FACULTY OF ENGINEERING OF UNIVERSITY OF PORTO
Department of Electrical and Computer Engineering

USING EVOLUTIONARY SWARMS (EPSO) IN POWER SYSTEM RELIABILITY INDICES CALCULATION

Leonel de Magalhães Carvalho

Dissertation submitted in partial fulfillment of the requirements
for the Degree of Master of Science,
in the specialization area of Power Systems

Dissertation written under the supervision of
Doctor Vladimiro Henrique Barrosa Pinto de Miranda
Full Professor

Porto, Portugal
July, 2008

U.PORTO

FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

MIEEC - MESTRADO INTEGRADO EM ENGENHARIA ELECTROTÉCNICA E DE COMPUTADORES 2007/2008

A Dissertação intitulada

"APLICAÇÃO DE ENXAMES EVOLUCIONÁRIOS (EPSO) NA DETERMINAÇÃO DE ÍNDICES DE FIABILIDADE DE UM SISTEMA DE POTÊNCIA"

foi aprovada em provas realizadas em 17/Julho/2008

o júri

Nuno Fidalgo

presidente Professor Doutor José Nuno Moura Marques Fidalgo
Professor Auxiliar da Faculdade de Engenharia da Universidade do Porto

Gerald Sheblé

Professor Doutor Gerald Sheblé
Investigador do INESC - Porto

V. Miranda

Professor Doutor Vladimiro Henrique Barrosa Pinto de Miranda
Professor Catedrático da Faculdade de Engenharia da Universidade do Porto

O autor declara que a presente dissertação (ou relatório de projecto) é da sua exclusiva autoria e foi escrita sem qualquer apoio externo não explicitamente autorizado. Os resultados, ideias, parágrafos, ou outros extractos tomados de ou inspirados em trabalhos de outros autores, e demais referências bibliográficas usadas, são correctamente citados.

Autor - Leonel de Magalhães Carvalho

Leonel Magalhães Carvalho.

Faculdade de Engenharia da Universidade do Porto

Abstract

This thesis presents an application of the Evolutionary Particle Swarm Optimization (EPSO) meta-heuristics in power system reliability analysis. The developed technique belongs to a new class of reliability assessment methods, the Population Based (PB) methods, which cleverly perform a search in the huge state space of the traditional power systems for the states considered as failure states.

Usually, the reliability of the power systems is evaluated according to two different approaches: the analytical methods and the simulation methods. The first type tries to describe the immensity of power systems in a closed mathematic model. This characteristic is its Achilles heel: the more complex is the power system under evaluation the more intricate it is to construct that model; the assumptions made and simplifications adopted for making it possible typically lead to lack of significance in the results provided. On the other hand the simulation methods, based on the statistical theory, estimate the reliability indices by random sampling of system states. Even though the problem of system size was relatively mitigated with the use of simulation methods, there was another problem: in view of the fact that normally power systems are very reliable, it is necessary to analyze a considerable number of states, including non-failure states, to compute estimates with a certain degree of confidence. The PB methods conjugate the best of the two disciplines: they are fairly immune to the system size and do not need to evaluate large amounts of states, especially if the characteristics of power systems are taken into account.

Several meta-heuristics were used to perform the search, namely Genetic Algorithms (GA), Particle Swarm Optimization (PSO) and a discrete version of PSO, the Binary Particle Swarm Optimization (BPSO). Nevertheless, EPSO was never used to perform that search. Moreover few efforts have been made to increase the PB methods efficacy, measured by how good the estimation of the reliability indices is, and efficiency, measured by the ratio of different states visited against the total number of states visited. Therefore, in this thesis, an EPSO based method for assessing the adequacy of power systems generating capacity is proposed, which tries to reduce the

computational effort and simultaneously increase the diversity in the population, by using different spreading techniques, to avoid visiting time and again states which were already visited.

The results obtained by EPSO will be compared with the work of other researchers in PB methods as well as with the results of the two traditional methods of reliability assessment.

Resumo

Nesta tese é apresentado uma aplicação da meta-heurística Evolutionary Particle Swarm Optimization (EPSO) na análise de fiabilidade de um sistema eléctrico de energia. A técnica desenvolvida pertence a uma nova classe de métodos de avaliação de fiabilidade, denominados métodos populacionais, que inteligentemente pesquisam o enorme espaço de estados dos sistemas de energia por aqueles que são considerados como estados de falha.

Geralmente, a fiabilidade dos sistemas de energia é avaliada segundo duas diferentes disciplinas: os métodos analíticos e os métodos de simulação. O primeiro tipo de métodos tenta descrever a imensidão do tamanho do sistema de energia por um rigoroso modelo matemático. Esta característica é o seu calcanhar de Aquiles: quanto maior é o sistema mais difícil é a sua modelização e, para o conseguir, normalmente realizam-se aproximações que podem por em questão a validade dos seus resultados. Por outro lado os métodos de simulação, baseados na estatística, fornecem estimativas dos índices de fiabilidade por amostragem aleatória de estados do sistema. Ainda que o problema do tamanho do sistema viesse bastante atenuado com a utilização dos métodos de simulação, existe outro contratempo: os modernos sistemas de energia são bastante fiáveis e, portanto, é necessário amostrar um número considerável de estados, incluindo estados que não contribuem para a formação das estimativas índices, os estados que não tem falha, com um certo grau de confiança. Os métodos populacionais conjugam o melhor das duas disciplinas: são amplamente imunes ao tamanho do sistema em análise e não necessitam de avaliar grandes quantidades de estados, sobretudo se for considerado na construção do modelo de base populacional, as características dos sistemas de energia.

Várias metaheurísticas foram utilizadas para realizar a pesquisa de estados, a saber, os Algoritmos Genéticos (GA), o Particle Swarm Optimization (PSO) e uma discreta versão do PSO, o Binary Particle Swarm Optimization (BPSO). No entanto a meta-heurística EPSO nunca foi utilizada para realizar esta pesquisa. Além disso poucos foram os esforços realizados para aumentar a eficácia dos métodos populacionais, medida pela diferença entre as estimativas dos índices

fornecidos por este tipo de método e o seu valor real, bem como a sua eficiência, medida pelo rácio entre o número de diferentes estados visitados e o número efectivo de estados visitados. Mediante esta análise é proposto nesta tese um método de base populacional baseado na meta-heurística EPSO para avaliar a capacidade de produção de um sistema de energia, tentando reduzir o esforço computacional e simultaneamente aumentar a diversidade na população, ou seja, evitar a visita a estados previamente registados, através da utilização de diferentes técnicas de repulsão.

Os resultados obtidos serão comparados não só com o trabalho de outros investigadores na área dos métodos populacionais mas também com os resultados dos dois métodos tradicionais de avaliação de fiabilidade.

Acknowledgments

First of all I would like to acknowledge my thesis supervisor, Professor Doctor Vladimiro Henrique Barrosa Pinto de Miranda, for the support, guidance and time dedicated in the refinement and improvement of this work. Without his help and passion probably this thesis would have never seen the day light. I also address a special thanks to Professor Doctor Manuel António Cerqueira da Costa Matos which exemplarily introduced me to the basic concepts of power systems reliability.

Next I would like to express my gratitude to the Power System Unit of INESC Porto for all the resources provided as well as all the good moments ("we are the champions, my friends"…). I address a special thanks to Mauro Augusto da Rosa for his effort in supplying useful material for the making of this thesis and especially for his companionship, availability and guidance in the earlier stages of this work.

Above all I must thank my family, in particular my parents, for all the sacrifices made to make me an engineer. From the bottom of my heart I thank them very much.

I also have to show gratitude to my brother Nuno for the help, presence and especially for showing me that we always must fight for our ideals. Patience, the best is yet to come.

This list of acknowledgements is not completed until I make a reference to the person who owns my heart: my beloved girlfriend Ana Catarina Rola. Her presence, understanding, enthusiasm and, most of all, her love were crucial for me to finish this thesis. What I am today I owe it to you.

Last but not least I would like to give special thanks to my friend Samuel who always supported me when I needed and for all the excellent moments that we have spent throughout these 11 years. The next year is your turn.

Contents

List of figures

List of tables

List of acronyms

DWA	Dynamic Weight Aggregation
EA	Evolutionary Algorithms
EIR	Energy Index of Reliability
EP	Evolutionary Programming
EPNS	Expected Power Not Supplied
ES	Evolution Strategies
EPSO	Evolutionary Particle Swarm Optimization
FEUP	Faculty of Engineering of University of Porto
FOR	Forced Outage Rate
F&D	Frequency and Duration
GA	Genetic Algorithms
IEEE	Institute of Electric and Electronic Engineers
INESC	Institute for Systems and Computer Engineering of Porto
LOEE	Loss of Energy Expectation
LOLD	Loss of Load Duration
LOLE	Loss of Load Expectation
LOLF	Loss of Load Frequency
LOLP	Loss of Load Probability
MSGA	Modified Simple Genetic Algorithm
MTTF	Mean Time to Failure
MTTR	Mean Time to Repair
PB	Population Based
PSO	Particle Swarm Optimization
RTS	Reliability Test System

Chapter 1

Introduction

In this Chapter the problem to be addressed will be explained as well as its context and the ideas that will be defended. Initially a general overview about the importance of the power systems reliability evaluation it will be performed. Then a distinction will be made about the two main categories of reliability assessment: adequacy and security. Subsequently the motivation for the work developed will be justified. Finally the organization of this thesis will be explained.

1.1. The importance of power systems reliability evaluation

Presently we can define a power system as a system that delivers to its costumer’s two different types of products: electric energy and reliability. As a matter of fact, the economic development of a country is strictly correlated with the reliability of its power system since most of its economic agents rely on this type of energy to boost up their activity. Therefore constant interruptions on electric energy supply can reduce dramatically their income forcing them to buy reliability, usually in the form of emergency generators. In the worst of scenarios the economic agent will be forced to move its activities to another country affecting not only the economic sector but also the social environment. Typically the short term government measures in order to avoid this situation are to offer these enterprises different types of benefits, like tax reductions, financial compensations or facilities in the acquisition of patrimonial assets. Nevertheless, in time, heavy investments in the electric power system will be needed since low reliability generally leads to an unsustainable development. The main aspect to be retained is that the power systems constitute a basic element for the improvement of both economic and social sectors of a modern society.

The first power systems were relatively small in size. Their first purpose was to supply the public illumination grid. The development of electric energy powered devices firstly for the industry and later for domestic use, lead to the widespread of electric energy consumption and consequently to the enlargement of the powers systems size. As a result the modern power systems are extremely complex, progressively more interconnected, with national or even continental dimensions. The high number of components, geographically distributed throughout a country or continent, coupled with the demand uncertainties and the availability of energy resources, make the design and operation of these systems a highly complex task.

The basic function of an electric power system is to supply the load demand as economically as possible within pre-defined continuity, quality and security patterns. However due to the enormous quantity of components in these systems, combined with their unique operation characteristics, there is a possibility of failure of the entire system simply by failure of a crucial or a group of crucial components. The good news is that these types of events have a low probability of occurrence. The most common security scenario is the strategic disconnection of a certain number of costumers in order to maintain the security of supply. However the same question arises: how much does the frequent failure of the system cost, considering all its possible consequences? In order to decrease the probability as well as the frequency and the duration of these events investments have to be made. However the tendency is to postpone those investments and to operate the electric systems in their limits. Managing all these contradictory requirements is the constant struggle of the decision makers when it comes to reinforce the electric system in order to increase its reliability.

Recently the institutional changes in the electric sector, such as the progressive deregulation or the former electric utilities privatization, with the purpose of creating an electric market, have given another degree of importance to the continuity of service: now it is the responsibility of the electricity provider to assure a continuous power supply, usually established in a contract, especially in the case of a very important client. Moreover the operation paradigm of the electric power systems is also changing. New concepts as distributed generation, micro-grids and the raising penetration of energy from intermittent sources, have brought the need to fully describe the entire energy system in order to correctly evaluate its reliability.

1.2. Adequacy vs. security

At this moment the necessity is clear for reliability studies whose main objective is to obtain performance indices for the behavior of the power systems, which can be used in the decision making process. However one must distinguish between reliability adequacy and reliability security. Reliability adequacy is related to the existence of sufficient resources within the system to meet the customer demand and the system operational requirements. This includes resources for generation, transmission and distribution needed to "carry" the energy to the individual consumption points. The adequacy evaluation is associated with the static conditions and does not include the dynamics of the system and the response to the transient perturbations: the different system states are evaluated without taking into consideration possible instability that may be introduced by failure of the system components. On the other hand the security is related to the system ability to respond to the dynamic or transient disturbances that might occur. Thus the security evaluation is associated with the reaction of the system to any disruption to which it can be subjected. This includes sudden loss of generation and/or transmission capability that can lead to transient instability, frequency instability, voltage instability among others. Most of the currently available techniques for assessing the reliability are in the field of adequacy evaluation. The ability to evaluate the security is still very limited mainly due to the complexity associated with the modeling of the dynamic behavior of the electric power systems. Most of the assessed reliability indices are, in fact, adequacy indices and not security indices, although they are commonly designated as so [1]. This thesis will be focused in the assessment of adequacy indices.

1.3. The purpose of this thesis

Monte Carlo remains the standard method to calculate estimates of reliability indices in power systems. This statistically based method has gained importance over analytic models since the emergence of enough computing power in the beginning of the 90's coupled with the adoption of efficient convergence acceleration techniques. The two basic advantages of Monte Carlo were:

- Allowing simulation of realistic characteristics of systems, even those not necessarily reducible to formal mathematical models;
- Allowing the calculation of distributions and not only of mean values (in its simplest form, allowing the estimation of variance).

Non-chronological models became successful then. However, as it is usual in such cases, the growth in computer power opened the way to the desire to perform chronological simulations and this became demanding of increased computing power. At the same time, even non-chronological models became more complex because of the availability of computing power at desktop level. As it happened in many other cases in the development of science and technology, the moment one has available more computing power this becomes almost at once insufficient for the new and more complex models one wishes to run.

Recently, an alternative to Monte Carlo started to emerge: Population Based (PB) methods. While Monte Carlo is a statistically based method, relying on the theorems of sampling to provide an estimate of a result plus some interval of confidence, PB methods are methods that try to search only for the meaningful subset of the state space and are enumeration methods. If all states contributing to a certain index could be identified and their probabilities known, the index would be calculated exactly. PB methods try therefore to discover, if not the totality, the majority of states so that a good approximation of the index is computed. Notice that visited states of the current developed PB methods are similar to those sampled by the non-chronological Monte Carlo: there is no sequential relation between them.

The methods are called PB because they rely on meta-heuristics that have a population of solutions (individuals, particles) as their core. In this class one may count Evolutionary Algorithms (EA) – Evolution Strategies (ES) / Evolutionary Programming (EP) or Genetic Algorithms (GA) – and Particle Swarm Optimization (PSO) algorithms. They all were traditionally developed to be an optimization tool but the problem now is the discovery of a set of states that have maximum contribution to the index to be calculated – so, some mechanism to generate diversity must be kept, otherwise all solutions would tend to converge to a maximizing state and space exploration would be hampered.

In this thesis a new PB method is presented, using a brand new meta-heuristic: Evolutionary Particle Swarm Optimization (EPSO). This meta-heuristic proved in the past its superiority in relation to the other types of EA [2] mostly due to its robustness in finding consecutively better solutions for a wide range of variation of its strategic parameters. Although this can be a good property for optimization problems, in search algorithms it is necessary to create some diversity in the population in order to maximize the number of visits of significant states and to avoid the visit to repeated ones. Therefore it is necessary to include in the traditional EPSO algorithm some sort of forgetting mechanism as well as a spreading instrument between the elements of the population (in EPSO they are called particles) when some type of convergence into a specific zone of the search space is detected. The first objective can be accomplished by moving constantly the population objective, for instance, using a bi-objective fitness assignment process [3] and regular forgetting of the best particle in the population, and using a fitness assignment method which penalizes the already visited states. The second objective can be achieved by adding an extra term to the particle velocity taking into consideration the proximity between them and the number of the similar cases. Additionally, in this thesis, another method will be used to maximize the diversity in the population taking into concern a new social cognitive model for the swarm: not only the position of the best particle in the population influences movement of a particle but also the position of its nearest best particle. With this feature a more intensive search is expected, in the local zones of the search space.

The proposed methodology will be tested in the reliability adequacy assessment of the generating capacity of a world-wide benchmark power system, to have a basis of comparison with the results of others PB methodologies. Also a comparison with the results provided by a non-chronological Monte Carlo simulation will be done.

1.4. Organization of this thesis

This thesis is organized in the following manner. In Chapter 2 a brief overview is performed about the traditional methods of reliability adequacy assessment. The analytical methods will be distinguished from the probabilistic methods and, among the probabilistic methods, the analytical approach and the simulation approach will be explained. Then the main lines of the most important EA will be introduced. Finally will be performed an overview of the new PB methods stating important works in the area and showing their particularities.

In Chapter 3 the traditional EPSO algorithm is described in detail. After this description the ideas of the proposed methodology will be detailed. Then a brief introduction to the reliability adequacy indices of the problem addressed in this thesis is presented as well as how they are assessed in the particular reliability adequacy evaluation. Next an introduction to the process of fitness assignment is presented, with the inclusion of the bi-objective approach. Subsequently a brief description of the diversity techniques and the way that they are incorporated in the proposed algorithm will be carried out. Finally the EPSO based search algorithm will be fully described in its particular steps for a systematic application.

In Chapter 4 the performance of the EPSO is illustrated, based on the search algorithm proposed in Chapter 2 in the evaluation of the generating capacity of a benchmark power system. To accomplish that, first the effect of each fitness function is evaluated, as well as the effect of each spreading technique in the improvement of the search efficacy and efficiency. Then the EPSO reliability algorithm, proposed in this thesis, which includes the best of these means to create diversity, will be compared with a search methodology proposed in the literature in order to exhibit its merits and eventually its demerits.

Finally, in Chapter 5, the main conclusions of this work will be summarized and some suggestions or guidelines for future work will be offered.

Chapter 2

State of the art

In this Chapter a brief review of the state of the art of power systems reliability assessment techniques will be made. As it is known this type of analysis is very complex and its intricacy increases with the growth in system size. On the other hand, these types of studies are fundamental to determine whether it is necessary to perform investment in order to optimize, from an economic point of view, the planning and operation of the power systems. To measure the power systems reliability the scientists developed several methods to solve each specific reliability problem. In this thesis the generating capacity adequacy assessment problem will be addressed. Therefore, the next subsections of this Chapter will be focused on the methods developed to solve this specific problem, with special attention to the ones which were recently developed since they are the vital subject of this thesis.

Firstly an overview of the general subject of power systems reliability assessment will be presented. Then two different approaches will be distinguished for evaluating the adequacy of the generating capacity: the deterministic approach and the probabilistic approach. Further, the probabilistic approach will be also divided in its main two schools: the enumeration methods and the simulation methods. Finally the new reliability assessment technique will be presented, which can be considered also a probabilistic enumeration method but with an increase of efficiency as it will be shown.

2.1. Power systems reliability adequacy assessment – an overview

Reliability assessment of real size electrical power systems is a complex problem. The historical reasons for this complexity are the enormous number of components that this type of systems possesses, the distinctive ways that these components may fail, and the singularity of the system operation.

As a matter of fact there are several ways to produce electricity, each one with its own characteristics. Moreover in some of these generators there is an uncertainty associated to the availability of the primary source of energy. Combined to these facts, the power flow through the electric network obeys to the 1st and 2nd Kirchhoff laws, unlike the common transportation problem, which has only to verify the 1st Kirchhoff law, and, to maintain the stability, the system power production must always equal the losses in the electric grid plus the randomness of the customer demand. Therefore understanding how the electric power systems works is essential to assess its reliability.

Taking into account the previous mentioned facts, it is usual to divide the electric power systems in their main functional zones. These are:

- Generation;
- Composite generation/transmission;
- Distribution.

This division was first proposed in [4] introducing the concept of hierarchical levels. The hierarchical level one, HLI, refers to generation facilities and their ability to supply the system demand; the hierarchical level two, HLII, refers to the composite generation/transmission systems and its capacity to deliver energy to the bulk supply points; the hierarchical level three, HLIII, refers to the complete system including distribution and its aptitude to assure the power and the energy demand of the individual consumers.

This partition allowed the development of specific techniques to quantify the reliability, according to the zone characteristics and the reliability study in question. For instance, one of the traditional reliability studies is the adequacy of the generation capacity. In this particular study it is frequent to ignore the influence of the network and to aggregate all of the system demand in one single bus powered by all system plants. These simplifications allow to assess the reliability of the power system generation subset and to draw conclusions on whether is necessary the construction of new power plants to enhance the security of supply. However, if the transmission system is not properly sized, a large amount of costumers may not be supplied, even though the generation subset is considered reliable. Therefore care must be taken in the application of a specific technique because the results provided are only valid in the scope of the problem formulation and its simplifications.

Nowadays we witness the progressive deregulation of the electric sector. In the past, utilities were vertically oriented, frequently owned and controlled by the governments, comprising power production, transmission and delivery. Hence the planning and operation of the electric power systems were made in a monopolistic scenario and the reliability concerns were basically focused in the security of supply. Scale economics was the rule. The division of these utilities in production, transmission, distribution and commercialization was made to increase competition, to give the electricity consumers the opportunity of choosing their electric provider, and to allow that the future of the electricity power system is in the hands of their agents. Electricity is now treated as a commodity and the concept of consumer is being replaced by the term costumer. Reliability is a responsibility of the system agents (except costumers, obviously), imposed by the market regulator in the form of targets that have to be satisfied or otherwise they will incur in monetary penalties. This fact combined with the increased amount of generation in the distribution subset from intermittent sources, makes more difficult to solve any reliability related problem and HLIII studies are now growing on importance.

This thesis addresses the static reserve problem by applying Population Based methods instead of the classical Monte Carlo simulation, namely adopting a special and new technique called Evolutionary Particle Swarm Optimization (EPSO). This problem is included in the HLI type of studies and measures the adequacy of the generating capacity considering future decommissioning of old power plants, the possibility of failure of the ones in service as well as outages due to scheduled maintenance, and the load growth estimates in a long-term horizon. It differs from the operating reserve problem, which evaluates the actual capacity to meet a given load level in a short-term horizon, being the fundamental difference between static and operating reserve, the period of time in study.

2.2. The deterministic approach

Several techniques were developed to tackle the power production adequacy problem. Two approaches can be identified: the deterministic approach and the probabilistic approach.

The deterministic approach is a simple method to measure the adequacy of the generating capacity and was widely used in the past by the electric utilities to support their decisions. In a few words, this approach quantifies the electrical power system reliability using a pre-specified rule based on the past experience of the utilities. Therefore, each utility adopted different criteria according to its internal organization and the electrical power system in question. Some of these criteria can be found in the specialized literature or in the utilities handbooks. A typical worldwideknown example of this approach is the Planning Generating Capacity [1], which determines the minimum necessary installed capacity, which is equal to the expected maximum demand plus a fixed percentage of the expected maximum demand. Also it is common to determine the static reserve, which is the difference between the generating capacity and the expected maximum demand, using as reference the capacity of the largest generating unit.

As the reader may have noticed, these deterministic criteria are not suitable for the reliability assessment of today's electrical power systems. From an economic point of view, this type of approach leads in most cases to solutions that waste financial resources without apparent justification, as this approach does not consider the stochastic behavior of the electrical power systems or, in other words, disregards the way in which this systems operates, the way that its components fail and the randomness of the system load. The main advantages of this approach are the straightforwardness and robustness of their results since the criteria used by the utilities were usually developed to be on the side of the security of supply. However, due to its limitations, this approach can also lead to under-investment solutions and probably to an unacceptable number of interruptions on load supply. Quantifying the cost of load curtailment is far behind the context of this thesis, but it is easy to understand that the modern society does not tolerate a too frequent failure of the electrical power systems. On the other hand this same society does questions the unjustified investment of large amounts of money to improve power system reliability. Therefore each dollar, euro or another currency invested to improve the system reliability has to be justified. For this reason the deterministic approach is being gradually replaced by probabilistic methods, although several utilities still use the deterministic approach (such as the n-1 criterion), especially in the transmission system.

2.3. The probabilistic approach

The probabilistic approach is the soundest way to assess power system reliability since this approach incorporates the fact that there is an uncertainty associated to the events that can occur in this type of systems. The most common types of uncertainties that can be found in the electric power systems are:

- The components state;
- The weather state;
- The hydrological resources state;
- The load state.

These types of uncertainties are incorporated in the probabilistic approach using stochastic models. The classical reference is the Markov model, which uses the exponential distribution to represent the duration of the system events, leading to constant transition rates between states [5]. For this reason this type of stochastic model is called the homogeneous Markov model and it is attractive because of its mathematical elegance, allowing the inclusion of different system states and the way that the system evolves from on state to another. The stationary probability, which is the probability of a state occurrence when the Markov process tends to the infinity, or, in other words, the expected value of the state probability, is calculated from the transition rates between different states.

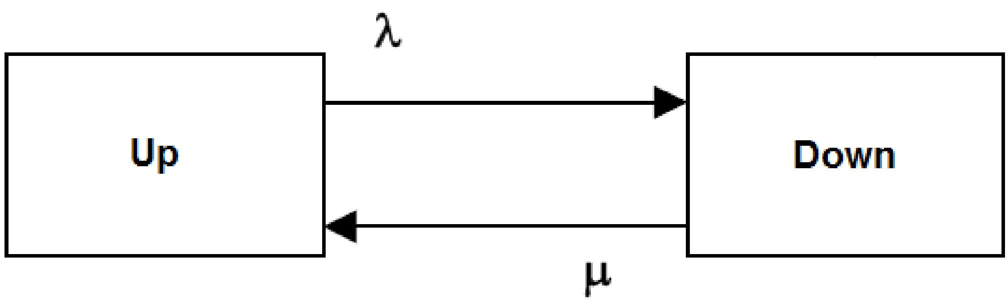


Figure 2.1. Two-state Markov model for a system component, where λ is the expected failure rate and μ is the expected repair rate.

Modeling the durations of system events by the exponential distribution is extremely helpful on a mathematical point of view. However the duration of a specific type of event may not follow this distribution, like the case of the duration of the components repair. Several efforts had been made to conceal this fact and in [6] is presented a technique that enables the use of bell-shaped duration distributions such as the Weibull distributions in the homogeneous Markov models. In [7] is discussed the use of others non-exponential distributions in the Markov process.

