**POLITECNICO DI MILANO**

**SCHOOL OF INDUSTRIAL ENGINEERING**

**DIPARTIMENTO DI INGEGNERIA GESTIONALE**

# Application of the Kriging Method in Energy Matrix Portfolio Optimization

Pedro Henrique Carvalho Araujo

Matricola: 818367

Tesi di Laurea Magistrale

Supervision: Prof. Sergio Mariotti

Milano

2015

# ABSTRACT

The remarkable growth of discussions about energy matrix in the last decades and the recognition of the necessity of the adoption of renewable energies led to the questioning of which would be the ideal energy matrix for a country, in terms of costs, benefits, and risks to the population. Hence, with the purpose of supporting politicians in hers decisions, this study proposed to create a method to define the optimal portfolio of energy matrix, which considers not only the generation costs of each technology, but also its risks. In order to do so, not just deviation risk measures (e.g. variance) were taken into consideration: tail measures were also used, for example the Value at Risk (VaR) and the Conditional Value at Risk (CVaR), capturing as well extreme events, which are very important to the analysis. Therefore, data on seven technologies of the United States was analyzed, Monte Carlo simulations were carried out, and with the support of the Kriging Method, the Paretto`s efficient frontier and the compositions of the optimal portfolio were finally obtained for the years of 2030, 2035, and 2040. The results, besides of assuring that tail risk measures are the most applicable in this kind of analysis, also pointed out a greater allocation in the future of renewable energies, such as wind and biomass technologies, revealing, hence, that environment aggressive technologies (e.g. coal and gas) should play a minimal role in future energy matrix.

# LIST OF FIGURES

# LIST OF TABLES

# SUMMARY

# 1 INTRODUCTION

An important issue that remains a matter of discussion in the engineering field is how to allocate resources in uncertainty conditions. From the portfolio manager stand point, it is interesting to determine allocations which guarantee a minimum financial return from their investment and also have a reasonable safety level. Hence, one important question in this field of knowledge is how to minimize risks.

The risk minimization theme never was so discussed as in the 2008 financial crisis. That is so because of the lack of regulation and the constant indiscipline on doing practices aiming the risk mitigation led the financial markets of the major developed countries (impacting, clearly, the emerging markets as well) into collapse. This fact may be easily noticed with a quick observation of the stock indexes, which reveal indirectly the investors risk perceptions and the results of the major listed companies of a country. Figure 1 shows the evolution during this tough period of one of the major stock indexes of the United States, the S&P 500.

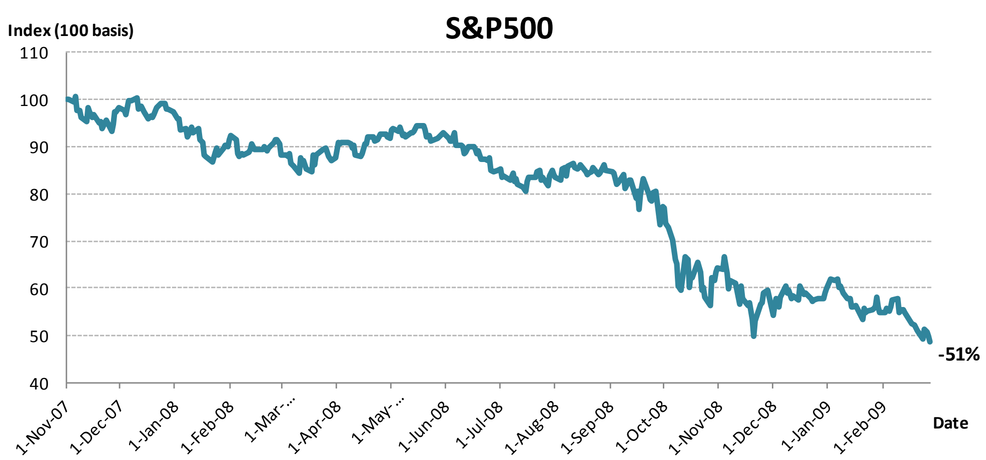



**Figure 1: S&P 500 Index**

It may be observed that during the year of 2008 the index plunged considerably, revealing a lower risk appetite from the investors. This lack of interest on allocating capital in companies (i.e. equity) generates a great social damage, since that in the capitalism way of living, the companies are in fact the economic agents that move the economy, and therefore, create social welfare.

Yet in the scope of risk minimization, a theme always very discussed by everyone and each time more present in the literature is the issue of allocating risks in the energy generation process. In the process of defining a country energy matrix, politicians and strategists should take in consideration not just the incurred costs in the energy generation, but also involved in this operation. A simple example of energy generation risk is the fuel price fluctuation, since, depending on the technology employed, this components may be a relevant part of the cost structure.

The oil, for example, is used as raw material in thermoelectric power plants. This input is commercialized in the commodity financial markets and presents high price volatility, what may be observed in Figure 2. This great price fluctuation shall turn some technology not interesting to use and, therefore, may be considered a risk to policy makers in the energy matrix decision process.

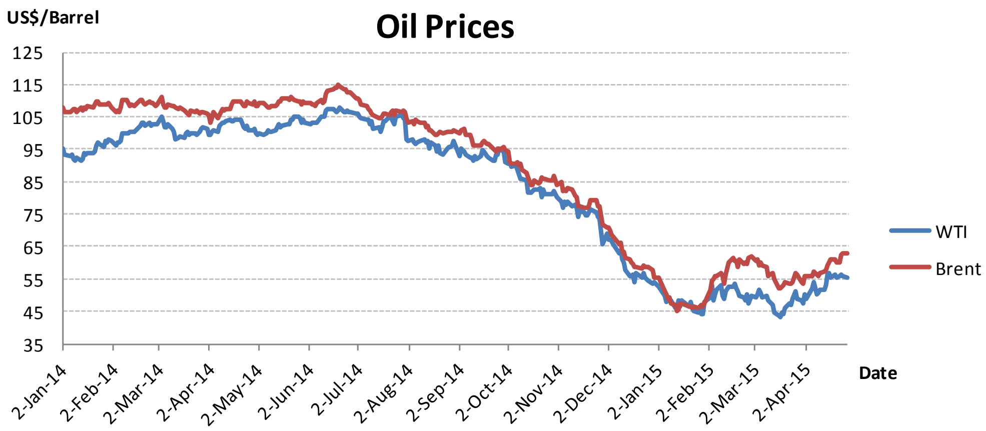



**Figure 2: Oil prices evolution**

## 1.1 Study`s Objectives

Politicians and energy policymakers need a tool which supports their decision making, so an energy matrix portfolio may be created taking in to consideration not only the generation costs of each technology, but also its risks. This is in fact the objective of this study: use a method of investment portfolio selection, something commonly used in the finance field, with the aim of optimizing the relation between risk and return of a country energy matrix portfolio.

Regarding the tool that will take place, this study proposes, initially, in using risk measures that capture, in its formulation, extraordinary event, since these events, as already mentioned, are capable of generating huge crises and recessions. Risk measures of this nature are known in the literature as tail risk measures (ROCKAFELLAR e URYASEV, 2000). Hence, in this study, it is going to be used not only deviation risk measures (e.g. Variance), but also two other tail risk measures, the Value at Risk (VaR) and the Conditional Value at Risk (CVaR).

These two risk measures, although manage to capture extreme events, present a great drawback: require high computational capacity. To solve this issue, this study proposes the use of the Kriging Method (RIBEIRO e FERREIRA, 2004). Even if this technique is not traditionally employed to portfolio selection, it is believed that with the application of this method it is possible, efficiently, to optimize energy matrix portfolios which take in consideration any risk measures, including the tail ones.

## 1.2 Study`s Structure

This study is divided in five parts. In the first chapter, the theme is presented. It is shown why it is important and relevant, situating it in the current macro landscape and exhibiting its objectives and structure.

The second chapter reveals the literature review indicating the concepts necessary to the study comprehension. Initially, it is established a conceptual base about portfolio management, naming its mark in the literature and describing some important definitions which distinguish a portfolio. Then, different risk definitions are exposed and three main risk measures, which will be used later, are described. After this, four models of portfolio selection are presented, being the Kriging Method one of them. Yet in this second chapter, finally, the models used in the energy sector are demonstrated. It is worth saying that the Kriging Method has not yet been used to such purpose, hence this study proposes an innovative application.

The following chapter refers to the methodology adopted in this study. To such effort, first of all, it is analyzed the Market here studied, the energy sector of the United States. After such definition, it is analyzed the LCOE (Levelized Cost of Energy), which will be the cost measure in this study. Later, data collected is described and treated through the Monte Carlo simulation.

In the fourth chapter, the results of the application of the proposed method (i.e. the Kriging Method) are revealed. Such results are demonstrated in both the forms of efficient frontier and optimal portfolio composition, concepts that will be extensively described in the literature review section.

Finally, the fifth chapters the conclusions and future extensions are presented. The appendix and the references finish the document.

# 2 LITERATURE REVIEW

In this chapter, it will be presented relevant fundamentals and concepts to the completely understanding of this study. First, the portfolio manager theme will be discussed, including its objectives, main characteristics and a preliminary analysis of the Modern Portfolio Theory. Next, it will be described different definitions of risk, which is a highly frequent concept in this study. Then, the portfolio selection models are presented, which will be later implemented in the energy matrix theme, what is by the way the topic discussed in the end of the literature review.

## 2.1 Portfolio Management

The portfolio manager is responsible for defining which is the best way of allocating capital so that the return expected is achieved, taking in consideration an acceptable risk level, or in other words, what degree of risk the investor is willing to incurre. In order to do so, the manager has basically two investment options:

1) Investments whose returns are previously known;
2) Investments whose returns are unknown and involves uncertainties.

