network with two new substations was excluded from consideration due to its major drawback, namely, requirement to lay too many feeders. In this case the traditional three 33 kV cables from 220/33 kV substation to 33/11 kV station must be replaced by about twenty to thirty 11 kV cables – depending on the number of feeders and customers. Instead, the utilization and development of 33 kV is considered.

The modified pre-selected alternatives are summarized in Table 7-3, where the description of reinforcement actions corresponding to every alternative as well as the results of criteria calculations are given.

Table 7-3 Summary of pre-selected alternatives

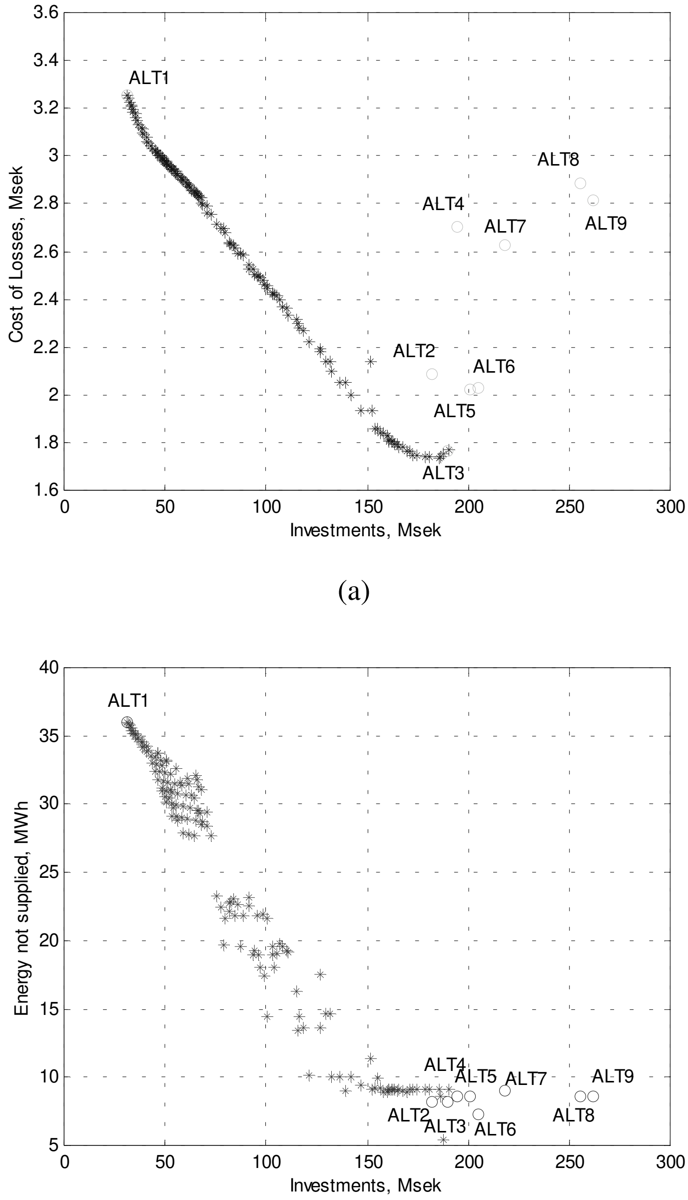
	Pre-selected alternatives (presented in order of investments augmentation)	Corresponding network configuration
1	Existing network • reinforcement of S2	
2	Elimination of 110 kV cables (a) • reinforcement of S2 • reinforcement of S1 • third 110 kV XLPE cable S1-S4 • L5 is fed from S1 • minor reinforcement of S5 • L3, L4 and L12 are fed from S2 on 33 kV • elimination of 110 kV oil-filled cables S2-S5-S4	
3	Elimination of 110 kV cables (b) • reinforcement of S2 • reinforcement of S1 • switchgear units in S3 • L6 is fed from S3 • three 33 kV cables S1-S4 • L5 is fed from S1 • minor reinforcement of S5 • L3, L4 and L12 are fed from S2 on 33 kV • elimination of 110 kV oil-filled cables	

4	New substation S20 (a) • S20 substation • minor reinforcement of S5 • L5, L4 and L3 are fed from S20 • reinforcement of S2 • L3 is fed from S2 • third 110 kV XLPE cable S1-S4 • elimination of 110 kV oil-filled cables S2-S5-S4 • switchgear units in S3 • L6 is fed from S3	
5	Elimination of 110 kV cables (c) the same as 3, but • third 110 kV XLPE cable S1-S4	
6	New substation S18 • S18 substation • L11, L5, L6 are fed from S18 • reinforcement of S2 • minor reinforcement of S5 • L3, L4 and L12 are fed from S2 on 33 kV • elimination of 110 kV oil-filled cables S2-S5-S4	
7	New substation S20 (b) • S20 substation • minor reinforcement of S5 • reinforcement of S2 • the loads of L3 are supplied by S20 and S2 • the loads of L4 are supplied by S20 and S3 • L12 and L5 are fed from S20 • third 110 kV XLPE cable S1-S4 • elimination of 110 kV oil-filled cables S2-S5-S4 • switchgear units in S3 • L6 is fed from S3	

8	Two new substations – S18 + S19 (a) • S18 substation • S19 substation • minor reinforcement of S5 • reinforcement of S2 • the loads of L3 are supplied by S19 and S2 • the loads of L4 are supplied by S19 and S3 • L12 is fed from S19 • L5 is fed from S18 • third 110 kV XLPE cable S1-S4 • elimination of 110 kV oil-filled cables S2-S5-S4	
9	Two new substations – S18 + S19 (b) • S18 substation • S19 substation • minor reinforcement of S5 • reinforcement of S2 • the loads of L3 are supplied by S19 and S2 • L4 and L12 are fed from S19 • L5 is fed from S18 • third 110 kV XLPE cable S1-S4 • elimination of 110 kV oil-filled cables S2-S5-S4 • switchgear units in S3 • L6 is fed from S3	

The values of the criteria for pre-selected alternatives in comparison with the results of dynamic optimization are depicted in Figure 7-10. The following conclusions can be made from this analysis:

- new substations considerably improve reliability, and result only in a minor improvement of power losses (due to new transformation losses)
- construction of two new substations in the area is unreasonably expensive
- elimination of 110 kV oil-filled cables leads to considerable decrease in the network losses and at the same time the corresponding actions for reinforcement of existing network substantially improve reliability and personnel safety.



(a)

(b)

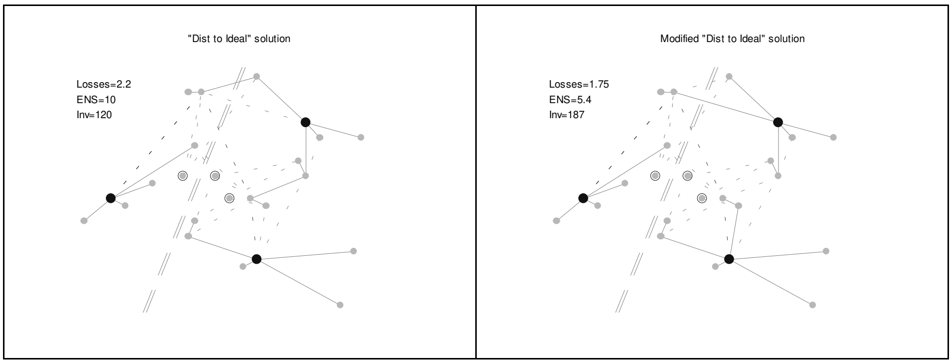
Figure 7-10 Pareto optimal set obtained as a result of dynamic optimization (a) Cost of losses and (b) Energy Not Supplied versus Investments and the set of pre-selected alternatives

7.1.6.4 Search for "the optimal" solution - "Distance to the Ideal" optimization

To check the adequacy of subjective choice of the possible candidates from the Pareto optimal set the optimization concept "Distance to the Ideal" was applied.

The results of optimization confirm the common trends identified so far. However, the resulting network configuration is not adequate enough (Table 7-4). It can be modified to the reasonable form by adding two cable connections – feeding L11 from S3 and feeding L12 from S2. The results of dynamic optimization and both original and modified "Distance to the Ideal" solutions are depicted in Figure 7-11.

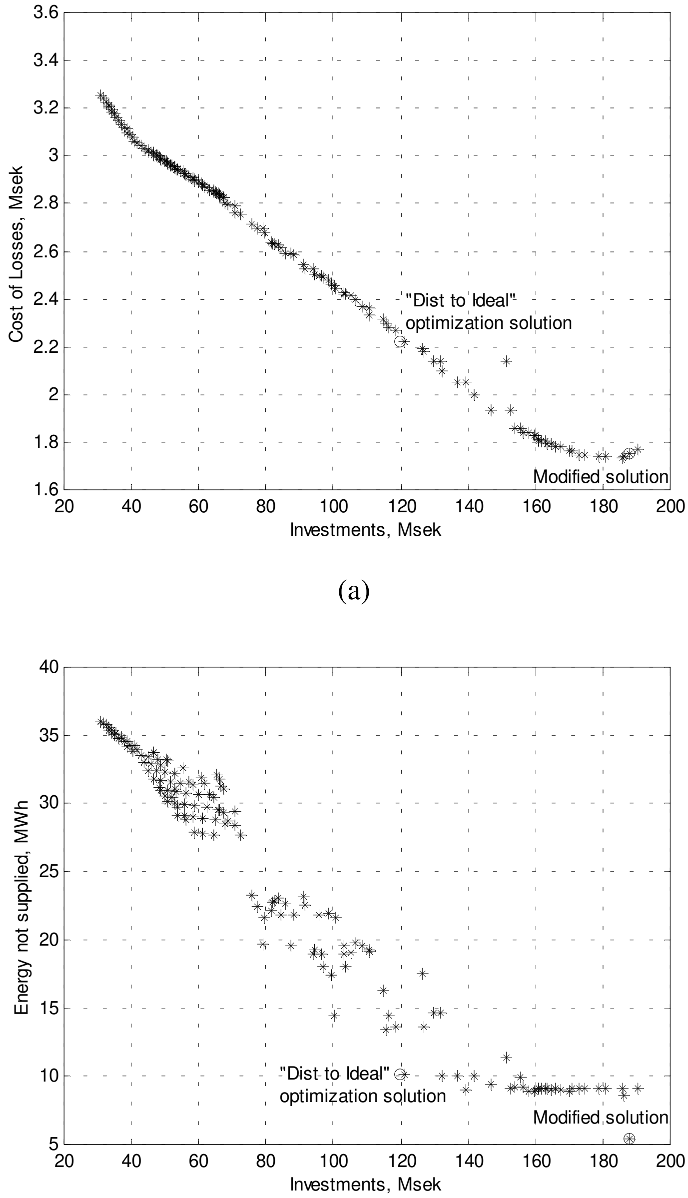
Table 7-4 Final configurations for the "Distance to the Ideal" solutions



7.1.6.5 Possible solutions – summary of the results

All results of different types of optimizations described in the previous sections result in similar alternatives. These alternatives can be summarized and several reasonable candidates for the final solution can be chosen.