The probabilistic approach is subdivided in analytical methods and simulation methods. The analytical methods describe the system behavior through a mathematical model and assess the system reliability by the numerical calculation of the mean values of the desired system reliability indices. This type of approach was used up to the 80´s basically for its low computational effort. However if a complex system is considered, several assumptions and simplifications have to be made for analytical tractability. Therefore there is a great possibility of these methods to provide unrealistic results.

The simulation methods, often called Monte Carlo simulation methods, estimate the reliability indices by the random sampling of scenarios. These types of methods have the advantage to incorporate multiple system dependencies and characteristics, electrical and nonelectrical, which is extremely difficult to represent in the analytical methods.

2.3.1. The analytical methods

In order to evaluate the adequacy of the generation capacity two methods can be clearly defined: the basic probability methods and the frequency and duration (F&D) methods [1].

The first one uses the concept of unavailability, which is the probability of finding the generation unit out of service, to construct, in a recursive manner, the so called Capacity Outage Probability Table. Usually, the unavailability of a generation unit, also known as Forced Outage Rate (FOR), is computed assuming a two state homogeneous Markov model to describe its operation cycle. This model is widely used on this type of study due to its simplicity. For instance, to completely describe generation system reliability it is only necessary to know the mean time to failure and the mean time to repair of the unit in question which can be obtained for analyzing the history of the unit in question. It is also possible to include in this type of study more detailed models to cover the unit's specific operation conditions like peaking service [8].

The calculation of the Capacity Outage Probability Table is no more no less than the enumeration of all system states and their probability of occurrence, each state represented by its outage capacity. The result is the discrete probability distribution of an outage occurrence. For very large systems it is common to truncate this table by rejecting the states which possess a probability inferior to a pre-specified threshold, with the purpose of reducing the computational effort. In [1] can also be found another approximation method to this table by a continuous distribution, valid for very large systems.

After obtaining all the entrances of this table a discrete convolution with the system load curve is made to obtain the loss of load risk. To do this mathematical operation first the individual peak loads of the load curve are arranged in a descending order creating the cumulative load model. Then, for each value of the Capacity Outage Probability Table the number of hours, days or weeks is computed (depending on the base of the load diagram) where the load exceeds the capacity in service. Dividing this number for maximum number hours, days or weeks of the load curve, the probability of the load being higher than the capacity in service is obtained for each particular state. This probability multiplied by the probability of the respective entry in the Capacity Outage Probability Table gives the loss of load probability for this particular state. The next step is to add the individual values of the loss of load probability to acquire the system loss of load probability. It is also possible to compute the loss of load risk multiplying the obtained value by the maximum number of hours, days or weeks of the load diagram. However if the diagram is in an hourly base, for example, the result of the risk of loss of load cannot be extrapolated to another base like days or weeks.

The described process can be summarized in the following mathematical formula:

$$LOLP = \sum_{i=1}^{n} p(X_i) \times p\big(L > (X_{max} - X_i)\big) \ , \tag{2.1}$$

where $LOLP$ is the Loss of Load Probability reliability index, $p(X_i)$ is the probability of the generation capacity being $X_i\ (MW)$, X_{max} is the maximum generation capacity, $p\big(L > (X_{max} - X_i)\big)$ is the probability of the load exceeds the capacity of the state i and n is the dimension of the Outage Probability Table.

This method also allows the calculation of energy indices. As a matter of fact the area bellow the load curve gives the total energy consumed in the period of study. Therefore, as it is easy to compute the number of hours, days or weeks that it will be a load curtailment, it is also easy to obtain the expected value of the loss of energy. This approach can also take into account in the indices calculation the effect of scheduled maintenance, the uncertainty in the load forecast and the FOR uncertainty [1].

The indices calculated by the basic probability methods are the expected value of the number of hours in which the load exceeds the generation capacity and the expected value of energy not supplied in a given period of time. The focus of the F&D perspective is to provide indices that indicate the frequency of occurrence of a generation outage and the expected duration of these interruptions. These are the main advantages of the F&D methods. Their main disadvantage is the more complicated mathematical concepts that this type of approach possesses. To apply and to master these types of methods it is crucial to understand the concept of frequency and the concept of state transition.

The F&D methods require the knowledge of the transition rates between the states that constitute the chosen homogeneous Markov model. Like the basic probability methods, the reliability indices are calculated through the convolution of the load model and the recursive constructed generation model. The F&D methods can also incorporate the uncertainty on the load forecast. In [1] and [9] the fundamental development of these types of methods can be found.

In [1] one can also found two different methods to analyze the adequacy of the generation capacity in interconnected systems, which are the Probability Array Method and the Equivalent Assisting Unit Method that can also be formulated in the basic probability approach and in the

F&D approach, analyzing the effect of tie line capacity. As it is known, the electric power systems are progressively more interconnected and the effect of adjacent areas in the reliability analysis of the generation capacity cannot be forgotten.

2.3.2. The simulation methods – Monte Carlo

Simulation techniques, often known as Monte Carlo simulation, estimate the reliability indices by simulating the random behavior of the system. There are two major types of Monte Carlo simulation: the non-chronological type and chronological type. In the non-chronological type the samples are obtained by producing “snapshots” of the system state, without any dependence on time between samples. Alternatively, in the chronological type, a virtual or fictitious clock is set in motion and, with the flow of time, sequences of events are randomly generated, like a “story of the life” of the simulated system.

The number of the needed samples for given level of accuracy is independent of the system size (depends on the variance of the variable under estimation), which makes Monte Carlo simulation appropriate to assess the reliability of very large systems. Also, Monte Carlo methods have the advantage to provide information about the variability of the reliability indices as they provide their underlying probability distributions [10]. Quoting [11] “the probability distribution provides both a pictorial representation of the way the indices vary and important information on significant outcomes, which, although they occur very infrequently, can have very serious system effects. These effects, which can easily occur in practice, may be neglected if only average values are available”.

Due to the incredible increase of the computational capabilities in the last two decades and the development of variance reduction techniques the Monte Carlo methods are the most commonly used methods for reliability assessment. However, in order to guarantee a certain degree of confidence in the estimates provided by these types of methods, a large number of samples have to be randomly obtained. Furthermore the number of samples needed depends on the system reliability which means that, for very reliable systems, the number of samples necessary to assure that the estimated indices belongs to the pre-specified confidence interval can be extremely large.

The Monte Carlo simulation methods can be divided in two approaches: the non-chronological approach and the chronological approach. In the non-chronological approach the system states are randomly sampled without any preoccupation with the chronology of the system operation. A non-chronological system state is obtained by sampling the state of all system components according to their probability of failure. Therefore it cannot model time correlations or sequential events. In the chronological approach, the up and down cycles of all components are sampled in accordance with their probability distribution and a system operating cycle is obtained by combining all the component cycles [12]. For that reason this technique allows to include in the reliability evaluation, chronological issues like the time-dependent load curve as well as the hydrological affluences or the sequential behavior of the system components. For instance, there is a correlation between the load curve and the unit's operation cycle. As a matter of fact some units are in service for long periods of time and others are only started when they are needed and normally operate for relatively short periods, usually when the system load is near its peak value. The first unit type is called base load unit and the second type peaking unit. This dependency cannot be easily incorporated in a non-chronological reliability evaluation scheme (although there is a specific Markov process to model the operation of peaking units [1] which requires more detailed data than the traditional two state Markov model, usually extremely difficult to obtain) making the flexibility the main advantage of the chronological approach. On the other hand the main disadvantage of the chronological approach in relation to the non-chronological approach is the enormous computing time and effort required to verify the same convergence criteria.

In [10] two methods for single-area generating system adequacy assessment can be found: the State Duration Sampling Method and the State Sampling Method. The first method belongs to the chronological Monte Carlo simulation type. The second method is a non-chronological Monte Carlo type.

2.3.2.1. State Duration Sampling method

The first operation of the State Duration Sampling method is to generate the unit operation history by sampling the time to failure and the time to repair, according to the probability distribution of these random variables, assuming a two state model to describe the unit operation. This concept can easily be extended to multiple states modeled units, by sampling the time for all possible transitions from the current state. The next residence state will be the one which have the lowest transition time. This process is repeated in order to obtain the sequence of duration of the unit's state. The next step is to superimpose the load curve with the system available capacity curve to calculate the reliability indices.

The chronological simulation stops when the coefficient of variation of an index, usually, the expected energy not supplied because of its lowest convergence "speed", is inferior to a threshold value. The coefficient of variation of an index X is defined as:

$$\beta = \sigma(X)/E(X) \ , \tag{2.2}$$

where $E(X)$ is the estimated expectation of the index and $\sigma(X)$ is the standard deviation of the estimated expectation.

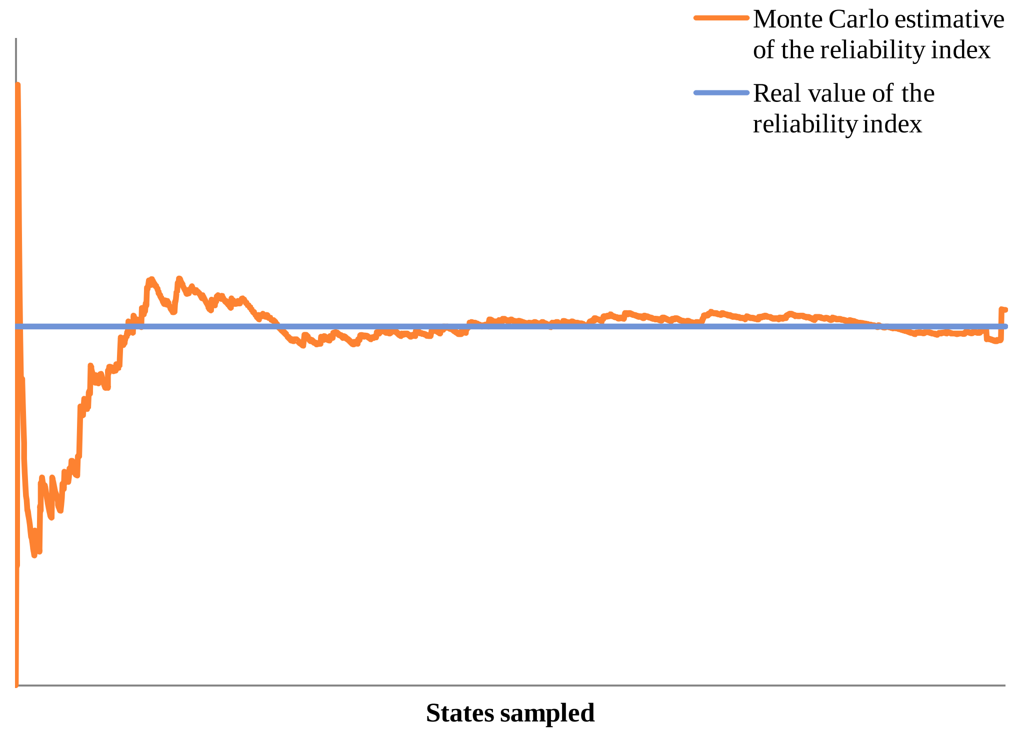


Figure 2.2. Illustration of the evolution of the estimative of a reliability index using Monte Carlo methods.

The main advantages of the State Duration Sampling method are:

- The easiness in the calculation of frequency indices;
- The use of non-exponential distributions to model unit state durations;
- The trouble-free inclusion of peaking unit.

2.3.2.2. State Sampling method

In the State Sampling method, the unit state is obtained by generating an uniformly distributed number in a [0,1] range which is compared with the unit FOR. If the random number is inferior to the unit FOR, the unit is considered unavailable otherwise it is considered available. Therefore the system state is a random combination of all generating unit system states. This process can be extended to a derated state unit model or to a multiple state unit model since this method only needs the probabilities data of the generating unit states.

In view of the fact that the non-chronological sample corresponds to a “snapshot” of the system state, the superimposition of the load curve is no longer valid. So, to obtain the reliability indices the sample must be compared to all periods of time that the chronological load curve is subdivided. This process requires a huge computational effort. To suppress this fact three approaches can be identified:

- Sample load states according to the load cumulative distribution function;
- Create a multistep model of the annual load curve using, for example, cluster techniques;
- Sample load states according to the multistep model cumulative distribution function.

The convergence of the State Sampling method is analyzed in the same way as in the State Duration Sampling method.

The State Sampling method has the advantage of accessing the reliability indices in a shorter computing time and memory storage than the State Duration Sampling method. However the calculation of frequency indices cannot be done as simply as they are obtained in the chronological approach. Also it is very difficult the use of non-exponential distributions to model unit state durations.

The use of simulation methods in the assessment of generating system adequacy can be very useful in modeling complex systems. However to obtain reasonable results it is necessary to draw an extremely large number of system samples, mainly in the case of very reliable systems. The known way to overcome this problem is to reduce the standard deviation of the estimated expectation. This can be achieved with the so called variance reduction techniques. In [10] five of these techniques can be found which can be applied in power systems reliability evaluation. Further in [13] the formulation of the Control Variates technique and the Importance Sampling technique to composite generation-transmission reliability evaluation is presented.

2.3.2.3. The Control Variates technique

Quoting [10] “variance reduction is simply a mean to use known information about the problem.” This is the concept used in the Control Variates technique. This method assumes that it is possible, by an analytical method, outside and independent of Monte Carlo, to calculate an approximation for the value that is to be determined. Thus the Monte Carlo simulation is only used to calculate the difference between the approximation and the solution of the problem. Choosing a correct control variable, which is the approximation for the desired value, is extremely important to obtain effective convergence acceleration. In the literature it is shown that in order to achieve a high convergence “speed”, the control variable and the value that is to be determined have to be strongly correlated.

2.3.2.4. The Importance Sampling technique

The Importance Sampling technique is based on the distortion of the probability density function of sampling in order to increase the probability of occurrence of relevant events and reduce it for those which are irrelevant. This method involves the use of a previous known auxiliary probability density function, obtained by an analytical method, from which the events are drawn. Like the Control Variates technique, the variance reduction depends on how similar is the shapes of the auxiliary probability function and the original probability density function. In other words, the knowledge contained in the auxiliary probability function allows to reduce the estimates variance obtained by the Monte Carlo process. It is as if the auxiliary function “explains" a good fraction of the variance, and therefore the efforts should focus specifically on the evaluation of the unexplained part.

In these last few paragraphs, the benefits of combining the "knowledge" of analytical models with the flexibility of Monte Carlo to achieve high convergence performances become evident. However the reliability analysis still depends in the drawing of a large number of random samples in order to verify the value of the coefficient of variation.

2.4. Evolutionary Algorithms

The Evolutionary Algorithms (EA) are inspired in the biological evolution in order to find the optimal solution of a problem. In this type of metaheuristic the optimization process begins with the establishment of an initial population which is the set of possible solutions (individuals). Then, to each individual, the genetic operators’ reproduction, mutation, recombination and selection are applied, to generate a new slightly different set of individuals that, in general, have better fitness than their parents. Then this evolutionary process is recursively applied to the successive generations until a stopping criterion is satisfied. In the end, the population is filled with individuals with better evaluation than the ones in the initial population and the supreme individual is selected as the solution of the optimization problem.

A general EA, which can be applied to all its variants, is defined as follows:

> **Procedure EA**
>
> Initialize a random population P of μ elements
>
> ***Do***
>
> Reproduction (introduce stochastic perturbations in the new population) – generate λ offspring…
>
> …by recombination
>
> …by mutation
>
> Evaluation - calculate the fitness of the individuals
>
> Perform selection - of μ survivors for the next generation, based on the fitness value
>
> ***Until the convergence criteria is satisfied*** (based on fitness, on number of generations, etc)
>
> **End EA**

The EA are distinguished from one another through the individual chromosome (an array of the problem variables) coding/decoding processes. Therefore there are two fundamental types of EA: the phenotype methods and the genotype methods. In the phenotype methods the individual is constituted directly by the set of the problem natural variables and there is a one-to-one mapping between chromosomes and system solutions. Alternatively in the genotype methods each solution is coded in a sequence whose interpretation implies the use of an external algorithm which allows computing the “natural” set of the problem solutions – but there is no one-to-one mapping to the reverse process (converting a solution into a unique chromosome) cannot be done directly. The Evolution Strategies (ES) [14], developed by I. Rechenberg e H.-P. Schwefel, and the Evolutionary Programming (EP) [15], proposed by Lawrence J. Fogel, are typical examples of the phenotype methods. Genetic Algorithms (GA) [16] are usually also organized as phenotype methods. However, biology follows a genotype principle: one cannot reconstruct a chromosome from a human body because the chromosome is a “program” with instructions to build a body and not a one-to-one mapping between genes and body characteristics.

2.4.1. Evolution Strategies / Evolutionary Programming vs. Genetic Algorithms

The ES and the EP are very similar. In both meta-heuristics the individual represents a possible solution of the problem in its “natural” variables, unlike GA, in which an individual is traditionally represented by some binary coding of the solution. This distinction has been, however, erased in recent times, especially with the emergence of the so called GA with real chromosomes. Therefore, the distinction between EP/ES and GA has been lately relying on the relative importance of the use of the operators recombination and mutation, with GA preferring the former and EP/ES preferring the latter (this however not totally true for the most recent ES models).

To maintain diversity and to push the population towards the optimum, the three algorithms rely in the following three operators:

- Selection;
- Mutation;
- Recombination.

The selection operator is what defines an algorithm as being “evolutionary”: is chooses those individuals that will survive into the next generation. It allows the process to favor the number of better performing individuals in the population throughout the evolutionary process. It has also been demonstrated that the application of elitism (the forced selection of the best individuals) favors in many cases the algorithm convergence. The selection can be done according to various techniques, being the following three the most used ones:

- Elitism (typical in ES);
- Stochastic tournament (typical in EP and GA);
- Spinning roulette (used in many GA models).

The mutation operator characteristics vary according to the representation of the individual. In ES/EP the mutation of an individual is made applying stochastic modifications in the vector of solutions. On the contrary, in GA, the mutation of an individual is accomplished by the random change of a pre-specified number of bits in the genotypic solution representation. Nevertheless the role of mutation is more important in ES/EP than in GA, where the main mechanism to generate new solutions is recombination. Self-adaptive mutation schemes an important feature of ES/EP (although also incorporated in GA but with not as much impact due to the lesser relevance of mutation in these methods), making this type of algorithm more “intelligent” and more independent form the initial setting of the strategic parameters. There is also an important quality of ES/EP that is worth to be mentioned (and that is beneficial in the GA with real-valued chromosomes). Working with the natural variables of the problem makes it easy to understand the neighborhood structure created by the genetic operators and eventually modify them to enhance the performance of the algorithm. In GA, this is not at all easy whenever the individuals have a binary (string of bits) representation of the possible solutions where tiny changes can lead the individual to a very far place from the actual search zone, increasing the diversity but compromising the convergence speed. This is why in many GA implementations the authors defend the use of a Gray Code representation of the binary values – to have neighboring values differing in one bit only.

2.5. Swarm Intelligence

The Swarm Intelligence is a Population Based computing technique inspired in the collective behavior of a population of simple individuals that coordinate using decentralized control and self-organization. This type of meta-heuristic takes advantage of the individual local interactions with one another and with their environment to tackle optimization and data analysis problems. Examples of systems studied by Swarm Intelligence are colonies of ants and termites, schools of fish, flocks of birds, and herds of land animals.

One of the most used swarm intelligence algorithms is Particle Swarm Optimization (PSO) [17] [18]. This methodology was proposed by James Kennedy and Russel Eberhardt in 1995, claiming to be inspired in the observation of the behavior of bird flocks in the search of food.

2.5.1. Particle Swarm Optimization

PSO is a Population Based optimization algorithm. The optimization process begins with a randomly crated population which is constituted by the so called particles. Each particle contains a position vector (a potential solution of the problem) a velocity vector and a memory vector of its previous best position. It also has a value of fitness of the current position and the value of fitness of its best position. New individuals are created through the application of the “movement equation”: each member of the population is moved in the search space according to three vectors called inertia, memory and cooperation. The first vector leads the particle in its previous direction. The second vector attracts the particle towards its previous best position. Finally, the third vector points the particle to the best solution ever found by the entire population. These movement "concepts" are summarized in the following equations:

$$V^{t+1} = V^t + U(0,1) \times W_m \times (b - X^t) + U(0,1) \times W_c \times (b_G - X^t) \ , \quad (2.3)$$

$$X^{t+1} = X^t + V^{t+1} \ , \quad (2.4)$$

where t is the iteration number, X is the position of the particle, b is the personal best position, b_G is the global best position, W_m is the memory weight, W_c is the cooperation weight and $U(0,1)$ is a random number drawn from the [0,1] interval.

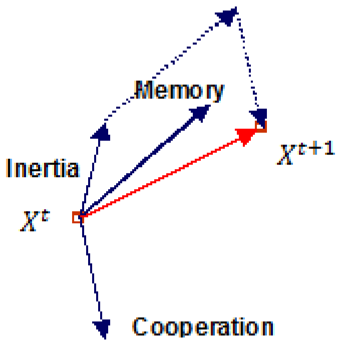


Figure 2.3. Illustrating the movement of a particle in PSO, influenced by the three terms: Inertia, Memory, and Cooperation [2].

It is also necessary to define in the beginning of the process the memory and the cooperation weights. The optimization process usually stops when a maximum number of iterations is reached.

The initial experiments with PSO shown that this algorithm has a fast convergence speed to the optimum region. However, due to the excessive velocity of the particles, PSO suffered from lack of precision. Therefore several actions were adopted in order to reduce the particles velocity, such as the application of constriction coefficients, decreasing functions with the growth of iterations, and clamping the velocity to an interval. The precision problem was partially solved but another problem related to the convergence of PSO subsisted: the dependence on the weights value and the requirement for fine tuning.

2.6. Evolutionary Particle Swarm Optimization as a best of two worlds

Unlike Evolutionary Algorithms, in PSO there is neither competition between particles nor self-adaptation of the strategic parameters. The progression towards the optimum is governed by the movement equation which is responsible for the creation of new particles. On the other hand in the traditional Evolutionary Algorithms the mutation and recombination operators dictate the characteristics of the next generation. However, the enrichment of the population with better individuals depends only on the selection operator. Furthermore, the convergence in PSO is strongly dependent on the value of the weights, contrarily to the Evolutionary Algorithms, were the self–adaptation of the strategic parameters is commonly used, giving a sort of intelligence to the evolutionary process.

The recognition of the different advantages of both methodologies led to the birth of the Evolutionary Particle Swarm Intelligence (EPSO) [19] [20] in 2002. This Population Based method uses the PSO equation of movement to recombine the particles instead of the traditional mutation/recombination operators. This fact combined with the selection operator, gives to EPSO a fast convergence “speed” as these two operators cooperate on the improvement of the population. In addition EPSO was developed with a sense of self-adaptation, reducing the dependence on the pre-set value of the weights. Therefore EPSO can provide feasible solutions for a wide range of the weight values. Also, EPSO possesses a particular feature. The global best position, instead of being a static point in the search space, is randomly moved, according to the Gaussian probability distribution. As a result, even when the particles have already converged to a specific zone of the search space, the population continues to be disturbed. Results show that EPSO is superior to the traditional approaches of PSO and Evolutionary Algorithms.

In powers systems EPSO was used with success in many applications, such as loss minimization and voltage control [19] [20], deriving optimal strategic decisions for an Energy Retailer, dimensioning of Power System Stabilizers, Double Fed Induction Generators PI controller tuning, among others. Therefore all the work on this thesis will be developed around this Evolutionary Algorithm.

2.7. Reliability adequacy assessment using the new Population Based methods

The Population Based (PB) methods evaluate the system reliability by enumeration of the system states. The main concept of this technique is to drive the individuals of the population in a guided search through the state space in order to find the most significant ones. Usually a state is considered significant if it is a failure state and its probability is superior to a threshold value. Then, the obtained set of states is convoluted with the load curve to provide the reliability indices.

From this point of view, this methodology is similar as the one used in the analytical methods (in fact, mainly due to the power systems dimension, the Capacity Outage Probability Table has also be truncated by the rejection of states which have a probability lower than a pre-specified value). However, the results provided by PB methods are underestimates of the correct value since only a subset of the total failure states is obtained. Eventually if the total number of the states which contribute to the formation of an index are within this subset then the PB methods give an exact value.

In the PB approach, the estimate of a reliability index is obtained from:

$$\hat{F} = \sum_{i \in D} p_i \times F_i \quad , \tag{2.5}$$

where $\hat{F}$ is the estimate of the index, p_i is the probability of the state i, F_i is the value of the index in the state i, and D ⊆ U, i.e. D is usually a subset of all possible states U.

The reader may now question the usefulness of this type of methods since there are analytical methods which have the advantage of theoretically assessing the correct values of the indices. However the analytical methods grow in complexity as the power system increases in size as well as the type of problem that is to be solved (for example assessing the reliability of the composite generation/transmission system). These facts act in favor of the PB methods for the reason that an individual is constituted by the state of all the components which the power system possesses. Therefore the complexity of the PB methods is widely immune to the system size and to the type of reliability study.

The PB methods also have an advantage over Monte Carlo. As it was previously mentioned, Monte Carlo is statistically based method, relying on the theorems of sampling to provide an estimate of a result plus some interval of confidence. Therefore in order to guarantee that the estimate belongs to the interval of confidence a large number of samples have to be drawn. In addition some of these samples are not failure states (characteristic of power systems) which also have to be evaluated. Thus reducing the number of evaluations, especially in the HLII and HLIII type of studies where the minimum load curtailment has to be determined by an Optimal Power Flow, can decrease considerably the computational effort. In PB methods this reduction is effective since it works with a state array with the most significant states.

However, in order to determine during the search process if the state is worth to be memorized, some methodology has to be defined. In [21] the adoption of intelligent pattern recognition methods such as neural networks to discriminate between failure and success is used. In the particular reliability problem addressed in this thesis a state is classified as a failure state if the total generation capacity cannot meet the peak of the load curve.