The first option above mentioned deals with fixed-income securities, indicating that the acquired asset yields the investor a fixed amount of money, pre-established in its contract. Hence, in general, the portfolio manager knowledge previously not just the capital allocation, but also the future returns of her investments. Therefore, one might say that the uncertainties involved in the process of investing are minimized, leaving just some uncertainties regarding the probability of default by the issuer (i.e. credit risk) and possible variations of the benchmark (market rate which some securities are commonly indexed, such as the LIBOR). Some examples are (LUENBERGER, 2008):

1) Certificates of Deposit (CD);
2) Treasury bonds and treasury inflation-protected securities;
3) Asset Backed Securities (ABS);
4) Collaterized Debt Obligations (CDOs).

Although fixed-income investments are broadly used all around the world, it is not the scope of this study to analyze it, being its citation and utilization merely illustrative.

The second form of investment above mentioned includes those whose initial amount of capital invested is known, with its returns uncertain though, what one may consider random events, from the stochastic stand points. Hence, it is possible to handle the price of an asset, in different time periods, as a random variable.

Both investments forms defined above are valued based on its expected returns, but in the second form, these return are unknown due to the characteristics of the random variable, turning necessary the analysis of the risks associated with the investments.

In such field, Harry Markowitz, in 1952, published an article called “Portfolio Selection”, what was a considered by many as the born of the modern finance economy (RUBISTEIN, 2002). This author’s theory was so hailed that it led the author to receive the Nobel Prize in economics in 1990.

From the beginning of the first paragraph of his article, MARKOWITZ (1952) affirms that the investment process and the asset selection consist in, actually, two separate parts: initially, one shall observe the available assets, considering theirs historical returns, using this data in order to estimate an expected future return. Then, the second part consists in choosing which assets are going to compose the portfolio of investments.

Through its article, the author frequently uses the concept of correlation. The correlation between two articles is non-dimensional and varies between -1 and +1(CSOTA NETO, 2002); may be defined as:

$$r_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii} * \sigma_{jj}}}$$

Where $\sigma_{ii}\ e\ \sigma_{jj}$ são os desvios padrão e $\sigma_{ij}$ é a covariância entre os retornos dos ativos.

Markowitz states that a portfolio manager should diversify its investments. The author demonstrates that a investment diversification (i.e. the capital allocation in more than one asset) generates better portfolio compositions in terms of the relation risk return when compared to investments allocated in just one asset (BREALEY-MEYERS, 2003).

According to the author, assets that have a low covariance between each other ends up generating such a protection to the portfolio, because it decreases the portfolio risk. On the flip side, assets that have a high correlation between each other, as it reacts in way similar to the markets, may destroy value, since many times the financial markets do not behavior as the investor desires.

Figure 3 shows, graphically, the main Idea of the Markowitz`s theory, the investment diversification.

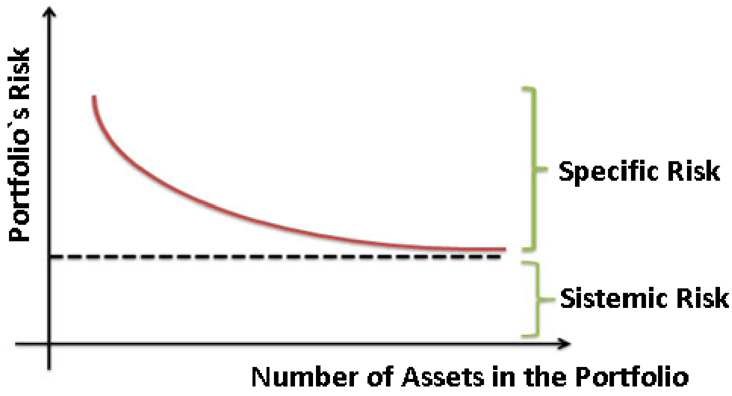



**Figure 3: Diversificatino effect on the risk mitigation**

It may be seen that as the numbers of assets increases the risk of the portfolio decreases, minimizing (eliminating, possibly) the Specific Risk, which is defined as the individual risk of an asset. In such way, the portfolio risk tends to be Systemic Risk, which cannot be eliminated through diversification. The latter is the risk associated to the market as a whole, being influenced by many aspects: political, socials, macroeconomics, among others.

Other factor that directly impacts the portfolio management is the investor risk profile. It is so, because not all the investors are willing to tolerate the same risk levels, what obligates the portfolio manager to always pay attention in this factor. GIUDICI (2010) defines, according to their risk profiles, three types of investors:

1. **Risk Averse**: an investor that chooses the lowest risk investment when faced with two investments with similar returns, with different risks though. For this kind of investor, the sense of unease associated with the loss of a determined amount of money is greater than the feeling proportioned by the gain of the same exactly amount of money.
2. **Risk Indifferent:** in this case, the investor does not have a preference in the moment of choosing between an investment with greater risk and greater return and an investment with lower risk and lower return.
3. **Risk Prone**: is the opposite from the risk averse investor, since this one prefers choosing an investment with greater return and greater risk to an investment with lower risk and lower return.

LUENBERGER (1998) proposes the concept of Utility Function, which is a manner of classifying the investments regarding the investor risk profile. Figure 4 presents a Return ($\mu$) *versus* Risk ($\sigma$) graph where it is illustrated the three curves correspondents to the Utility Functions for the three different risk profiles already mentioned. The curve $\rho_1$ represents the behavior of the risk averse investor, in which to an increment of risk Δσ, demands an increase of return such as$\Delta\mu > \Delta\sigma$. The curve $\rho_2$ illustrates the behavior of a risk indifferent investor, in which to an increment of risk Δσ demands just that$\Delta\mu = \Delta\sigma$. The curve $\rho_3$ represents the behavior of a risk prone investor, which behavior is the opposite from that one described for the curve$\rho_1$.

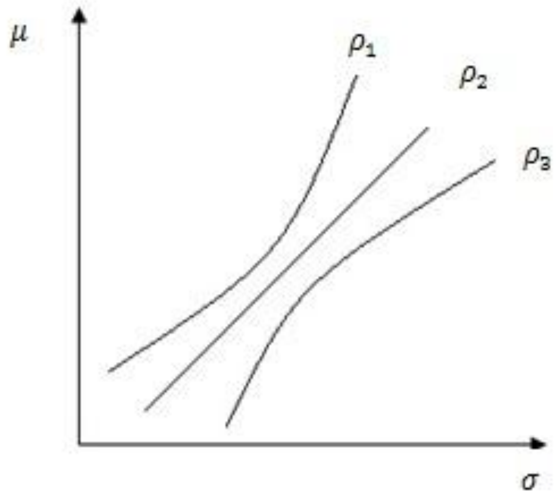



**Figure 4: Utility function for three different risk profiles**

It could be seen through this session that the theme risk is much discussed during the process of asset selection and allocation, and therefore, will be detailed even further in the next session, where it will be revealed a historical perspective of the risk measures present in literature, as well as its definitions, advantages and limitations.

## 2.2 Risk

Throughout history, many undesired situations from the financial stand point influenced the scientific and economic community to study and develop tools and metrics aiming to guarantee safer investments, in other words, decreases its risk. Some examples of these occurrences are: the 1929 financial crisis, the 1973 oil crisis, and more recently, the 2008 sub-prime crisis. All of them, although with its peculiarities, imposed catastrophic consequences on the global economy, negatively impacting the population well fare.

As said before, these financials crisis may be seen in a positive way, as they motivated studies on the risk control. Such fact occurred more heavily after the nineteen seventies, a period when many changes happened in the global scenario, such as the extinction of the currency fix rate regime and the implementation of floating exchange rate in some countries. Besides all of that, the rising globalization, which affected the economic, technologic and political parameters of such time, turned the countries more dependable from each other. Because of these reasons, the war regional effects, inflation differentials, changes in politics (such as the fall of the socialist world in the nineteen nineties), and natural disasters started to be reflected, in greater extent, in other economies, including on countries located at different continents. Such trend contributed even further for the increase the necessity of external risk control by financial institutions, encouraging the risk measurement and the study of new risk measures.

The most recent example of such movement was the enhancement of the Basel Indexes after the 2008 financial crisis. In 2010, a committee composed by the major political and economic authorities of the world gathered in Basel, in Switzerland, in order to create more strict rules to be applied into financial markets, more specifically, in the banks. An event of such global importance revealed that this theme continues to be relevant in the actual financial landscape.

But after all, how is risk defined? BARROSA (2015) states that risk, in its general form, is the product of an undesired result and its probability of occurrence, measured by monetary values. The determination of this undesired event and the knowledge of its probability of occurrence are what represent the focus of the definition of many risk measures.

More specifically, JORION (1997) indicates that risk may be defined as the variability of unexpected results (e.g. stock prices, exchange rates, interest rates and etc). ARTZNER (1999) emphasises that risk is related to the variability of future values of the portfolio positions due to market moves and efects on random variables which compose and charaterize an investment. Hence, it is confirmed that all operations are exposed to risk em greater or lesses degree.

It is interest to note that JORION (1997), besides of revealing a generalized concept of risk, also list and classify the many types of risk; presented in Table 1.

**Table 1: Risk definitions according to JORION (1997)**

| Type of Risk | Definition according to JORION (1997) |
|---|---|
| **Operational Risk** | It is related to the probability of loss caused by fail or inefficiency of internal process or even human errors. |

| | |
|---|---|
| **Liquidity Risk** | It is related to the capacity of the institutions to raise money and turn resources available to respect the cash flow and cover all the illiquid assets. |
| **Market Risk** | Risk related to the volatility of the assets prices. It may be directional (when related to the portfolio exposure to determined types of investments) or non-directional. |
| **Credit Risk** | Due to the possibility PF the counter-parts to not honor its debt obligations. |

There has been lately an intensification of studies about more robust risk metrics, motivated by the sub-prime financial crisis of 2008, but observing the literature, is comes clear that this theme is not recent. The first author that created a risk metric was Bernoulli in 1738. The author proposes the named Utility Function, which may be defined as a relative satisfaction measure of an economic agent ((BERNOULLI, 1738). This measure was utilized years later to characterize different investor profiles (LUENBERGER, 1998), as it was already mentioned. According to Bernoulli, from the analysis of its variation it is possible to explain the behavior of such agent, which in turn results in options chosen by the same in a way of increasing its satisfaction degree. It is, by the way, a very frequent measure in Economy to investigate the decision of consumption of goods and services. In economic terms, one may consider this measure revolutionary, since it was the first one to quantify, in fact, the expectation of the economic agents. However, its application in measuring investment risks presents a great drawback: its degree of subjectivity. It is so because the utility functions may assume many forms, such as quadratic, logarithm, exponential, among other, varying according to the economic agent.