The final set of candidate solution (conditional decision set) is depicted in Figure 7-12. Four dynamic alternatives are included into the set of candidates, as well as both the original and modified "Distance to the Ideal" solutions denoted respectively IdealMin and IdealMod. Furthermore, allowing for the tolerance margin (10%), a number of alternatives close to the one minimizing distance to the Ideal can be obtained. Five such alternatives are also included into the set. These are the points situated close to the IdealMin in Figure 7-12. The attributes for all the candidates are summarized in Table 7-5.



(a)

(b)

Figure 7-11 Pareto optimal set obtained as a result of dynamic optimization (a) Cost of losses and (b) Energy Not Supplied versus Investments and the solution of “Distance to Ideal” optimization

Table 7-5 Summary of the attribute values for the conditional decision set

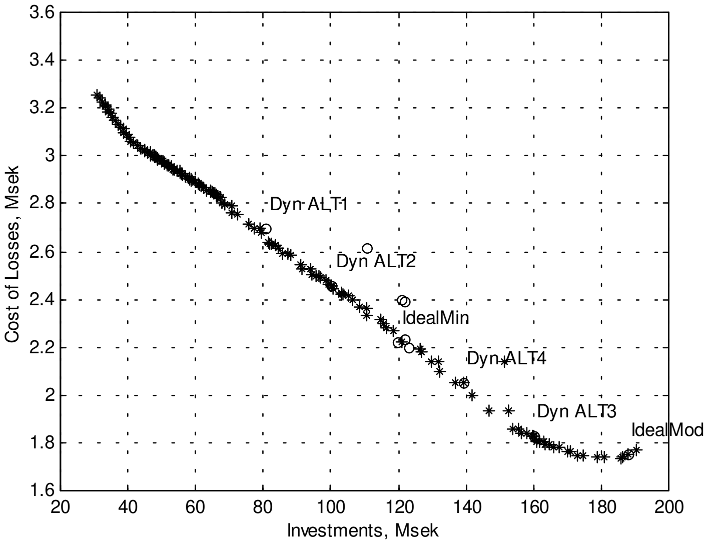
Alternative	Losses, Msek	ENS, MWh	Investments, Msek
Dyn ALT1	2.72	20.64	80.8
Dyn ALT2	2.50	14.47	100.4
Dyn ALT3	1.82	9.01	159.9
Dyn ALT4	2.01	9.00	139.2
Ideal_Min	2.20	10.13	119.0
Ideal _Mod	**1.74**	**5.41**	**187.8**
Ideal ALT1	2.23	10.06	122.0
Ideal ALT2	2.19	10.12	123.2
Ideal ALT3	2.38	9.17	121.9
Ideal ALT4	2.62	10.81	110.7
Ideal ALT5	2.39	9.17	120.9

The solution minimizing losses and ENS is Ideal_Mod (bold in Table 7-5). Unfortunately, this alternative is unacceptably expensive. According to the utility management the investments into this project should not exceed 100÷120 Msek. Since losses are expressed in monetary terms and represent costs for the same shareholder as the investments, they can be directly compared. The comparison clearly indicates that no benefit can be made investing money in order to reduce losses. However, there is an intention to improve reliability in the network and keep it at the certain level. This level can set by the limit of 25 MWh in Figure 7-12, which should not be exceeded.

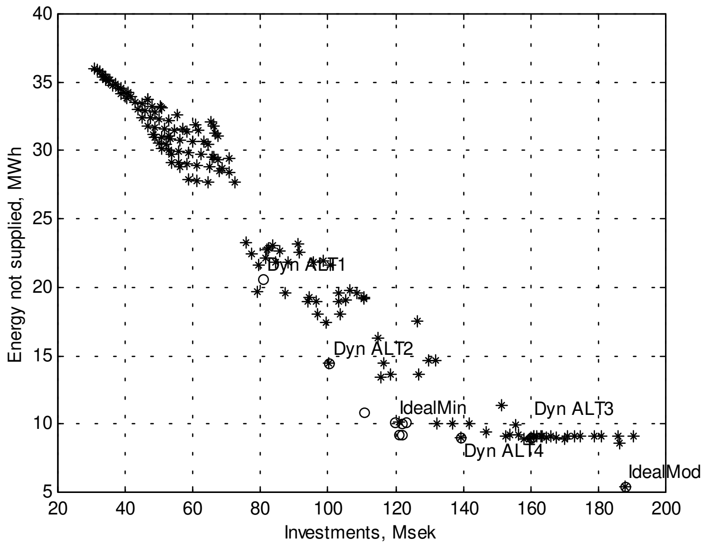
7.1.7 Analysis under Uncertainty

The results of investigation of load uncertainty influence on planning attributes are presented below. Taking into account quadratic dependence between power losses and loads, the most interesting results can be obtained investigating the behavior of cost of losses varying nodal loads.

Assume that the future development of loads is known. The variation of nodal loads in this case is caused by short-term uncertainties and modeling imprecision. Table 7-6 contains the total cost of losses for the conditional decision set for deterministic case, 4% load deviation and 8% load deviation in each node. Furthermore, in presence of uncertainty the additional characteristic reflecting risk – EMV – is used. EMV is calculated with 95% confidence limit. The results were obtained applying Monte-Carlo simulation with 1000 trials to each candidate alternative.



(a)



(b)

Figure 7-12 The set of candidates for the final decision and Pareto optimal- result from dynamic optimization (a) Cost of losses and (b) Energy Not Supplied versus Investments

Table 7-6 Mean and EMV of losses for the candidate solutions for two levels of load variation, Msek

Alternative	Deterministic	Mean value		Expected Max Value (95% confidence)	
		σ =4%	σ =8%	σ =4%	σ =8%
Dyn ALT1	2.72	2.83	3.04	3.23	3.83
Dyn ALT2	2.50	2.60	2.81	2.99	3.71
Dyn ALT3	1.82	1.87	1.99	2.17	2.73
Dyn ALT4	2.01	2.06	2.20	2.36	2.97
Ideal_Min	2.20	2.24	2.38	2.58	3.25
Ideal _Mod	**1.74**	**1.78**	**1.93**	**2.02**	**2.67**
Ideal ALT1	2.23	2.27	2.4	2.59	3.26
Ideal ALT2	2.19	2.23	2.38	2.54	3.26
Ideal ALT3	2.38	2.41	2.59	2.72	3.52
Ideal ALT4	2.62	2.66	2.82	3.04	3.73
Ideal ALT5	2.39	2.42	2.55	2.74	3.31

Table 7-7 Mean and EMV of losses for the candidate solutions for single and multi rate tariff and σ =4% , Msek

Alternative	Deterministic	Mean value		Expected Max Value (95% confidence)	
		Single rate	Multi rate tariff	Single rate	Multi rate tariff
Dyn ALT1	2.72	2.83	2.89	3.23	3.30
Dyn ALT2	2.50	2.60	2.67	2.99	3.20
Dyn ALT3	1.82	1.87	1.90	2.17	2.28
Dyn ALT4	2.01	2.06	2.09	2.36	2.50
Ideal_Min	2.20	2.24	2.30	2.58	2.74
Ideal _Mod	**1.74**	**1.78**	**1.80**	**2.02**	**2.18**
Ideal ALT1	2.23	2.27	2.33	2.59	2.79
Ideal ALT2	2.19	2.23	2.28	2.54	2.73
Ideal ALT3	2.38	2.41	2.46	2.72	2.90
Ideal ALT4	2.62	2.66	2.72	3.04	3.21
Ideal ALT5	2.39	2.42	2.48	2.74	2.90

The most important observation, which can be made from the analysis of the results is that not only risk value increases with the level of uncertainty, but so does the mean value even if the mean of nodal loads remains the same.

In this case study the single rate for the losses was used to estimate costs of present and future losses. In reality losses during the peak hours are more expensive. Table 7-7 illustrates the influence of multi rate tariff on total losses. Again, it can be observed that both the mean value and EMV move towards higher values. These results are consistent, since such differentiation in price places higher weight on losses during peak hours.

Previous calculations were performed accepting deterministic forecast for future load growth, however this is random parameter. There is no statistical information on future trends, therefore it is convenient to represent the interval of possible load growth and the corresponding membership function. The results obtained using Monte-Carlo simulation for both probability and fuzzy distributions are presented in Table 7-8.

Table 7-8 Mean and EMV of losses for the candidate solutions for deterministic and fuzzy load growth (load variation σ =4%), Msek

Alternative	Deterministic	Mean value		Expected Max Value (95% confidence)	
		Determi-nistic growth	Fuzzy growth	Determi-nistic growth	Fuzzy growth
Dyn ALT1	2.72	2.83	3.18	3.23	3.97
Dyn ALT2	2.50	2.60	2.93	2.99	3.85
Dyn ALT3	1.82	1.87	2.18	2.17	2.91
Dyn ALT4	2.01	2.06	2.33	2.36	3.21
Ideal_Min	2.20	2.24	2.57	2.58	3.49
Ideal _Mod	**1.74**	**1.78**	**2.09**	**2.02**	**2.90**
Ideal ALT1	2.23	2.27	2.60	2.59	3.52
Ideal ALT2	2.19	2.23	2.57	2.54	3.43
Ideal ALT3	2.38	2.41	2.76	2.72	3.62
Ideal ALT4	2.62	2.66	2.98	3.04	3.97
Ideal ALT5	2.39	2.42	2.75	2.74	3.61

Moreover, if the membership functions are unknown or if there is a possibility of appearance of discrete future events influencing the nodal loads, it would be convenient apply scenarios approach. In this particular case there is a possibility that one side of the railway will develop more rapidly than another. Thus, the following three scenarios can be identified:

- Realistic – the forecasted load growth is given in Appendix A
- Sc1 Left side of the railway develops very rapidly
- Sc2 Right side of the railway develops very rapidly

The mean and EMV for losses for all three scenarios are presented in Table 7-9.

Table 7-9 Mean and EMV of losses for the candidate solutions for three scenarios (load variation σ =4%), Msek

Alternative	Mean value			Expected Max Value (95% confidence)		
	Realistic	Sc1	Sc2	Realistic	Sc1	Sc2
Dyn ALT1	2.83	3.18	3.44	3.23	3.65	3.84
Dyn ALT2	2.60	2.91	3.24	2.99	3.39	3.71
Dyn ALT3	1.87	2.13	2.46	2.17	2.50	2.78
Dyn ALT4	2.06	2.34	2.69	2.36	2.68	3.05
Ideal_Min	2.24	2.53	2.90	2.58	2.91	3.30
Ideal _Mod	**1.78**	**2.02**	**2.36**	**2.02**	**2.31**	**2.70**
Ideal ALT1	2.27	2.54	2.94	2.59	2.92	3.32
Ideal ALT2	2.23	2.51	2.90	2.54	2.89	3.31
Ideal ALT3	2.41	2.77	3.15	2.72	3.18	3.57
Ideal ALT4	2.66	3.06	3.40	3.04	3.48	3.84
Ideal ALT5	2.42	2.79	3.13	2.74	3.15	3.49

It can be concluded that the behavior of losses is very predictable – whatever the future is more expensive alternatives have lower losses. Comparing the values of cost of losses and the required investments under all possible future outcomes allow us to conclude, that it is not worthwhile to invest with a purpose to reduce losses in this network.