In [22] a simple modified GA is used to evaluate the generation capacity, not as an optimization tool but "as a search tool to truncate the probability states space and to track the most probable failure states". This methodology takes advantage of the chromosome concept allowing a binary representation of the system state according to the homogeneous two state Markov unit model. Also the authors use the fact that some generators have the same characteristics (in this case, the same generation capacity and the same FOR) to calculate from one particular state the number of states which have the same probability and the same load curtailment, discarding the need to visit all states. In [23] GA is also used to the assessment of the annual frequency and duration indices in composite system reliability, with the same search philosophy, modeling the transmission lines by the same two state Markov process. To determine if a state is worth to be saved the fitness function uses a linear programming module in order to minimize load curtailment without violating system constraints. The load at each load bus is considered fixed and equal to its yearly maximum value. In [24] the application of a PSO based method, the Binary Particle Swarm Optimization (BPSO [25]) is presented for reliability evaluation of power-generating systems including time-dependent sources. The authors used BPSO for the reason that it allows the coding of the generators states in a vector of binary numbers according to the homogeneous two state Markov model. BPSO, in its formulation is quite similar to PSO. However, unlike the typical PSO, in BPSO the velocity is used as a probability to determine whether a bit will be 1 or 0. Therefore after calculating the actual velocity with the same equation used in the traditional PSO, its value is squashed using a logistic function. Then if a randomly generated number within $[0,1]$ is less than the squashed value, the bit is set to be 1, otherwise is set to be 0.

In the previous mentioned works two different methodologies of fitness assignment can be defined. One is based on the maximization of the state probability for the states which are classified as failure states and which were not previously saved. Therefore the population is driven to the zone of the space state which possesses a smaller number of failure states. The other methodology is based in the state severity by maximizing the product between the state probability and the respective load curtailment. This one takes into account the "weight" of the state in the reliability indices calculation. Nevertheless none of these works used a technique to enhance the search. As a result the algorithm may visit time and again the same states before the exploration of a new zone of the state space.

In [3] a multi-objective version of PSO is used in the search for failure states to evaluate the composite system reliability. As it is known, in the classical PSO formulation there is a single objective or center of attraction. Consequently, if this optimization approach is used the search can be prematurely stopped with a high probability of missing important states and repeating previous saved ones. To overcome this problem the proposed method defines conflicting objectives (maximizing probability and maximizing the curtailment of the loss of load state) so the particles never converge into any specific point of the state space and thereby end the search. The authors claim that this methodology allows a better control over the particle dynamics although they also reefer that "the convergence behavior showed considerable sensitivity to the size of the swarm" as well as to the value of the PSO weighs.

2.8. Conclusions

In this Chapter an overview of the actual state of art in the power systems reliability evaluation was presented, specially focused in the generation capacity adequacy problem. It discussed the validity of the deterministic approach. As it was seen, this type of approach is dropping out of use since nowadays many data exist associated to the cycle of operation of the system components. Furthermore, the increase of the computational capabilities allowed the reliability analysis of very large power systems. Therefore the probabilistic methods are now the most widely used methods for assessing power systems reliability. The calculation of the reliability indices in this type of methods can be accomplished by two different philosophies: the enumeration, corresponding to the analytical methods, and the simulation, related to the Monte Carlo methods. The first one has the theoretical advantage of determining the exact value of the reliability indices although it is almost impossible to use it in the analysis of very large systems. The second one estimates the value of the reliability indices as well as their probability distributions. The major drawback of these latter methods is the large number of samples needed to assure statistical validity of the results, especially in the case of very reliable systems. The new PB methods provide a guided search through the state space for the ones which contribute the most to indices formation. These intelligent methods of sampling can reduce dramatically the computational effort and provide a set of the most severe states. This set, convoluted with the load model, provides an estimative of the desired reliability indices.

Nevertheless few efforts have been made to enhance the search process, or, in other words, to increase its efficiency, measured by the ratio of different states visited against the total number of states visited, and efficacy, measured by the difference between the index exact value and the estimative provided by the PB method. In this thesis these questions will be explored in detail using EPSO as the PB method reference.

Chapter 3

Modeling the problem with EPSO

In this Chapter an EPSO based algorithm to evaluate the adequacy of the electric power system generating capacity will be presented. This method belongs to the PB methods family described earlier in Chapter 2.

As it was seen, the main objective of these type of methods is to restrict as much as possible the search to a subspace of the huge state space, knowing in advance that the number of failure states is extremely small compared to the dimension of the space (characteristic of the power systems), being the search result a set of the meaningful states. Subsequently this set is convoluted with the system load model in order to assess the desired reliability indices. These are the main lines from which a PB method can be briefly described.

First, the traditional EPSO algorithm will be described in detail, evidencing the similarities with the classic PSO and with the EA. This description will follow narrowly reference [2]. Second, the main lines of the proposed algorithm, EPSO reliability, will be drawn as well as the assumptions and the scope of the application of this methodology. Finally two approaches to enhance the search performance will be presented: the use of distinct fitness functions and the use of spreading techniques between particles.

3.1. Formal description of EPSO

In Chapter 2 the general lines which differentiate EPSO form the traditional EA and a particular Swarm Intelligence one called PSO were described. It was said that EPSO, instead of using the classical mutation and recombination operators relies on the general scheme of the movement rule of PSO to produce new individuals. As it is known, the movement rule of PSO is the only mechanism that pushes the swarm towards the optimum. Therefore, if joined together with a selection operator, one may expect that these cumulative effects may improve the performance of an optimizing algorithm. Further, if the weights of the traditional PSO rule of movement have a considerable effect on the efficiency of the optimizing process, a self-adaptive scheme should include a mechanism for selecting weights in order to give to the algorithm the best performance possible in the progress towards the optimum. EPSO is the practical realization of these concepts.

In EPSO each individual is called a particle, as in PSO. Each particle is constituted by the object and the strategic parameters (this terminology is inherited from the traditional ES). The object parameters correspond to the variables of the optimization problem. In contrast, the weights of the PSO equation of reproduction are the strategic parameters since they define the reproduction/recombination strategy of each particle. Given a population with a set of particles, the general scheme of EPSO is the following:

Procedure EPSO

Initialize a random population P of μ particles

Do

Replication – each particle is replicated r times

Mutation – each particle has its strategic parameters mutated

Reproduction - each mutated particle generates an offspring through recombination, according to the particle movement rule, described below

Selection - by stochastic tournament or other selection procedure, the best particles survive to form a new generation, composed of a selected descendant from every individual in the previous generation

Until the convergence criteria is satisfied (based on fitness, on number of generations, etc)

End EPSO

In EPSO, the mutation of a strategic parameter w into w^* is accomplished using the following equation:

$$w^* = w \times [logN(0,1)]^{\tau} \quad , \tag{3.1}$$

where $logN(0,1)$ is a random variable which follows the Lognormal distribution obtained from the Gaussian distribution with the mean value equal to 0 and the variance equal to 1, and τ is the externally fixed learning parameter which controls the amplitude of the mutations.

The Lognormal distribution is widely used in the ES as an multiplicative scheme of mutation since, accordingly to this distribution properties, the probability of obtaining a new value multiplied by m is the same as the one obtaining a new value multiplied by $1/m$.

Alternatively a multiplicative and an additive mutation scheme can be defined using the following equations:

$$w^* = w \times \left(1 + \sigma \times N(0,1)\right) \ , \tag{3.2a}$$

$$w^* = w + \sigma \times N(0,1) \ , \tag{3.2b}$$

where $N(0,1)$ is a random variable drawn from the Gaussian distribution with mean value equal to 0 and variance equal to 1, and σ, like τ, is a fixed parameter to control the amplitude of mutations. However, in these particular cases, σ must be small enough to avoid the strategic parameter become negative.

Another characteristic of EPSO is the treatment of the global best. Instead of being attracted to the best point ever found by the population, the particles are driven to a sort of “foggy best-so-far region” where it is more likely to find the real global best solution. This is accomplished by randomly disturb the global best solution as follows:

$$\boldsymbol{b}_{\boldsymbol{g}}^* = \boldsymbol{b}_{\boldsymbol{g}} + w_{i4}^* \times N(0,1) \ , \tag{3.3}$$

where $\boldsymbol{b}_{\boldsymbol{g}}$ is the global best position, $N(0,1)$ is a random variable drawn from the Gaussian distribution with 0 mean vale and variance equal to 1, and w_{i4}^* is the weight conditioning the amplitude of the disturbance which has also to be mutated.

With this procedure the population continues to be “agitated” even when the particles have converged to the same region of the search space.

The next step in the EPSO algorithm is reproduction. The birth of a new particle is as follows (see figure 3.1.):

$$\boldsymbol{X}_i^{(k+1)} = \boldsymbol{X}_i^{(k)} + \boldsymbol{V}_i^{(k+1)} \ , \tag{3.4}$$

$$\boldsymbol{V}_i^{(k+1)} = w_{i1}^* \times \boldsymbol{V}_i^{(k)} + w_{i2}^* \times \left(\boldsymbol{b}_i - \boldsymbol{X}_i^{(k)}\right) + w_{i3}^* \times \left(\boldsymbol{b}_{\boldsymbol{g}}^* - \boldsymbol{X}_i^{(k)}\right) \times \boldsymbol{P} \ , \tag{3.5}$$

where $\boldsymbol{b}_i$ is the best point found by particle i in its past life up to the current generation, $\boldsymbol{b}_g^*$ is the best overall point found by the swarm of particles in their past life up to the current generation $\boldsymbol{X}_i^{(k)}$ is the location of particle i at generation k, $\boldsymbol{V}_i^{(k)}$ is the velocity of particle i at generation k, w_{i1}^* is the weight conditioning the inertia term, w_{i2}^* is the weight conditioning the memory term, w_{i3}^* is the weight conditioning the cooperation or information exchange term, and $\boldsymbol{P}$ is the communication factor (discussed below).

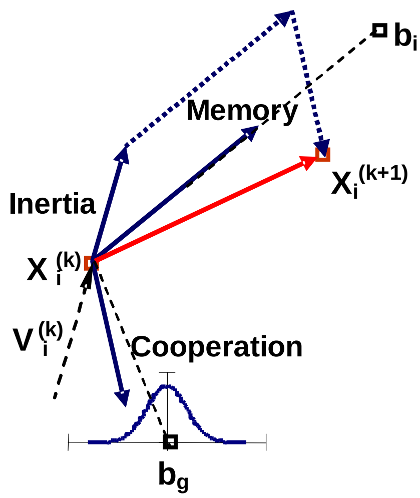


Figure 3.1. Illustration of the EPSO movement rule [2].

The communication factor $\boldsymbol{P}$ induces a stochastic star topology for the communication among particles. It is a diagonal matrix affecting all dimensions of an individual, containing binary variables of value 1 with probability p and value 0 with probability $(1 - p)$; the p value controls the passage of information within the swarm and is 1 in classical formulations (this is the star). Therefore this stochastic scheme oscillates between the star arrangement and a selfish version called cognitive model where no communication exists and a descendent of an individual is built only of contributions from its ancestor line [2].

Selection is modeled from the Stochastic Tournament concept: among the offspring of each particle, one compares the best one with another particle randomly sampled, and the best is selected with probability (1 – luck), where the luck parameter is defined in $[0,1]$ but is usually small. If luck is equal to 0 we have elitist selection [2].

EPSO can also be interpreted as a PSO based algorithm with a self-adaptive scheme of the strategic parameters. This particularity gives to EPSO a global drift more adjusted to the landscape. Furthermore, because these weights are subject to mutation, this may give an extra chance for the swarm to escape local minima (i.e., having particles that still explore other regions of space, because they may gain enough speed). On the other hand, EPSO also shows ability to focus and zoom in the optimum, precisely because mutations in the weights may favor the selection of the cooperation factor and reduce the importance of inertia and memory, if this strategy proves successful. This may in part explain why EPSO has shown, in many tests, robustness by consistently reaching the same optimum in a number of runs [2].

3.2. Introduction to the EPSO reliability algorithm

To describe the proposed algorithm firstly it is necessary to refer to the adopted Markov model describing the behavior of the generating units. In this study it is assumed that all units are base load units and their state transitions follow the exponential distribution. Therefore, and considering these facts, the suitable Markov model is the two-state homogeneous model. This process was previously illustrated in the figure 2.1 in Chapter 2.

According to this model the probability of finding the unit up, is as follows:

$$P_{up} = \mu/(\mu + \lambda) \ , \tag{3.6}$$

On the other hand, the probability of finding the unit down is defined as:

$$P_{down} = \lambda/(\mu + \lambda) \ , \tag{3.7}$$

where λ is the expected failure rate and μ is the expected repair rate.

These rates are computed according to the statistical data obtained by analyzing the "story of life" of the specific unit. According to the exponential distribution:

$$\lambda = 1/MTTF \ , \tag{3.8}$$

$$r = 1/MTTR \ , \tag{3.9}$$

where $MTTF$ is the Mean Time To Failure and $MTTR$ is the Mean Time To Repair.

In the problem addressed in is common to call to the probability of finding the unit down as the FOR (Forced Outage Rate).

Another aspect, usual in electric power systems, is that some units are, according to this model, identical, i.e., possess the same generating capacity as well as the same expected failure and expected repair rates. This information, if carefully incorporated in the program, can lead to a decrease in computational effort. The algorithms developed in [22] and [23] use this property to calculate, from a single state, all the possible states that are identical to a given one.

This leads us to adopt the following definitions:

- A (system) state: a vector that corresponds to a particular instantiation of all units in their states of failure or in operation;
- A case: corresponds to a set of (system) states comprising all combinations of n equal generators when only p are in a failure state.

The state vector has in principle a dimension equal to the number of units of the system in question, since each dimension represents the state of the particular generating unit, which, according to the two-state Markov model, is up or down, represented by the number 1 and 0, respectively. However, this state vector can be further simplified if, instead of trying to find the state of a particular unit, the search is driven to find the number of equal units that are in the up state. This reduces considerably the “size” of the particle. For instance, if the power system contains 32 generators, which can be arranged in 9 different groups of equal generators, the dimension of the problem is reduced in 23.

Each particle will be represented in this way:

$$\boldsymbol{X} = \begin{bmatrix} G_1 \\ \vdots \\ G_n \end{bmatrix} , \tag{3.10}$$

where n is the number of different units and G_n is a real number which belongs to the interval $]-0.5, max(G_n) + 0.5[$ where $max(G_n)$ is the maximum number of equal units in the state up.

According to this characterization, the position of the particle in the search space is a set of real values. Nevertheless, the definition of the problem suggests that a particle should be constituted by a set of integer values. Thus, in order to achieve this requirement, each dimension is rounded to its nearest integer value.

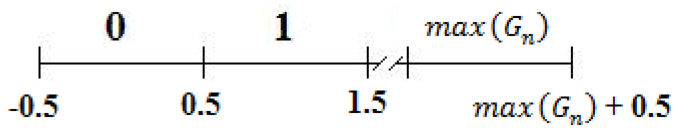


Figure 3.2. Illustration of the rounding process.

For instance if $G_1 = -0.4$, the correspondent integer value is $G_1 = 0$; if $G_1 = 4.6$ the number is rounded to 5. Note that the interval extremes are not included to avoid obtaining a negative number or a higher number than the maximum number of equal units.

To complete the proposed model, it is necessary to define the limits for the velocity variation. The chosen scheme is as follows:

$$\boldsymbol{V} = \begin{bmatrix} V_1 \\ \vdots \\ V_n \end{bmatrix}, \tag{3.11}$$

where n is the number of different units and V_n is a real number which belongs to the $[\alpha \times -1 \times (max(G_n) + 1.0), \alpha \times (max(G_n) + 1.0)]$ interval in which α takes a value between $]0,1]$.

This allows the particle to "fly" forward and backwards in the search space without any preference for some direction. The value of α has to be carefully chosen. If the value of the velocity is too high the particles may bounce (explained in the next paragraph) from the extremes of the position vector, missing important system states. On the other hand, if this value is too low the population may be trapped in a particular region compromising the convergence velocity of the search process. The suggestion is to assume a number near 0.5 for the value of α.

This method has another interesting feature. When some dimension of a particle reaches or surpasses any of its limits of variation, its value is set to the limit violated and its current trajectory is reversed, i.e., the velocity is multiplied by -1 with the purpose of bouncing back the particle to the search space. With this characteristic, the imprisonment of the particle at any of its limits is avoided (at least until the velocity naturally changes its signal, which can take several generations).

3.3. Measuring the importance of the generating state

The main objective of a reliability assessment study is to compute the so called indices which are the measure of the power system reliability. Until now, the focus was simply on the state of the system generating units. However, even when a great number of units are out of service it is possible to have no load curtailment. Therefore the assessment of the reliability indices depends not only of the unit's state but also on the state of the system load.

It is usual to evaluate one year of system operation. As it is known, the system load fluctuates with time. Each hour, week, or month has its correspondent value of load (notice that in these types of reliability studies, the system load curve is a set of forecasted hourly, weekly or monthly peak load). If the system load has this kind of behavior, how is the importance of a unit being down defined? Or, in other words, how is, in a "non-chronological way of the units state sampling", the effect of the yearly variation of the load included?

In the non-chronological Monte Carlo method there are two ways of including the load model in the reliability analysis: the comparison of the sampled generating state with every period of time that of the load curve is subdivided (which requires a tremendous computational effort) or sampling the state of the load according to its cumulative probability distribution. Only the first methodology is possible to implement in the proposed algorithm since the PB methods are not statistically based. In [22] and in [23] a state is classified as a failure state if the sum of the capacity of the individual units in service cannot meet the system peak load. If this requirement is verified and the probability of the state is above a specified threshold, the state is considered worth to be saved for posterior convolution with the annual load curve in order to assess the desired reliability indices – this will be referred to as a saved state. This characteristic adds another advantage to the PB methods over Monte Carlo methods: like the analytical methods, the construction of the generation model is independent of the system load. This is valid as long as the state load shedding is obtained using the system peak load [22][23].

3.4. Assessing the reliability indices

Almost in every "corner" of this thesis one refers to the term "reliability indices". They were introduced as a measure of the system reliability.

In the generating capacity adequacy assessment problem it is common to assess the following indices:

- Loss of Load Expectation (LOLE);
- Loss of Load Frequency (LOLF);
- Loss of Load Duration (LOLD);
- Loss of Energy Expectation (LOEE).

The LOLE (hour/year, day/year or week/year) index is the average number of hours, days or weeks (depending on the basis of the load model) in the evaluation period (usually a year) that the hourly, daily or weekly peak load is expected to exceed the available generating capacity. Its mathematical definition is as follows [10]:

$$LOLE = \sum_{i \in S} p_i \times T \quad , \tag{3.12}$$

where p_i is the probability of system state i, T is the evaluation period, and S is the set of all system states associated with loss of load.

From this index, it can be calculated another one, which is the so called Loss of Load Probability (LOLP).

$$LOLP = LOLE/T \quad , \tag{3.13}$$

Usually the LOLE index (in hours, days or weeks per year) is preferred to the LOLP index due to its understandability (the LOLP index results from a sum of probabilities and therefore its value has no units).

As it was shown, in the PB methods, first a set of meaningful states is found. Then this set is convoluted with the system load curve. Therefore, for each hour, day or week is necessary to analyze if all saved states result in load curtailment, or, in other words, it is necessary to compute the hourly, daily or weekly LOLP index. The yearly LOLE index is obtained by adding all the individual values of LOLP.

$$LOLE = \sum_{i=1}^{T} \sum_{j \in S'} p_j \quad , \tag{3.14}$$

where i is the i^{th} hour, day or week of the system load curve, T is the total number of hours, days or weeks of the system load curve, and S' is the set of failure states for the hour, day or week i.

The annual LOLP index can be calculated from equation (3.13).

The LOLE index does not indicate the severity of the loss of load nor the frequency or the duration of these interruptions on supply. Nevertheless, the LOLE index is the most widely used probabilistic criterion in generating capacity planning studies [10].

The LOLF (occurrence/year) measures the number of times, during the evaluation period, that the occurrence of a load curtailment is expected. In the enumeration methods, such as the PB methods, is not easy to assess this type of index. This is one of the advantages of the Monte Carlo methods over the enumeration methods in which the concept of frequency is easily incorporated.

However, with the development of the F&D methods, the assessment of this type of index is no longer a hard task.

For this particular problem, the LOLF is obtained like this [22][23]:

Step 1) Calculate the frequency of all saved states using:

$$f_i = p_i \times \left(\sum_{j=1}^{n} \mu_j - \sum_{k=1}^{m} \lambda_k\right) \ , \tag{3.15}$$

where p_i is the probability of the state i, n is the number of units down in the state i, m is the number of units up in the state i, μ_j is the expected repair rate of the unit j, and λ_k is the expected failure rate of the unit k;

Step 2) LOLF includes two components: frequency of generating capacity "FG" and frequency due to load change "FL":

$$FG = \sum_{i=1}^{T} LOLF_i \ , \tag{3.16}$$

$$LOLF_i = \sum_{j \in S'} f_j \ , \tag{3.17}$$

where T is the total number of hours, days or weeks of the system load curve, S' is the set of failure states for the hour, day or week i, and f_j is the frequency of the failure state j of the hour, day or week i.

$$FG = \sum_{i=2}^{T} V_i \times (LOLP_i - LOLP_{i-1}) \ , \tag{3.18}$$

where T is the total number of hours, days or weeks of the system load curve, $LOLP_i$ is the value of the index LOLP at the hour, day or week i, and V_i is a binary variable which takes the value 1 if the value between brackets is positive and the value 0 if the value between brackets is negative.

Step 3) Finally, the annual LOLF is calculated as:

$$LOLF = FG + FL \ , \tag{3.19}$$

The LOLD (hour/occurrence, day/occurrence or week/occurrence) index measures the expected duration of a failure. It is easily calculated if the LOLF and the LOLE indices are available.

$$LOLD = LOLE/LOLF \ , \tag{3.20}$$

The final index is the LOEE (MWh/year). Although the previous indices can provide important information about the expected number of hours, days or weeks that the load cannot be supplied as well as the frequency and duration of these interruptions, there is no information about the severity of the system expected unavailability. The LOEE measures the expected energy that will not be supplied due to generation failure. Knowing that the area below the load curve corresponds to the annual energy demanded, it is easy to assess the LOEE index. Firstly it is necessary to compute, for each load value, the Expected Power Not Supplied (EPNS). Then each individual EPNS value is added to compute the annual LOEE index

$$LOEE = \sum_{i=1}^{T} \sum_{j \in S'} p_j \times \left(L_i - Cap_j\right) \ , \tag{3.21}$$

where p_j is the probability of the failure state j, L_i is the value of load at the hour, day or week i, Cap_j is the generating capacity of the failure state j, and S' is the set of failure states for the hour, day or week i;

Another useful index can be calculated form LOEE which measures the capability of the power system to meet its annual demanded energy: the so called Energy Index of Reliability (EIR). The EIR index can be considered a normalized value of LOEE and is commonly used to compare the adequacy of systems that are very different in size.

$$EIR = 1 - LOEE/E \ , \tag{3.22}$$

where E (MWh/year) is the annual energy demanded by the power system in study.

3.5. Description of the distinct objective functions

In the literature one may find the use two distinct fitness functions: maximization of the state probability and the maximization of the state probability multiplied by the load curtailment in relation to the peak of the annual system load curve. However, the fitness of a particle takes also in account if the particle has been previously saved and if the particle does not meet the requirements to enter the list where the most important system states are recorded.

At this moment it is necessary to recall what a particle and a system state represent. The particle, as it was defined, represents the number of similar units which are in the state up.

However it may correspond to a great number of system states with equal probability and load curtailments as the single state that the particle seems to be. Therefore the rounded representation of the particle position will now be called a case which represents many system states. The number of similar states, $copy$, is calculated as follows:

$$copy = \binom{N_1}{M_1} \times \binom{N_2}{M_2} \times \ldots \times \binom{N_n}{M_n} \ , \tag{3.23}$$

where $\binom{N_n}{M_n}$ is the number of combinations $\binom{n}{k} = \frac{n!}{k! \times (n-k)!}$, N_n is the maximum number of equal units in the state up, M_n is the actual number of equal units in the state up, given by the rounded representation of the particle position and n is the total number of different units.

When a case is saved, the number of equal states is also saved. Therefore, it is possible to obtain a large number of system states saving only a few cases. However, care must be taken when calculating the system reliability indices. In every equation the set of all saved states, not cases, has to be considered. This is accomplished by multiplying the value of the case probability by the number of copies, reducing dramatically the necessary number of calculations. Each case can now be seen as a "single state" with its generating capacity and the correspondent probability of occurrence.

This explanation was necessary to describe the process of fitness assignment which is described as follows:

Step 1) Compute the probability of occurrence of the case:

$$p = \prod_{i=1}^{n} (1 - FOR_i)^{a_i} \times (FOR_i)^{b_i} \ , \tag{3.24}$$

where p is the case probability, n is the total number of different units, FOR_i is the common Forced Outage Rate, characteristic of the i^{th} set of equal units, a_i is the number of units up of the i^{th} set of equal units, and b_i is the number of units down of the i^{th} set of equal units.

Step 2) Compute the number of the case copies using equation (3.23).

Step 3) Compute the generating capacity of the case:

$$Cap = \sum_{i=1}^{n} a_i \times \overline{G_i} \ , \tag{3.25}$$

where Cap (MW) is the generating capacity of the case, n is the total number of different units, a_i is the number of units up of the i^{th} set of equal units, and $\overline{G_i}$ (MW) is the generating capacity of the Markov model up state of the i^{th} set of equal units.

Step 4) If the case probability is superior to a threshold and the generating capacity is inferior to the power system yearly peak load, go to the **Step 5)**; else go to the **Step 8)**.

Step 5) Search if the case has not been previously saved (direct comparison between vectors). If the case has already been saved, go to the **Step 7)**; else go to the **Step 6)**.

Step 6) The case fitness is calculated by one of the following functions:

$$\max \quad probability \ , \tag{3.26}$$

$$\max \quad probability \times copy \ , \tag{3.27}$$

$$\max \quad probability \times (\overline{L} - Cap) \ , \tag{3.28}$$

$$\max \quad probability \times copy \ \times (\overline{L} - Cap) \ , \tag{3.29}$$

where $probability$ is the case probability, $copy$ is the number of equal states which can be obtained from the case, Cap (MW) is the generating capacity of the case, and $\overline{L}$ (MW) is the system yearly peak load.