This subjectivity drawback was partially overcame in mid XX Century by, the already mentioned, Harry Markowitz. In 1952, the author creates the so famous Modern Portfolio Theory, a piece of work considered a milestone in the Finance world and that served as the starting point of many other modern studies of risk measurement. MARKOWITZ (1952) created the Mean-Variance (which will be described deeply later on this study) to valuate investment portfolios.

In additional to introducing the concept of variance (also known as volatility), which allowed a standardization and conceptual alignment of the risk measure, the author also spread the idea that the covariance between two assets influences the overall portfolio return. He demonstrates that portfolios composed by assets of negative covariance present lower risk, to a certain risk level, when compared to portfolios that posses assets with positive covariance.

Simultaneously to the work of Markowitz, ROY (1952) developed the Safety First Criterion, in which the risk is measured as the probability of the return of certain portfolio being below a pre-established level, considered disastrous. This study is interesting because it is the first one to mention the concept of risk measured named Bellow-Target Models, being the introduction to the study of tail distribution of returns as a form of evaluating risk (ROMAN, 2008).

In Roy`s model, R is considered the return of a given investment and $\tau$ the level of return defined as a disaster. Hence, the Safety First Criterion is formally represented by:

$$\text{SFC} = P\,(R \leq \tau),$$

Being, therefore, considered a measure of probability. However, its application was limited to the arbitrariness of the definition of which would be the level of return taken as reference. So, in the end it was not broadly utilized in practice.

Despite of the fact that the Safety First Criterion did not have its practical development amplified, Roy introduced new concepts which were essential to the creation of new risk measures, mainly on the aspect of the observation of asymmetric distribution of probability of return, emphasizing one of the sides of the distribution, the one that represents the loss for the investor, called downside risk. Hence, lending continuity to the theme, MARKOWITZ (1970) developed a model named Partial Moments, in which the sample semi-variance is considered the risk measure.

The sample semi-variance may be defined as (ANDRADE, 2006):

$$\varsigma_i = \int Ma[0,(E(R_{it}) - R_t)]^2 dt$$

In which:

$\varsigma i$ is the semi-variance of the asset $i$,

$Rit$ is the return of the asset $i$ at the moment $t$,

$(Ri)$ is the average return of the asset $i$.

This definition turned the undesired side of the probability distribution of returns as the only side considered when analyzing a distribution, as defined by Markowitz.

Simultaneously, FISHBURN (1977) and BAWA (1978) enhanced the research through the creation of the model ($\alpha$,$\tau$), utilizing, this time, the Lower Partial Moment as risk measure, in a project that sums up and gathers the concepts of risk measures previously developed.

Differently from nowadays, during the nineteen seventies and the nineteen eighties, there were no methods and sophisticated computational tools to obtain solutions to the quadratic (or any non-linear) optimization problems of great size (PEROLD, 1984). It suited as a stimulus for academics to develop linear metrics for analyzing risk. Having this in mind, KONNO and YAMAZAKI (1991) proposed the utilization of the first absolute moment of return distribution as a risk metric (RIBEIRO, 2004), through the model named Mean Absolute Deviation (MAD), enhanced by the innovative work of SHARPE (1971). This risk measured is defined as:

$$MAD(Ri)= E[|Ri-\mu|],$$

which turns the portfolio optimization problem a linear programming problem, presenting an alternative to the Mean-Variance. However, its optimization is not simple, since it is an absolute function and, therefore, presents discontinuities in its derivative. Hence, both analytical and numeric methods turn to be non-practical, given the quantity of restriction intrinsic to the optimization problem.

Given sequence to the development of risk measures that consider extremely undesired events (as the case of the financial crisis) caused on the tail distribution of loss probability, the G-30 proposed, in 1994, a risk measure whose objective is to answer one simple question: “How big may the loss of an investment in a certain period of time and a probability?” (ROMAN, 2008). This risk measure is called Value at Risk – VaR, defined as:

$$VaR(R_i)= -q_\alpha(R_i)=q_{1-\alpha}(R_i)$$

In which:

$Ri$ is the return of a certain risk $i$

$q\alpha$ is the percentile defined in a given confidence level $\alpha$ (G-30, 1994).

Besides of being applicable as a tool on the portfolio optimization decision making, the Value at Risk is also used, throughout the planet, as a regulatory measure. This regulatory process came through especially after the notorious publication of the article “Risk Metrics: Technical Report”, by the north-American bank JP Morgan, in 1994 (ROMAN, 2008).

According to the criteria defined by ARTZNER (1999), the Value at Risk (VaR) is not considered a consistent measure of risk. It is due, mainly, by the fact that it does not attend the following probability:

$$Va(R_1+R_2)\leq VaR(R_1)+VaR(R_2))$$

In other words, this risk measure does not have the subjectivity property, making the diversification not necessarily awarded. Again, it is not possible to guarantee that the risk of a portfolio composed by two assets, each other with a risk $VaR_1$ and $VaR_2$, respectively, is equal or lower than $VaR_1 + VaR_2$. Moreover, the Value at Risk (VaR), in its non-parametric form, presents a variety of quantities of local minimums, turning its optimization a hard task (QUARANTA and ZAFFARONI, 2008). Even so, the broad dissemination of the Value at Risk as a risk measure and its convenience of conceptual comprehension make it selected for further analysis in this study.

Since the VaR presents such controversial aspects as risk measure, it was proposed in 200 the creation of a new measure, the named Conditional Value at Risk (CVaR). ROCKAFELLAR and URYASEV (2000) defined the CVaR as:

$$CVaR(R_i)=E\{(R_i)|R_i\leq v\}$$

In which:

$R_i$ represents the return f the asset $i$,

$v$ represents the Value at Risk (VaR) of the probability distribution of returns of this same asset.

In other words, the CVaR is the average of the values which exceed the Value at Risk. The CVaR, in turn, attends the properties defined by ARTZNER (1999), being therefore coherent (LIM, 2011). However, as its definition is representative of the modeling of the tail probability distribution of returns, it presents great complexity, turning its optimization very hard and by no means trivial (RIBEIRO, 2004).

ROCKAFELLAR e URYASEV (2000) proposed, aside from the risk measure itself, sophisticated techniques of optimization of this function. Hence, for this reason and for being the goal of many researches around the globe, the CVaR, in addition to the Variance and the VaR is selected to deeper analysis in this study.

Figure 5 gives a historical perspective of risk measures available on literature until nowadays, which were already described in this study.

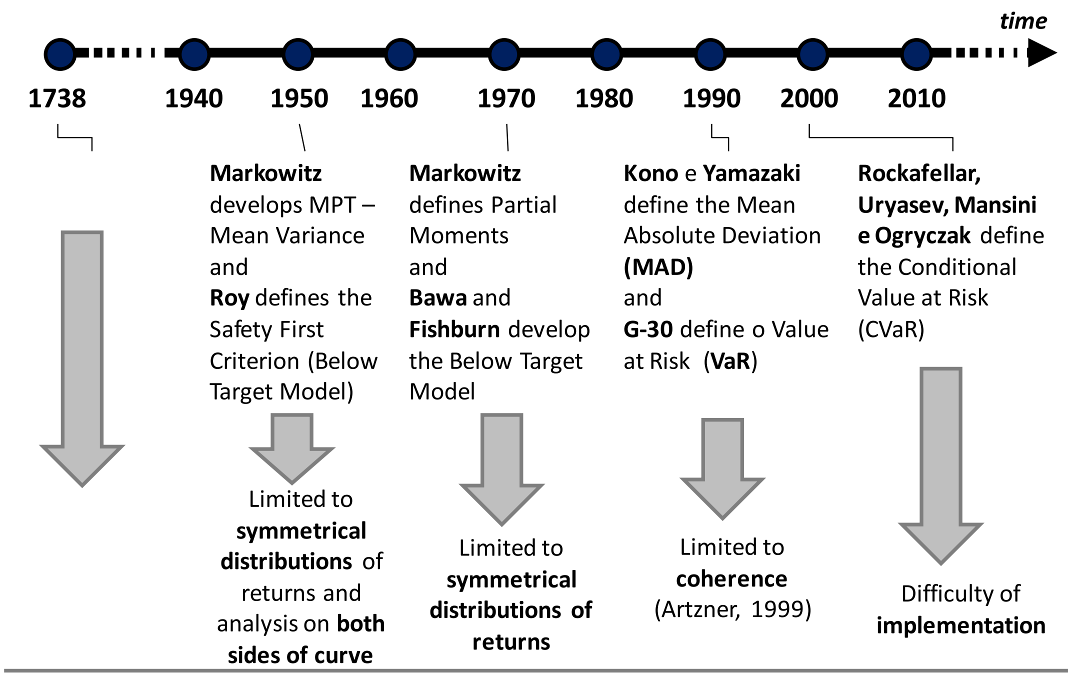



**Figure 5: Historical perspective of the risk measures**

From what has been revealed in this study, it is possible to classify the main risk measures proposed in literature in two categories:

1) **Deviation from Target**

2) **Tail Measures**

Given that the first category may be sub-divided in two other categories:

1.1) **Symmetrics**: consider both sides of the probability distribution of returns.

1.2) **Non-symmetrics**: consider just the side of the losses of the probability distribution of returns.

Table 2 classifies each risk measure described in this session within this classification.