Table 7-10 Mean and EMV of ENS for the candidate solutions for deterministic case and fuzzy load growth with load variation σ =4%, MWh

Alternative	Deterministic	Mean value	Expected Max Value (95% confidence)
Dyn ALT1	20.64	21.5	**28.5**
Dyn ALT2	14.47	15.9	21.0
Dyn ALT3	9.01	9.1	11.5
Dyn ALT4	9.00	9.1	11.1
Ideal_Min	10.13	10.1	12.5
Ideal _Mod	5.41	5.4	6.3
Ideal ALT1	10.06	10.1	12.5
Ideal ALT2	10.12	10.1	12.6
Ideal ALT3	9.17	9.2	11.6
Ideal ALT4	10.81	10.8	13.3
Ideal ALT5	9.17	9.0	11.6

Finally, it is important to verify the performance of reliability under uncertainty. Assuming fuzzy load growth and probabilistic load variation, the

Monte-Carlo simulation was applied in order to calculate ENS for the conditional decision set. The results are summarized in Table 7-10. The mean value practically coincide with ENS value calculated in deterministic case. However, there is a risk that Dyn ALT1 will not meet the reliability requirements (as said the ENS should not exceed 25 MWh).

In presence of uncertainty the number of criteria increase – in this case study the risk measures for cost of losses and ENS were added. For graphical representation of the results it is convenient to use PCA analysis (see section 6.2.4). The corresponding scores plot and the loadings are depicted in Figure 7-13.

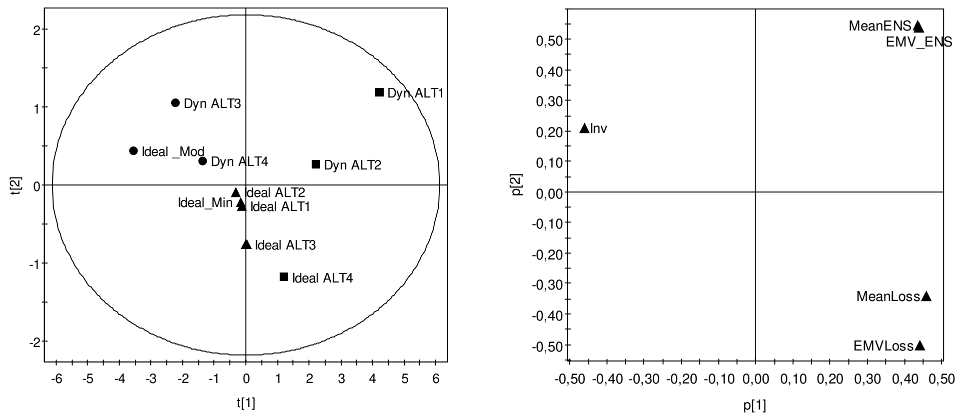


Figure 7-13 Scores plot (left) and loadings (right) for the conditional decision set

The conditional decision set can be divided into three sub-sets. The first one is depicted by dots and consists of IdealMin, Dyn ALT3, Dyn ALT4 (see scores plot in Figure 7-13). These alternatives are the most expensive, but they perform best in terms of reliability and power losses. The second group is depicted by boxes and consists of Dyn ALT1, Dyn ALT2 and Ideal ALT4. They are much cheaper, but perform worse on other criteria. The third group is intermediate between the first two and is depicted by triangles.

The interesting information can be revealed from the plot of loadings in Figure 7-13. First of all, it confirms the conflict between the objectives – as they are situated in the opposite segments of the plot. Furthermore, the plot demonstrates that the objectives to minimize losses and risk of extensive losses agree. Similarly perform the objectives to minimize ENS and the corresponding risk.

7.2 Rural Network Planning Project

7.2.1 The Commercial Software Used for the Studies

7.2.1.1 Latvian Dynamic Model (LDM) Software

The program LDM [34],[71] is foreseen for reinforcement planning of MV and LV distribution network under information uncertainty. The program has the following main functions:

- technical and economic estimation of the present state of the network
- define economically appropriate actions from the given set of alternative actions (construction, reconstruction or elimination of network elements) and terms of their realization
- estimate power supply quality and define the most effective actions to improve it
- provide risk analysis as a decision-making tool under information uncertainty.

Moreover, calculations such as load flows, voltage drops, fault currents, energy and power losses, reliability estimates, annual and total costs for the particular network as well as investments pay-back times are also provided.

There are two main options in the program: dynamic optimization and analysis. Dynamic optimization uses as an input data from the database together with development conditions and alternative actions.

Net present value of total network costs (investments, cost of losses, cost of energy not supplied) is used as an optimization criterion.

As a result of optimization except for the optimal strategy, one gets the set of "good" strategies. The final decision can be made taking into account uncertainties according to the minimal risk criterion.

7.2.1.2 Swednet Software

Swednet is a network planning program aimed for expansion as well as reinforcement planning of LV distribution networks [8].

The program has the following main functions:

- technical estimation of the present state of the network (load flow, fault currents)
- search for the optimal sites for substations, their sizes, loading and supply areas
- search for the optimal network configuration (feeders routes), sizes of feeders and loading.

The basic planning data consists of information about the existing network of substations and feeder system, possible new sites for substations and feeder routes and the loading points.

The cost function (optimization criterion) consists of the sum of investments and fixed and operation related costs of the network during the planning period.

Optimization proceeds uses a horizon year basis. This means, that the network is once and for all constructed to meet the demand over the whole utilization period.

7.2.2 Rural Network in Jelgava

In Figure 7-14 a 0.4 kV network in a rural area is presented. Presently, the network is supplied by one substation of 250 kVA. Calculations indicate that even in the present state voltage drops in the network exceed allowable limits (voltage drop threshold is 5%). Therefore a reinforcement strategy is required.

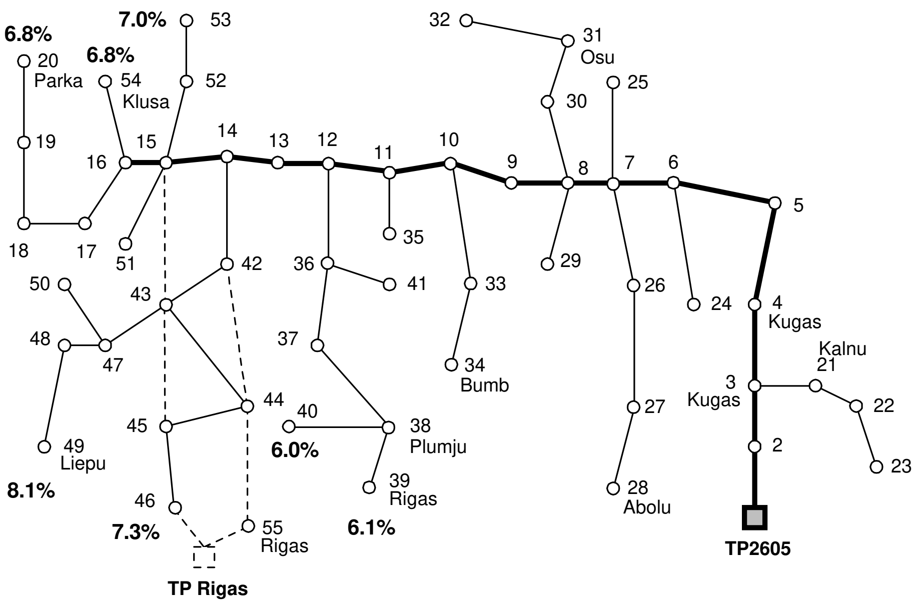


Figure 7-14 0.4 kV network in Jelgava

The following actions for network reinforcement are suggested for further analysis (Figure 7-14):

- Introduction of new substation TP Rigas
- New lines (doted lines in Figure 7-14)
- Replacement of conductors in several lines.

The task was studied with both Swednet and LDM programs. The planning period is 15 years forward. Interest rate is 10% and decision-making period for

LDM program is year 2000.

Preliminary trials showed that in this case the main factor, which influences the solution is forecasted load growth. The following load growth scenarios were used in the calculations:

- Basic 0 % annual load growth
- Slight Load Growth 2 % annual load growth
- Average Load Growth 5 % annual load growth
- Rapid Load Growth 7 % annual load growth.

For the Basic scenario both programs gave almost the same results. The suggestion for the optimal configuration in this case is presented in Figure 7-15. No need for reinforcement of lines was detected.

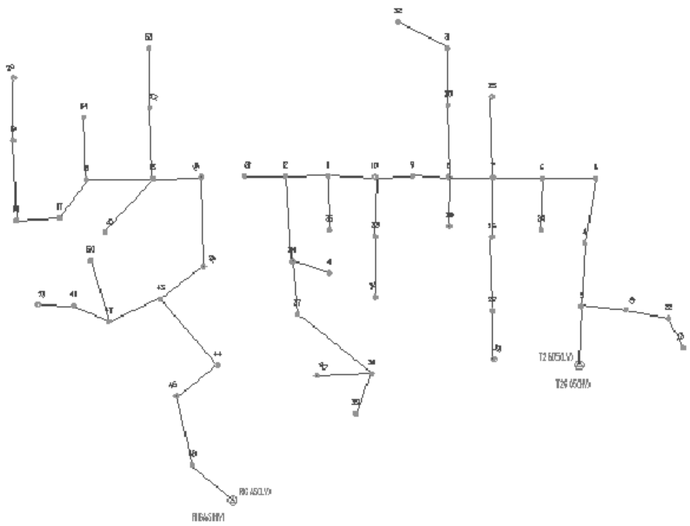


Figure 7-15 Optimal configuration for the Basic Load Growth scenario (by Swednet)

Optimal configuration for the Slight Load Growth by Swednet program looks exactly the same as for the Basic case. However there are slight differences in suggestions for conductor reinforcement. For load growth according to the scenarios Average Load Growth and Rapid Load Growth even configurations for the solutions look different (Figure 7-17, Figure 7-16).

If the uncertainty is high and all the scenarios taken into consideration are realistic, which is exactly the case in this example, then it is difficult to make a decision based only on the results presented above.

The interesting results were obtained after calculations by LDM program and analysis under uncertainty.