The value of the load curtailment calculated by $\overline{L} - Cap$ can be normalized, simply dividing this value by $\overline{L}$.

Step 7) The case fitness is assigned a small value in order to decrease the chances of previous saved cases occurring in the next generations. As it will be discussed later in this thesis, only the current position of the particle is a set of real values. The particle best position ever and the population best position are a set of integer values. Therefore, it is legitimate to search the memory of all of the population particles in order to determine if the repeated case is the best case for any particle. If so the value of fitness of the particle best position is also assigned a small value to avoid the continuous pushing to a previous saved case, avoiding repetition. This idea was also implemented in [3].

Step 8) The case fitness is assigned a small value since these category of cases does not fulfill the requirements for entering the list of cases.

As it is shown, the process of fitness assignment tries to decrease the number of repeated cases. However, none of these fitness functions improve the search procedure. The only difference between maximizing the case probability and maximizing the case probability x load curtailment is that in the second one the severity of the case is taken into account, i.e., there is a "bonus" if the case generating capacity is low.

The search procedure must discover the cases with high probabilities, which are the cases with a low number of units out of service, and simultaneously the cases with high load curtailment, which are the cases with low probability. According to this definition, the problem can be seen as a maximization of these two conflicting objectives: maximization of the case probability and maximization of the case load shedding. Therefore, a multi-objective fitness function can be formulated, to drive the search towards the decision space zones near the Pareto Front defined by the two objectives, where probably the most contributing cases for the construction of the reliability indices are located.

In this thesis, will be used the "*Dynamic Weighted Aggregation*" approach proposed in [26].

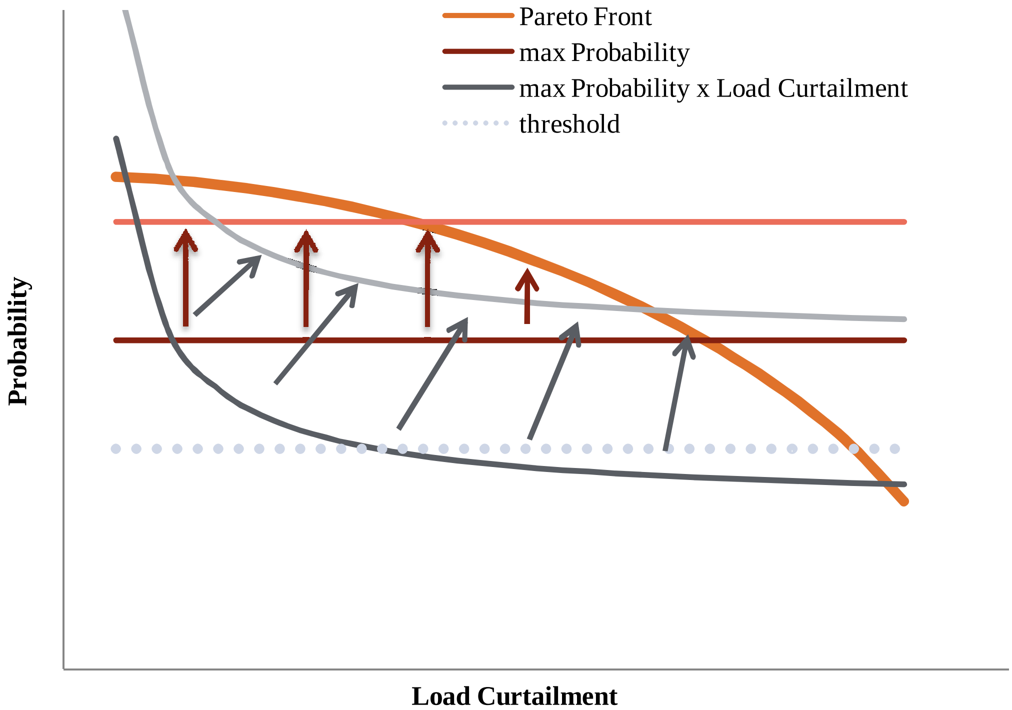


Figure 3.3. Illustration of the optimizing procedure of two different fitness functions in the decision space.

This approach aggregates the two objectives in a single objective function and dynamically changes the weights used for this aggregation using the following equations:

$$\max \quad w_1(t) \times A + w_2(t) \times B \ , \tag{3.30}$$

$$w_1(t) = |sin((2 \times \pi \times t)/T)| \ , \tag{3.31}$$

$$w_2(t) = 1 - w_1(t) \ , \tag{3.32}$$

where $w_1(t)$ is the first objective, A, weight, $w_2(t)$ is the second objective, B, weight, t is the EPSO reliability iteration's index, and T is the weight's change frequency.

Obviously this method is only valid for bi-objective problems. Notice that both objectives have to be equally scaled. Therefore, is necessary to determine the lowest and the higher value that each objective can take and convert them to a common scale using a utility function.

The utility function selected is the linear function. However since the case probability can take very low values, a modification to the scale must be taken. The decimal logarithmic function is one very adequate way of performing this modification.

The following steps describe the scaling procedure for a generic objective, ζ, which cannot take the value 0:

Step 1) Calculate the parameters of the linear transformation:

$$m = 1.0/(\log_{10}\left(max(\zeta)\right) - \log_{10}\left(min(\zeta)\right) \ , \tag{3.33}$$

$$b = -m \times \log_{10}\left(min(\zeta)\right) \ , \tag{3.34}$$

Step 2) Calculate the scaled value of ζ using:

$$scaled\left(\log_{10}(\zeta)\right) = m \times \log_{10}(\zeta) + b \ , \tag{3.35}$$

where $\log_{10}\left(max(\zeta)\right)$ is the decimal logarithmic of the maximum value of the objective ζ, and $\log_{10}\left(min(\zeta)\right)$ is the decimal logarithmic of the minimum value of the objective ζ.

Notice that in the case of the load curtailment maximization this utility function cannot be used since the decimal logarithmic function is not defined for the value 0. Instead the following utility function is used:

$$\text{scaled}(\text{load curtailment}) = \frac{\bar{\text{L}} - \text{Cap}}{\bar{\text{L}}} \ , \qquad (3.36)$$

where Cap (MW) is the generating capacity of the case, and $\bar{\text{L}}$ (MW) is the system yearly load peak.

For the probability x copies x load curtailment objective, its maximum and minimum (non zero, explained below) value is only obtained by a previous maximization procedure, using, for instance, the traditional formulation of EPSO with this single objective in other to find the fittest case. Nevertheless, when this objective is used in any fitness function in this thesis, single-objective or multi-objective, the load curtailment will be normalized using the equation (3.36). Therefore this singularity demands a new definition of the utility function for the probability x copies x load curtailment objective, described as follows, using also ζ for the representation of a generic objective which cannot take the value 0 (in this case, probability x number of copies):

Step 1) Calculate the parameters of the linear transformation:

$$\text{m} = 1.0/(\log_{10}\left(\max\left(\zeta \times \frac{\bar{\text{L}} - \text{Cap}}{\bar{\text{L}}}\right)\right) - \log_{10}\left(\min\left(\zeta \times \frac{\bar{\text{L}} - \text{Cap}}{\bar{\text{L}}}\right)\right) \ , \qquad (3.37)$$

$$\text{b} = -\text{m} \times \log_{10}\left(\min\left(\zeta \times \frac{\bar{\text{L}} - \text{Cap}}{\bar{\text{L}}}\right)\right) \ , \qquad (3.38)$$

Step 2) Calculate the scaled value of $\zeta \times \frac{\bar{\text{L}} - \text{Cap}}{\bar{\text{L}}}$ using:

$$\text{scaled}\left(\zeta \times \frac{\bar{\text{L}} - \text{Cap}}{\bar{\text{L}}}\right) = \begin{cases} 0, & \bar{\text{L}} - \text{Cap} = 0 \\ \text{m} \times \log_{10}\left(\zeta \times \frac{\bar{\text{L}} - \text{Cap}}{\bar{\text{L}}}\right) + \text{b}, & \bar{\text{L}} - \text{Cap} > 0 \end{cases} , \qquad (3.39)$$

where Cap (MW) is the generating capacity of the case, $\bar{\text{L}}$ (MW) is the system yearly load peak, $\log_{10}\left(\max\left(\zeta \times \frac{\bar{\text{L}} - \text{Cap}}{\bar{\text{L}}}\right)\right)$ is the decimal logarithmic value of the maximum value of the objective $\zeta \times \frac{\bar{\text{L}} - \text{Cap}}{\bar{\text{L}}}$, and $\log_{10}\left(\min\left(\zeta \times \frac{\bar{\text{L}} - \text{Cap}}{\bar{\text{L}}}\right)\right)$ is the decimal logarithmic value of the minimum value of the objective $\zeta \times \frac{\bar{\text{L}} - \text{Cap}}{\bar{\text{L}}}$.

3.6. Description of the spreading techniques

Although it may seem that the EPSO reliability has sufficient features to perform a satisfactory search, it is necessary to include some mechanisms to increase diversity or, in other words, to spread the population in the decision space.

First of all, as the search procedure evolves, the best position is always changing, especially in the case of the bi-objective optimization. The main reason for that is the fitness assignment procedure. For instance, if a case has been the best one in the previous generation which was never been saved and survived to the current generation, its current fitness will be assigned a small value since it already belongs to the list of significant cases. Therefore, the search procedure can be enhanced if the best particle is changed in each generation according to the fitness of the particles in the current population. Additionally, if the best particle is maintained during a certain number of generations the population is encouraged to search the zone around it. Nevertheless if the number of these generations is high the process can be stuck in a given region. Both techniques will be addressed.

The second technique is inspired in [27]. Often the particles are very “close” or “overlapped”, especially in the end of the search process where the particles velocity is insufficient to drive the search into another region. To avoid this situation, each case, obtained by rounding the particle’s position, is compared to the others in the current population. Then, accordingly to the distance to the other cases and their relative position, a value is calculated which will be added in the velocity update equation of the correspondent particle. The process is as follows:

Do

Step 1) Initialize an auxiliary vector with the length of the particle’s position with 0 in each dimension.

Do

Step 2) Measure the distance to the other particles in the population:

$$dist_{ab} = \|round(X_a) - round(X_b)\|_2 \ , \tag{3.40}$$

where $\|round(X_a) - round(X_b)\|_2$ is the Euclidian norm of the vector defined by the difference of the rounded position of the particle a and the rounded position of the particle b.

Step 3) If the particles represent the same case is added to each dimension of the auxiliary vector the values 1 or -1, randomly chosen.

Step 4) If the distance is within a specified threshold previously chosen, the auxiliary vector is added the difference between the two rounded positions of the particles.

Step 5) If the distance is beyond this threshold the auxiliary vector remains equal.

While all the distance between the actual particle and all the others in the population is not measured.

Step 6) Scale each auxiliary vector dimension to avoid excessive velocity grown:

Do

Step 7) If the dimension of the auxiliary vector is equal to 0, then do nothing.

Step 8) If the dimension of the auxiliary vector is lower than -1 scale the value of the auxiliary vector dimension according to the following linear transformation:

$$m = \big(-\alpha \times (minVel[i] - maxVel[i])\big) / \big(-1 - \min\,(\text{auxvect})\big)\ , \tag{3.41}$$

$$b = -1 - m \times\ -\alpha\ \times minVel[i]\ , \tag{3.42}$$

$$scaled(\text{auxvect[i]}) = m \times \text{auxvect[i]} + b\ , \tag{3.43}$$

where $minVel[i]$ is the pre defined velocity minimum value for the dimension i, $maxVel[i]$ is the pre defined velocity maximum value for the dimension i, $\min(auxvect)$ is the minimum value that each dimension of the auxiliary vector, $auxvect$, can take (defined at the beginning of the algorithm) and α is a reduction factor belonging to $]0,1]$ interval(defined at the beginning of the algorithm).

Step 9) If the dimension of the auxiliary vector is higher than 1 scale the value of the auxiliary vector dimension according to the following linear transformation:

$$m = \left(\alpha \times (minVel[i] - maxVel[i])\right) / \left(1 - \max(\text{auxvect})\right) \ , \tag{3.44}$$

$$b = 1 - m \times \ \alpha \ \times minVel[i] \ \ , \tag{3.45}$$

$$scaled(\text{auxvect[i]}) = m \times \text{auxvect[i]} + b \ \ , \tag{3.46}$$

where $minVel[i]$ is the pre defined velocity minimum value for the dimension i, $maxVel[i]$ is the pre defined velocity maximum value for the dimension i, $\max(auxvect)$ is the maximum value that each dimension of the auxiliary vector, $auxvect$, can take (defined at the beginning of the algorithm),and α is a reduction factor belonging to $]0,1]$ interval(defined at the beginning of the algorithm). Note that the reduction factor is equal for both “negative” and “positive” scaling to avoid any preference in a particular direction.

While all dimensions of the auxiliary vector are not evaluated.

Step 10) Sum the correspondent dimension of this vector to the velocity vector of the particle.

While all particles are not evaluated.

The idea in this technique is to let the relative distance between the particles decide if the velocity will be decreased or increased and how deeply will be this variation. Furthermore, this spreading technique will only be applied to the particles “very close”. The other ones will not be “harmed” and therefore the search procedure will not be significantly agitated ensuring a careful search of the actual zone in which the population is within. Further, the reduction factor is set in order to avoid excessive velocity variations, ruining the effect of the others components of the particle’s velocity on the next position computation.

The third technique is based in [28]. The socio-cognitive learning process defined in the standard PSO an also in EPSO is based on the particle’s own experience and the experience of the most successful particle. In [28] is added to PSO a new dimension to this approach: each particle also learns from the experience of the neighboring particles that have a better fitness than itself.

The results shown by the authors proved that this methodology allows to obtain better results than the traditional formulation of PSO.

The idea implemented in this thesis is to add a new term to the EPSO equation, which is used to calculate the particle´s new velocity, called neighbor term. This term is similar to the memory term, possessing its own weight which is also mutated thorough the generations, as it is represented in the following equation:

$$\boldsymbol{V}_i^{(k+1)} = w_{i1}^* \times \boldsymbol{V}_i^{(k)} + w_{i2}^* \times \left(\boldsymbol{b}_i - \boldsymbol{X}_i^{(k)}\right) + w_{i3}^* \times \left(\boldsymbol{b}_{\boldsymbol{g}}^* - \boldsymbol{X}_i^{(k)}\right) \times \boldsymbol{P}$$

$$+ w_{i4}^* \times \left(\boldsymbol{b}_{neigbor} - \boldsymbol{X}_i^{(k)}\right) \ , \qquad (3.47)$$

where $\boldsymbol{b}_i$ is the best point found by particle i in its past life up to the current generation, $\boldsymbol{b}_{\boldsymbol{g}}^*$ is the best overall point found by the swarm of particles in their past life up to the current generation, $\boldsymbol{b}_{neigbor}$ is the particle current generation nearest best position, $\boldsymbol{X}_i^{(k)}$ is the location of particle i at generation k, $\boldsymbol{V}_i^{(k)}$ is the velocity of particle i at generation k, w_{i1}^* is the weight conditioning the inertia term, w_{i2}^* is the weight conditioning the memory term, w_{i3}^* is the weight conditioning the cooperation or information exchange term, w_{i4}^* is the weight conditioning the new neighbor term, and $\boldsymbol{P}$ is the communication factor.

As it is clear in equation (3.47), to include the neighbor term in the calculation of the new velocity it is necessary to discover the nearest best particle in the current generation. To accomplish that, it is necessary to define a measure of the distance between particles and a measure of how much the neighbors are better than the particle in evaluation. In [28] a particle is considered to be the nearest best one if, in the whole population, it maximizes the following equation:

$$(Fitness(a) - Fitness(b))/|X_i(a) - X_i(b)| \ , \qquad (3.48)$$

where b is the particle to be assessed the nearest best particle, a is a general particle of the swarm, $Fitness(a)$ is the fitness of the particle a, $Fitness(b)$ is the fitness of the particle b, $X_i(a)$ is i^{th} dimension of the position of the particle a, and $X_i(b)$ is i^{th} dimension of the position of the particle b.

However, in [28], the particle position is not seen as whole and the selection of the nearest best particle is performed for each dimension of the position vector. In this thesis it is assumed that the nearest best particle gives it experience of “flight” in the form of the reliability case that it represents. Therefore, the “distance” between particles is measured using the equation (3.40) since now the distance between vectors is calculated. As a result, instead of selecting a nearest best particle for each dimension, is selected only the one which maximizes the following equation:

$$(Fitness(a) - Fitness(b))/\|round(X_a) - round(X_b)\|_2 \quad , \tag{3.49}$$

where b is the particle to be assessed the nearest best particle, a is a general particle of the swarm, $Fitness(a)$ is the fitness of the particle a, $Fitness(b)$ is the fitness of the particle b, and $\|round(X_a) - round(X_b)\|_2$ is the Euclidian norm of the vector defined by the difference of the rounded position of the particle a and the rounded position of the particle b.

Care must be taken when both particles represent the same case since, in these conditions, the Euclidian norm value is 0. When that happens, it is usual to assume a small constant value for the Euclidian norm, lower than the other possible distances (in this problem, the lowest distance, excluding "equal" particles, is 1).

3.7. Stopping criteria

Unlike the Monte Carlo methods, the PB methods are not statistical based. Therefore, it cannot be defined a degree of confidence which assures that the correct value of the index being estimated is within the correspondent interval of confidence (usually the estimates obtained by the Monte Carlo methods have a degree of confidence equal or higher than 95%). As it was previously shown in Chapter 2 the computation of the estimate of the reliability indices in the PB methods is made according to the equation (2.5). Therefore, the value of the index estimate grows as the number of the generations increases until a considerable quantity of the significant states for the index formation is not saved. When this occur the variation of the estimate in between generations is insignificant and the search procedure can be stopped.

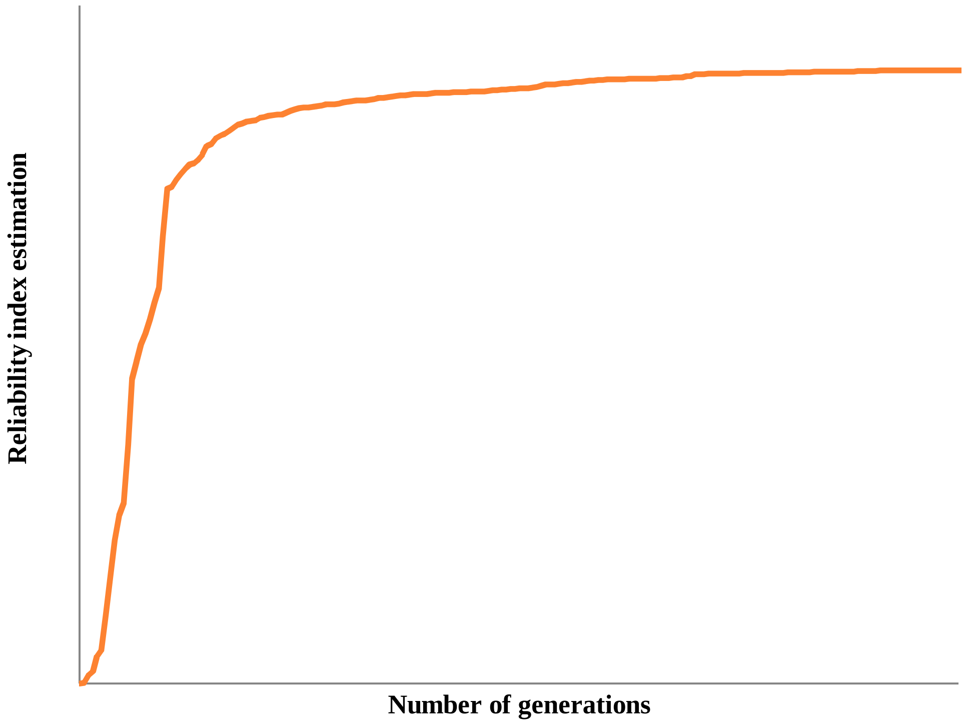


Figure 3.4. Illustration of the formation of the estimate of the desired reliability index as the number of generations increases.

In this particular problem, the estimates of the reliability indices are computed after obtaining a set of significant cases (group of similar states). This set is achieved using a fixed value of load in order to know if the particular case has load curtailment. Therefore, during the search procedure an estimate of an index is computed, which is only used to analyze the convergence of the process. The chosen index is the Expected Power Not Supplied (EPNS) and it is calculated as follows:

$$EPNS = \sum_{j \in S'} p_j \times \left(\bar{L} - Cap_j\right) \ , \tag{3.50}$$

where p_j is the probability of the failure case j, $\bar{L}$ is the value of the annual load curve peak, Cap_j is the generating capacity of the failure case j, and S' is the set of failure cases.

All in all the search process converges when the EPNS index stabilizes. This is said to be true if the relative change in the index stays below a threshold for a consecutive number of iterations:

$$\sum_k \Delta EPNS/EPNS < \epsilon \ , \tag{3.51}$$

where $\sum_k \Delta EPNS$ is the total change in the obtained in the EPNS estimate over k consecutive iterations, k being a predetermined integer, and ϵ is a predetermined threshold.

Another criterion is the maximum number of generations allowed, usually set considering the experience or sensitivity of the user. For a commercial application of a PB based method, this criterion is insufficient since the convergence depends on the initialization of the algorithm parameters (for instance, the dimension of the population). In this thesis the maximum number of generations criterion will be used for the reason that this is an academic work and above all for comparison with other methods which use this same convergence criterion.

3.8. Description of the EPSO reliability algorithm

After the introduction of the EPSO based algorithm core ideas as well as the basic reliability indices, now the steps of this methodology are explained in detail. This algorithm will follow closely the algorithm proposed in [22].

Step 1)

- Initialize the size of the population;
- Define the dimension of the problem;
- Obtain the following reliability parameters of the generators:
    - FOR;
    - MTTF;
    - MTTR.
- Obtain the power system annual load curve and figure its peak value;
- Define the threshold probability for case rejection;
- Initialize randomly the value of the weights, with a number sorted from the uniform distribution, $U(0,1)$;
- Initialize the learning parameter, τ;
- Initialize the probability of communication between particles;
- Define the range for the position variation;
- Define the range for the velocity variation;
- Define the fitness function to use:
    - If it is the ones define in (3.26), (3.27), (3.28) and (3.29) there is no setback;
    - If it is the one defined in (3.30) initialize the weight's change frequency T and determine the highest as well as the lowest value for each objective.
- Define the spreading techniques to use;

- Define the maximum number of iterations which the global best position is not changed;
- Define the reduction factor of the additional velocity as well as the maximum and minimum value that each dimension of the auxiliary vector can acquire;
- Initialize randomly the value of the neighbor term weight, with a number sorted from the uniform distribution, $U(0,1)$.

Step 2) Randomly fill the population with particles and compute for each case its probability, the number of copies and the generating capacity using the equations (3.23), (3.22), and (3.24), respectively.

Step 3) Create a case list and analyze if any particle of the current population meet the requirements to enter in it as described in the subsection 3.3. If the particle verifies the requirements, save the particle's case, as well its probability, number of copies and generating capacity.

Step 4) Update the fitness of all particles in the current population using the algorithm described in the subsection 3.5..

Do

Step 5) Update each particle's its nearest best particle using the methodology described in the subsection 3.6..

Step 6) Update each particle's auxiliary vector which will be used in the computation of its new velocity using the methodology described in the subsection 3.6..

Step 7) If the maximum number of generations without changing the global best position is reached, update its value using the methodology described in the subsection 3.6..

Step 8) Replicate the population one of time.

Step 9) Mutate the strategic parameters of the replicated particles using equation (3.2b).

Step 10) Each particle generates an offspring according to (3.47) and (3.4).

Step 11) Compute for each particles case its probability, the number of copies and the generating capacity using the equations (3.23), (3.22), and (3.24), respectively.

Step 12) Analyze if any particle meets the requirements for entering the list of significant cases as described in the subsection 3.3..

Step 13) Update the fitness of all particles using the algorithm described in the subsection 3.5..

Step 14) Perform elitist selection among each particle's both offspring.

Until the maximum generation criteria, presented in subsection 3.7., is verified.

Step 15) Compute the power system reliability indices suing the methodology described in the subsection 3.4.. Do not forget that each case represent a number of system states. Therefore when the calculation of the reliability indices the probability of each case must be multiplied by the number of copies of the case in other to account all the effective saved states.

3.9. Conclusions

In this Chapter the main lines of the EPSO based reliability algorithm and the validity of its application were presented. The reliability problem that will be addressed is the generating capacity adequacy assessment and therefore the proposed methodology was assembled around the particularities of this specific issue. Nevertheless is has to be mentioned that the spinal column of this methodology is based on the one proposed in [22].

One has shown that several fitness functions could be built in order to guide the search of the algorithm in the system state space. One aims at exploring the region that concentrates cases (sets of systems states) with high probability and leading at the same time to high load curtailment, because these most likely give a higher contribution to the reliability indices sought. From a theoretical point of view it is not easy to tell which one will work better in specific cases, and testing in practical cases must be done.

The next Chapter is devoted to applying and testing the competing models to the adequacy analysis of the generating capacity of a widely-known power system in order to discuss their advantages as well as drawbacks.

Chapter 4

Solving the problem with EPSO

In this Chapter the fundamental results of the methodology proposed in Chapter 3 will be presented: the EPSO reliability algorithm, which uses different schemes for creating diversity in the population in order to increase the number of feasible cases visited as well as to decrease the number of repeated ones. This can be accomplished by using different fitness functions in the fitness assignment procedure, by controlling the velocity of the particles, introducing new terms in the equation with which the particles new velocity is assessed (with the objective of spreading "similar" particles and to search more intensively the zone around each particle), and by constantly moving the population objective as well as erasing the memory of each particle every time that it represents a previous saved case.

First of all a brief reference will be made to the power system which will have its generating capacity evaluated. Then the effect of each fitness function proposed in Chapter 3 will be analyzed in detail, in their efficacy and the efficiency characteristics associated with the search of the meaningful cases. This research led to the selection of the best performing fitness function, under a mix of techniques to generate diversity in the search. The best performing fitness function was then tested in different experiments, isolating each factor contributing to the dispersion of the swarm. This hopefully allows one to improve the strategies to generate dispersion in a fashion that leads the search in the most promising region. The results of the EPSO reliability algorithm will be presented and a comparison with the results provided by another search method, the MSGA [22] approach, will be made, to demonstrate the advantages of the ideas defended in this thesis.