**Table 2: Categories and descriptions of the risk measures**

| Category | Description | Examples |
|---|---|---|
| *Deviation from Target* | Symmetrical | Variance<br>MAD |
| | Asymmetrical | *Central Semi deviation*<br>*Lower Partial Moments* |
| *Tail Risk Measures (seriousness of potential loss)* | Considers the worst case scenario at α confidence level | VaR<br>CVaR |

### 2.2.1 Variance

Proposed by MARKOWITZ (1952), the variance indicates the mean of the quadratic deviation of a random variable and the distribution average. Regarding the portfolio management subject, it common to say that the variance measures the degree of deviation between the expected returns of the assets.

For a random variable, the variance is defined as:

$$var(x) = \sigma^2(x) = E[(x - \bar{x})^2]$$

In which:

$$E(x) = Expected\ value\ of\ x$$

According to COSTA NETO (2002), depending on the nature of the random variable (i.e. if it is discrete or continuous), the variance is calculated as the following:

$\sigma^2 = \sum (x - E(x))^2 \times p(x)$ , for discrete variables

$\sigma^2 = \int_{-\infty}^{\infty} (x - E(x))^2 \times f(x) d(x)$ , for continuous variables

In which:

$$p(x) = \ probability\ of\ x$$

$$f(x) = \ density\ probability\ function\ of\ x$$

This statiscal measure is broadly utilized for the risk measurement, and, in the specific case of a portfolio composed by different assets, it may be applied with the support of a covariance matrix. However, the variance has some restrictions regarding it use and may be utilized in asymmetrical probability distributions (SZEGÖ, 2002). Therefore, it is possible to use the variance model merely to analyze elipitical distributions, such as normal and t Student distributions, which does not represent the majority of the existing distributions.

There are two main characteristics of the variance that hamper its utilization as an efficient risk measure in the portfolio management:

a. The variance does not consider the difference between negative and positives returns in relation to the expected returns, which have opposite impacts on the return of the investors and in their perception, who gives priority to those assets that present a greater return than the expected value;
b. This risk measure also does not analyze the tail distribution (RIBEIRO and FERREIRA, 2004); what may represent great losses in stressed scenarios.

As such flaws started to be highlighted, came up in the literature other studies regarding more robust risk measures which solved the problems revealed by the variance. Hence, in 1994, it was figured out the concept of Value at Risk (VaR) (SZEGÖ, 2002).

### 2.2.2 Value at Risk (VaR)

It is a risk measure used by many economic agents: regulatory agencies, financial institutions, portfolio managers and central banks (HULL, 1999). According to this author, the Value at Risk was created as an attempt to summarize, in just one number, the risk involved in a certain portfolio of financial assets.

This metric involves the definition of level of reliability, time horizon and percentiles. The VaR may be defined as the value that represents the great loss that will occur with a probability α% in a certain time horizon.

Depending on the sector that the company is placed and on the portfolio to be analyzed, the time horizon for the VaR analysis will vary. For a company that has a great asset turnover, the time horizon will b short, for example, a month. Now, in companies whose assets are traded at a slow pace, the time horizon could be a year (RIBEIRO and FERREIRA, 2004).

According to Quaranta e Zaffaroni (2008), considering K a random variable and F its distribution function, $F(h) = P\{K \leq h\}$ e $F^{-1}(w) = \min\{h: F(h) \geq w\}$, for a fixed level of reliability ($\alpha$), it is given that:

$$VaR_{\alpha}(K) = F^{-1}(\alpha)$$

**Erro! Fonte de referência não encontrada.** presents the VaR`s graphic definition.

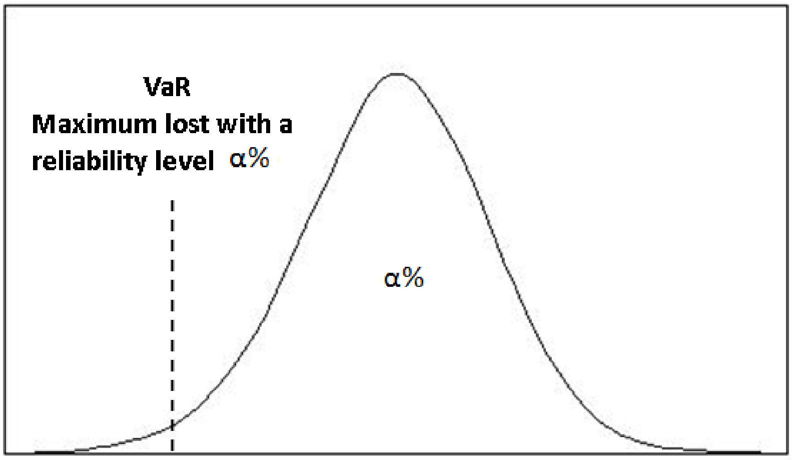



**Figure 6: VaR graphic representation**

JORION (1997) demonstrate two methods to calculate the VaR:

1) **Parametric**: considers that the portfolio return presents a normal distribution what simplifies its calculation. It receives such name because the parameters are estimated instead of identifying its percentiles; in this method, the VaR derives directly from the standard deviation using a multiplication factor that depends on the level of reliability

$$VaR = \mu - Z_{\alpha} * \sigma$$

In which:

$\mu$ is the average of the returns

$\sigma$ is the standard deviation of the returns

$Z_{\alpha}$ is the value that representd the inverse of the normal cumulative distribution.

2) **Non-parametric**: also known as the historical series method, considers the N portfolio returns sorted, so that the VaR consists on the ((1-α)-N)th worst value of the series. This method assumes as assumption that the future return is linked to the past return.

Besides the methods proposed by Jorion, this measure may be calculated by Monte Carlo simulation. This method uses this simulation to construct a variety of scenarios to generate a prediction if future returns for each one of them, based on the historical series of returns (RIBEIRO, 2004).

BARROSA (2015) performed, as an example, a study in order to demonstrate the VaR`s behavior, both the parametric and the non-parametric. The author utilized a portfolio composed by two assets, Itaú-Unibanco (ITUB4) and Petrobrás (PETR4), both traded in the BOVESPA (i.e. the Brazilian stock exchange) and a sample of the historical series between May 27th, 2009 and May 11th, 2012, totalizing 718 observations. To facilitate the visualization of the results, the author takes $\lambda = x1$ and $x2 = 1 - \lambda$, being $\lambda$ a portfolio. The result of such study is demonstrated on Figure 7.

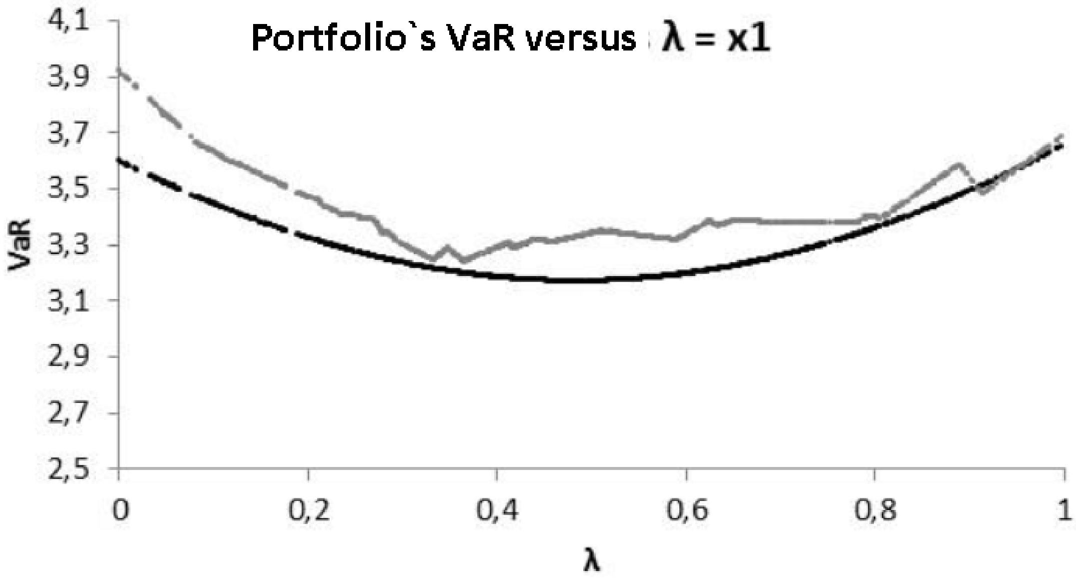



**Figure 7: Exemple of application of the VaR**

When analyzing Figure 7, it is possible to notice that the normality hypothesis of the returns assumed on the application of the parametric method is not necessarily true, so that the estimators obtained through the historical series samples of the assets returns, assuming normality, present error. In this sense, COSTA and BAIDYA (2001) verified empirically the non-conformity of many Brazilian assets with the hypothesis of symmetry on the probability distribution of returns. The assessment of the surface VaR(x) through its broadest method of calculation, the non-parametric, relaxes this hypothesis, but its optimization becomes significantly more complex, mainly due to the existence of many local minimums.

Even knowing that the VaR provides information about the tail distribution of returns, SEIGÖ (2002) presents many problems related to the application if this risk measure. Among them, it may be highlighted:

a. It does not measure the losses that exceed the VaR, in other words, this risk measure does not accomplishes in given information about the dispersion o the tail distribution beyond its value, given a certain level of reliability;

b. It may generate conflicting results for different levels of reliability;
c. It is not considered a coherent risk measure.

ARTZNER (1999) defines the properties for a risk measure to be considered coherent:

a) Invariation of translations: $\rho(x + \alpha \bullet r') = \rho(x) - \alpha, \forall\ \alpha \in R, \forall\ X \in G$
b) Subadditionality: $\rho(x_1 + x_2) = \rho(x_1) + \rho(x_2), \forall\ x_1\ e\ x_2 \in G$
c) Positive homogeneity: $\rho(\lambda x) = \lambda\rho(x), \forall\ \lambda \geq 0, \forall\ x \in G$
d) Monotonicity: $\rho(y) \leq \rho(x), \forall\ y, x \in G, y \leq x$

The VaR does not present the property of subadditionality, in other words, does not guarantee that the risk of a portfolio composed by two assets with $VaR_1$ and $VaR_2$is equal or lower than $VaR_1 + VaR_2$. The risk of this portfolio cannot be predicted, and this complicates its optimization (QUARANTA and ZAFFARONI, 2008).