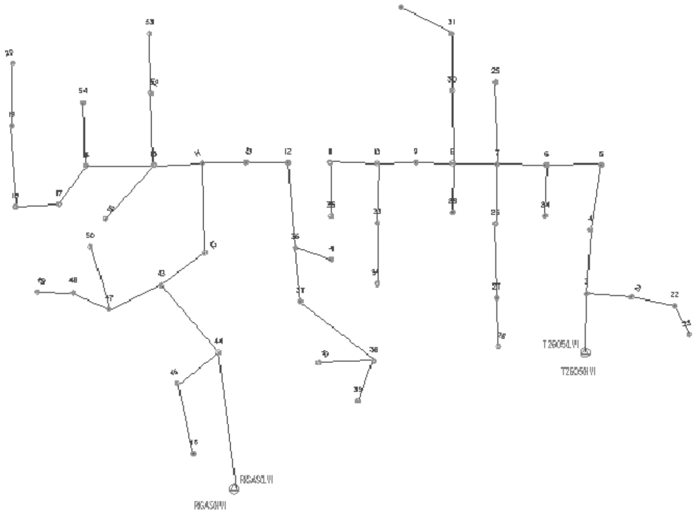


Figure 7-16 Optimal configuration for the Average Load Growth scenario (by Swednet)

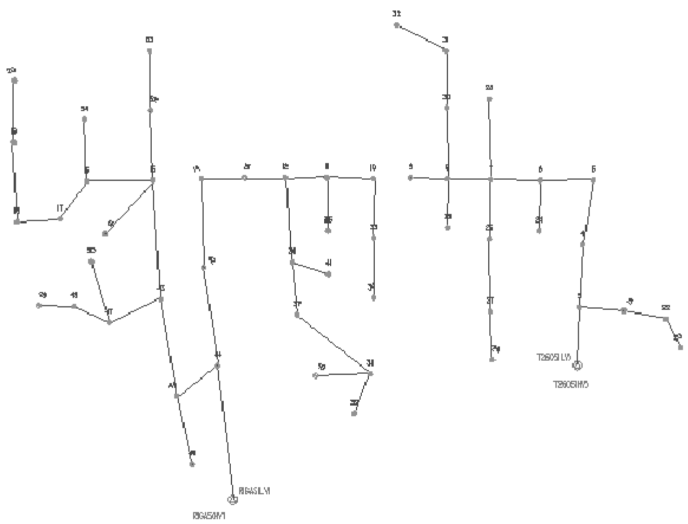


Figure 7-17 Optimal configuration for the Rapid Load Growth scenario (by Swednet)

As a result of dynamic optimization the program, except for the optimal solution, suggests several "good" alternative solutions. The optimal solution for the Basic scenario coincides with the one obtained by Swednet (alternative 7 in Table 7-11, Figure 7-15). Table 7-11 illustrates the results after analysis under uncertainty and presents the regret matrix. Zero regret means that the alternative is optimal for the particular scenario and asterisks in table cells mean that the alternative is technically unfeasible. Alternative 7 corresponds to the solution depicted in Figure 7-15, alternative 4 to Figure 7-16 and alternative 5 to Figure 7-17. This means, that for each scenario both programs detect the same optimal solutions.

Table 7-11 Matrix of regrets, 10^3sek

Scenario	Alternatives						
	1	2	3	4	5	6	7
Basic	122.9	11.4	1.6	6.4	180.1	182.4	0.0
Load Growth 2 %	122.7	11.3	1.6	6.3	179.9	182.1	0.0
Load Growth 5 %	116.2	5.0	338.5	0.0	173.4	175.7	571.8
Load Growth 7 %	45.5	63.4	***	84.9	0.0	2.3	301.1
Maximal Risk	122.9	63.4	***	84.9	180.1	182.4	571.8

However, it can be seen from the regret matrix (Table 7-11), that according to the minimal regret criterion, the decision to be made is outside the set of optimal solutions for different scenarios. The lowest risk corresponds to the alternative 2, which differs from alternative 7 mainly by additional conductor replacement actions, which are not illustrated here. The minimal risk alternative, which is not optimal for any particular scenario, however provides economic and reliable resolution no matter which scenario occurs.

Consider, however, the total costs for the respective alternatives and scenarios (Table 7-12). The table also the EMV of each alternative contains assuming 100% confidence. Optimal values are shaded. From the Table 7-12 follows that according to EMV criterion the first alternative should be chosen. However, it should be observed, that the values of the objective function for the alternatives 1,2 and 4 vary inconsiderably. Furthermore, the difference between the alternative 6 and 1 is less than 4%. Therefore, only the alternative 7 is somewhat more risky than the other alternatives.

The same task was solved by the methods presented in this dissertation. The set of Pareto optimal solutions for the Basic scenario is presented in Figure 6-7. The same figure reveals the relation between the criteria in the objective function. Clearly, in this task losses are dominating. Therefore, if all the criteria are aggregated into singe objective function, the optimal solution will correspond to the one minimizing losses.

Table 7-12 The total costs for respective alternatives and scenarios, 10^3sek

Scenarios	Alternatives						
	1	2	3	4	5	6	7
Basic	1570.2	1458.7	1448.9	1453.7	1627.4	1629.7	1447.3
Load Growth 2 %	1640.0	1528.6	1518.9	1523.6	1697.2	1699.4	1517.3
Load Growth 5 %	1713.5	1602.3	1935.8	1597.3	1770.7	1773.0	2169.1
Load Growth 7 %	1702.8	1720.7	***	1742.2	1657.3	1659.6	1958.4
EMV	1713.5	1720.7	***	1742.2	1770.7	1773.0	2169.1

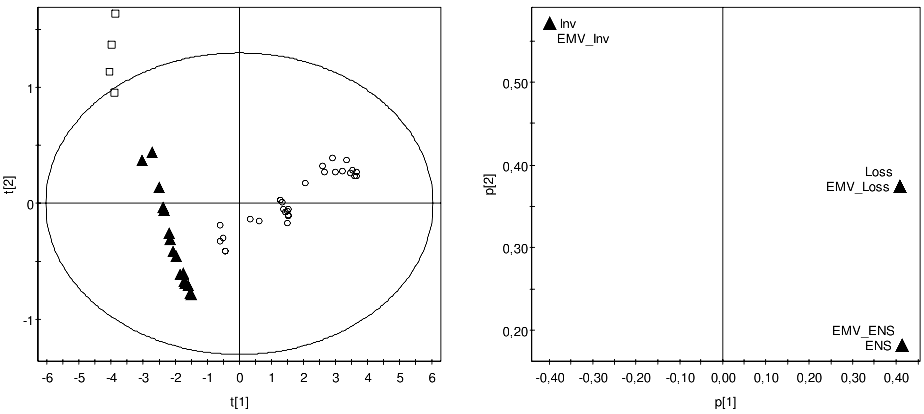


Figure 7-18 Scores plot (left) and loadings (right) for network in Jelgava

Given four scenarios with unknown probabilities the problem can be solved using simultaneous optimization of expected values of the attributes and the corresponding risks measured by EMV. The results are presented by the scores plot in Figure 7-18. The alternatives depicted by the filled triangles represent the best trade-off between all the criteria. The alternatives depicted by the empty boxed can be considered too expensive in terms of required capital investments. The alternatives 1 and 2 from Table 7-12 can be found in this group. Finally, the empty circles represent the alternatives, which perform purely in terms of losses and reliability.

7.3 Conclusions

- The analysis performed on the "Large Swedish City" model leads to the following main conclusions:
  - ✓ New substations considerably improve reliability, but result only in a minor improvement of power losses (due to new transformation losses). Construction of new substations in the area is unreasonably expensive.
  - ✓ Elimination of 110 kV oil-filled cables leads to considerable decrease in the network losses and at the same time the corresponding actions for reinforcement of existing network substantially improve reliability and personnel safety.
  - ✓ Several candidate alternatives have been suggested. The alternatives satisfy all the main objectives of the planning problem, namely: the

environmental criterion is satisfied after elimination of 110 kV oil-filled cables, the needs of operation are taken into account by reinforcing of S2, the level of personal safety is improved after reinforcement of older substations and losses and ENS are reduced to a different extend depending on the level of the investments.

- The possible influence of load uncertainty modeling on power losses was studied on the "Large City" example. Larger load variations result in increase in power losses. However, it appears that in primary distribution networks investments considerably dominate cost of losses and therefore this increase is not decisive.

- The opposite situation – when losses dominate the investments – can be observed in secondary distribution networks.

- Three different software applied for the Rural network example result in comparable solutions. However, the drawback of regret as a measure of risk is revealed – it results in comparison of only the relative quantities and disregard the absolute values.

- The results and possible solutions obtained by the software presented in this dissertation can serve as an aid in further planning studies, an expert opinion of an experienced decision-maker is needed in order to evaluate all possible consequences from each candidate alternative. Furthermore, the information about the expected budget for the projects is vital for making a final decision.

8 Closure

8.1 Conclusions

1 Latest tendencies in development of power systems – deregulation of electricity market, introduction of new technologies and increase of local generation – influence considerably the process of distribution networks planning. New methods able to facilitate the decisions resulting in reduced capital investments and power losses and improved reliability and power quality are required.

2 The problem of optimal planning of reinforcements in distribution networks is a multi-criteria and dynamic task with a large number of state and decision variables. This task must be formulated taking into consideration possible influence of random and uncertain parameters. Solution of such a problem is associated with considerable mathematical, computational and informational difficulties.

3 Analysis of the planning problem in its Bayesian formulation validates the necessity to use multi-criteria approach and to employ additional criteria characterizing risk associated with possible alternatives.

4 The development of electricity distribution networks can be described by the following principal types of factors and parameters:

- Deterministic
- Probabilistic
- Fuzzy
- Truly uncertain

The dissertation first provides the review of the methods for modeling of uncertain parameters, then suggests the model, which includes all four informational conditions. To consider the influence of all types of parameters it is necessary to use the methods of Monte-Carlo, powerful methods of optimization, scenario approach and game-theoretic decision-making criteria.

5 Simplified methods maintain their significance in network planning, for the reason that they can be used for preliminary analysis and contribute to the choice of candidate alternatives.

6 Modeling of distribution networks in order to determine power losses, reliability indices and required investments can be based on known and well-developed methods and algorithms, but providing additional program elements reflecting the influence of uncertain and random factors. In the most general case the efficiency criteria can be reflected by the numerical

parameters of the probability distribution of the expected revenues (or losses), which can be calculated by integration of the multidimensional probability distribution function.

7 Presence of strict dependence between the membership functions of fuzzy variables and probability distribution functions allows us to choose one of two possible methods to calculate the values of planning attributes based on

- Application of Fuzzy Arithmetic
- Methods of Monte-Carlo.

Utilization of Monte-Carlo methods implies additional complexity due to the conflict with most of available optimization techniques. In order to enhance the performance of the algorithms it is suggested to use the method of importance sampling, the method of common random numbers and the modification of the GA, which allows for the simultaneous search of the whole Pareto optimal set.