All the results presented were obtained with a C++ application developed for this purpose, which uses some pre-programmed modules of the traditional EPSO algorithm.

4.1. Institute of Electric and Electronic Engineers Reliability Test System 79

The proposed methodology will be tested in the evaluation of the adequacy of the generating capacity of the IEEE RTS -79 [29]. This power system was developed to satisfy the need for a standardized data base to test and compare results from different power system reliability evaluation methodologies such as the one proposed in this thesis. Although the reliability information for the transmission network is described, only the generation data is needed as well as the load model, in order to solve the problem addressed in this thesis.

The IEEE RTS-79 generation system is composed of 32 units. However there are only 9 different types of units. Therefore, with the proposed methodology, the dimension of the position vector of an EPSO particle will be 9, resulting in a decrease in length of 23 in relation to the dimension of a GA individual in the method proposed in [22].

The annual peak load for this system is 2850 MW. Moreover the system load model can be described in a weekly, daily or hourly basis. In this thesis the hourly basis will be adopted since the results published in [22] come from the same type of load model.

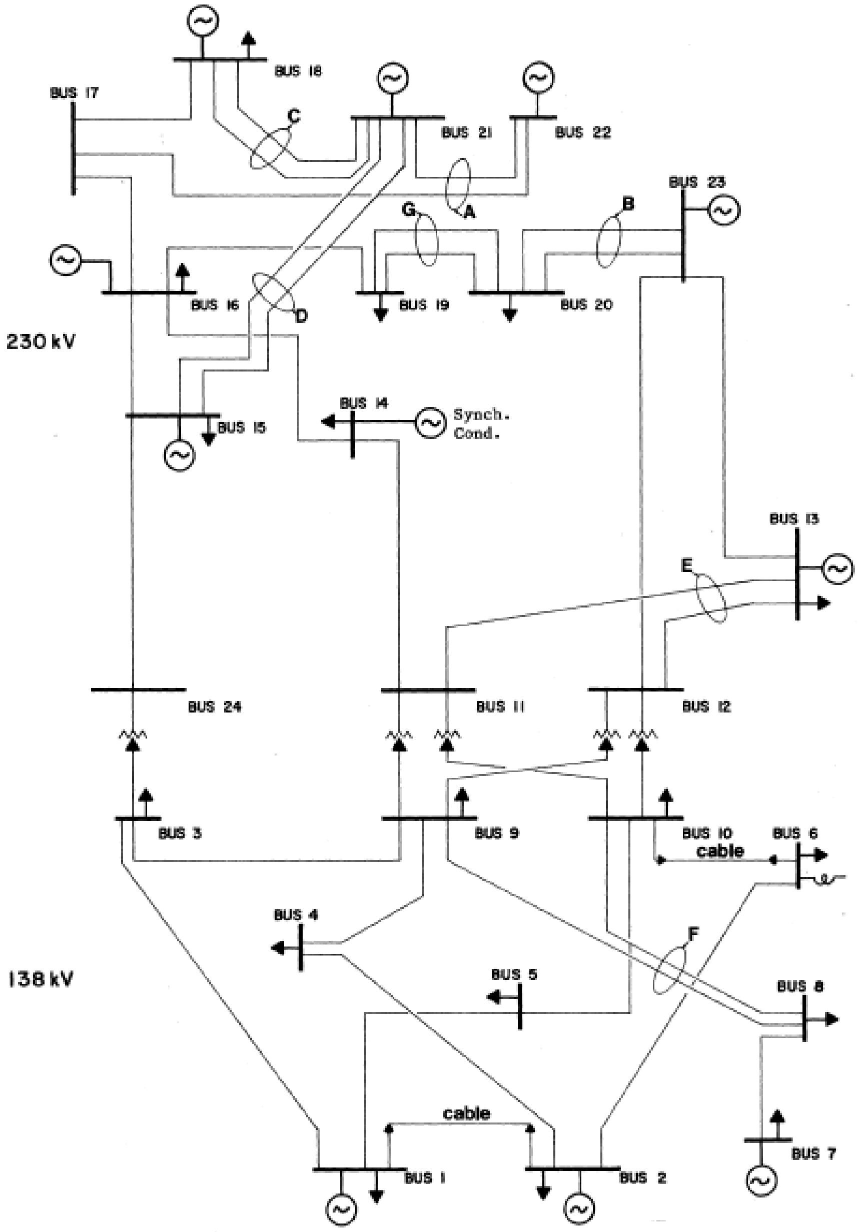


Figure 4.1. Topology of the IEEE RTS – 76 [29].

The value of the reliability indices for the generation capacity adequacy assessment problem assuming an hourly load model are known with good precision. The following table summarizes the value of the most important ones, obtained from [9] with analytical calculation.

Table 4.1. IEEE RTS -76 generating capacity adequacy reliability indices.

Adequacy reliability indices	
LOLE (hour/year)	9.394179
LOLF (occurrence/year)	2.019717
LOLD (hour/occurrence)	4.651236
LOEE (MWh/year)	1176.3

4.2. Evaluation of the performance of the different fitness functions

As it was seen in the previous Chapter, there are two most used types of fitness function in the PB models for reliability assessment: the maximization of the case probability and the maximization of the case probability x the load curtailment. Nevertheless these fitness functions can also take in account the number of states that an each case represents. Moreover, in Chapter 3, a new fitness function based on a bi-objective formulation has been presented, where simultaneously a dynamic weighted aggregation of two distinct objectives is maximized. Therefore one can identify four different types of fitness function which will be compared in this subsection: the two single objective pre-described fitness functions with the inclusion of the number of states that each case represents and two bi-objective formulations corresponding to the maximization of the aggregation of the two used single objective fitness functions and the maximization of the aggregation of the case probability x number of copies x load curtailment. The parameters of this simulation can be found in the Annex B.

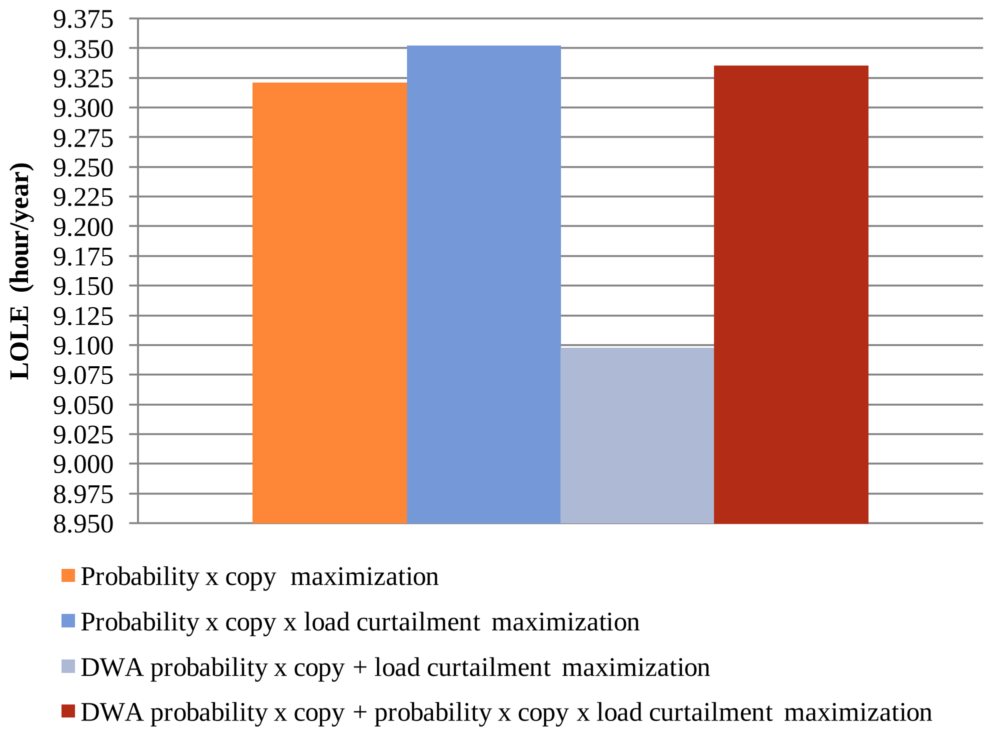
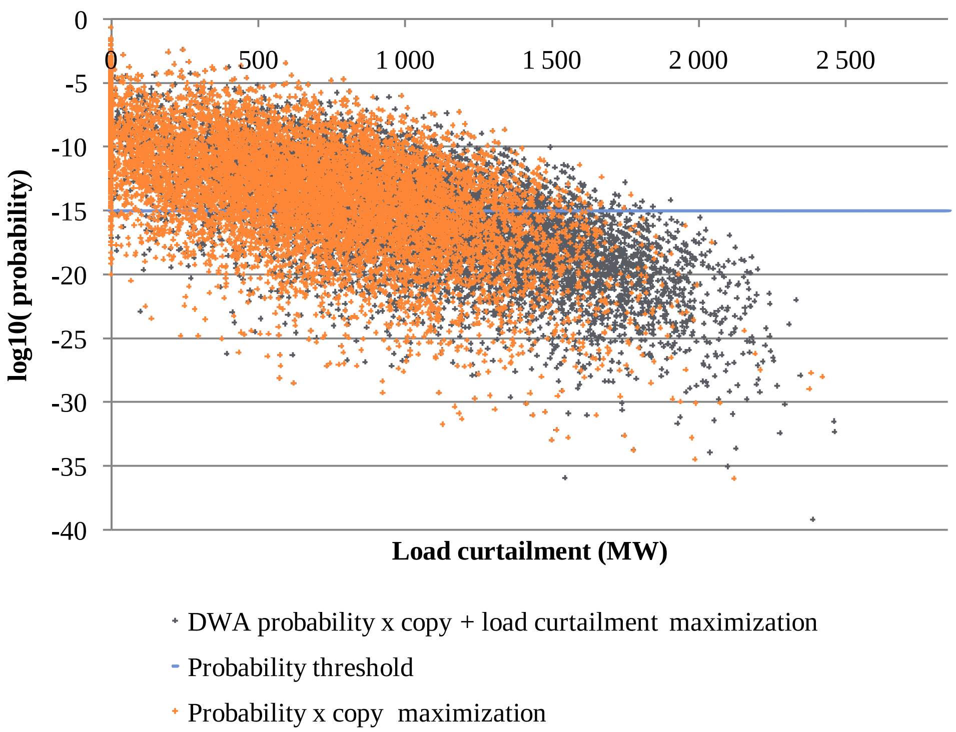


Figure 4.2. Comparison of the LOLE estimation by different objective functions.

The figure 4.2. illustrates the efficacy of the four different fitness functions. It seems that the maximization of the probability x number of copies x the load curtailment is the best fitness function. On the contrary, the DWA approach using the maximization of probability x number of copies + the load curtailment is the poorest one. In fact when the search is driven to the state space zone where the highest failures cases are, there is a decrease in the number of cases which verify the conditions to enter in the reliability list since their probabilities do not pass the threshold probability limit defined in the beginning of the search. The next figure illustrates this fact.



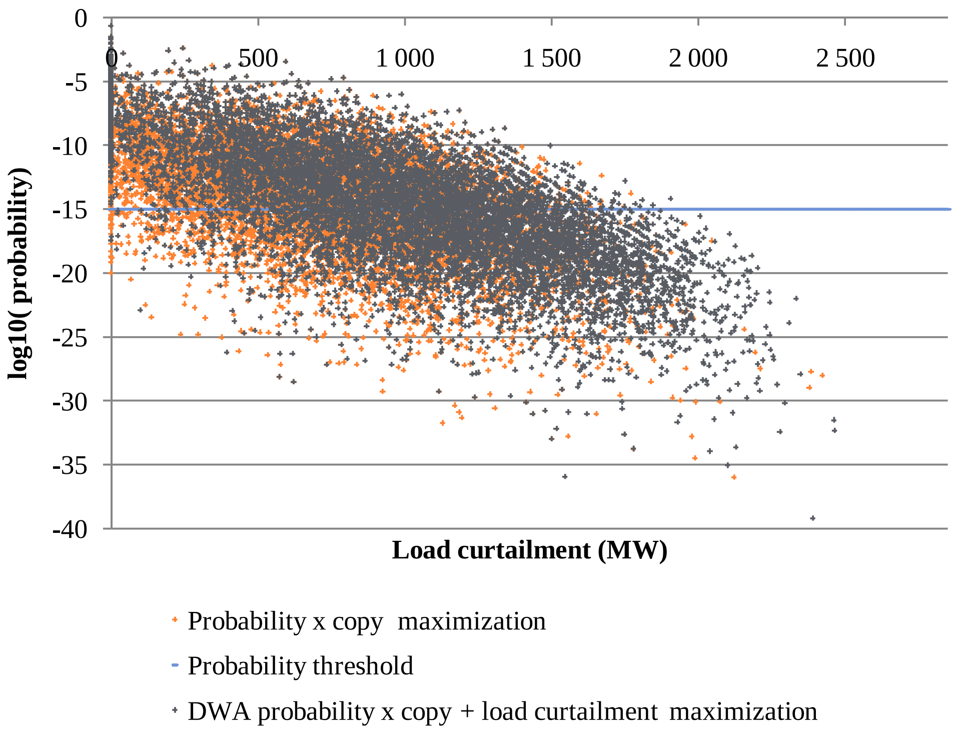


Figure 4.3. Representation of the attributes probability and load curtailment of the cases visited during the search. Notice that case probability is in a decimal logarithmic scale for better display.

Figure 4.3., constituted by two graphics to illustrate the superposition of common visited cases in the different two methods of fitness assignment, shows why maximizing the load curtailment leads to a fewer number of cases counted as significant (saved): most of the cases visited are beyond the probability threshold. Moreover this figure also confirms the idea that there is a clear trade-off between the case probability and its load curtailment. Then we can assume that the most contributing cases for the calculation of the reliability indices are near the Pareto front defined by those two objectives. However, the case probability has a remarkable effect in the weight of the correspondent case in the reliability indices calculation. As a result, the search must not be blindly conducted to the Pareto front: instead the search must be done in the region defined by the Pareto front and the probability threshold and conducted gradually from the cases with high probability to the ones with a low probability to achieve a maximum convergence performance, i.e., increasing the number of the most important cases visited and simultaneously increasing the approximation of the estimated indices to its real value.

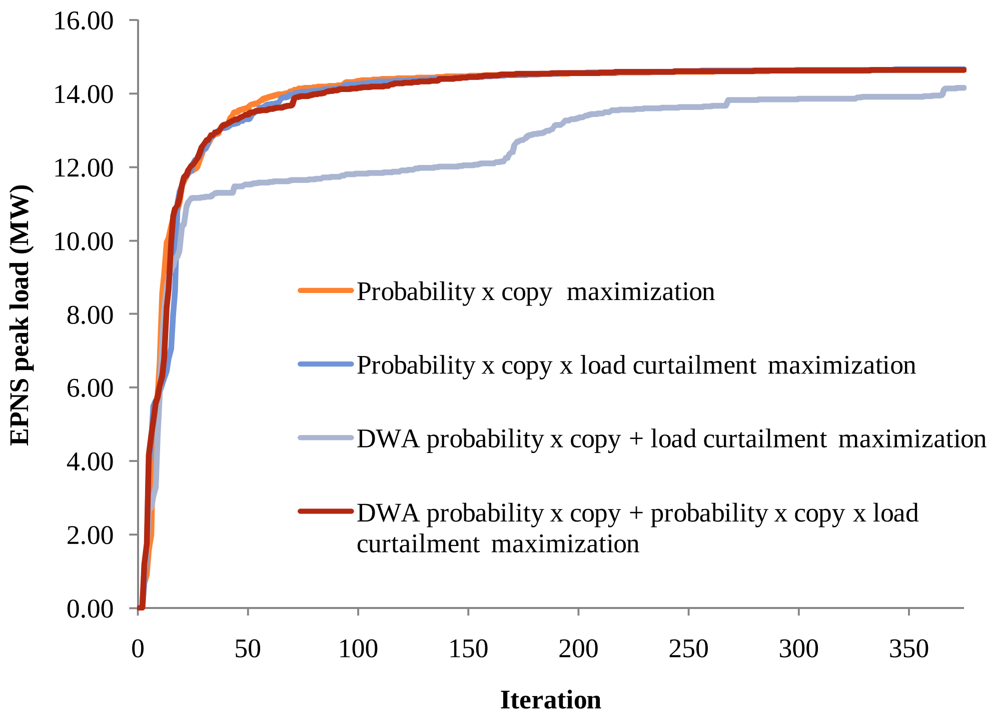


Figure 4.4. Evolution of the EPNS index for the peak load value as the number of the iterations increases.

Now is becomes evident why the DWA approach with using as objective the maximization of the case probability x its number of copies + the case load curtailment has the poorest behavior. If the reader cares to read the simulation parameters described in Annex B, he will notice that the DWA weights change frequency was set to 375, precisely to have only two periods where the probability weight is higher than the weight of the load curtailment. It was also chosen that the weight probability would begin with the number one. At this time the reader must have already discovered why the curve of the DWA probability x number of copies + load curtailment has this behavior: when the weight of the probability objective is higher that the weight of the load curtailment objective there is a sudden increase of the EPNS estimated value; on the contrary, when the weight of the load curtailment is higher than the weight of the probability the estimation of the EPNS indices almost stops, suggesting that the search was conducted to unimportant zones of the space state.

The other fitness functions have a similar convergence behavior. However why there is such a difference on their efficacy measured in figure 4.2.? It appears that when the search is conducted to the zones where there are the most probable states the process forgets the influence of the case generating capacity or, if the reader prefers, the case load curtailment. As it was demonstrated in Chapter 2, the reliability indices are calculated by convolution of the cases saved, counting, obviously, the number of copies each case, with the case generating capacity. For each hour of the load diagram, each case generating capacity is compared with the corresponding load value in order to identify the subset of the saved cases which must be considered as failure cases for the particular load value or, in other words, which has a generating capacity lower that the specific hour load value. One believes that the maximization of the case probability fails to encounter the cases with sufficiently low generating capacity in order to include the effect of the valley hours of the load diagram, having also a significant effect in the quality of the approximation to the loss of load probability during peak hours.

Table 4.2. Number of hours of the load model where LOLP is 0 for two different sampling strategies.

Fitness function	Number of hours
Probability x copy	464
Probability x copy x load curtailment	150

Table 4.2. illustrates a comparison of efficiency resulting from the application of two fitness functions guiding the search in the state space. The cases sampled and saved when using the probability x copy function lead to no load disconnection, when considering all the possible load steps of the load curve, in a much larger number of hours than if the cases were sampled and saved from using the other function.

The number of hours where the LOLP index obtained with the probability x number of copies x load curtailment objective is superior to the one obtained with the probability x number of copies objective, is equal to 7209. There is no further need to justify why the probability x number of copies x load curtailment fitness function performs better than the probability x number of copies objective.

This explanation also justifies why the DWA approach using these two objectives has an efficacy slightly superior to the probability x number of copies objective and somewhat inferior efficacy concerning the probability x number of copies x load curtailment objective: according to the weights change frequency, there are four different periods in the search process: in two of them the probability x copies objective has a weight superiority in relation to the probability x copies x load curtailment objective. Therefore, the search efficacy is degenerated mostly due to the two search periods where the population is driven to a search space region with high probability but with very low load curtailment values at peak level; thus the global contribution of these cases is uninteresting because for lower load values there will be no load curtailment.

In relation to the efficiency of the four fitness function approaches, in figure 4.5. is shown that the probability x number of copies x load curtailment is the one with the lowest ratio of unfeasible cases visited versus the total number of cases visited and with the highest ratio of the significant cases saved versus the number of the total cases visited despite having the second highest ratio of repetitions.

The significant cases ratio is almost similar in, at least, three of the four approaches. This reinforces the inference that the "quality" of the cases visited is perhaps more important than the quantity when it comes to the efficacy of the method. Actually the DWA approach using the probability x number of copies objective and the probability x number of copies x load curtailment has a lower significant cases ratio than the probability x number of copies maximization and still the DWA approach has a better efficacy than this one.

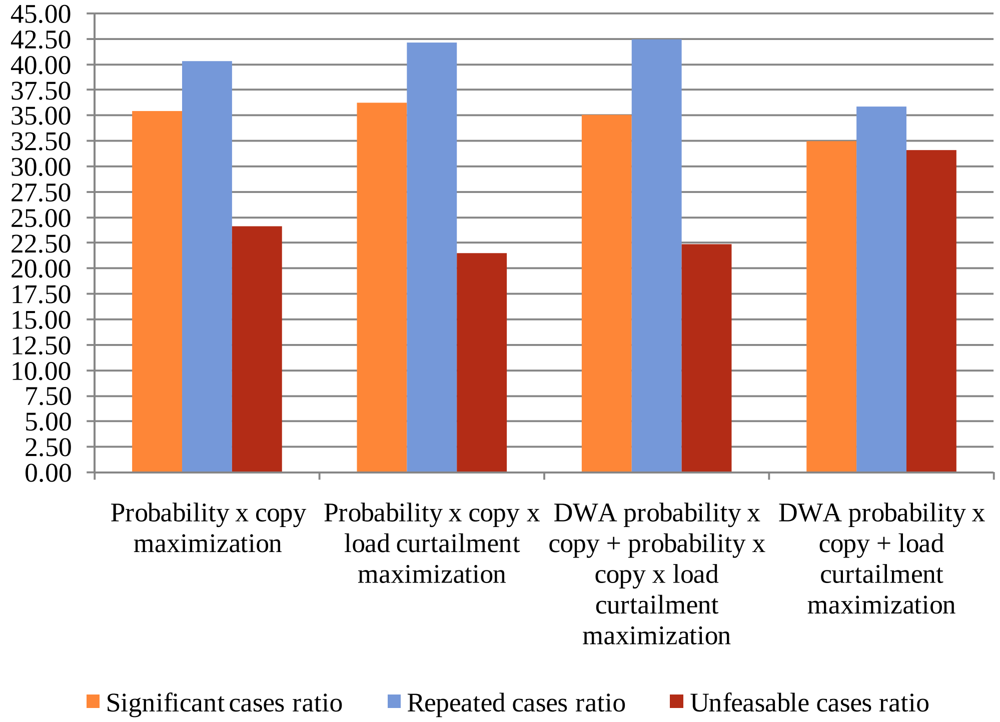


Figure 4.5. Ratio of the number of significant cases saved versus the total number of cases visited and the repeated cases versus the total number of cases visited and the unfeasible cases visited versus the total number of cases visited, all represented in percentage.

This figure also demonstrates that the DWA approach using load curtailment as one of the objectives has the poorest efficiency. In fact the highest unfeasible cases ratio of this approach is consistent to the fact that maximizing the cases load curtailment usually leads to zones of the search space where the correspondent case probability is lower that the pre-specified threshold. However the reader may say that this approach has the lowest ratio of repeated cases. This fact results from the changing of the search objective between two distant zones in the search space: the zone where the cases are more probable and the zone where the cases have the more load curtailment. Therefore is more probable that the cases visited are different from each other decreasing the ratio of repetitions.

To put it briefly the case repetition is a consequence of the type of search: if the population is confined to a tight search space, such as the one that it is driven the population by three of the four fitness functions, is obvious that there will be more repetitions. On the contrary, if the search space is wide, like the one defined by the DWA fitness function using the load curtailment as one of its objectives (in a certain moment of the search is told to the population that the fittest cases are the ones with high load curtailment, or, if the reader prefers, with low probability), the repetitions will decrease along with the ratio of significant cases saved.

In order to complete this evaluation, the effect of the weights change frequency must be also discussed. To accomplish that was selected the DWA probability x number of copies x load curtailment objective plus the probability x number of copies objective and five simulations were performed with the following weights change frequency:

- $T = 37.5$;
- $T = 75$;
- $T = 375$;
- $T = 1875$;
- $T = 3750$.

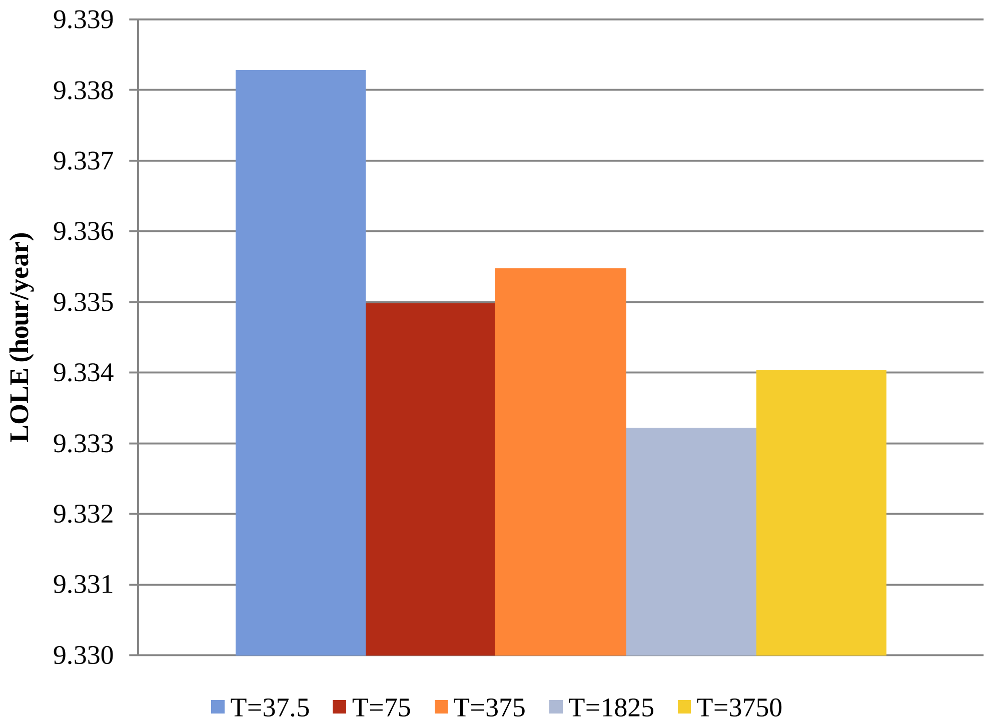


Figure 4.6. Comparison of the LOLE estimation for different weights change frequency, T.