### 2.2.3 Conditional Value at Risk (CVaR)

The risk measures presented so far are not convex, or in other words, if those are applied to non-elliptical distributions it provides inconsistent results. Furthermore, these measures do not analyze the tail distribution for extreme scenarios.

Within the objective of solving these problems, the literature has been given in the last decade a great importance to the Conditional Value at Risk (CVaR), a coherent risk measure that may be defined as the average of the values that exceed the VaR, for a certain level of reliability. In other words, considering the worst case scenario, this measure provides the mean value of the tail. This definition guarantees that the VaR is never greater than the CVaR in absolute value (ROCKAFELLAR and URYASEV, 2000).

Having $\mathbf{x} \in \mathrm{X} \subset R^N$, a decision vector representing the portfolio and $\mathbf{y} \in \mathrm{Y} \subset R^N$the future values of returns of those assets which compose the portfolio, $z = f(X, Y)$the function of the portfolio losses; the CVaR is given by (QUARANTA and ZAFFARONI, 2008):

$$\Psi(x,\mathrm{a}) = P\{y|f(x,y) \leq \mathrm{a}\}$$

In which:

a is the portfolio`s VaR.

**Erro! Fonte de referência não encontrada.** presents the graphic definition of the CVaR.

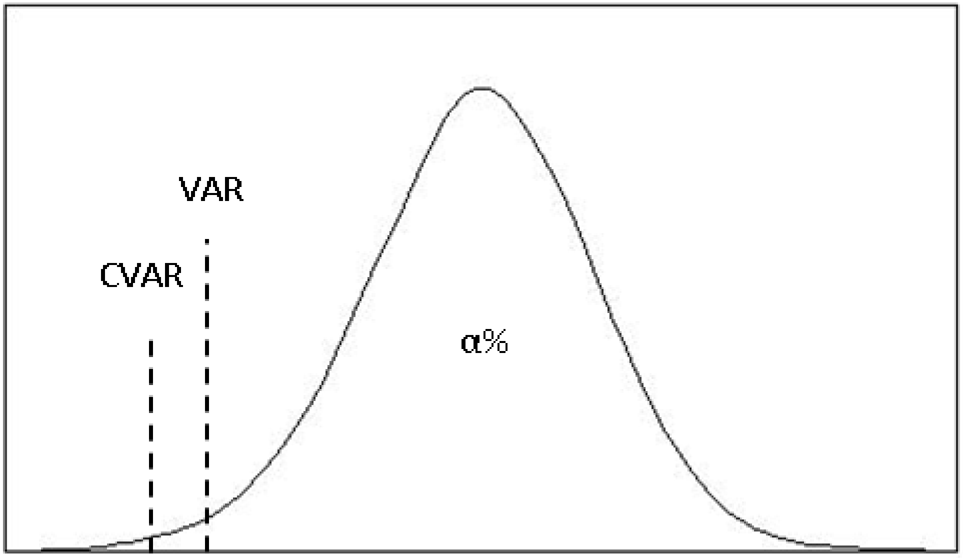



**Figure 8: CVaR graphic representation**

One may observe that the VaR and the CVaR measure different properties of the distribution, since the former refers to the percentile and the latter to the tail average (PFLUG, 2000). The CVaR presents consistency with the VaR only in normal distributions (or elliptical) (ROCKAFELLAR and URYASEV, 2002).

Despite of the fact that the CVaR depends on the VaR determination, it is possible to define simultaneously these two risk measures through the following function (ROCKAFELLAR and URYASEV, 2002):

$$F_{\alpha}(x,\text{a}) = \text{a} + \frac{1}{1-\alpha} E\{[f(x,y) - \text{a}]^{+}\}$$

In which:

$$[t]^{+} = max\{0, t\}.$$

Hence, one may affirm that the CVaR presents some advantages over the VaR:

- It is a coherent risk measure, respecting all the axioms proposed by ARTZNER (1999);
- It provides information about the distributions` tail, analyzing stressed scenarios.
- It may be algebraically expressed by a formulation that pursues to change the portfolio composition problem in a linear programming problem, which will be presented in the next section.

Again, BARROSA (2015), as an example, perform a similar study: the author considers the same example from the previous section (i.e. a portfolio compose by Itaú-Unibanco (ITUB4) and Petrobrás (PETR4), both traded in the BOVESPA and a sample of the historical series between May 27th, 2009 and May 11th, 2012) in order to calculate the portfolio`s CVaR. The author takes $\lambda = x1$ and $x2 = 1 - \lambda$, being $\lambda$ a portfolio. Results are presented in Figure 9.

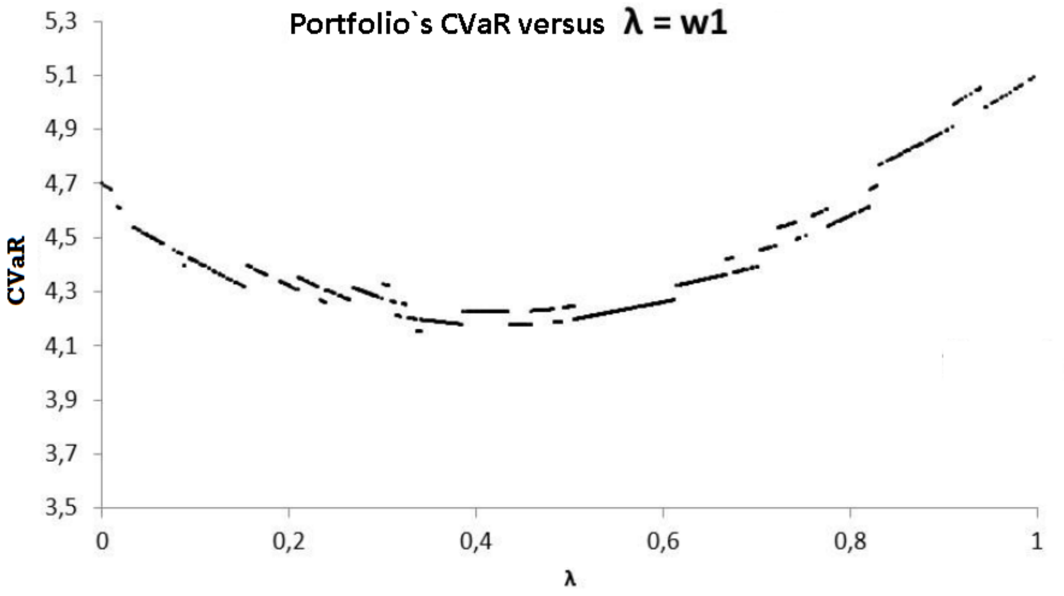



**Figure 9: Example of application of the CVaR**

The author (BARROSA, 2015) notes that the behavior of the presented curve, including once again the existence o local minimums and discontinuities, making substantially harder its optimization through conventional methods, turning its large scale application not favored.

Nowadays, there is no consensus on literature about which is the most adequate risk measure on practice, since the CVaR, even being coherent, is not easy to be applied (as demonstrated by Barrosa), demanding a great computational effort.

Hence, academics concentrate their attentions in two main research topics:

1) The definition of new risk measures, so that better characterize the probability distribution of returns;
2) Enhancing the studies on existing risk measures, focusing not just on those used in practical terms, as the VaR, but also on those coherent and considered innovative, as the CVaR. Both having the purpose of analyzing their behavior, and finally, enabling their use as an object function in decision processes.

Table 3 presents the three risk measures which are presented in literature and that will be used in this study.

**Table 3: Coherence and limitations of the risk measures**

| Risk Measure | Category | Coherency | Limitations |
|---|---|---|---|
| **Variance** | Deviation | Yes | -Applicable just in symmetric distributions<br><br>-Does not different positive returns from negative returns<br><br>-Does not consider the distribution tail |
| **VaR** | Tail | No | -Does not provide information of events that exceed the VaR<br><br>-Not coherent |
| **CVaR** | Tail | Yes | -Tough practical application |

## 2.3 Portfolio Selection Models

The determination of the portfolio composition is directly linked to the risk associated with this portfolio and the return yielded. The objective of a portfolio manager is, for a given level of risk, maximize the return, or similarly, for a given return, minimize its risk.

However, this is a non-trivial issue, since the financial assets are exposed to many types of risk, such as market risk, liquidity risk, credit and/or operational risk. For example, one single stock is subject to risks related to the company image, its directors' reputation, or even market and liquidity risks. In the case of the portfolio composition, the problem is hampered by the existing correlations of the many assets involved (MARKOWITZ, 1952).

Given $A_0$a fix amount of money available to the investment allocation, measured in monetary terms, and $n$ the number of pre-selected assets, the portfolio is defined as the asset obtained through the allocation from $A_0$ to $n$. It is the same as stating:

$$A_{0i} = x_i A_0$$

With: $$\sum_{i=1}^{n} x_i = 1$$

The vector $x = x_1\ x_2 \ldots\ x_n$ represents the allocation of the amount $A_0$ in each asset $i$ which compose the portfolio, or the weight of each asset, and more importantly, it represents the variables of the portfolio optimization problem.

The portfolio return is calculated by the weighted average of the individual returns of each asset the compose the portfolio, what may be easily expressed by:

$$R_c = \frac{\sum_{i=1}^{n} R_i x_i A_0}{A_0} = \sum_{i=1}^{n} R_i x_i$$

Hence, the portfolio management problem has the purpose to minimize certain function *Risk(x),* which represents the portfolio risk as a function of its composition subject to restrictions regarding the expected return $R_c$. In order to define more precisely this problem, two elementary assumptions of the human behavior are necessary (LUENBERGER, 2008):

(1) **Non-satiety**: the investor prefers more money over less money;

(2) **Risk aversion**: given two portfolios with the same return and different risks, the investor opts for the lower risk.