8 To capture the statistical behavior of the measured load it is suggested to use the empirical Pearson's charts. The statistical model reflects both the uncertainties due to external factors such as weather, and approximation in modeling – too rough time intervals. The process of load variation in time can be described by fuzzy numbers.

9 The suggested system for distribution network planning consists of the following principal stages:

- Form conditional decision set (the set of candidate alternatives) based on simplified deterministic approach.
- Perform the detailed analysis under uncertainty on the conditional decision set based on fuzzy-probabilistic approach.

The first stage comprises both "wide search" involving modification of the GA, which allows for searching simultaneously for the set of Pareto optimal solutions, and "deep search" using the novel method obtained from combination of GA with Dynamic Programming. The efficiency of the suggested algorithms is demonstrated on several examples.

At the second stage it is appropriate to use the methods of Monte-Carlo and stochastic GA, which exploits the Monte-Carlo method of common random numbers.

10 In presence of uncertainty the planner aims at finding the robust and flexible plans to reduce the risk of considerable losses. Dissertation recommends the criteria, characterizing risks of the alternative solutions,

and shows the necessity and rationale of these criteria.

11 If the number of attributes in multi-criteria optimization is more than three, it may be difficult to interpret the obtained results. In order to simplify the trade-off between the alternatives obtained as a result of optimization it is suggested to use the Principal Component Analysis (PCA).

12 Application of the suggested methodologies for reinforcement planning of the distribution network both in the large Swedish City and rural area in Latvia confirm their feasibility for practical application. The obtained results can serve as a base for the final decision-making.

8.2 Future Work

From the analysis of the research presented in this dissertation, considering the experience obtained during the assessment of the real network problems and from the interaction with power utility specialists, the following two main conclusions can be drawn:

1 Modern mathematical and computational tools provide the possibility to solve the planning problems of large distribution network in its general formulation, accounting for:
   - Uncertain and random factors
   - Multiple criteria
   - Dynamic development process.

2 Practical application of the suggested methods and algorithms would require considerable efforts for elaboration of powerful and user-friendly software providing:
   - Data gathering and processing – both from the data bases, which may be shared with other applications and in direct interaction with the planner. The last may involve the considerable amounts of information and therefore require both high qualifications of the planer and convenient software. There is a large potential for creation a new tool via integration of SCADA, DMS and Geographical Information Systems (GIS) under single framework.
   - Presentation of the results – the suggested methodology results in set of alternatives, therefore the tools for the convenient trade-off between the alternatives and their further analysis should be provided.

The recommended directions for the future work are mostly defined by the second conclusion. In real applications the mathematical tools and employed algorithms are hidden from user by the “black box” of the interface. Convenient, habitual, nice user-friendly interface in many cases may promote the choice of the one software package instead of another. The choice is made by the specialists of the power utility and depends on their experience and qualification. In this case, the experience may influence negatively, since the bias will be towards habitual interface.

The task of the planner is to provide the initial information including the set of planning options. This complicated but very important issue was practically excluded from the scope of the dissertation. The exception is Chapter 2, which concludes, that the development of new technologies increases the number of alternatives. Possible, that elaboration of expert systems able to generate automatically technically feasible alternatives would be needed. Such tools would process the databases and suggest possible options and cost associated with their realization. The databases must constantly be updated. Realization of the tool for generation of the alternatives would engage elaboration of the corresponding algorithms, possibly based on Artificial Intelligence concepts.

Collection and pre-processing of the essential initial information may also require significant efforts. The presence of SCADA provides only the possibility to obtain the large amount of data. The actual realization is needed.

Previously the conclusion was made, that the problem of distribution network planning can be solved in most general form without avoidable simplifications. It does not mean, that more powerful and efficient algorithms or more accurate models cannot be found. On the contrary, the author of this dissertation hopes for it and wishes good luck to all the researchers in this field.

Bibliography

[1] R.N. Adams, M.A. Laughton, “A Dynamic Programming/Network Flow Procedure for Distribution System Planning”, *Proceedings of the IEEE Power Industry Conference (PICA)*, 1973.

[2] R. Allan, “Predictive Assesment and Reliability Worth”, *IEE North Eastern Centre Power Section Symposium on the Reliability*, Security and Power Quality of Distribution Systems, 1992, pp. 3/1-3/19.

[3] R.N. Allan, M.G. Da Silva, “Evaluation of Reliability Indices and Outage Costs in Distribution Systems”, *IEEE Transactions on Power Systems*, Vol. 10 1, February 1995, pp.413-419.

[4] C.J. Andrews, “Evauating Risk Management Strategies in Resource Planning”, *IEEE PES Summer Meeting*, San Francisco, July 1994.

[5] G. Anders, R. Entriken, P. Nitu, “Risk Assessment and Financial Management”, *Tutorial at IEEE PES Winter Meeting*, 1999.

[6] D.A. Arzamastev, A.V. Lipes, A.L. Mizin, *Models of Optimization of Power System Development*, Vissaja Skola, 1987 (in Russian).

[7] B. M. Ayyub, M.M. Gupta, *Uncertainty Analysis in Engineering and Sciences: Fuzzy Logic, Statistics and Neural Network Approach*, Kluwer Academic Publishers, 1997.

[8] Y. Backlund, *Hierarhical Modelling of Regional Electric Systems*, Thesis for the degree of Doctor of Technology at the Royal Institute of Technology, Stockholm, 1978.

[9] M. E. Baran, F. F. Wu, “Optimal Sizing of Capacitors Placed on a Radial Distribution System”, *IEEE Transactions on Power Delivery*, Vol.4, No. 1, January 1989, pp. 735-742.

[10] R. Billinton, R.N. Allan, *Reliability Evaluation of Power Systems*, Pitman Publishing Inc., 1984.

[11] R. Billinton, A. Sankarakrishan, “A Comparison of Monte-Carlo Simulation Techniques for Composite Power System Reliability Assessment”, *IEEE WESCANEX’95 Proceedings*, 1995, pp.145-150.

[12] M. Blanchard, L. Delorme, C. Simard, Y. Nadeau, “Experience with Optimization Software for Distribution System Planning”, *IEEE/PES Winter Meeting*, 1996.

[13] V. Bochkareva, *Development Planning of Electricity Distribution Networks*, Licentiate Thesis at the Royal Institute of Technology, ISSN 1100-1607, Stockholm, Sweden, 1997.

[14] G. Boone, H.-D. Chiang, “Optimal Capacitor placement in Distribution Systems by Genetic Algorithm”, *Electrical Power&Energy System*, Vol.15, No.3, 1993.

[15] J.P. Brans, B. Mareschal, P.Vincke, “PROMETHEE: a New Famili of Outranking methods in Multicriteria Analysis”, *Operational Research*, Elsevier, Amsterdam, 1984, pp.477-490.

[16] D.L. Brooks, R.C. Dugan, M. Waclawiak, A. Sundaram, “indices for Assessing Utility Distribution System RMS Variation Performance”, IEEE Transactions on Power Delivery, Vol.13, No.1, January 1998, pp.254-259.

[17] R.E. Brown, S. Gupta, R.D. Christie, S.S. Venkata, R. Fletcher, "Automated Primary Distribution System Design: Reliability and Cost Optimization", *IEEE Transactions on Power Delivery*, Vol.12, No.2, April 1997, pp.1017-1022.

[18] W.J. Burke, H.M. Merrill, F.C. Schweppe, et al, "Trade off Methods in System Planning", *IEEE Transactions on Power Systems*, Vol.3, No.3, August 1988.

[19] N.P. Buslenko, *Monte-Carlo Methods*, Fizmatgiz, 1962 (in Russian).

[20] P.M.S. Carvalho, L.A.F.M. Ferreira, "Planning Large-Scale Distribution Networks for Robust Expansion under Deregulation", *IEEE/PES Summer Meeting*, 2000.

[21] P.M.S. Carvalho, L.A.F.M. Ferreira, L.M.F. Barruncho, "Hedging Large-Scale Distribution System Investments Against Uncertainty", *IEEE/PES Winter Meeting*, 1999, Vol. 2, pp. 901-906.

[22] K.H. Cheong, P. Forsgren, "In Search of an Implementation Strategy for Distribution Management Functionality*", Proceedings of the DA/DSM Europe 96 Conference*, Vienna, Austria, October 8-10 1996.

[23] P.-H. Chen, H.-C. Chang, "Large-Scale Economic Dispatch by Genetic Algorithm", *IEEE Transactions on Power Systems*, Vol.10, No. 4, November 1995 , pp. 1919 – 1926.

[24] R.L. Chen, K. Allen, R. Billington, "Value-Based Distribution Assessment and Planning", *IEEE Transactions on Power Delivery*, Vol.10, No.1, January 1995, pp. 421-429.

[25] T.H. Chen, M.S. Chen, "Distribution System Power Flow Analysis – A Rigid Approach", *IEEE Transactions on Power Delivery*, Vol.6, No.3, July 1991, pp. 1146-1152.

[26] R. Cherkaoui, A. Bart, A.J. Germond, "Optimal Configuration of Electrical Distribution Networks Using Heuristic Methods", *11th Power Systems Computation Conference (PSCC)*, Avignon, Switzerland, 1993.

[27] H.-D. Chiang, J.-C. Wang, J. Tong, G. Darling, "Optimal Capacitor Placement, Replacement and Control in Large-Scale Unbalanced Distribution Systems: Modeling and a New Formulation", *IEEE Transactions on Power Systems,* Vol. 10, No.1, February 1995, pp. 356 –362.

[28] H.-D. Chiang, J.-C. Wang, J. Tong, G. Darling, "Optimal Capacitor Placement, Replacement and Control in Large-Scale Unbalanced Distribution Systems: System Solution Algorithms and Numerical Studies", *IEEE Transactions on Power Systems,* Vol. 10, No.1, February 1995, pp. 363 –369.

[29] CIGRE Working Group 37.10, "Methods for Planning under Uncertainty: Towards Flexibility in Power System Development", *Electra,* No.161, August 1995, pp. 143-163.

[30] R.Cicoria, G.L. Fracassi, G.Gambelli, M.Mazzoni, W.Palenzona, E.Ricci, "Load Flow Calculations on Distribution Networks by Using a Statistical Approach", *CIRED'97,* 2-5 June 1997, pp. 6.13.1-6.13.5.

[31] C.A.C. Coello, "An Updated Survey of Evolutionary Multiobjective Optimization Techniques: State of the Art and the Future Trends", *Proceedings of the 1999 Congress on Evolutionary Computation*, Vol.1, 1999.

[32] M.B.D. Coutto Filho, A.M. Leite Da Silva, V.L. Arienti, S.M.P. Ribeiro, "Probabilistic

Load Modelling for Power System Expansion Planning", *Third International Conference on Probabilistic Methods Applied to Electric Power Systems*, 1991, pp. 203 –207.