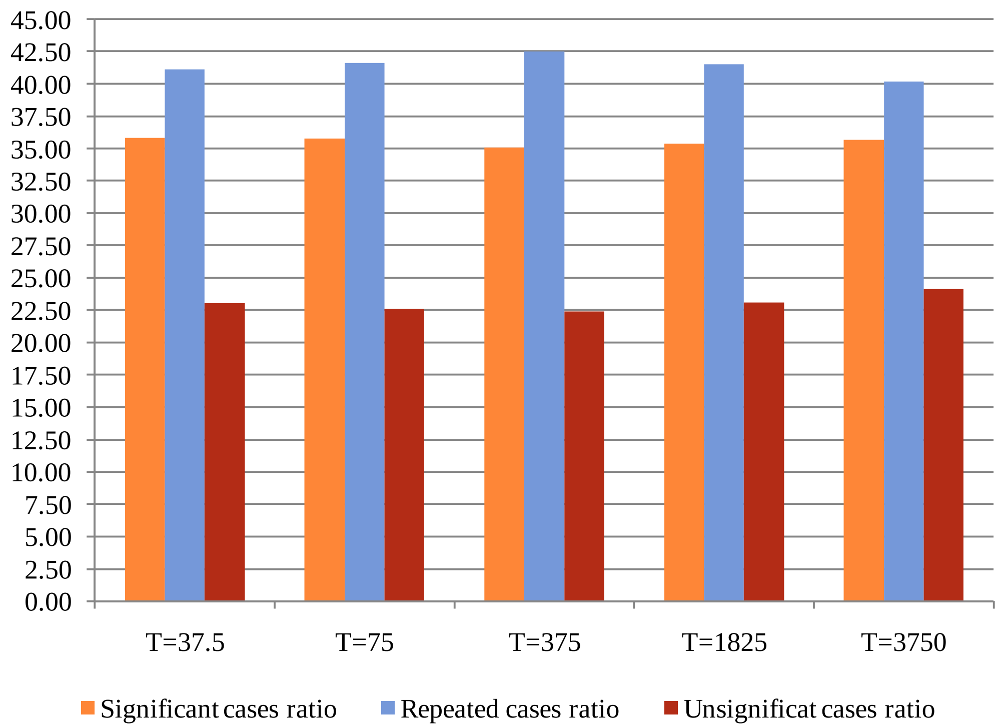


Figure 4.7. Ratio of the number of significant cases saved versus the total number of cases visited and the repeated cases versus the total number of cases visited and the unfeasible cases visited versus the total number of cases visited, all represented in percentage for different weights change frequency, T.

As it is shown in figure 4.6., it seems that a high rate of weights changing helps the DWA objective function efficacy. This result is easily explained if it is taken into account that the probability x the number of copies objective weight starts with the value 1, making this objective the first driving force. Additionally, as it was previously shown, maximizing the probability x number of copies x load curtailment gives a better result in terms of efficacy. Combining these two facts it is now evident why the efficacy degenerates as the weights change frequency increases: if the search remains for too long dominated by the probability x number of copies objective the reliability indices estimated are less accurate. This also explains why the indices estimates obtained by this DWA method are not as good as the ones obtained with the maximization of the probability x number of copies x load curtailment fitness function: there is always an influence of the probability x number of copies objective in the search process.

Then again the efficiency seems not to be strongly affected with the variation of the weights change frequency. Nevertheless the same old question arises: is the "quality" of the saved states equal for every frequency? Obviously not, as figure 4.6. demonstrates it.

4.3. Evaluation of the performance of the different spreading techniques

In order to evaluate the performance of the different spreading techniques first a search without any of them will be performed. This means that the best position of the population will be updated as in the traditional EPSO optimization algorithm as well as the best position ever found by particle in its past life (in the proposed search algorithm in this thesis, when a case is repeated it is searched in the memory of all particles in the current population if the repeated case is any personal best: if that is the case, the memory of the correspondent particles is erased) and there will be neither neighbor term nor additional speed. Then each individual spreading technique will be progressively added to the search algorithm, to make evident their advantages as well as their disadvantages.

The simulation parameters are equal to the ones described in Annex C. Nevertheless, only one fitness function will be used in all simulations, the one previously identified as the most promising: the maximization of the probability x number of copies x load curtailment.

4.3.1. Impact of forgetting the global best

Firstly the impact in efficacy and the efficiency of forgetting the best particle throughout the search procedure will be measured as well as the variation of these performance indicators with the variation of the number of iterations where the global best position is maintained equal. Notice that the memory of the particles is updated according to the proposed algorithm in Chapter 3. Therefore two modifications to the original EPSO algorithm can be identified: in the way that the population best particle is updated and the reset of the particle's memory every time that it represents a previous saved case.

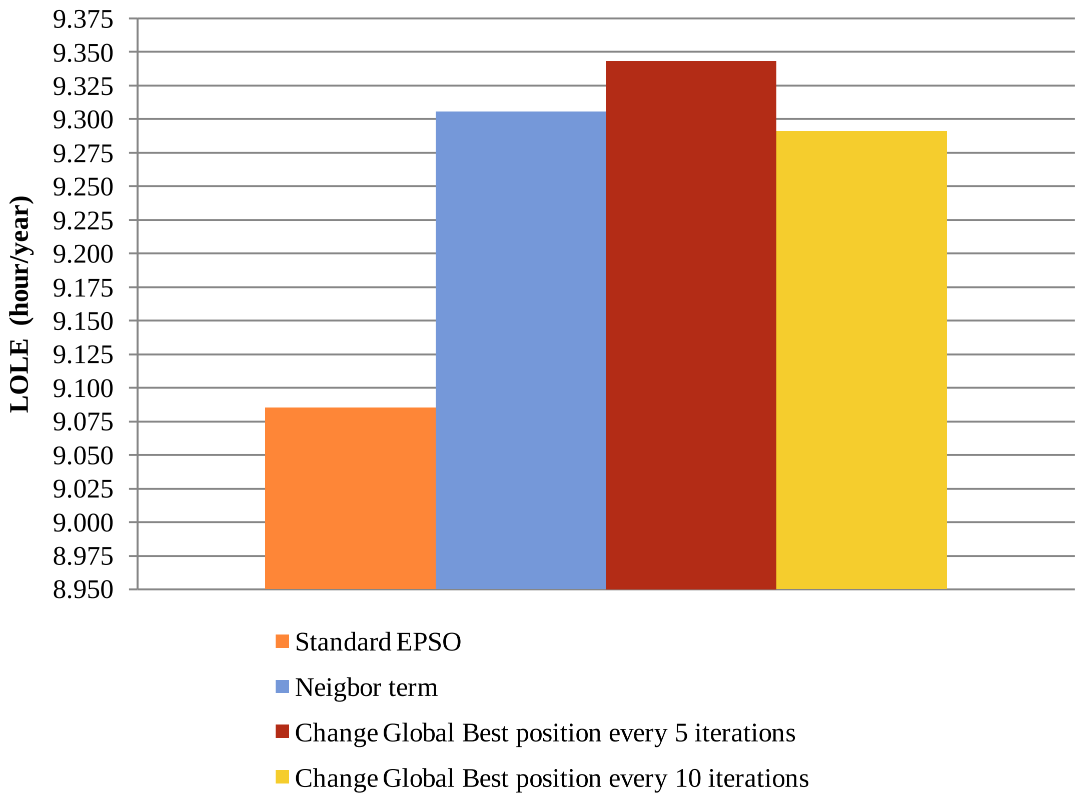


Figure 4.8. Comparison of the LOLE estimation for different number of iterations where the global best position is unchanged.

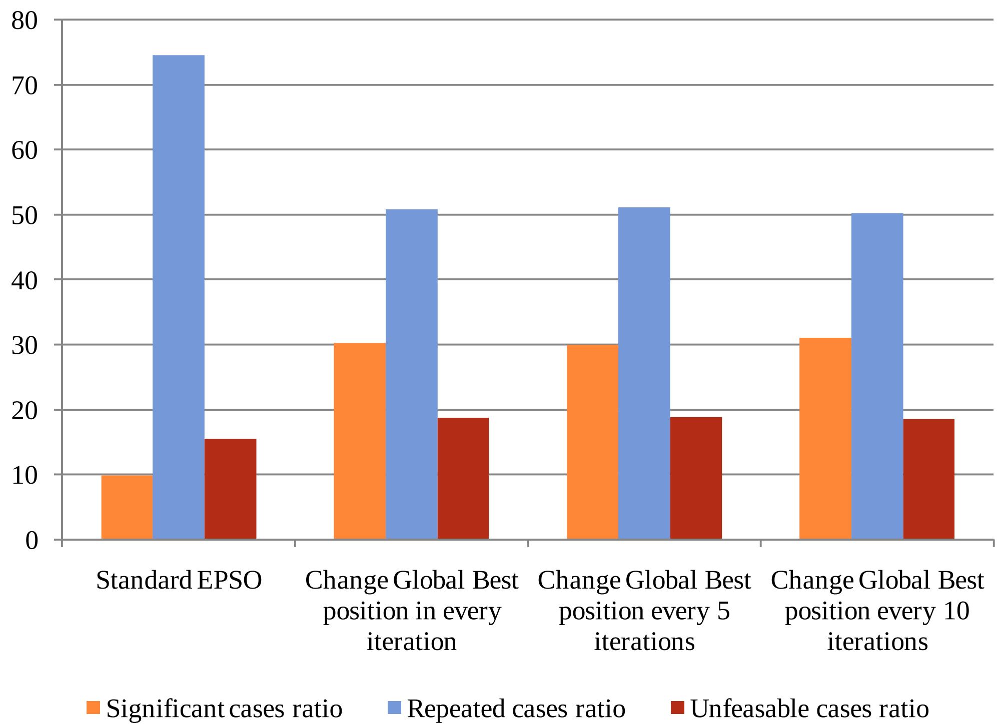


Figure 4.9. Efficiency ratios, in percentage of the total number of cases visited, for different number of iterations where the global best position is unchanged.

As it is shown in these last two figures, the first spreading technique is a success. Not only the efficacy was improved but also the efficiency of the search was increased dramatically especially in the ratio of significant cases saved as well as the ratio of repeated cases. This result is consistent with the principle that the population objective is always changing due to the process of fitness assignment: the cases already visited are given a small fitness to reduce the probability of appearance in the next generations. Therefore changing the population best particle according to the fitness of the particles of each generation is the same as changing the population global attractor encouraging the exploitation of new zones of the search space. This clearly favors the diversity in the population, unlike in traditional EPSO, were the high rate of repeated cases is mainly due to the restriction of the search to the zone of the particle ever found by the population. Moreover, changing the global attractor improved noticeably the efficacy of the search, since not only the number ratio of significant cases has improved but also its “quality” in terms of the reliability indices estimation.

In relation to the effect of maintaining the global best position for a certain number of generations, it seems that this number cannot be too high. If that occurs, the penalty is the decrease of the search efficacy since the search is trapped in a zone for a too long time. On the other hand always changing the objective can also decrease the efficacy since the population has not enough time to perform an effective search in a determined zone. Nevertheless there is a slight difference between the efficacy of constant global best position changing and the efficacy of changing the global best position every five generations. Therefore any number between 1 and 5 is considered a feasible choice for the number of generations without updating the population global attractor.

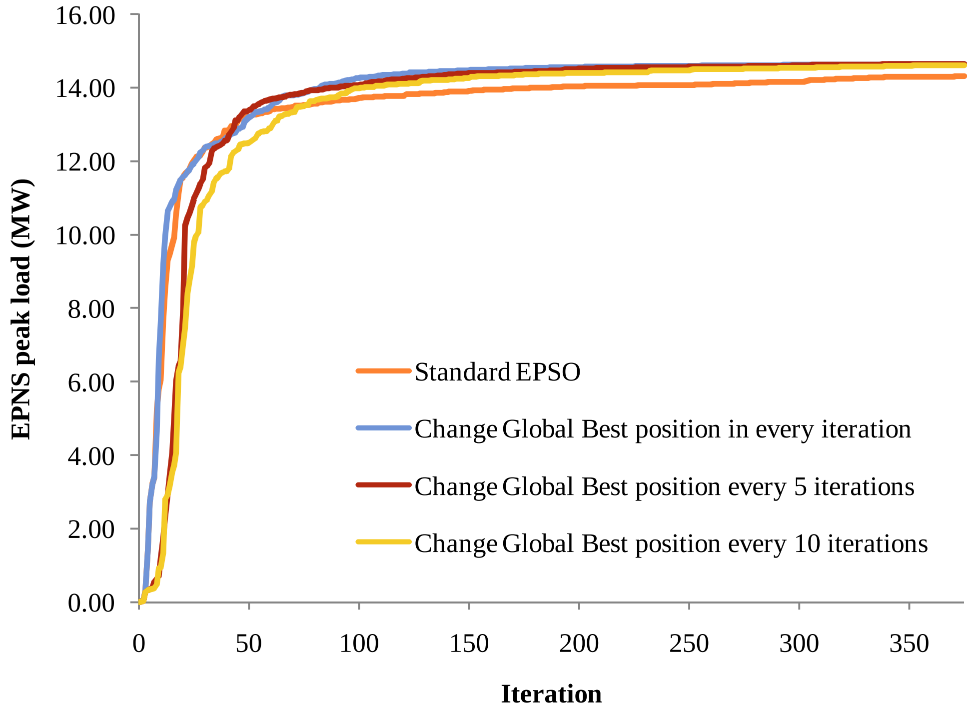


Figure 4.10. Evolution of the EPNS index for the system peak load considering different number of generations where the global best position is unchanged.

The figure 4.10. shows that, although the standard EPSO has a remarkable initial convergence speed, probably due to the aggregation of significant cases on the path to the best position in the search space, this same convergence speed slows considerably when the population reaches the best position in the search space. The same figure also shows that there is a decrease in the convergence speed when the number of iterations with the best position remaining unchanged increases. Nevertheless in the end of the pre-set number of generations, the estimate of the EPNS index for the peak load seems to be very close for the three different rates of changing the global best position, the most similar being the constant changing of the global best and the changing of this position every five generations.

4.3.2. Extra velocity terms

The next spreading technique to be presented is the one which adds an extra velocity to the particles when they are overlapped or very close. As it was explained in Chapter 3, this technique aims not only to diminish the number of repeated cases but also to push similar particles to another zone of the search space to increase the rate of feasible particles saved.

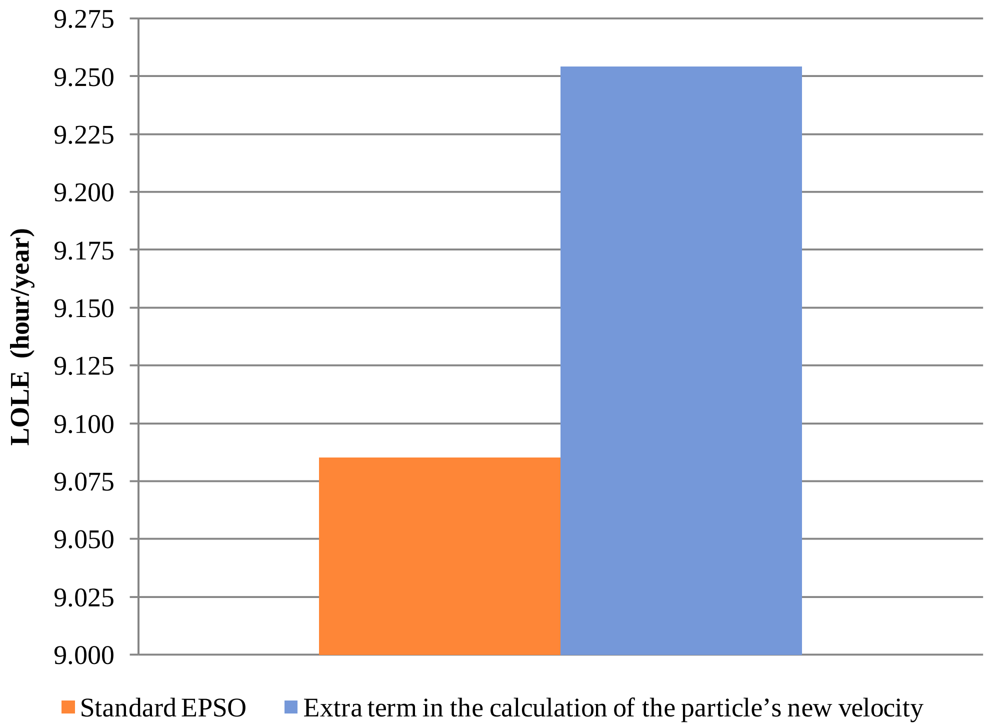


Figure 4.11. Comparison of the LOLE estimation between standard EPSO methodology and modified version using an extra term in the calculation of the particle's new velocity.

As it is shown in figure 4.11., this spreading technique, as it was expected, leads to a better efficacy than standard EPSO formulation. Notice that, in this simulation, there is neither forgetting of the population best particle nor resetting of the particles memory. The success of this spreading technique comes from the perception that, after a given number of generations, the population is confined to the zone of the best position without sufficient velocity to escape or to find other feasible cases. Therefore, adding this velocity term proportional to the number of "similar" particles and with a direction determined by the relative position of these particles, leads to a run off of the concentration zone, hopefully to a zone where there are other cases which are important to the reliability indices estimation. Nevertheless, it is essential to understand that the global attractor is not changed. After a few generations the particles are attracted to this point and its velocity reduced. Then, when they are again very close the additional velocity troughs back the particles outside the best zone. This process is repeated over and over along the search process until the maximum number of generations is reached.

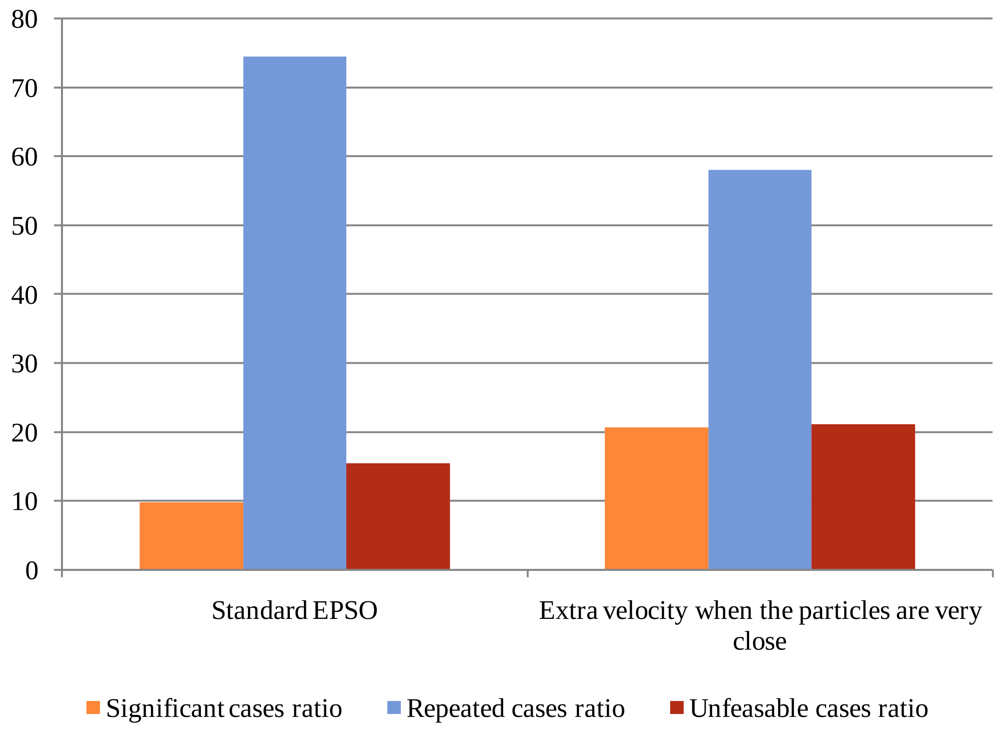


Figure 4.12. Efficiency ratios, in percentage of the total number of cases visited of the standard EPSO methodology and a modified version using an extra term in the calculation of the particles new velocity.

The limitation of the value that this additional velocity term can acquire has the following objective: since the feasible particles are near the zone of the best case, there is a higher chance of finding more feasible cases by adding a small value of velocity than a high value. This affirmation is consistent with the results shown in figure 4.12.. As it is exposed, this spreading technique not only improves the number of feasible cases saved but also decreases the number of repeated ones supporting the hypothesis that the number of repeated cases results from a restriction of the population to a specific zone of the search space.

Figure 4.13. also demonstrates how this technique improves the efficacy of the search. After a few number of generations, the population of the standard EPSO is stagnated at the zone of the best case, resulting in a high number of repeated cases. This affects not only the efficiency but also the efficacy. Adding the proposed velocity term, there is an increase of significant cases saved being these cases extremely important for the reliability indices computation.

Figure 4.13. also shows an initial gain in convergence velocity of the standard EPSO method in relation to the proposed spreading technique probably due to the interference of the new velocity term in the convergence to the optimum (it seems that the path of the population towards this optimum in the traditional EPSO formulation point has a positive effect in the efficacy of the search). Nevertheless as figure 4.14. shows, after a few number of generations the ratio of

repetitions decreases in the proposed spreading technique and as a consequence the number of important cases saved increases, improving the efficiency of the search.

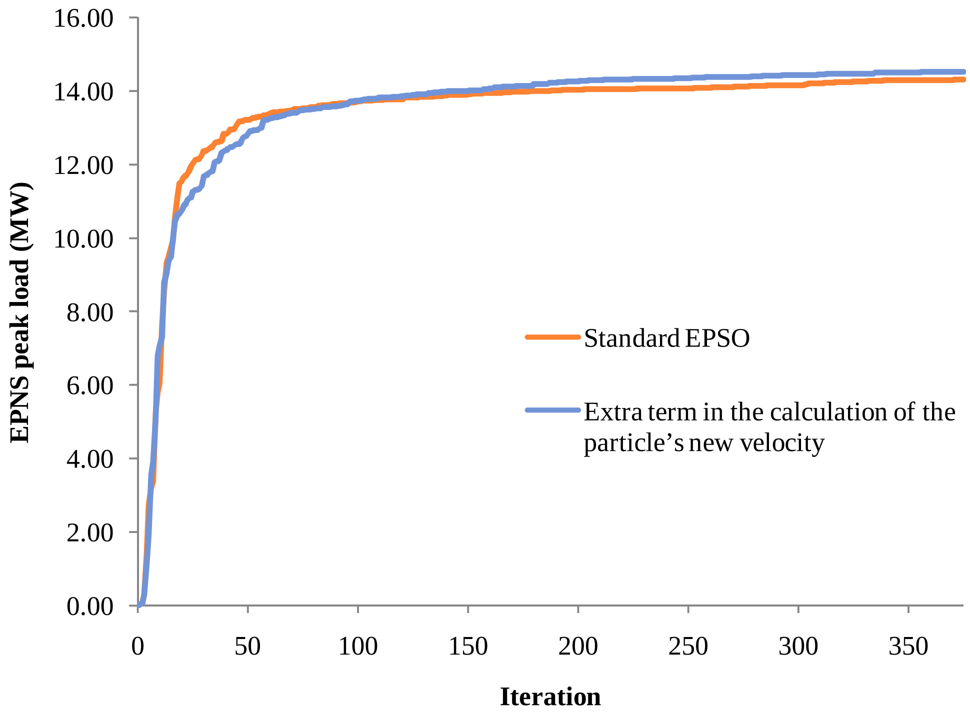


Figure 4.13. Evolution of the EPNS index for the system peak load between the standard EPSO methodology and a modified version using an extra term in the calculation of the particles new velocity.

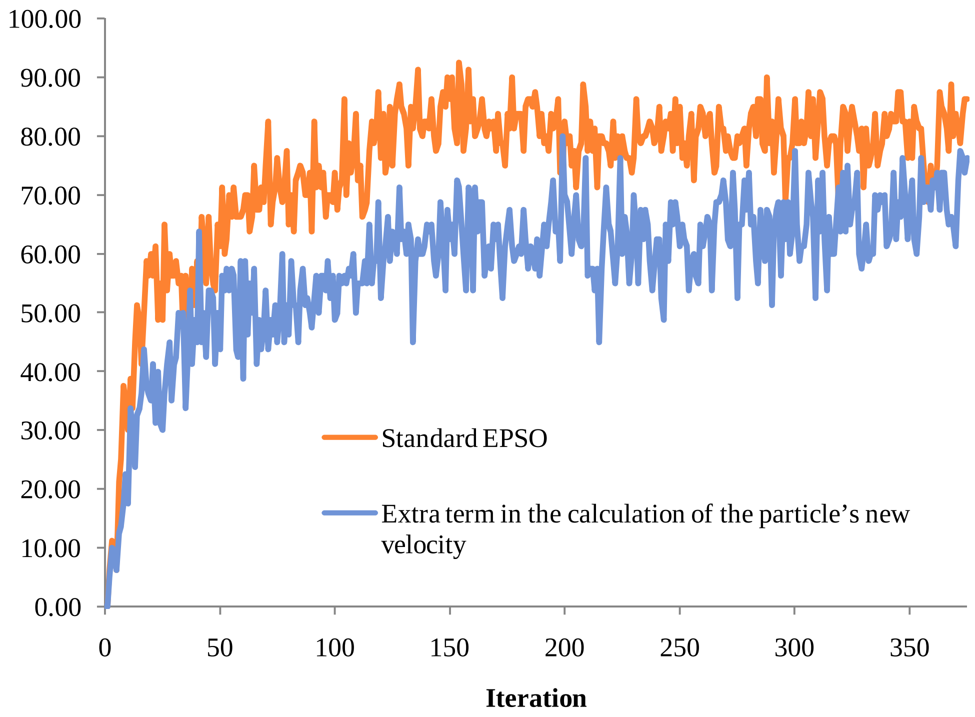


Figure 4.14. Evolution of the ratio of case repetitions in percentage of the total number of case visited in each generation for the standard EPSO methodology and for a modified version using an extra term in the calculation of the particles new velocity.

4.3.3. Use of neighbor terms

Finally the last spreading technique which needs to be evaluated is the use of the neighbor term introduced in Chapter 3. The idea implemented results from the observation of the actual social learning behavior of the particles: often they gain knowledge by observing the trajectory of the nearest ones. The objective of this methodology is to provide a more intensive search of the zone around the particles by adding an attractive force that may point to the best nearest one.