Therefore, the portfolio optimization problem, in its general form, is defined as(BARROSA, 2015):

**Minimize:** $f(X) = Risk(x)$

**Subject to:**

$$\sum_{i=1}^{n} E(R_i) x_i \geq G$$

$$\sum_{i=1}^{n} x_i = 1$$

$$x_i \geq 0; i = 1, \ldots, n$$

The first constrain refers to the model`s parameterization regarding the average expected return of the portfolio, which is necessarily greater or equal to the minimum value stipulated *G* by the investor.

The second constraint guarantees that the exact same amount of resources available for investing is allocated, what is surely accomplished, since the sum of the weights of all the assets which composed the portfolio is equal to one.

Finally, the third constraint guarantee that all portfolio allocations are non-negative. In other others, this restriction does not allow short selling.

The problem was presented, in this section, in its general form and, additionally, in the next sections, will be revealed its specific application using the following risk measures: the Variance (Markowitz Model), the Value at Risk (VaR), the Conditional Value at Risk (CVaR), and at last, the method hear proposed, the Kriging Method, which allow its application within any risk measure.

### 2.3.1 The Markowitz Model

Harry Markowitz, through his famous publication in 1952, called Modern Portfolio Theory, attempted to develop a universal metric of market risk for a given investment. The risk measured used is the Variance, which will be detailed in this study.

Now, let Σ be the covariance matrix between the assets that compose a certain portfolio, in which:

$$\sum_{ij} = cov(R_i, R_j) = \sigma_{RiRj} = E[(R_i - \mu_{Ri})(R_j - R\mu_{Rj})] \ ,$$

Or:
$$\sum = \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \cdots & \sigma_n^2 \end{bmatrix},$$

which represents a symmetric matrix with individual variances from the historical series of each asset that compose the main diagonal, and with the covariance between these assets, for every $i \neq j$, it is given that the risk of this portfolio is:

$$Risk\ (x) = \sigma_c^2 = x' \sum x \ ,$$

which is a quadratic function that represents the portfolio variance depending on its composition (MARKOWITZ, 1952).

This definition led the creation of the original concept of diversification of Markowitz, considered a milestone in his time and continues until now being a dogma among the portfolio manager. This concept says that the portfolio composed by assets with negative correlation may presents a better risk-return relation when compared to an investment in one single assets or in assets with a positive correlation (LUENBERGER, 2008).

Thus, the portfolio composition problem using the variance as the risk measure may be expressed in the following manner:

**Minimize:** $Risk(X) = x' \sum x$

**Subject to:**

$$\sum_{i=1}^{n} E(R_i)x_i \geq G$$

$$\sum_{i=1}^{n} x_i = 1$$

$$x_i \geq 0; i = 1, \dots, n$$

Given that the three constraints here presented are equal to those described in the former section of this study and the objective function, which represents the portfolio risk, is the variance in itself.

In practice, the parameters of the function $x' \sum x$ are estimated through a sample study of the historical returns of each of the assets that compose a certain portfolio.

BARROSA (2015) creates an example to clarify the above described concepts. In Figure 10 it is possible to see the curve Risk x Return for a simulation of a portfolio composed by three assets, providing different portfolio compositions.

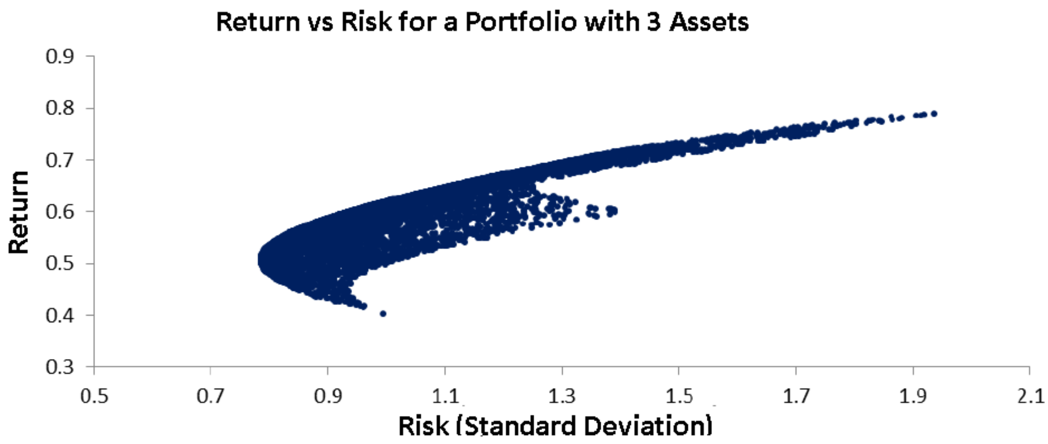



**Figure 10: Return x risk relation for a three asset portfolio**

It stays clear through the graph observation that the Risk (x) function assumed in the example (in this case, the variance) is a quadratic function. It is worth noting also that there are three bordering points, which represent the three portfolio composed by an asset weighted one, or in other words, portfolio composed by one single asset.

Considering that a portfolio optimization problem presents assumptions of non-satiety, return maximization, and risk aversion, minimizing the potential losses, it stays clear with the support of Figure 10 that only one small portion of the different portfolios with different compositions would in fact satisfy both assumptions.

The region of the graph that, in fact, is capable of satisfying both simultaneously is obtained through an optimization method (e.g. the Kuhn-Tucker method), minimizing therefore the function $Ri(\boldsymbol{x})=\boldsymbol{x}'\Sigma\boldsymbol{x}$ and parameterizing the minimum return desired by the investor through the restriction $R_c \geq G$. The results of these optimization problems assuming different values o $G$ are presented in Figure 11. This curve is denominated Paretto's Optimality Frontier for conflicting objectives ((PAPALAMBROS, 2000).

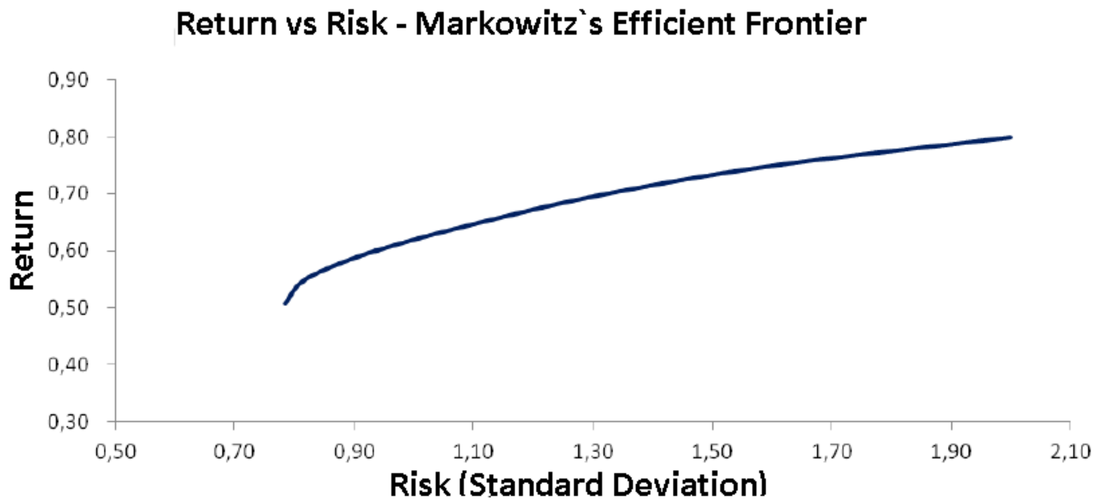



**Figure 11: Markowitz efficent frontier example**

It is important to note that any portfolio compositions that are not placed on this curve present, necessarily, a relation Risk-Return inferior to the points on such curve. In other words, the points that were generated with a greater risk level for a given return, or with a lower return for a given risk, are not, thus, considered in the optimal compositions.

At this stage of the study, it is important to utilize the concepts of Utility Function and different investor profiles regarding the risk. It is so, because this function sorts the investments by the investor risk profile (LUENBERGER, 2008), and therefore, it is it that, considering the basic idea that greater risks imply in greater return and vice-versa, will define which point on the Efficient Frontier curve the investor will pick.

Since it presents a quadratic function optimization problem, subject to linear constraints, the Markowitz model is not perfect, in a way that it is based upon the hypothesis of symmetry of the probability distribution of returns of the assets which compose the portfolio. Moreover, this model also deals with another weakness: its fragility on representing stressed scenarios (tail risk), which will be surpassed with the utilization of the models VaR and CVaR

### 2.3.2 The VaR Model

Given the weakness of the Mean-Variance model in not considering the analysis of the tail distribution (i.e. scenarios with significant losses, such as losses during financial crisis), it is necessary to use other risk measures when dealing with portfolio optimization problems. This was, thus, one of the motivations for the development of the Value at Risk (VaR).

This model generates the same efficient frontier of the former model when faced with normal distributions. The difference comes up in the case of non-normal and asymmetric distributions, which represents the behavior of many random variables. Another great difference between this model and the Markowitz one consists on the fact that the later consider the mean deviations, for both positive and negative values, what does not match with the investor way of thinking. In the VaR model, it is considered just the tail distribution that represents a loss in risk analysis.

Similarly to the Markowitz model, to solve this kind of problem using the VaR as risk measure, it is common to perform a sample of the historical series of the return in order to obtain estimators for the main parameters of the probability distribution. Analogous to the parametric calculation method for the portfolio variance, the portfolio VaR may be calculated through the following relation (JANABI, 2012):

$$VaR\ (x) = VaR_c = [\nu^T \Gamma \nu]^{1/2}$$

Where $\boldsymbol{\nu}$ represents the vector of the individual VaRs of each asset, in function of its individual assets, $\boldsymbol{x}$ and $\boldsymbol{\Gamma},$ represent the correlation matrix between the assets that compose the portfolio, being:

$$\Gamma = \begin{bmatrix} 1 & \cdots & \rho_{1n} \\ \vdots & \ddots & \vdots \\ \rho_{n1} & \cdots & 1 \end{bmatrix}$$

The matrix $\boldsymbol{\Gamma}$ is symmetric to the main unit diagonal, indicating the correlation between the historical series of returns of the asset *i* and itself. The other values of the correlation matrix represent the Pearson Coefficient of Correlation, obtained through the sample of historical returns between each asset *i* and *j*.