[33] E.O. Crousillat, P. Dörfner, P. Alvarado, H.M. Merrill, " Conflicting Objectives and Risk in Power System Planning", IEEE Transactions on Power Systems, Vol.8, No.3, August 1993, pp.887-893.

[34] V. Dale, Z. Krishan, O. Paegle, *Dynamic Optimization of Power Networks Development*, Zinatne, 1990 (in Russian).

[35] K.P. Dahal, G.M. Burt, J.R. McDonald, S.J. Galloway, "GA/SA-based Hybrid Techniques for the Scheduling of Generator Maintenance in Power Systems", *Proceedings of the 2000 Congress on Evolutionary Computation*, Vol.1, 2000, pp.567-574.

[36] M.H. DeGroot, *Optimal Statistical Decisions*, McGraw-Hill Series in Probability and Statistics, 1970.

[37] T. De la Torre, J.W. Feltes, T.S.R. Gomez, H.M. Merril, "Deregulation, Privatization and Competition: Transmission Planning Under Uncertainty", IEEE Transactions on Power Systems, Vol.14, No.2, May 1999, pp.460-465.

[38] D. Dubois, H. Prade, "Fuzzy Sets and Probability: Misunderstandings, Bridges and Gaps*", Second IEEE International Conference on Fuzzy Systems*, vol.2, pp. 1059 – 1068, 1993.

[39] D. Dubois, H. Prade, S. Sandri, "On Possibility/Probability Transformations", *Fuzzy Logic, Kluwer Academic Publishers*, pp.103-112, 1993.

[40] R.C. Dugan, D.L. Brooks, T.E. McDermott, A. Sundaram, "Using Voltage Sag and Interruption Indices in Distribution Planning", IEEE PES 1999 Winter Meeting, Vol.2, pp.1164-1169.

[41] EPRI TR-107938, *Reliability Benchmarking Methodology*, May 1997.

[42] A.G. Esposito, E.R. Ramos, "Reliable Load Flow Technique for Radial Distribution Networks", *IEEE Transactions on Power Systems*, Vol.14, No.3, August 1999, pp.1063-1069.

[43] G.W. Evans, B. Stuckman, M. Mollaghasemi, "Multicriteria Optimization of Simulation Models*", Proceedings of the 1991 Winter Simulation Conference*, 1991, pp.894-900.

[44] T.H. Fawzi, K.F. Ali, S.M. El-Sobki, "A New Planning Model for Distribution Systems", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-102, No.9, September 1983.

[45] P.Finlay, "Steps Towards Scenario Planning", *Engineering Management Journal*, October 1998, pp.243-246.

[46] C.M. Fonseca, P.J. Fleming, "Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization", *Proceedings of the Fifth International Conference on Genetic Algorithms*, pp.416-423, San Mateo, California, 1993.

[47] F. Glover, "Future paths for Integer Programming and Links to Artificial Intelligence", *Computer and Operations Research*, Vol. 13, No.5, 1986, pp. 533-549.

[48] D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*,

Addison-Wesley, 1989.

[49] T. Gomez, J. Rivier, "Distribution and Power Quality Regulation under Electricity Competition. A Comparative Study", *Proceedings. Ninth International Conference on Harmonics and Quality of Power*, 2000, Vol.2, pp. 462-468.

[50] B. Gorenstin, N. Campodonico, J. Costa, M. Pereira, "Power System Expansion Planning Under Uncertainty", *IEEE Transactions on Power Systems*, PWRS-8(1), February 1993, pp.129-136.

[51] J. Grönqvist, "Status of the SICEPS Project*", Final Report, Royal Institute of Technology*, Stockholm, 1997.

[52] H. Grønly, "Customer Consequences from Separating Electricity Distribution and Supply - Nordic Experience", *Proceedings of the DA/DSM DistribuTECH 97 Conference*, Amsterdam, The Netherlands, October 14-16, 1997.

[53] T. Gönen, I.J. Ramirez-Rosado, "Optimal Multi-Stage Planning of Power Distribution Systems", *IEEE Transactions on Power Delivery*, Vol. PWRD-2, No.2, April 1987, pp. 512-519.

[54] G.J. Hahn, S.S. Shapiro, *Statistical Models in Engineering*, John Wiley&Sons, Inc., 1967.

[55] M. Hendriks, J.H. de Boer, A.K. Smilde, D.A. Doornbos, "Multicriteria decision making", *Chemometrics and Intelligent Laboratory Systems*, No.16, 1992, pp.175-191.

[56] B.F. Hobbs, P.M. Meier, "Multicriteria Methods for Resource Planning: An Experimental Comparison", IEEE Transactions on Power Systems, Vol.9, No.4, November 1994, pp.1811-1817.

[57] J. Horn, N. Nafpliotis, D.E. Goldberg, "A Nished Pareto Genetic Algorithm for Multiobjective Optimization", Proceedings of the First IEEE Conference on Evolutionary Computation, 1994, pp. 82-87.

[58] Y.Y. Hsu, Y. Jwo-Hwu, "Planning of Distribution Substations, Feeders and Sectionalizing Switches Using Heuristic Algorithms", *Electrical Power & Energy Systems*, Vol.18, No.5, 1996, pp. 315-322.

[59] L. Jacobsson, "Impact of IT and DA for Utility Efficiency", *Proceedings of the DA/DSM Europe 96 Conference*, Vienna, Austria, October 8-10 1996.

[60] Y.-J. Jeon, J.-C. Kim, "An Efficient Algorithm for Network Reconfiguration in Large Scale Distribution Systems", *IEEE/PES Summer Meeting*, 1999, Vol.1, pp.243-247.

[61] S. Jonnavithula, R. Billinton, "Minimum Cost Analysis of Feeder Routing in Distribution System Planning", *IEEE/PES Winter Meeting*, 1996.

[62] N. Kagan, R.N. Adams, "Application of Benders Decomposition Technique to the Distribution Planning Problem", *Tenth Power System Computation Conference (PSCC)*, Graz, Austria, 19-24 August 1990.

[63] K. Kauhaniemi, *Long-Term Investment Planning of Electricity Distribution Networks Considering the Uncertainty of Load Forecasts*, Tampere University of Technology, Thesis for the degree of Doctor of Technology, 1993.

[64] K. Kauhaniemi, *Genetic Algorithms in Power Systems Computation*, Tampere University of Technology, Report 4-95, 1995.

[65] K. Kauhaniemi, "A Probabilistic Approach to the Long-Term Planning of Public

Electricity Distribution Networks", *Third International Conference on Probabilistic Methods Applied to Power Systems*, pp.311-316.

[66] R.L. Keeney, *Value-Focused Thinking: A Path to Creative Decision-Making*, Harvard University Press, 1992.

[67] R.L. Keeny, H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*, John Wiley & Sons, 1976.

[68] K.-H. Kim, S.-K. You, "Voltage Profile Improvement by Capacitor Placement and Control in Unbalanced Distribution Systems Using GA", *IEEE/PES Summer Meeting*, 1999, Vol. 2, pp. 800-805.

[69] Y.-C. Kim, B.-H. Ahn, "Multi-Criteria Generation-Expansion Planning with Global Environmental Considerations", *IEEE Transactions on Engineering Management*, Vol.40, No.2, May 1993, pp.154-161.

[70] Z. Krishans, V. Bochkareva, G. Andersson, "Dynamic Model for Reinforcement Planning of Distribution System under Uncertainty", *Universities Power Engineering Conference (UPEC'97),* Manchester, UK, September 1997, pp.1089-1092.

[71] Z. Krishans, V. Neimane, G. Andersson, "Dynamic model for planning of reinforcement investments in distribution networks", *Power Systems Computation Conference (PSCC'99),* Trondheim, Norway, June 1999.

[72] H. Kuwabara, K. Nara, "Multi-Year and Multi-State Distribution System Expansion Planning by Multi-Stage Branch Exchange", *IEEE/PES Summer Meeting*, 1996.

[73] E. Lakervi, E.J. Holmes, *Electricity Distribution Network Design*, Peter Peregrinus Ltd., 1989.

[74] V. Lasserre, G. Mauris, Y. Foulloy, "Fuzzy Modelling of Measurements Acquired by an Intelligent Ultrasonic Telemeter", *IEEE Instrumentation and Measurement Technology Conference, IMTC/98 Conference Proceedings*, vol. 2 , pp. 837 –842, 1998.

[75] T. Laine, M. Lehtonen, E. Antila, M. Seppänen, "Feasibility Study of DA in Rural Distribution Company", *Proceedings of the DA/DSM Europe 96 Conference*, Vienna, Austria, October 8-10 1996.

[76] A.M. Leite da Silva, S.M.P. Ribeiro, V.L. Arienti, R.N. Allan, M.B. Do Coutto Filho, "Probabilistic Load Flow Techniques Applied to Power System Expansion Planning", *IEEE Transactions on Power Systems*, Vol.5, No.4, November 1990, pp.1047-1053.

[77] P. L'Ecuyer, "Efficiency Improvement and Variance Reduction", Proceedings of the 1994 Winter Simulation Conference, pp.122-132.

[78] F.S. Ma, L. Isaksen, "Impact of Dispersed Storage and Generation on Distribution Planning and Operation Practices", IEEE/PES Transmission and Distribution Conference and Exposition, April 1979, pp.29-36.

[79] E. Masud, "An Interactive Procedure for Sizing and Timing Distribution Substations Using Optimization Techniques", *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-93, No.5, September 1974.

[80] Materials of the EES-UETP Course "*Electricity Distribution in Open Market Conditions*", Helsinki, Finland, November 19-22, 1997.

[81] H.M. Merrill, F.C. Schweppe, "Strategic Planning for Electric Utilities: Problems and

Analytic Methods", *Interfaces*, Vol.14, No.1, January&February 1984, pp.72-83.

[82] V. Miranda, L.M. Proenca, "Genetic Algorithms and Fuzzy Models - An Application to Gas and Electricity Distribution Planning Under Uncertainty", *RSSC'94*, San Jose, California, USA, November 10-12, 1994, pp. 43-50.

[83] V. Miranda, L.M. Proenca, "Probabilistic Choice vs. Risk Analysis - Conflicts and Synthesis in Power System Planning", *20th International Conference on Power Industry Computer Applications (PICA)*, Columbus, Ohio, May 11 - 16, 1997.

[84] V. Miranda, L.M. Proenca, L. Oliveira, L. Carvalho, "Dynamic Planning of Distribution Networks for Minimum Regret Strategies", *13th International Conference on Electricity Distribution* (*CIRED)*, Belgium, Liege, 1995.

[85] V. Miranda, L.M. Proenca, "A general Methodology for Distribution Planning under Uncertainty, Including Genetic Algorithms and Fuzzy Models in a Muti-Criteria Environment", *IEEE/KTH Stockholm Power Tech Conference*, Stockholm, Sweden, June 18-22, 1995, pp. 832-837.