As it is illustrated in figure 4.15., this spreading technique improved considerably the quality of the reliability indices estimation. Furthermore, the estimative obtained is even better than the one calculated with the previous spreading technique.

The reason for that success can be assigned to the concept of a meticulous search in the zones near each particle. Nonetheless the meaning of distance and the meaning of best in this spreading technique must be understood. The determination of the closest best particle is necessary to find out a distance and a measure which establishes how good the neighbor is. The interesting particularity is that the distance is measured in the decision space, which corresponds to the space where the number of equal generators which are in the state on is represented. On the other hand the measure of how good the neighbor is, is determined in the attribute space or, if the reader prefers, in the space where the case probability is represented as well as its load curtailment. Therefore this concept of neighboring, which relates the decision and the attributes space leads to an significant increase not only of the efficacy of the search but also of the convergence speed which is clearly represented in the figure 4.17..

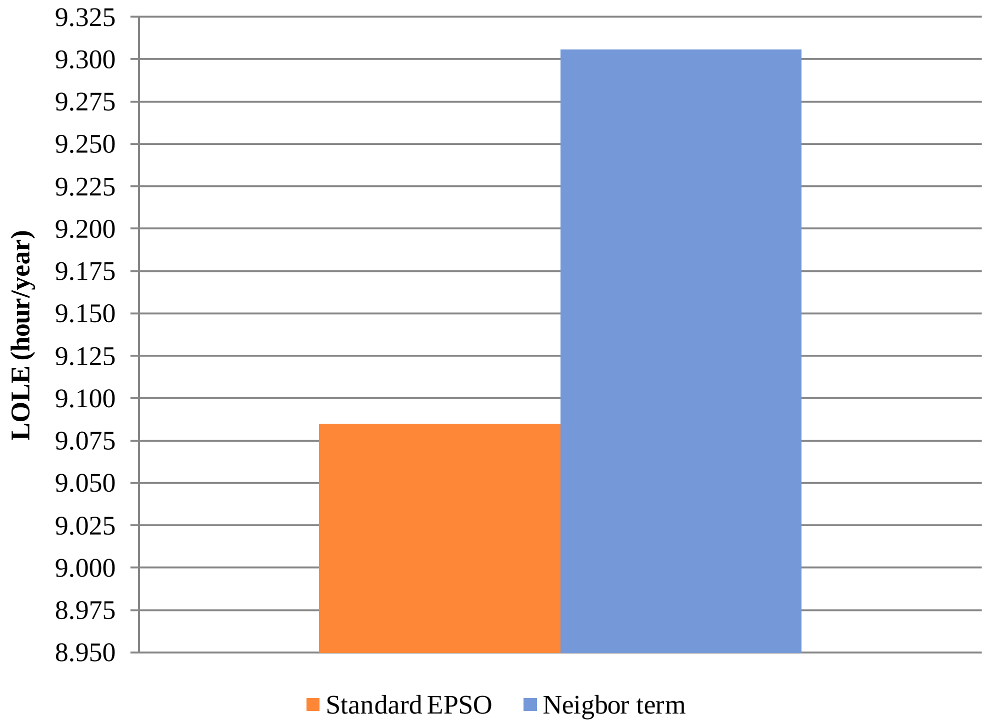


Figure 4.15. Comparison of the LOLE estimation between standard EPSO methodology and modified version of this methodology using an extra term in the calculation of the particle's new velocity named neighbor term.

Another interesting feature is associated to the concept of neighborhood. Remember that, after classifying a case as a repeated one, its fitness is assigned a very low value in order to decrease the probability of appearance in the subsequent generations. This is always valid even if the position of the best particle in the population is forgotten or even when the memory of the particles is erased. This concept of fitness assignment leads to a constant change of the neighbor structure of the search: for instance if in the previous generation a feasible case was visited for the first time and in the current generation this same case is visited again, it will be considered a repeated case and will be assigned a very small value as fitness. Therefore, in the previous generation this case was a point of attraction differing for the current generation where this case is no longer classified as a feasible one. Considering these facts the structure of neighboring is always changing as well as the nearest best neighbor of the particles. This seems to be the main reason why this spreading technique works so well: the rate of unfeasible cases visited stays equal, the rate of significant cases doubles and there is a decrease of the number of repeated cases. Moreover this technique is the only one in which it is evident an effective acceleration of the convergence speed right from the beginning of the search process.

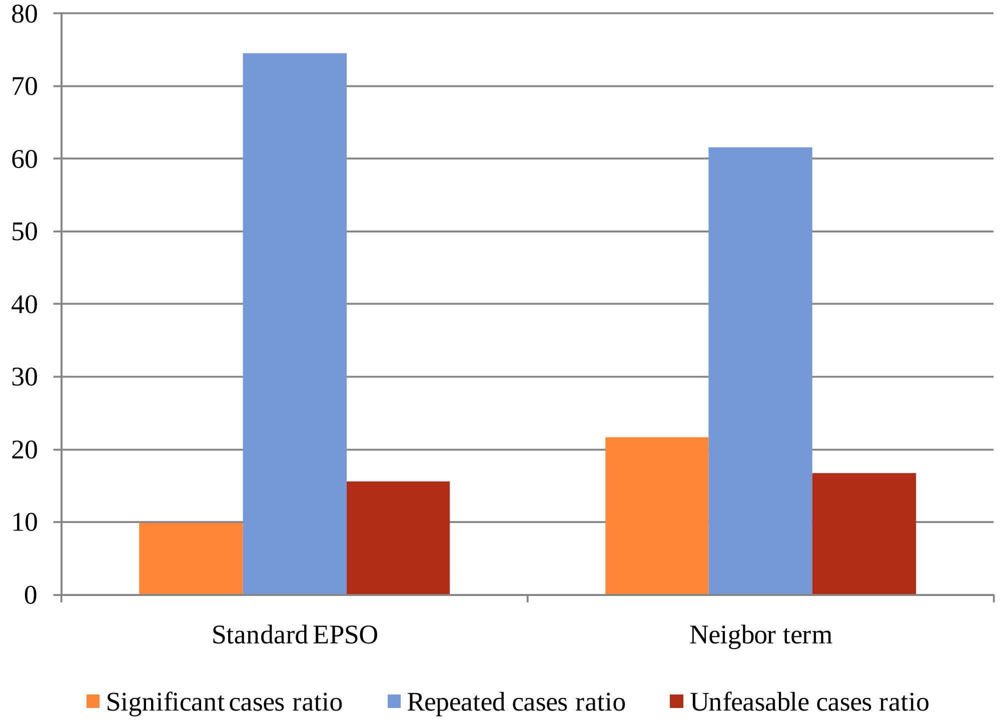


Figure 4.16. Efficiency ratios, in percentage of the total number of cases visited of the standard EPSO methodology and a modified version of this methodology using an extra term in the calculation of the particles new velocity named neighbor term.

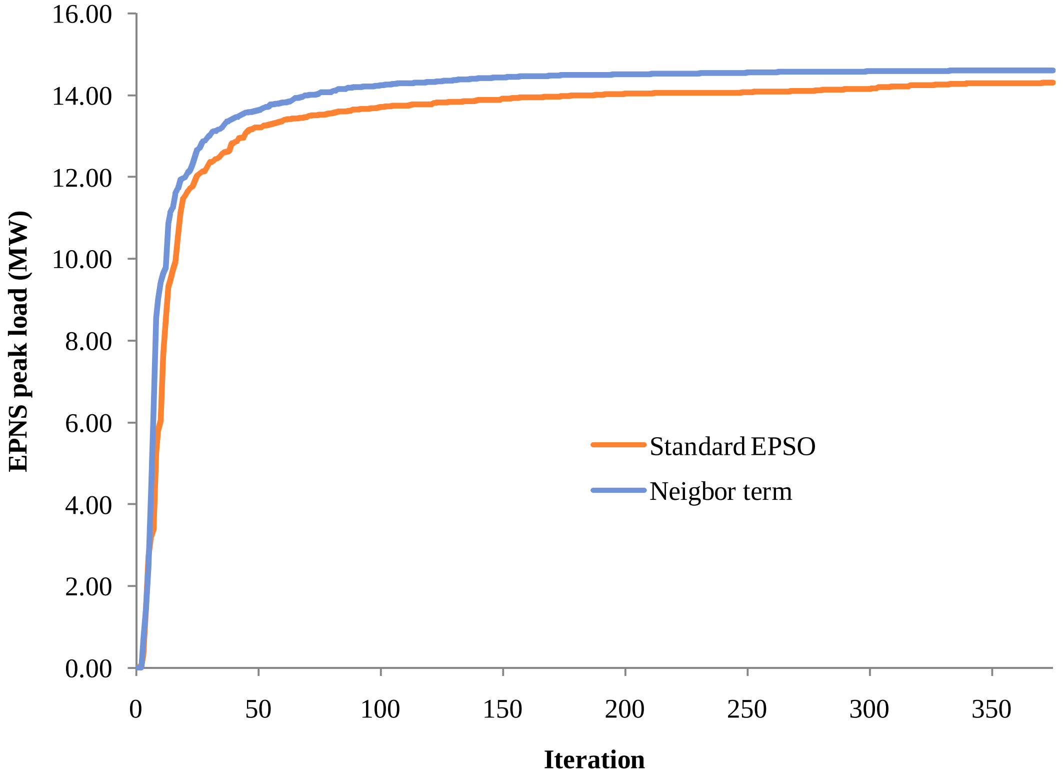


Figure 4.17. Evolution of the EPNS index for the system peak load between the standard EPSO methodology and a modified version of this methodology using an extra term in the calculation of the particles new velocity named neighbor term.

In these last two figures is obvious the gain in convergence speed and the gain in efficacy that this spreading technique provides.

4.4. Discussion of the best methodology

In this section the results of the new EPSO based method will be presented and discussed with the objective of finding the states contributing the most for the reliability indices computation. This method will include the mentioned three spreading techniques since, as it was proved, they improve the performance of the search.

First, the conditions or the parameters of the three spreading techniques must be defined. To be coherent with the analysis previously made, the parameters of this simulation will be equal to the ones in the earlier section, except for the fact that they will all be included at the same time in the search. The simulation parameters can be found in the Annex C. Furthermore, it is necessary to define the number of generations, or if the reader prefers, the number of iterations where the global best position remains unchanged. This parameter was set to 1.

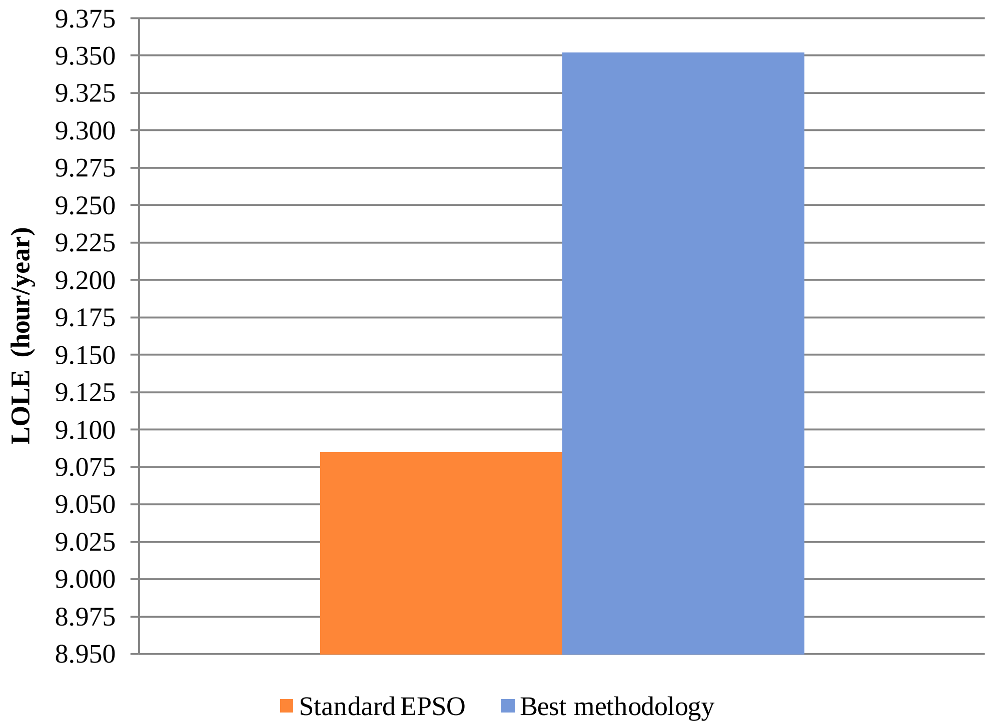


Figure 4.18. Comparison of the LOLE estimation between standard EPSO methodology and modified version considered the best search methodology which incorporates the tree spreading techniques.

In figure 4.18. the efficacy benefits of using the combination of the three spreading methods become evident. Notice that there is a significant gain in the estimation of the reliability index LOLE by using the so called “best methodology”. Therefore it can be concluded that the three methods for improving the diversity combined with the maximization of the probability x number of copies x load curtailment increase the efficacy of the search. This is also evident in figure 4.20. in which the evolution of the index EPNS along the search procedure for the two discussed search methods is illustrated: after a few generations the convergence speed of the EPSO methodology with swarm spreading starts to increase due to the effect of these methods which prevent stagnation of the search procedure in one zone of the state space, avoids the repetition of the previous saved cases and provides a more intensive search in the zones near each particle.

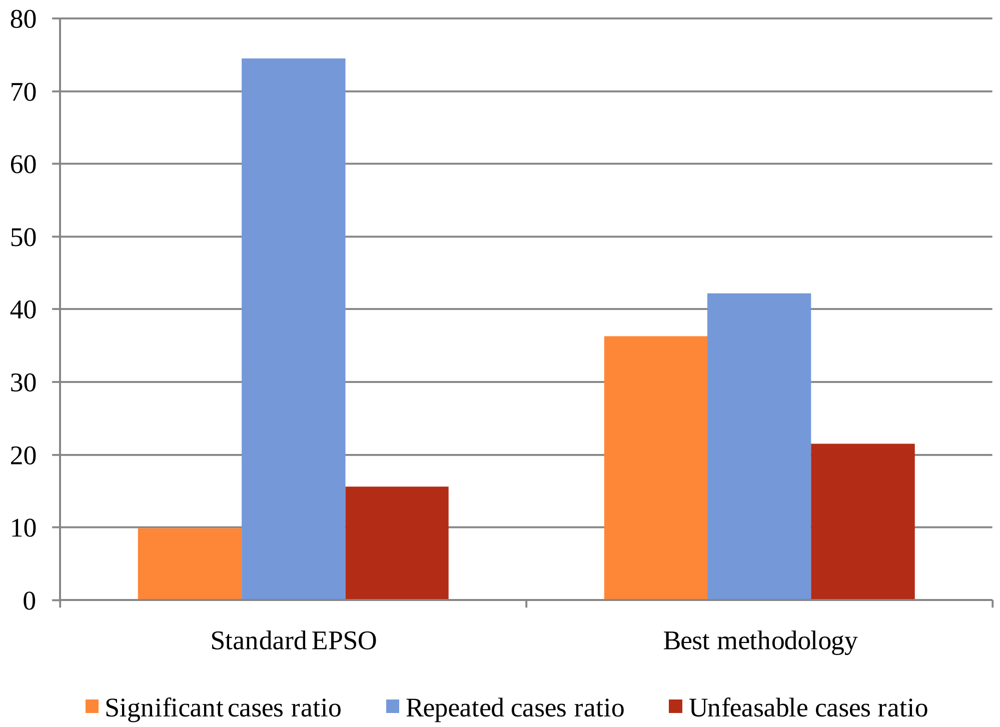


Figure 4.19. Evolution of the EPNS index for the system peak load between the standard EPSO methodology and a modified version considered the best search methodology which incorporates the tree spreading techniques.

As it was also expected the efficiency of the search was also incremented. In figure 4.19. it is shown that with the use of the combined methods of spreading the significant cases ratio have increased more or less 27 % in relation to the traditional EPSO methodology. Moreover the number of repetitions decreased almost 33 %. And the good news is that the new saved cases were extremely significant for quality of the estimative of the LOLE index. Note that the number of unfeasible cases visited has increased. Probably this is a consequence of the additional velocity term due to cases repetitions when the search is confined near the threshold rejection probability value. Nonetheless the other two rates became optimized. Therefore it seems that this methodology increased the efficiency performance of the search.

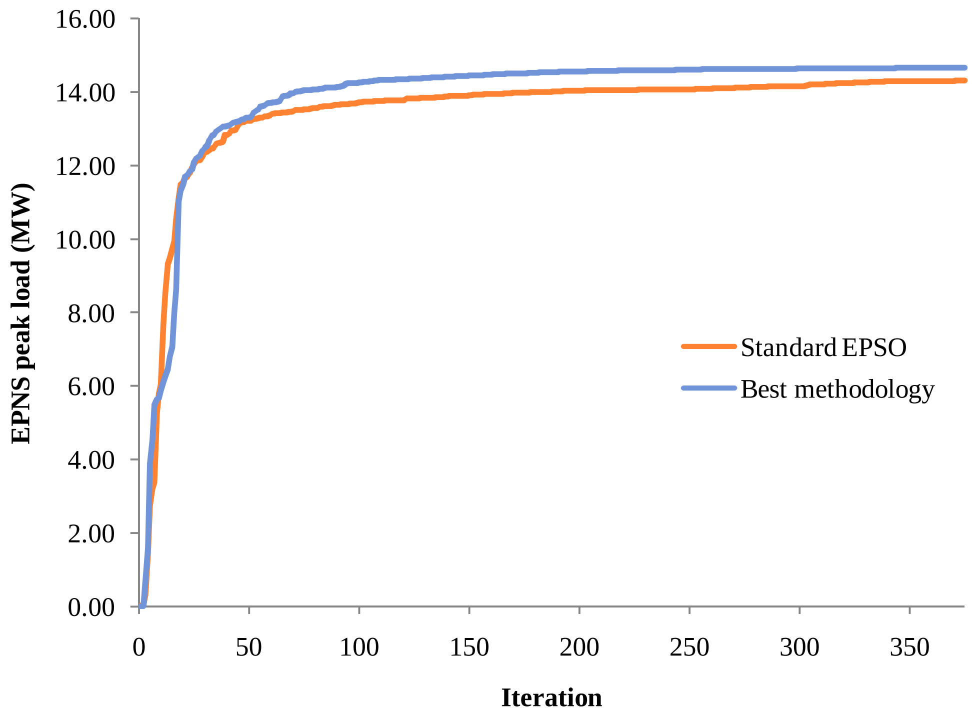
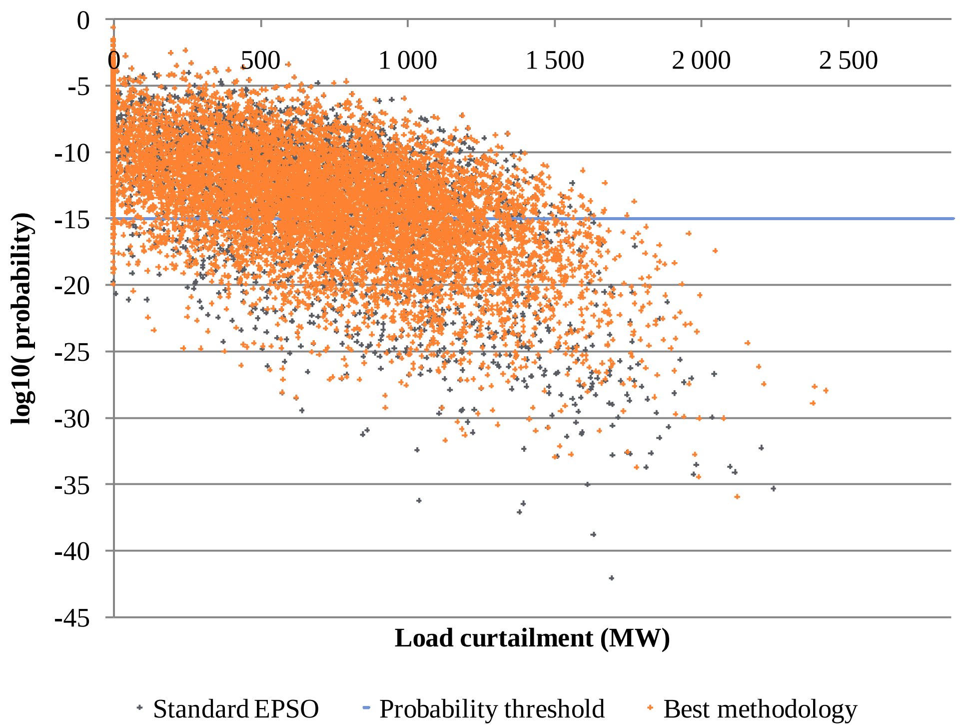


Figure 4.20. Evolution of the EPNS index for the system peak load between the standard EPSO methodology and a modified version considered the best search methodology which incorporates the tree spreading techniques.



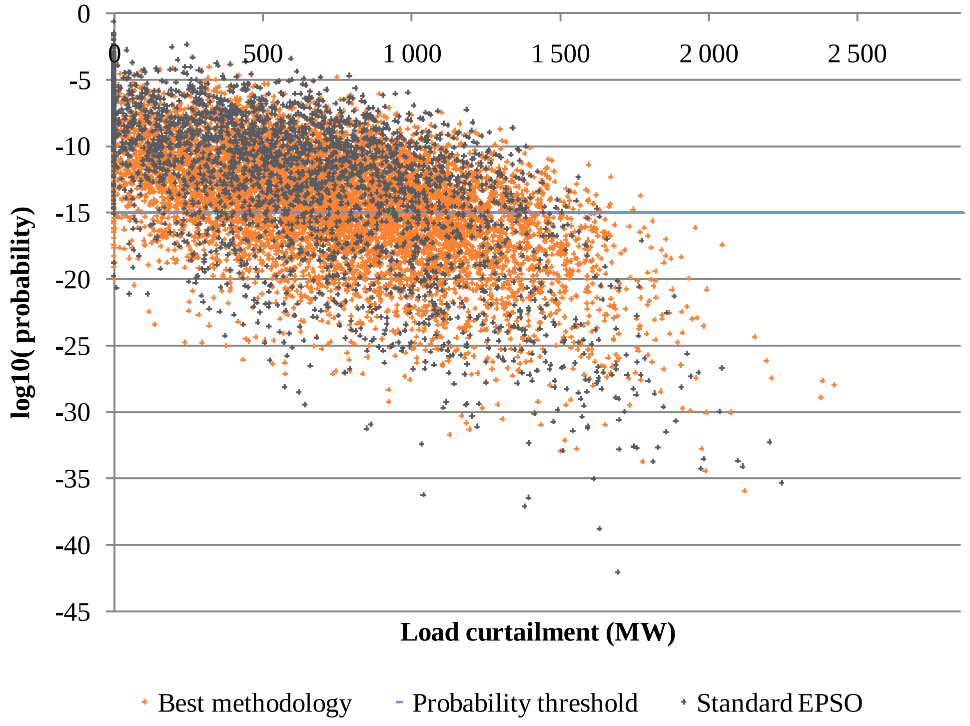


Figure 4.21. Representation of the attributes probability and load curtailment of the cases visited during the search procedure. Notice that case probability is in a decimal logarithmic scale for better display.

These last two figures illustrate the difference in the attributes space of the visited cases between the traditional EPSO methodology and the proposed algorithm. There are two conclusions to be drawn from this figure. First the standard EPSO concentrates the search near the attribute space zone where the most probable cases with high load curtailment exist. This reinforces the idea that the search is referred to an area from a certain generation. The second conclusion is the extension of the search that the spreading techniques provide. This benefits the number of different cases visited and as well as the approximation of the estimative of the reliability indices to the real value.

However figure 4.21. must be analyzed with care: it only represents the probability of each different case visited as well as its load curtailment. As it is known, each case represents a number of equal system states, in terms of reliability evaluation. Therefore the probability of the case must be multiplied by its number of copies to know its real contribution for the calculation of the reliability indices. The objective of this explanation is to alert the reader that although a case may have a probability below the threshold line, the number of system states that it represents may classify him as relevant. This characteristic should be the criterion for case rejection: if the probability x the number of copies that it represents is above a threshold value then the case would be classified as relevant and countable. The reason for the adoption in this thesis other criterion than this one was to have results that could be compared with the ones obtained by other methodology, namely the MSGA proposed in [22] which has the same criterion for classifying the feasibility of a case.

We now present a quantitative comparison of the proposed search methodology, named EPSO reliability, with the results of other researchers, namely the MSGA approach [22]. The errors resulting from the MSGA approach in [22] have been recalculated to make them relative to the exact results, known form [9] and referred to in [30] which are also replicated in table 4.1.. The errors deriving from the application of the EPSO reliability have also been calculated relative to this same reference. The comparison is made in tables 4.3. to 4.5..

In these tables one may compare the estimation of the principal reliability indices obtained by these two methods in the evaluation of the adequacy of the generating capacity for the IEEE RTS-79 with the annual load curve, with the same number of cases visited (in the MSGA methodology the number of individuals of a generation is 40 and it is performed 750 generations evaluating 30040 cases, although phenotypic representation of the case differs from the proposed methodology in this thesis; in the EPSO reliability algorithm in each generation 80 cases are evaluated, as a result of the replication of the particles of the previous generation (there are two offspring) and it takes 375 generations resulting in a total of 30040 cases visited) and with the same criterion of case rejection. The results in table 4.3., 4.4., and 4.5. speak for themselves.

Table 4.3. LOLE index comparison between the MSGA and the EPSO reliability algorithm.

	LOLE (hour/year)	Absolute error (%)
MSGA	9.324000	0.75%
EPSO reliability	9.352507	0.44%

Table 4.4. LOLF index comparison between the MSGA and the EPSO reliability algorithm.

	LOLF (occurrence/year)	Absolute error (%)
MSGA	2.003700	0.79%
EPSO reliability	2.010145	0.47%

Table 4.5. LOEE index comparison between the MSGA and the EPSO reliability algorithm.

	LOEE (MWh/year)	Absolute error (%)
MSGA	1163.00	1.13%
EPSO reliability	1169.18	0.61%

The absolute errors are calculated by comparison with the analytical results of the reliability indices for this power system reproduced in table 4.1..