And now, regarding the VaR calculation by the non-parametric method, the methodology is simple: the ordination of the asset VaR is substituted by the ordination of the different portfolio VaRs and, hence, the technique is applied as it was presented before.

In such way, the portfolio optimization problem considering the VaR as risk measure may be written in the following manner:

**Minimize:** $Risk\ (x) = VaR\ (x) = [\nu^T \Gamma \nu]^{1/2}$

**Subject to:**

$$\sum_{i=1}^{n} E(R_i)x_i \geq G$$

$$\sum_{i=1}^{n} x_i = 1$$

$$x_i \geq 0; i = 1, \ldots, n$$

Again, the constraints of this problem are the same of those in the base models, but this time the objective function to be minimized is the function that calculates the portfolio VaR.

However, as it has already been noted, the VaR has a few limitations, among them: it does not offer information about the dispersion of the distribution tail beyond a certain value depending on the level of reliability; and it is not considered a coherent risk measure (ARTZNER, 1999). Aiming on solving these drawbacks, ROCKAFELLAR and URYASEV (2000) developed and applied the Conditional Value at Risk (CVaR) in portfolio optimization problems.

### 2.3.3 The CVaR Model

The critics to the Mean-Variance model (Markowitz model) in addition to the search for a coherent risk measure (attempting to overcome one of the VaR's limitations), made the CVaR model notorious in the literature. This model is based on the portfolio CVaR and leads to more reliable results regarding the portfolio risk, since it considers the tail risk, and moreover, the values that surpass the VaR.

As previously presented, the CVaR calculation depends on the determination of the portfolio VaR, what may seem complex in practice. However, ROCKAFELLAR and URYASEV (2000) proposed a simpler approach to this problem, in which the VaR is calculated and at the same time the CVaR is minimized.

Yet according to the same authors, being $f(X, Y)$ the loss function associated to a decision vector X $\epsilon$ $R^n$and a random vector Y $R^m$, for each vector X, the loss $f(X, Y)$ is a distribution random variable in R inducted by the vector Y, which has density $p(Y)$.

The portfolio return is calculated through the sum of the product of the weights and returns of the individual assets. The function loss is the opposite of this return

$$f(X, Y) = -[x_1y_1 + \cdots + x_ny_n] = -X^TY$$

And when negative represents a gain.

Therefore, the mean and the variance of the loss function associated to the portfolio X may be defined in terms of the average $m$ and the covariance matrix $\sum$ of the returns:

$$\mu = -X^T m$$

$$\sigma^2 = X^T \textstyle\sum X$$

The probability that $f(X, Y)$ does not exceed a level $a = VaR$ is:

$$\Psi(X,\text{a}) = \int_{f(X,Y) \leq \text{a}} p(Y)\, dy$$

It is assumed that $\Psi(X, \text{a})$ is non-decreasing and continuous regarding the VaR, in order to simplify the mathematical formulation which follows the CVaR calculation. This function determines the behavior of the random variable and is fundamental to the determination of the risk.

For a given level of probability α between (0,1), in which α may assume values as $\alpha = 0.90$ or $\alpha = 0.99$, for example, the VaR and the CVaR may be defined as:

$$\text{a} = VaR(X, \alpha) = \min \{\text{a} \in R \colon \Psi(X,\text{a}) \geq \alpha\}$$

$$F_\alpha(X,\text{a}) = CVaR(X,\text{a}) = \text{a} + \frac{1}{(1-\alpha)} \int_{Y \epsilon R^m} [f(X,Y) - \text{a}]^+ p(Y) dy$$

In which:

$$[f(X,Y) - \text{a}]^+ = [t]^+ = max\{0, t\}.$$

ROCKAFELLAR and URYASEV (2002) proposed a mathematical formulation which transforms the CVaR calculation problem into a linear programming problem. What the model proposes is a manner of discretizing the integer to facilitate an approximation of the CVaR. For this purpose, the authors suggest the use of samples of the probability distributions of Y, according to its density $p(Y)$, generating many vectors $y_1, y_2, \dots, y_q$. Moreover, associated to the creation of base cases scenarios, it may be applied for analyzing and optimizing the risk of a portfolio composed by a great number of assets, without many computational resources.

Therefore, taking in consideration the quantity of scenarios generated (q), an approximation to the function $F_\alpha(X,\mathrm{a})$ is given by:

$$F_\alpha(X,\mathrm{a}) = CVaR(X,\mathrm{a}) = \mathrm{a} + \frac{1}{q(1-\alpha)}\sum_{k=1}^{q}[f(X,Y_k) - \mathrm{a}]^+$$

Substituting the term $[f(X,Y_k) - \mathrm{a}]^+$ by auxiliary variables $\mu_k$ that comply with constraints that guarantee that its value is equal to $max\{0,t\}$, transforming the model resolution into a linear programming problem.

It may be described in the following manner:

$$Min \qquad \tilde{F}_\alpha\,(X,\mathrm{a}) = \mathrm{a} + \frac{1}{q(1-\alpha)}\sum_{k=1}^{q}\mu_k$$

Subject to:

$$x_j \geq 0\; for\; j = 1, \dots, n$$

$$\sum\nolimits_{j=1}^{n} x_j = 1$$

$$X^T Y \geq G$$

$$\mu_k + X^T Y_k + \mathrm{a} \geq 0$$

$$\mu_k \geq 0\,, \quad k \in \{1,2,\dots,q\}$$

In which:

$q$ is the number of scenarios generated

$G$ is the minimum accepted return

$\mu_k$ are auxiliary variables that substitute $[X^T Y_k - \mathrm{a}]^+$

The first constraint refers to the requirement that the asset allocation is positive, not considering thus, short selling positions.

The second one guarantees that all capital available will be invested.

The third one imposes that just the portfolios with a minimum return R are considered.

At last, the fourth and fifth constraints deal with the variable $\mu_k$, which must be positive, satisfying the restriction described in the fourth constraint.

The solution for the problem is the approximation of $F_\alpha(X, a)$ by $\tilde{F}_\alpha(X,\mathrm{a})$ and later minimization of the same, which is a function convex, linear, and differentiated in relation to X and to the VaR, and also may be minimized using regular methods of linear programming, what makes its implementation attractive.

Although the attention is not aimed directly to VaR itself, since the CVaR ≥ VaR, the portfolio which minimizes the former tends to be a good solution for the minimization problem of the later.

The transformation for the linear programming problem of the CVaR optimization does not depend that Y have a normal distribution previously known, which brings the model preferred over the Markowitz one. In addition, it is also considered a more robust model than the VaR, since the CVaR analyzes better the tail of the probability of the returns, and moreover, includes the calculation of the VaR itself implicitly.

### 2.3.4 Proposed Method: Kriging

The model revealed above, the CVaR model, presents a drawback: its hardship on the practical application. It happens, because in order to turn the model into a linear programming problem it is necessary to include more variables and constraints according to the number of scenarios; a number that turns greater as bigger the sample generated by the Monte Carlo simulation is.

This study has the objective to propose an optimization problem that aims to decrease the number of variables and the computational work to obtain the optimal portfolio composition. This method seeks to create an approximate surface of the function to be minimized.

In the models described in the previous sections, the past behavior of the returns of the assets is used in order to predict what is going to happen in the future, considering a covariance matrix in the generation and analysis of scenarios. This basic principle is used in the proposed method as well, since it models the tail of the distribution, proposing an approximation of the surface of the function, based on the historical values which suit as input to the problem solution.

It is known that, in order to approximate a function, one shall choose the appropriate points of the grid, or in other words, points that will represent the data in the space. There are many techniques to approximate a function, but in this study it will be proposed the Kriging Method, also known as DACE fit (Design and Analysis of Computer Experiments). This technique has its origins in the study of geology problems and is known as Kriging (RIBEIRO and

FERREIRA, 2004), which is a regression model used in geostatistics to interpolate data (YIN, J and NG, 2011).

Although this method is not traditionally applied to financial problems and initially had been used for geology problems, it is believed that are some similarities in both applications that justify its use in this study. In the case of the land composition, there is no hindering on obtaining the real values of the composition in all the researched area. In the case of the CVaR problem, despite of being possible to determine its value for a great part of the portfolio conFiguretions, the behavior of this theoretical function hampers its optimization.

The Kriging method proposes a fit on the surface of answer of the collected data, valuating the objective function and the problem restrictions in a few determined points. This answer surface is used to analyze the relations between the inputs and outputs of the problem, as well as the estimation of its optimal value (JONES, SCHONLAU e WELCH, 1998).

This technique has its objective function treated as the result of a stochastic process previously defined, characterized by a correlation function between the calculated values in different pair of points (JONES, SCHONLAU and WELCH, 1998). It is usually used in cases such that the computational cost of the objective function is high, not necessarily, in the case of the CVaR.

Considering the vector $X^i \in R^n$ $(i = 1, 2, \dots, n)$ and the vector $Y = \{y^1, y^2, \dots, y^q\} \in R^q$, in which n represents the number of assets which compose the portfolio and q the number of points observed in the grid, the DACE fit provides a polynomial approximation of the function $y^i = f(X^i)$, interpolating the points observed through the equation:

$$y(X^i) = \mu + e(X^i)$$

Where $e(X)$ are the random errors, correlated, normally distributed, with mean zero and constant variance, $\sigma^2$.

The correlation between $e(X^i)$ and $e(X^j)$, cited above, depends on the distance between the points. It will be greater when $X^i$ e $X^j$ are close to each other, what is the same as saying that it will tend to one when the distance is small and to zero when these grid points are distant from each other. The errors covariance is given by:

$$\mathrm{cov}\left(e(X^i), e(X^j)\right) = \sigma^2 \Sigma_{ij}$$

In which $\Sigma_{ij}$ is the correlation between the two errors ($\Sigma_{ij} = R(\theta, d_h) = \mathrm{corr}(X^j, X^j)$).