[86] V. Miranda, J.V. Ranito, L.M. Proenca, "Genetic Algorithms in Optimal Multistage Distribution Network Planning", *IEEE Transactions on Power Systems*, Vol.9, No.4, November 1994, pp. 1927-1933.

[87] V. Miranda, D. Srinivasan, L.M. Proenca, "Evolutioary Computation in Power Systems", *Power Systems Computation Conference (PSCC'96)*, Dresden, 1996.

[88] M. Mollaghasemi, G. Evans, "Multicriteria Design of Manufacturing Systems Through Simulation Optimization", IEEE Transactions on Systems, Man and Cybernetics, Vol.24, No.9, September 1994, pp.1407-1411.

[89] I.J. Nagrath, D.P. Kothary, *Modern Power System Analysis*, New Dehli, 1983.

[90] K. Nara, "Genetic Algorithm for Power Systems Planning", *4th International Conference on Advances in Power System Control, Operation and Management (APSCOM-97)*, Hong Kong, November 1997, pp.60-65.

[91] K. Nara, "State of the Art of the Modern Heuristics Application in Power Systems", *IEEE/PES Winter Meeting,* 2000.

[92] K. Nara, J. Hasegawa, "Configuration of New Power Delivery System for Reliable Power Supply", *IEEE/PES Summer Meeting*, 1999, Vol. 1, pp. 248-253.

[93] K. Nara, A. Shiose, M. Kitagawa, T. Ishihara, "Implementation of Genetic Algorithm for Distribution Systems Loss Minimum Reconfiguration", *IEEE Transactions on Power Systems*, Vol.7, No.3, August 1992, pp. 1044-1051.

[94] V. Neimane, G. Andersson, Z. Krishans, "Multi-Criteria Analysis Application in Distribution Network Planning", *33rd Universities Power Engineering Conference (UPEC'98),* Edinburgh, UK, September 1998, pp.791-794.

[95] V. Neimane, G. Andersson, "Distribution Networks Reinforcement Planning: A Dynamic Multi-Criteria Approach", *Budapest Power Tech'99*, Budapest, Hungary, August 1999.

[96] V. Neimane, Z. Krishans, "Multi-criteria analysis in electric power network planning", *Latvian Journal of Physiscs and Technical Sciences*, 1997.

[97] J. Neumann, "Various Techniques Used in Connection with Random Digits", *NBS Applied Mathematics Series*, 1951, No.12, pp.36-38.

[98] V. Nissen, *Evolutionary Algorithms in Management Science, An Overview and List of References*, Univetsity of Goettingen, Germany, 1993.

[99] V. Nissen, J. Propach, “On the Robustness of Population-Based Versus Point-Based Optimization in the Presence of Noise”, *IEEE Transactions on Evolutionary Computation*, Vol.2, No.3, September 1998, pp.107-119.

[100] V. Pareto, *Cours D’Ecomomie Politique*, Vol.I and II. F.Rouge, Lausanne, 1896.

[101] J.-B. Park, Y.-M. Park, J.-R. Won, K.Y. Lee, “Least Cost Generation Expansion Planning Based on an Improved Genetic Algorithm”, *IEEE/PES Summer Meeting*, 1999, Vol. 2, pp. 1043-1047.

[102] J.-B. Park, Y.-M. Park, J.-R. Won, “A Hybrid Genetic Algorithm/Dynamic Programming Approach to Optimal Long-Term Generation Expansion Planning”, *Electrical Power & Energy Systems*, Vol.20, No.4, pp.295-303, 1998.

[103] J. Partanen, *A PC-based Information and Design System For Electricity Distribution Networks*, Tampere University of Technology, Thesis for the degree of Doctor of Technology, 1991.

[104] 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, pp. 277-283.

[105] J. Peco, E.F. Sanchez-Ubeda, T. Gomez, “*Enhansing Optimal Transmission or Subtransmission Planning by Using Decision Trees*”, Proceedings of Budapest Power Tech, Budapest, September 1999.

[106] M.V.F. Pereira, M.F. McCoy, H.M. Merrill, “Managing Risk in the New Power Business”, *IEEE Computer Application in Power*, April 2000, pp.18-24.

[107] M. Ponnavaikko, K.S. Prakasa-Rao, S.S. Venkata, “Distribution System Planning trough a Quadratic Mixed-Integer Programming Approach”, *IEEE Transactions on Power Delivery*, Vol. PWRD-2, No.4, October 1987.

[108] M.T. Ponce de Leao, J.T. Saraiva, “A Simulated Annealing Approach to Evaluate Long Term Marginal Costs and Investment Decisions”, *IEEE/PES Summer Meeting,* 2000.

[109] D.S. Popovic, R.M. Ciric, “A Multi-Objective Algorithm for Distribution Networks Restoration*”, IEEE Transactions on Power Delivery*, Vol.14 3, July 1999, pp. 1134-1141.

[110] J. W. Pratt, “Risk Aversion in the Small and in the Large”, *Econometrica* 32, 1964, pp.122-136.

[111] L.M. Proenca, V. Bochkareva, V. Miranda, “Towards a Comprehensive Methodology for Distribution System Planning”, ELAB’96, Porto, 1996.

[112] L.M. Proenca, *Towards a Comprehensive Methodology for Power System Planning*, Doctoral dissertation at the University of Porto, Portugal, November 1997.

[113] I.J. Ramirez-Rosado, J.L. Bernal-Agustin, “Genetic Algorithms Applied to the Design of Large Power Distribution Systems”, *IEEE Transactions on Power Systems*, Vol.13, No.2, May 1998, pp. 696-703.

[114] I.J Ramirez-Rosado, J.A. Dominguez-Navarro, J.M. Yusta-Loyo, “A New Model for Optimal Electricity Distribution Planning Based on Fuzzy Set Techniques”, *IEEE/PES Summer Meeting*, 1999, Vol. 2, pp. 1048-1054.

[115] I.J. Ramirez-Rosado, T. Gönen, "Pseudodynamic Planning For Expansion of Power Distribution Systems", *IEEE Transactions on Power Systems*, Vol.6, No. 1, February 1991, pp. 245-253.

[116] S. Rana, D. Whitley, R. Cogswell, "Searching in the Presence of Noise", *Proceedings of the 4th International Conference on Parallel Problem Solving from Nature,* H.-M. Voigt, W. Ebeling, I. Rechenberg, H.-P. Schwefel, Eds. Berlin, Germany, Springer, 1996, pp. 198-207.

[117] J. Roman, T. Gomez, A. Munoz, J. Peco, "Regulation of Distribution Network Business", *IEEE Transactions on Power Delivery*, Vol. 14 2, April 1999, pp. 662-669.

[118] J.J.Saade, "A Unifying Approach to Defuzzification and Comparison of the Outputs of Fuzzy Controllers", *IEEE Transactions on Fuzzy Systems*, Vol.4, No.3, August 1996, pp.227-237.

[119] J.J. Saade, H. Schwarzlander, "Ordering Fuzzy Sets Over the Real Line: an Approach Based on Decision-Making Under Uncertainty", *Fuzzy Sets Systems*, Vol.50, pp.237-246, September 1992.

[120] T.L. Saaty, *Decision Making: The Analytic Hierarchy Process. Planning, Priority Setting, Resource Allocation*, RWS Publications, Pittsburgh, PA, 1990.

[121] J.D.Shaffer, "Multiple Objective Optimization with Vector Evaluated Genetic Algorithms", *Genetic Algorithms and their Applications: Proceedings of the First International Conference on Genetic Algorithms*, pp.93-100, Pittsburgh, PA, 1985.

[122] A.L. Shenkman, "Energy Loss Computation by Using Statistical Techniques", *IEEE Transactions on Power Delivery*, Vol.5, No.1, January 1990, pp.254-258.

[123] D. Shirmohammadi, H.W. Hong, A. Semyen, G. X. Luo, "A Compensation Based Power Flow Method for Weakly Meshed Distribution and Transmission Networks", *IEEE Transactions on Power Systems*, Vol.3, No.1, May 1988, pp.753-762.

[124] M.S. Srinivas, "Distribution Load Flow: A Brief Review", IEES PES *Winter Meeting* 2000.

[125] L. Söder, *Vindkraftens Effektvärde*, Final Rapport, Royal Institute of Technology, Stockholm, 1997.

[126] *The Swedish Electricity Market Reform and its Implementation for Svenska Kraftnät*, Rapport by Svenska Kraftnät, 1997.

[127] Y. Tang, "Power Distribution System Planning with Reliability Modeling and Optimization", *IEEE Transactions on Power Systems*, Vol.11, No.1, February 1996, pp. 181-187.

[128] A. Thommesen, "Regulatory Framework and Network Tariffs to Promote Market Competition Achieve Efficient Operation and Provide Return of New Investment Costs", *IEEE PES Winter Meeting*, 2000, Vol.2, pp.831-834.

[129] *Utvecklingen på Elmarknaden*, Rapport by NUTEK, Oktober, 1996.

[130] E. Van Geert, "Increased Uncertainty a New Challenge for Power System Planners", *IEE Colloquium on Tools and Techniques for Dealing With Uncertainty* (Digest No.1998/200), 1998, pp.1831-1837.

[131] S. Velander, *Energidistributionens economiska lagar, Stockholn,* 1947 (in Swedish).

[132] G. Wacker, E. Wojczynski, R. Billinton, "Comprehensive Bibliography of Electical

Service Interruption Costs”, IEEE Transactions on Power Apparatus, 1983, PAS-102(6), pp.1831-1837.

[133] D.L. Wall, G.L. Thompson, J.E.D. Northcote-Green, “An Optimization Model for Planning Radial Distribution Networks”, *IEEE Transactions on Power Apparatus and Systems*, Vol.PAS-98, No.3, May/June 1979.

[134] M. Wall, *GAlib: A C++ Library of Genetic Algorithm Components*, Software documentation, Massachusetts Institute of Technology, http://lancet.mit.edu/ga/.

[135] C. Walton, “Development and Implementation of an Integrated Technology Strategy*”, Proceedings of the DA/DSM DistribuTECH 97 Conference*, Amsterdam, The Netherlands, October 14-16, 1997.

[136] H.L. Willis, G.S. Scott, “Advanced Engineering Methods for Optimizing and Integrating Distribution Planning with Demand-Side Management and Dispersed Generation”, *38th Annual Rural Electric Power Conference*, 1994, pp. A1/1 -A1/9.

[137] H.L. Willis, H. Tram, V. Engel, L. Finley, “Selecting and Applying Distribution Optimization Methods”, *Computer Applications in Power*, Volume 9, January 1996, pp. 12-17.

[138] H.L. Willis, *Power Distribution Planning Reference book*, Marcel Dekker, Inc., 1997.