Table 4.6. presents a comparison between the two methodologies on the number of significant cases saved and the number of m significant system states visited. Note that although the number of saved cases is higher in the EPSO reliability algorithm, the number of significant system states is lower. However the estimation of the reliability indices provided by the EPSO reliability algorithm shows that this number of system states is more important for the accuracy of the indices estimative than the systems states visited by the MSGA algorithm.

Table 4.6. LOEE index comparison between the MSGA and the EPSO reliability algorithm.

	Nº of cases saved	Nº of their repetitions
MSGA	10428	19198310
EPSO reliability	10906	18987932

To gain statistical confidence in the results obtained, table 4.7. provides the results of 250 runs of the EPSO reliability algorithm with the same conditions of the earlier runs.

Table 4.7. Statistical data for the 250 runs of the EPSO reliability algorithm.

	Average	Standard Deviation
LOLE (hour/year)	9.337799	0.013395
LOLF (occurrence/year)	2.007116	0.002742
LOEE (MWh/year)	1166.38	1.93
Nº of cases saved	10699	94
Nº of their repetitions	18665090	519461

The aspect that is to be retained is the robustness of the results obtained by the EPSO reliability algorithm, expressed by its standard deviation, and the average value of the estimation of the reliability indices that, which however are still superior to the reliability indices estimative obtained in [22].

It would be interesting to benchmark the efficiency of the method presented in this thesis against a classical non-chronological Monte Carlo simulation. This work has not been done. Nevertheless, other researchers have reported a remarkable acceleration in computing time of PB methods over Monte Carlo. As the EPSO reliability algorithm proved to be more efficient that other published PB approaches, the advantage over Monte Carlo is considered established.

4.5. Conclusions

In this Chapter the main results of the EPSO reliability algorithm were presented, in the evaluation of the generating capacity of the IEEE RTS – 79. This presentation was performed step by step analyzing the effect on the efficacy and efficiency of the search of distinct fitness functions and different methods to achieve diversity.

As it was demonstrated, the best guide, in terms of efficacy, to conduct the search is the maximization of the probability x number of copies x load curtailment. With this fitness function not only the probability of the case determines its relevance but also the load curtailment, calculated for the system peak load, which is a measure of the case generating capacity, becomes important. Therefore, with this fitness function it is guaranteed that the search is guided to the cases which contribute the most for the calculation of the reliability indices: the cases with low generating capacity and with high probability. This fact is essential to take into account in the system reliability indices calculation not only the effect of the hours where the load is at a low value but also to describe correctly all the ways the system can fail in the hours of peak load. This is the main advantage of this fitness function. On the other hand, maximizing only the case load shedding drives the search to ineffectual zones of the search space reducing dramatically the number of entries in the feasible cases list. The DWA approach has a performance between the two objectives selected for the construction of the bi-objective function. It was also seen that the weights change frequency has an important effect in the efficiency and the efficacy of the search. Moreover, when the rate of changing of the aggregation weights is very low, meaning that the search will be guided to a specific objective for long periods, the weight of the objective which starts with the value one determines the behavior of the search. This means that at the end of the search procedure the efficiency and the efficacy will be similar to maximizing only this dominant objective.

On the other hand all the proposed spreading techniques performed satisfactorily. In all three both the efficacy and the efficiency of the search is enhanced. Attending to that, the idea of incorporating them in the same methodology was successfully tested, with the expectation of achieving a superior performance than the one that the methods proposed in the literature possess. A comparative study between the EPSO reliability algorithm and the MSGA [22] method was conducted for the same reliability problem and with the same conditions and it was shown that EPSO reliability has a better performance than MSGA. Moreover it was demonstrated that EPSO reliability has consistently superior results than MSGA, in terms of efficacy and, in a particular point of view, in terms of efficacy, since it visits more feasible cases (though corresponding to a fewer number of system states) for a higher number of runs.

Chapter 5

Conclusions and future work

This Chapter will summarize the previously drawn conclusions, when the main results in Chapter 4 were analyzed, in a context of assessment if the objectives proposed for this thesis were attained. Then some guidelines to continue and improve the work on this thesis will be detailed.

5.1. Objectives achieved

In this thesis a new method was presented, which belongs to the new type of methods for assessing power system reliability, the PB (Population Based) methods, for evaluating the reliability of power systems generating capacity, namely its adequacy.

First an overview of the traditional methods for solving this specific reliability problem was presented, namely the deterministic and the probabilistic methods with special focus on the two approaches of the last one: the analytical approach and the simulation approach. This overview was made not with the objective of detailing each specific characteristic of these methods but to demonstrate their advantages and disadvantages establishing a starting point to introduce the new type of methodologies for assessing power systems reliability: the PB methods. As it was described these methods rely on the use of meta-heuristics, namely EA, for constructing a list with the system states considered important for the computation of the reliability indices. Therefore an introduction to the most used EA was presented. Finally the work of some researchers on this area was detailed, introducing the main aspects of their search algorithms, which lead to the essence of this thesis: since most of the search methods did not use specific techniques to enhance the search, probably their performance could be improved using methods for creating diversity.

The first idea in this thesis was to use the traditional EPSO algorithm with a slight modification of the phenotypic representation of a particle in relation to the known PB methods to perform the search, which, per se, was a new application of this method constituting an additional contribution of the actual state of the art. Nevertheless, using the ideas exposed on the last paragraph and after analyzing the convergence behavior and the EPSO population dynamics during the search procedure, one realized that the diversity in the EPSO population could be created by the fitness assignment procedure and by the control of the particles velocity. To accomplish this, two main philosophies of fitness assignment were formulated : the single objective and the bi-objective using the DWA approach; and three methods were proposed for creating diversity: the continuous changing of the population objective as well as forgetting the best position of a particle if it represents a previous visited one; adding an extra velocity in case of convergence of the particles to a specific point of the search space; and introducing a new term in the equation from which the new velocity is calculated, to enhance the search around each particle.

The results show that the best fitness function was a single objective one, which maximizes the probability x the number of copies of each case x the normalized load curtailment. Moreover it was also verified in Chapter 4 that each method for creating diversity increased both measures of the search quality: efficacy and efficiency. Then the idea came of aggregating this information in order attain a superior performance methodology which was called EPSO reliability algorithm. Its advantages, namely in the efficacy and in the efficiency (although the efficiency measure was only superior in part) was confirmed by comparing the results obtained with EPSO with the ones of a search methodology named MSGA [22]. Moreover it was confirmed that the EPSO reliability algorithm provides robust solutions since the standard deviation of the obtained estimative of the reliability indices after a considerable number of runs was relatively low compared to its average value. However no conclusions were possible to be drawn, regarding the number of system states visited and the accuracy of the estimations, of the proposed method when compared with a non-chronological Monte Carlo simulation due to the lack of data.

To sum up in this thesis it was proved that EPSO with spreading techniques can be used as a search algorithm, extending its repertoire of applications in power systems problems, and also that the quality of the search can be enhanced in all of its aspects by using methods for creating diversity in the population.

5.2. Future work

The results of the work developed in this thesis may be the inspiration for other research studies. Therefore a list of possible studies that can be made is now presented:

- Give some sort of intelligence to the mechanism of changing the best position of the population. Instead of moving constantly the global attractor, the search would probably benefit if this point is only changed if the population is confined to a specific zone of the search space;
- Apply the developed method in the generating capacity analysis of a real power system and compare its results with the ones provided by other PB methodologies;
- Apply the developed system state representation in a HLII type of study. For this specific problem it would probably help in the construction of the particles position if the number of equal generators in each bus bar is taken in account. However the representation of the states of each transmission line cannot be forgotten;
- Provide a spreading mechanism that detects if the particles are in a collision course in order to avoid repetitions. To accomplish that, one needs to analyze not only the particle position but also its velocity vector;
- Apply in the fitness assignment method a fitness function based on the Pareto front. With this feature it will no longer be necessary to scale the value of each objective and perform an aggregation based approach. More of this subject can be found in [30].

References

[1] Billinton, Roy and Allan, Ronald N., “Reliability Evaluation of Power Systems”, Second Edition, New York, Plenum Press, 1996

[2] Miranda, V., Keko, H. and Jaramillo, A., “EPSO: Evolutionary Particle Swarms, Advances in Evolutionary Computing for System Design”, Book Chapter, ISBN 978-3-540-72376-9, Springer Berlin / Heidelberg, vol. 66, 2007

[3] Patra, Shashi B., Mitra, Joydeep, Earla Ramesh, “A New Intelligent Search Method for Composite System Reliability Analysis”, IEEE PES Transmission and Distribution Conference and Exhibition, 2005, pp. 803-807

[4] Billinton, Roy and Allan, Ronald N., “Power system reliability in perspective”, IEE J. Electronics Power, vol. 30, 1984, pp. 231-236

[5] Endrenyi, J., “Reliability Modeling in Electric Power Systems”, New York, Wiley, 1978

[6] Jasper F. L. van Casteren, Math H. J. Bollen, Martin E. Schmieg, “Reliability Assessment in Electrical Power Systems: The Weibull–Markov Stochastic Model”, IEEE Transactions on Industry Aplications, vol. 36, nº 3, 2000

[7] Singh, Chanan and Billinton, Roy, “System Reliability Modeling and Evaluation”, Hutchinson, London, 1977

[8] IEEE Task Group on Model Peaking Units of the Application of Probability Methods Subcommittee, “A four state model for estimation of outage risk for units in peaking service”, IEEE Trans., PAS-91, 1972, pp. 618-27

[9] Leite da Silva, A.M., Melo, A.C.G., Cunha, S.H.F., “Frequency and duration method for reliability evaluation of large-scale hydrothermal generating systems”, IEE Proceedings-Generation, Transmission and Distribution, vol. 138, 1991, pp. 94-102

[10] Billinton, Roy and Li, Wenyuan, “Reliability Assessment of Electric Power Systems Using Monte Carlo Methods”, New York, Plenum Press, 1994

[11] Billinton, Roy and Wang, P., “Teaching distribution system reliability evaluation using Monte Carlo simulation,” IEEE Trans. Power Syst., vol. 14, 1999, pp. 397-403

[12] Sankarakrishnan, A. and Billinton, Roy, "Sequential Monte Carlo Simulation for Composite Power System Reliability Analysis with Time Varying Loads", IEEE Trans. Power Syst., vol. 10, nº 3, 1995, pp. 1540-1545

[13] Pereira, M.V.F., Maceira, M. E. P., Oliveira, G. C., Pinto, L. M. V. G., “Combining Analytical Models and Monte Carlo Techniques in Probabilistic Power System Analysis”, IEEE Trans. on Power Systems, vol. 7, nº 1, 1992

[14] Schwefel, H.-P., “Evolution and Optimum Seeking”, Ed. Wiley, New York, 1995

[15] Fogel, D.B. “Evolving Artificial Intelligence”, Ph.D. Thesis, University of California, San Diego, 1992

[16] Goldberg, D.E., “Genetic Algorithms In Search, Optimization And Machine Learning”, Addison-Wesley, 1989

[17] Kennedy, J., Eberhart, R.C. “Particle Swarm Optimization”, Proceedings of the 1995 IEEE International Conference on Neural Networks, Perth (Australia), 1995, pp. 1942-1948

[18] Clerc, M., and Kennedy, J., “The particle swarm: explosion, stability, and convergence in a multi-dimensional complex space”, IEEE Transactions on Evolutionary Computation, vol. 6, 2002, pp. 58-73

[19] Miranda, V. and Fonseca, N., "EPSO – Best-of-Two-Worlds Meta-Heuristic Applied To Power System Problems ", Proceedings of WCCI 2002 - World Congress on Computational Intelligence - CEC -Conference on Evolutionary Computing, Honolulu, Hawaii, U.S.A., 2002

[20] Miranda, V. and Fonseca, N., “EPSO - Evolutionary Particle Swarm Optimization, a New Algorithm with Applications in Power Systems”, Proceedings of the IEEE Transmission and Distribution Asia-Pacific Conference 2002, Yokohama, Japan, 2002

[21] Leite da Silva, A. M., Resende, L. C., Manso, L. A. F., Miranda, V., “Composite Reliability Assessment Based on Monte Carlo Simulation and Artificial Neural Networks”, IEEE Transactions on PWRS, vol. 17, nº 4, 2002, pp. 974-981

[22] Samaan, Nader, and Singh, Chanan, “Adequacy Assessment of Power System Generation Using a Modified Simple Genetic Algorithm”, IEEE Trans. on Power Systems, vol. 17, nº 4, 2002

[23] Samaan, Nader, and Singh, Chanan, “Assessment of the annual frequency and duration indices in composite system reliability using genetic algorithms”, IEEE Power Engineering Society General Meeting, vol. 2, 2003

[24] Wang, Lingfeng, Singh, Chanan, Tan, Kay Chen, “Reliability Evaluation of Power-Generating Systems Including Time dependent Sources Based on Binary Particle Swarm Optimization”, IEEE Congress on Evolutionary Computation, 2007

[25] Kennedy, J. and Eberhart, R. “A discrete binary version of the particle swarm optimization”, IEEE Proceedings of the International Conference on Systems, Man, and Cybernetics, Orlando, 1997, pp. 4104-4108

[26] Parsopoulos, K. E. and Vrahatis, M. N., “Particle Swarm Optimization Method in Multiobjective Problems”, Symposium on Applied Computing: Proceedings of the 2002 ACM Symposium on Applied Computing, 2002

[27] Blackwell, T. M. and Bentley, P, J., “Dynamic Search with Charged Swarms”, Proceedings of the Genetic and Evolutionary Computation Conference”, 2002, pp. 19-26

[28] Peram, T., Veeramachaneni, K., Mohan, C.K. , “Fitness-Distance-Ratio Based Particle Swarm Optimization”, Proceedings of the 2003 IEEE Swarm Intelligence Symposium, 2003, pp. 174-181

[29] IEEE Reliability Test System Task Force of the Application of Probability Methods Subcommittee, “IEEE reliability test system", IEEE on Power Apparatus and Systems, vol. 1, PAS-98, 1979, pp. 2047-2054

[30] Allan, Ronald N., Billinton, Roy, Abdel-Gawad, N.M.K., "The IEEE reliability test system - extensions to and evaluation of the generation system", IEEE Transactions on Power Systems, vol. PWRS-1, nº 4, 1986

[31] Y. Lee, Kwang and El-Sharkawi, Mohamed A., “Modern Heuristic Optimization Techniques”, ISBN 978-0-471-45711-4, Wiley-IEEE Press, 2008

Annex A - IEEE RTS-76

The unit parameters used in the evaluation of the IEEE RTS-76 generating capacity are described in the following table:

Table A1 IEEE RTS-76 unit reliability data.

Unit	Unit size (MW)	MTTF (hour)	MTTR (hour)	FOR
1	12	2940	60	0.020
2	12	2940	60	0.020
3	12	2940	60	0.020
4	12	2940	60	0.020
5	12	2940	60	0.020
6	20	450	50	0.100
7	20	450	50	0.100
8	20	450	50	0.100
9	20	450	50	0.100
10	50	1980	20	0.010
11	50	1980	20	0.010
12	50	1980	20	0.010
13	50	1980	20	0.010
14	50	1980	20	0.010
15	50	1980	20	0.010
16	76	1960	40	0.020
17	76	1960	40	0.020
18	76	1960	40	0.020
19	76	1960	40	0.020
20	100	1200	50	0.040
21	100	1200	50	0.040
22	100	1200	50	0.040
23	155	960	40	0.040
24	155	960	40	0.040
25	155	960	40	0.040
26	155	960	40	0.040
27	197	950	50	0.050
28	197	950	50	0.050
29	197	950	50	0.050
30	350	1150	100	0.080

31	400	1100	150	0.120
32	400	1100	150	0.120

The hourly load model of the IEEE RTS-79 for an annual peak load of 2850 MW is calculated according to the following tables:

Table A2 Weekly Peak Load in Percent of Annual Peak [29].

Week	Peak Load	Week	Peak Load
1	86.2	**27**	75.5
2	90	**28**	81.6
3	87.8	**29**	80.1
4	83.4	**30**	88
5	88	**31**	72.2
6	84.1	**32**	77.6
7	83.2	**33**	80
8	80.6	**34**	72.9
9	74	**35**	72.6
10	73.7	**36**	70.5
11	71.5	**37**	78
12	72.7	**38**	69.5
13	70.4	**39**	72.4
14	75	**40**	72.4
15	72.1	**41**	74.3
16	80	**42**	74.4
17	75.4	**43**	80
18	83.7	**44**	88.1
19	87	**45**	88.5
20	88	**46**	90.9
21	85.6	**47**	94
22	81.1	**48**	89
23	90	**49**	94.2
24	88.7	**50**	97
25	89.6	**51**	100
26	86.1	**52**	95.2

Table A3 Daily Peak Load in Percent of Weekly Peak [29].

Day	Peak Load
Monday	93
Tuesday	100
Wednesday	98
Thursday	96
Friday	94
Saturday	77
Sunday	75

Table A4 Hourly Peak Load in Percent of Daily Peak [29].

Hour	Winter Weeks Weeks 1-8 & 44-52		Summer Weeks 18-30		Spring/Fall Weeks 9-17 & 31-43	
	Weekday	Weekend	Weekday	Weekend	Weekday	Weekend
12 pm -1 am	67	78	64	74	63	75
1 am -2 am	63	72	60	70	62	73
2 am -3 am	60	68	58	66	60	69
3 am -4 am	59	66	56	65	58	66
4 am -5 am	59	64	56	64	59	65
5 am -6 am	60	65	58	62	65	65
6 am -7 am	74	66	64	62	72	68
7 am -8 am	86	70	76	66	85	74
8 am -9 am	95	80	87	81	95	83
9 am -10 am	96	88	95	86	99	89
10 am -11 am	96	90	99	91	100	92
11 - Noon	95	91	100	93	99	94
Noon - 1 pm	95	90	99	93	93	91
1 pm - 2 pm	95	88	100	92	92	90
2 pm - 3 pm	93	87	100	91	90	90
3 pm - 4 pm	94	87	97	91	88	86
4 pm - 5 pm	99	91	96	92	90	85
5 pm - 6 pm	100	100	96	94	92	88
6 pm - 7 pm	100	99	93	95	96	92
7 pm - 8 pm	96	97	92	95	98	100
8 pm - 9 pm	91	94	92	100	96	97
9 pm - 10 pm	83	92	93	93	90	95
10 pm - 11 pm	73	87	87	88	80	90
11 pm - 12 pm	63	81	72	80	70	85

Annex B - Evaluation of the performance of the different fitness functions simulation data

In this Annex is enumerated the conditions of the referred simulations.

Table B1 EPSO reliability unit data

Unit type	Unit size (MW)	MTTF (hour)	MTTR (hour)	FOR	Number of equal units
1	12	2940	60	0.02	5
2	20	450	50	0.1	4
3	50	1980	20	0.01	6
4	76	1960	40	0.02	4
5	100	1200	50	0.04	3
6	155	960	40	0.04	4
7	197	950	50	0.05	3
8	350	1150	100	0.08	1
9	400	1100	150	0.12	2

General EPSO reliability algorithm parameters:

- Maximum value of the particles position vector:

$$\overline{\boldsymbol{X}} = \begin{bmatrix} 5.5 \\ 4.5 \\ 6.5 \\ 4.5 \\ 3.5 \\ 4.5 \\ 3.5 \\ 1.5 \\ 2.5 \end{bmatrix}$$

- Minimum value of the particles position vector:

$$\underline{\boldsymbol{X}} = \begin{bmatrix} -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \end{bmatrix}$$

- Velocity reduction factor: 0.5;
- Population size: 40;
- Learning parameter: $\tau = 0.3$;
- Probability of communication between particles: 0.6;
- Threshold probability for case rejection: 1.0^{-15};
- System yearly peak load: 2850 MW.
- Maximum number of iterations: 375.

Spreading techniques:

- Maximum number of iterations which the global best position is not changed: 1;
- Reduction factor of the additional velocity: 0.225;
- Maximum value of the auxiliary vector: 5;
- Minimum value of the auxiliary vector: -5;
- Use of the neighbor term.

Fitness function:

- Weight's change frequency: $T = 375$;
- Lowest value for the probability x copy objective:

$$\prod_{i=1}^{n} FOR_i = 1.20796^{-48} \quad , \tag{B1}$$

where n is the total number of units of the IEEE RTS-79 and FOR_i is the Forced Outage Rate of the unit i.

- Highest value for the probability x copy objective:

$$\prod_{i=1}^{n}(1 - FOR_i) = 2.36396^{-1} \quad , \tag{B2}$$

where n is the total number of units of the IEEE RTS-79 and FOR_i is the Forced Outage Rate of the unit i.

- Lowest value for the probability x copy x normalized load curtailment objective (normalization performed using the system peak load): 0. Nevertheless, since the decimal logarithmic is not defined for the value 0, the next minimum value is selected for the construction of the linear transformation. After using EPSO as an optimization tool it appears that this value is equal to the one obtained with the equation (B1);

- Highest value for the probability x copy x load curtailment objective: 3.83583^{-4}.

Annex C - Evaluation of the performance of the different spreading techniques simulation data

In this Annex is enumerated the conditions of the referred simulations.

Table C1 EPSO reliability unit data

Unit type	Unit size (MW)	MTTF (hour)	MTTR (hour)	FOR	Number of equal units
1	12	2940	60	0.02	5
2	20	450	50	0.1	4
3	50	1980	20	0.01	6
4	76	1960	40	0.02	4
5	100	1200	50	0.04	3
6	155	960	40	0.04	4
7	197	950	50	0.05	3
8	350	1150	100	0.08	1
9	400	1100	150	0.12	2

General EPSO reliability algorithm parameters:

- Maximum value of the particles position vector:

$$\overline{\boldsymbol{X}} = \begin{bmatrix} 5.5 \\ 4.5 \\ 6.5 \\ 4.5 \\ 3.5 \\ 4.5 \\ 3.5 \\ 1.5 \\ 2.5 \end{bmatrix}$$

- Minimum value of the particles position vector:

$$\underline{\boldsymbol{X}} = \begin{bmatrix} -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \\ -0.5 \end{bmatrix}$$

- Velocity reduction factor: 0.5;
- Population size: 40;
- Learning parameter: $\tau = 0.3$;
- Probability of communication between particles: 0.6;
- Threshold probability for case rejection: 1.0^{-15};
- System yearly peak load: 2850 MW.
- Maximum number of iterations: 375.

Spreading techniques:

- Maximum number of iterations which the global best position is not changed:
    - 1;
    - 5;
    - 10.

- Reduction factor of the additional velocity: 0.225;
- Maximum value of the auxiliary vector: 5;
- Minimum value of the auxiliary vector: -5;
- Use of the neighbor term.

Fitness function:

- Maximization of the probability x the number of copies x the normalized load curtailment (normalization performed using the system peak load).

Annex D - Submitted paper to publication in IEEE Transactions

As a result of the advantages achieved by the developed strategy in this work, it was decided to write a paper where the main achievements and ideas of such strategy are presented. In this Annex, one will do a brief and general presentation of the main topics that are focused in the referred paper. However it is important to underline that this paper is, at the moment, confidential, having been submitted for future publication in IEEE Transactions. Therefore only the first page is reproduced below.

Improving power system reliability calculation efficiency with EPSO variants

Vladimiro Miranda, *Fellow, IEEE*, and Leonel Carvalho

Abstract – This paper reports the application of a population based method (EPSO – Evolutionary Particle Swarm Optimization) to calculate power system reliability. Population based methods appear as competitors to the traditional Monte Carlo simulation because they can be much more computationally efficient in estimating a number of reliability indices. The work reported in this paper demonstrates that EPSO variants, suited to the problem of exploring a zone in the state space instead of searching for a single optimizing state, are efficient in calculating a number of system reliabilitu indices such as power not supplied. The results obtained with EPSO are compared with Monte Carlo and with work of other researchers in population based methods.

Index Terms **– Reliability. Monte Carlo, evolutionary algorithms, particle swarms**

I. INTRODUCTION

MONTE CARLO remains the standard method to calculate estimates of reliability indices in Power Systems.

This statistically based method has gained importance over analytic models since the emergence of enough computing power in the beginning of the 90's coupled with the adoption of efficient convergence acceleration techniques. The two basic advantages of Monte Carlo were: a) allowing simulation of realistic characteristics of systems, even those not necessarily reducible to formal mathematical models, and b) allowing the calculation of distributions and not only of Mean values (in its simplest form, allowing the estimation of Variance). Non-chronological models became successful then.

However, as it is usual in such cases, the growth in computer power opened the way to the desire to perform chronological simulations and this became demanding of increased computing power. At the same time, even non-chronological models became more complex because of the availability of computing power at desktop level. As happened in many other cases in the development of science and technology, the moment one has available more computing power this becomes almost at once insufficient for the new and more complex models one wishes to run.

Recently, an alternative to Monte Carlo started to emerge: *population based* methods. While Monte Carlo is statistically based method, relying on the theorems of sampling to provide an estimate of a result plus some interval of confidence, population based methods are methods that try to search only for the meaningful subset of the state space and are enumeration methods. If all states contributing to a certain index could be identified and their probabilities known, the index would be calculated exactly. Population based methods try therefore to discover, if not the totality, the majority of states so that a good approximation of the index is computed.

The methods are called *population based* because they rely on meta-heuristics that have a population of solutions (individuals, particles) as their core. In this class one may count evolutionary algorithms – evolutionary programming (EP) or genetic algorithms (GA) – and particle swarm optimization algorithms (PSO). They all were traditionally developed to be an optimization tool but the problem now is the discovery of a set of states that have maximum contribution to the index to be calculated – so, some mechanism to generate diversity must be kept, otherwise all solutions would tend to converge to a maximizing state and space exploration would be hampered.

This paper presents new results confirming the efficiency of a population based method – EPSO, Evolutionary Particle Swarm optimization, over Monte Carlo to calculate reliability indices in a Power System. The results obtained will be compared with the results from other researchers and conclusions drawn from the experiments designed.

V. Miranda is with INESC Porto and also with FEUP, Faculty of Engineering of the University of Porto, Portugal (email: vmiranda@inescporto.pt).

L. Carvalho is with INESC Porto and also with FEUP, Faculty of Engineering of the University of Porto, Portugal (email: lcarvalho@inescporto.pt).