[139] H.-T. Yang, C.-M. Huang, C.-L. Huang, “Identification of ARMAX Model for Short Term Load Forecasting: an Evolutionary Programming Approach”, *IEEE Transactions on Power Systems*, Vol. 11, No. 1, February 1996, pp. 403-408.

[140] H.-T. Yang, P.-C. Yang, C.-L. Huang, “A Parallel Genetic Algorithm Approach to Solving the Unit Commitment Problem: Implementation on the Transputer Networks”, *IEEE Transactions on Power Systems*, Vol. 12, No. 2, May 1997, pp. 661-668.

[141] J. Yen, J.C. Liao, L. Bogju, D. Randolph, “A hybrid Approach to Modelling Metabolic Systems Using a Genetic Algorithm and Simplex Method”, *IEEE Transactions on Systems, Man and Cybernetics*, Part B, Vol.28 2, April 1998, pp.173-191.

[142] E.-C. Yeh, S.S. Venkata, Z. Sumic, “Improved Distribution System Planning Using Computational Evolution”, *IEEE Transactions on Power Systems*, Vol.11, No.2, May 1996, pp.668-674.

[143] L.A. Zadeh, “Fuzzy Sets”, *Information and Control*, Vol.8, pp.338-353, 1965.

[144] M. Zeleny, “A Concept of Compromise Solutions and the Method of Displaced Ideal”, *Computational Operation Research*, Vol.1, 1974, pp.479-496.

Appendix A Summary of the Network Data Used in the "Large Swedish City" Project

A.1 Nodes

The nodes are described first by their type - load point of substation. The most important information about each load point is the corresponding active peak load. The load growth during the planning period is also taken into account.

Table A- 1 Node data

Number	Node	Pmax, MW	Load growth during the planning period, MW							
			2004	2007	2010	2013	2016	2019	2022	2025
1	S1	0								
2	S2	0								
3	S3	0								
4	S4	0								
5	S5	0								
6	L1	36.7	3	4	5	5.5	5.5	6	6	6
7	L2	52.1								
8	L3	20.4								
9	L3'	20.4								
10	L4	15.0	2	2.5	3	3.5	4	4.5	5	5
11	L4'	15.0	2	2.5	3	3.5	4	4.5	5	5
12	L5	28.7	3	4	5	6	6	6	6	6
13	L6	37.8								
14	L7	43.3								
15	L8	37.7	2	2	2	2	2	2	2	2
16	L9	52.0								
17	L10	17.5	6	12	12	12	12	12	12	12
18	S18	0								
19	S19	0								
20	S20	0								
21	L11	39.4	2	3	3	4	4	4	4	4
22	L12	6.5								
23	load of S1	18.8								
24	load of S2	4.0								

A.2 Cable lines

Each line is described by two nodes it connects, the length, type of the cable and the nominal voltage. The corresponding parameters for each cable type are given in section 0. Status “one” describes the existing line, while status “zero” describes a possible allocation for a new cable.

Table A- 2 Cable line data

Nr	From	To	Length, km	Cable	Status	Voltage, kV
1	1	4	7.16	1	1	110
2	2	5	3.4	2	1	110
3	5	4	3.5	2	1	110
4	2	14	4.1	3	1	33
5	2	15	3.3	3	1	33
6	5	10	3	3	1	33
7	5	8	4	5	1	33
8	5	12	3.3	6	1	33
9	1	7	1.9	3	1	33
10	1	6	1.3	3	1	33
11	4	13	2.4	3	1	33
12	3	17	3.9	5	1	33
13	3	16	5	3	1	33
14	10	11	0.1	6	1	33
15	8	9	0.1	5	1	33
16	18	12	1.2	7	0	33
17	18	4	3.3	7	0	33
18	18	13	4.3	7	0	33
19	2	8	3.8	7	0	33
20	2	9	4.4	7	0	33
21	19	9	2.5	7	0	33
22	19	5	1	7	0	33
23	19	11	3.5	7	0	33
24	3	10	1.9	7	0	33
25	3	13	3.6	7	0	33
26	3	4	3.9	7	0	33
27	1	12	4.4	7	0	33
28	17	10	1.9	7	0	33
29	12	4	2.6	7	0	33
30	1	4	7.16	8	0	110
31	2	22	3.4	7	0	33
32	4	21	0.1	3	1	33
33	5	22	0.1	3	1	33
34	1	23	0.1	3	1	33
35	2	24	0.1	3	1	33
36	20	9	3.5	7	0	33
37	20	22	2.1	7	0	33
38	20	12	1.2	7	0	33
39	20	11	4.5	7	0	33
40	1	4	7.16	9	0	33

A.3 Parameters of power cables

The essential information about the cables (or about the overhead lines in other applications) includes their resistances and reactances, as well as reliability indices – failure rate and duration. All the parameters are calculated for a number of cables in parallel (two or three). For existing cables cost is assumed to be zero. For new cables the given cost includes both cost for the cable itself and excavation and labor costs.

Table A- 3 Cable (overhead line) parameters

Nr	Cable[1]	R_km, Ω/km per phase, 20°C	X_km, Ω/km per phase, 20°C	Cost, KSEK/km	Failure rate fail/h/km	Failure duration, hours	Voltage level, kV
1	Al-3x1x500	0.0209	0.0408	0	$4.83e^{-10}$	84	110
2	Cu-3x300	0.0208	0.0337	0	$4.83e^{-10}$	84	110
3	Al-3x1x400	0.0266	0.0359	0	$2.33e^{-14}$	16	33
4	Al-3x1x300	0.0338	0.0375	0	$2.33e^{-14}$	16	33
5	Cu-3x150	0.0417	0.0380	0	$2.33e^{-14}$	16	33
6	Cu-3x300	0.0211	0.0268	0	$2.33e^{-14}$	16	33
7	Al-3x1x500	0.0202	0.0333	4350	$2.33e^{-14}$	16	33
8	Al-3x1x500	0.0202	0.0333	4800	$1.46e^{-13}$	56	110
9	Al-3x1x500	0.0202	0.0333	3450	$2.33e^{-14}$	16	33

A.4 Substations

Table A- 4 Substation data

Sub-station	Trans-formation stage, kV	Number of trans-formers	P_{rated}, MVA	Load losses, kW	No load losses, kW	Cost, KSEK
S1[2]	220/33 220/110	3 2	50 75	150*3 260*2	68*3	0
S2	220/33 220/110	3	50 75	150*3 290*3	75*3	0
S3	220/33	2	120	400*2	40.5*2	0
S4	110/33	3	60	234*3	22*3	0
S5	110/33	3	60	234*3	22*3	0
S18	220/33	3	110	400*3	40*3	74000
S19	220/33	3	110	400*3	40*3	74000
S20	220/33	3	110	400*3	40*3	74000

[1] The difference in cable parameters between Cable 1 and Cable 8 can be explained by the different number of cables in parallel – two for Cable 1 and three for Cable 8.
[2] Totally three 220/110/33 kV transformers. Only two windings 220/110 kV are in use.

The most important substation information includes the data about the transformers. It is essential to identify the knowledge about transformation stages, number of transformers, their nominal powers, as well as load and no-load losses. Similarly to the cables, we associate the cost only with new substations planned to be build. The cost of an existing substation is assumed to be zero. Furthermore, the cost presented here includes the cost of the whole substation including transformers, switchgears and other equipment.

A.5 Additional reliability data

In addition to the cable data the reliability indices for some of the substation equipment are needed. The indices used in this study are summarized in Table A-5.

Table A- 5 Reliability data

Type of component	Voltage, kV	Failure rate, fail/hour	Failure duration, hours
Breaker	220	$3.92e^{-6}$	168
Breaker	110	$2.62e^{-6}$	168
Breaker	33	$8.9e^{-6}$	72
Transformer	220/110	$5.3e^{-6}$	504
Transformer	220/33	$5.3e^{-6}$	504
Transformer	110/33	$5.3e^{-6}$	504

A.6 Economic indices

Economic indices used in calculations are presented below. Depreciation time is the economic life time of most of the electric equipment, which is assumed to be equal for different types of installations. Planning period, which is set to 24 years, in dynamic optimization is divided into 8 time stages, 3 years each.

Table A- 6 Economic indices

Depreciation time, years	40
Planning period, years	24
Interest rate, %	6
Utilization time, hours	8760
Loss utilization time, hours	3500
Cost of losses, SEK/MWh	165

A.7 Additional reinforcement costs

The costs corresponding to the identified reinforcement actions associated with the existing substations are given in this section. The problem under consideration is complicated in terms of numerous logical conditions and appropriate order of the events. The corresponding logic of the events is described in the next section.

Table A- 7 Reinforcement costs

Object	Reinforcement action	Cost, MSEK
S1	• reinforcement of substation (additional connections) • new transformers	6 24
S2	• reinforcement of substation (additional connections) • new transformers • new 33 kV switchgear for existing connections	8 24 35
S3	• new transformer 220/33 kV • GIS switchgear unit, 220 kV • 3x3 switchgear units, 33 kV • 2x3 switchgear units, 33 kV	8 7 15 9
S5	• minor reinforcement of substation (if 110 kV is eliminated) • new 33 kV switchgear to keep all the present connections	2 40
Cable S1 – S4	reinforcement of S1 for a new connection (additional switching equipment)	3

A.8 Logical Conditions

The planning problem comprises a number of logical conditions imposed on connections and sequence of the events, which must be taken into account in calculations and subsequently in decision-making. The logical conditions are summarized in Table A-8. In the first column the main actions, which require preceding actions are listed. The preceding actions can be related either by logical condition OR – one option out of several, or by condition AND – all listed action must be performed. The costs corresponding to each action can be found in Table A-3 and Table A-7.

Table A- 8 Logical conditions

Main action	Condition		Preceding actions – essential condition
Elimination of 110 kV cables S2-S5-S4	AND	OR	Reinforcement of S5 • minor reinforcements if S5 becomes a 33/11 kV load point • new 33 kV switchgear to keep all the present connections
		OR	Feeding of L11 and L6 • third cable S1-S4 • three 33 kV cables between S1-S4 • L11 and L6 are fed from S3
Three 33 kV cables between S1-S4		AND	Reinforcement of S1: • switchgear units • new transformers
Feeding of L11, L6 and L4 from S3		AND	Reinforcement of S3: • 33 kV switchgear units • 220 kV switchgear units • third transformer
Feeding of L11 or L6 or L4 (in any combination, but not more than two of three) from S3		-	Reinforcement of S3: • 33 kV switchgear units
Feeding of L5 from S1		AND	Reinforcement of S1: • switchgear units • new transformers
Feeding of L4 or L3 from S2		AND	Reinforcement of S2 • new 33 kV switchgear for present connections • switchgear units • new transformers
Feeding of L12 from S2 on 33 kV		AND	Reinforcement of S2 • new 33 kV switchgear for present connections • new transformers