In the Kriging Model, it is considered the following correlation functions (LOPHAVEN, NIELSEN and SONDERGAARD, 2002):

**Table 4: Correlation functions available for the Kriging Method application**

| Correlation | Function |
|---|---|
| **Exponential** | $R(\theta, d_h) = \exp\,(-\theta_h\lvert d_h\rvert)$ |
| **Gaussian** | $R(\theta, d_h) = \exp\,(-\theta_h {d_h}^2)$ |
| **Linear** | $R(\theta, d_h) = \max\,\{0, 1 - \theta_h\lvert d_h\rvert)$ |
| **Spherical** | $R(\theta, d_h) = 1 - 1{,}5\xi_h + 0{,}5{\xi_h}^3$, $\xi_h = min\{1, \theta_h\lvert d_h\rvert)$ |
| **Spline** | $R(\theta, d_h) = \varsigma(\xi_h)$, $\xi_h = \theta_h\lvert d_h\rvert$ |

The term $\mathrm{d}(\mathrm{X}^{\mathrm{i}}, \mathrm{X}^{\mathrm{j}})$ refers to the distance between the points and are not based on the Euclidian model ($\sqrt{\sum_1^n \left(x_h^i - x_h^j\right)^2}$), as a form of dealing with all points on the same weigh.

The measure of this distance between the two points is a function of the parameters $\theta_h$ and $p_h$.

$$\mathrm{d}(\mathrm{X}_i, \mathrm{X}_j) = \sum_{h=1}^{n} \theta_h \lvert x_h^i - x_h^j\rvert^{\,p_h}$$

According to JONES et al (1998), the parameter $\theta_h$ measures the influence of the variable $x_h$, or in other words, if the variable is active, it means that even small values of $\lvert x_h^i - x_h^j\rvert$ may influence on all major differences of the values of the function in $X^i$ e $X^j$.Statiscally, it means that the same small values of $|x_h^i - x_h^j|$ must imply in a lower correlation between the points $X^i$ e $X^j$as greater the value of $\theta_h$ is.

Yet according to the same author, the exponent $p_h$is related to the softness of the function regarding the points h. values of $p_h = 1$ correspond to less soft functions and $p_h = 2$ to more soft functions.

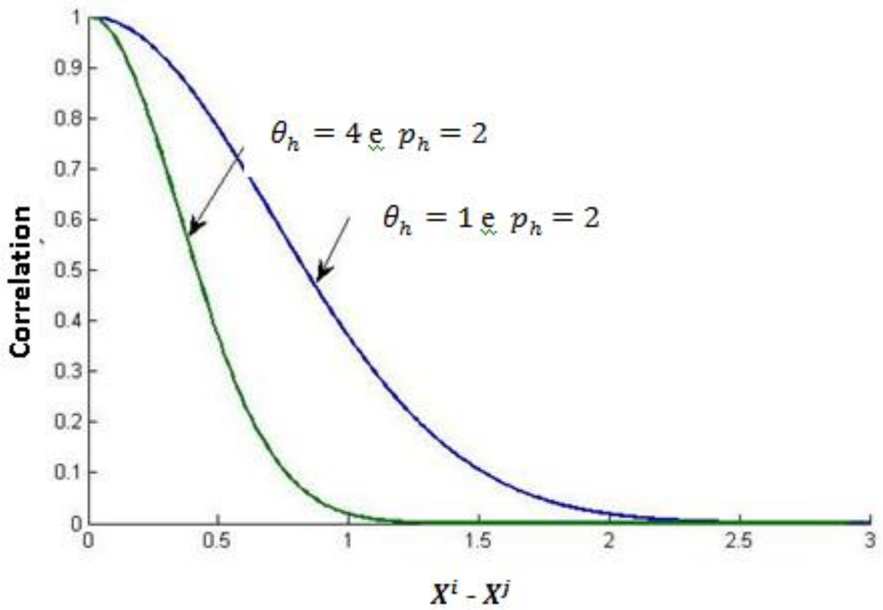



**Figure 12: Parameters applied for the Kriging Method**

In a similar approach to QUEIPO et al. (2002), it is adopted $\theta_h = 1$ and $p_h = 2$. And, therefore, the estimator non-biased of quadratic minimum quadratics for $\hat{f}(X^*)$ is given by (RIBEIRO & FERREIRA, 2004), (LOPHAVEN, NIELSEN, & SONDERGAARD, 2002):

$$\hat{f}(X^*) = \sum_{j=1}^{m} \beta_j^* f^j (X^*) + r'\Sigma^{-1}(y - F\beta^*)$$

In which:

$$\beta = (F^T\Sigma^{-1}F)^{-1} F^T\Sigma^{-1}y$$

$r$ is the vector of correlations between the error regarding the point $X^*$and the others points of the sample

$\Sigma$ is the correlation matrix between the sample points y is the vector of values observed to the CVaR

F is the matrix with the values of the functions calculated on the sample points

The first step to the application of the proposed model is related to the obtaining of the adequate sample for the experiment. Since the study`s objective is to analyze the portfolio composition problem, the grid points represent the percentage of capital allocated in each asset. To do so, these grid points assume values such as $x_j \epsilon$ [0,1].

The decision regarding the generation of the points that will be used for analysis is important as a manner of increasing the method`s efficiency and of decreasing its statistical uncertainty. There are three major methods for generating the sample:

1) **Random generation**: the points generated are normally distributed in the range $[0; 1]$, having the mean sequence equals to zero and unitary variance;
2) **Deterministic generation**: each of the hypercube $[0; 1]^n$ is subdivided in a certain number of ranges which origin other hypercube whose vertices are the sample points (RIBEIRO & FERREIRA, 2004);
3) **Latin hypercube generation**: guarantee that all portions of the space are represented. Initially, it is determined the $m$ ranges non-overlapped and with the same probability, and then, it is generated a random sample, uniformly distributed, in each range and in all the dimensions for further sample selection of these to compose the group of points for analysis.

LOPHAVEN, NIELSEN & SONDEGAARD (2002) present three regression models which may be used to approximate the problem`s response surface. In the first model, one approximates the surface to the value of a constant through a zero degree polynomial. The second alternative is to approximate it to a polynomial with degree of one, representing a liner regression, and finally, the third one, a quadratic regression, using a polynomial of second degree. The approach proposed in this study will use the linear regression, in order to decrease the problem complexity and facilitating its graphic representation. Therefore:

$$f_1(x) = 1, \qquad f_2(x) = x_1, \dots , \qquad f_{n+1}(x) = x_n$$

Defined the methodology to the sample generation, to the correlation and to the regression which will be used, the model continues with the following steps:

a) The set of points $\{X^i\}_{i=1}^{q}$ is generated, according to one of the methods above presented, subject to the following constraints:

$$X^i \epsilon \mathcal{R}^n \ e \ \ 0 \leq x_j^i \leq 1, \qquad i = 1, \dots, n$$

b) For each vector $X^i$ it is calculated the $y^i = \text{CVaR}$;

c) For each set of points$X^i$ e $y^i$, it is determined the approximate function through the Kriging Method, according to the regression and correlation models chosen, and also the parameters$\theta_h, p_h$ and $\sigma^2$;

d) It is generated a new set of points$\{X^i\}_{i=1}^{q}$, subject to the same constraints so that it is possible to analyze the estimation error between the value provided by the approximate function $\hat{f}$ and the value of the portfolio`s CVaR.

According to QUEIPO et al (2002), the benefits of using this probabilistic approach for modeling deterministic functions rely on the fact that the model uses an impartial estimator for representing a problem and providing the approximation estimated error.

## 2.4 Portfolio Selection on the Energy Sector

The Fundamentals of the Modern Portfolio Theory (i.e. a portfolio composed by assets which have negative correlated returns offers a better risk-return relation than a portfolio composed by just one asset), introduced by Harry Markowitz in 1952, also started to be used on the energy sector.

The utilization of the Markowitz theory in a different field from that originally created (i.e. the financial markets) is not recent: it was introduced for the first time by BAR-LEV & KATZ (1976). However, more solid results were obtained by AWERBUCH & BERGER (2003), AWERBUCH (2006) & KREY and ZWEIFEL (2006) and ended up becoming reference on the literature. According to these authors, the study`s objective is to select the optimal composition of country`s energy matrix portfolio (or maybe of a continent, in the case of Europe), formed by different generating technologies, such as wind, gas, nuclear, among others. In this context, the unit cost of energy production [kWh/$] is considered the return of an energy portfolio and the standard deviation of this return is given as a risk measure.

LOSEKANN et al (2013) and DELARUE et al (2011) perform a similar approach: these consider as the objective function of the problem the unit cost of the energy production [$/kWh], and the standard deviation of such cost, expressed as a percentage of the average cost, with the risk associated to each technology, which is considered an asset. The general form of this problem may be expressed as:

$$\text{Min} \qquad COST\ (x) = \textstyle\sum_1^n x_i UTCO_i$$

Subject to (1) $[x^t \sum x]^{1/2} \leq R$

(2) $\sum_1^n x_i = 1$

(3) $x_n \geq 0, i = 1, \dots, n$

The objective function COST (x) represents the total unit cost of energy production in terms of the decision vector x, which represent the allocation in each technology of the energy matrix portfolio. Notably, the total unit cost is expressed as the sum of the average unit cost of each technology, symbolized as UTCOi (DELARUE ET AL, 2011).

The first constraint to the problem represents the standard deviation (risk measure used) of the portfolio in terms of x (allocation in each technology), given that Σ refers to the covariance matrix between the historical values of the unit costs. This constraint is parameterized in R, which is the greatest acceptable portfolio risk.

The second constraint guarantees the total allocation of the energy supply provided by the different studied technologies.

At last, the third constraint assures that there is no negative allocation in the portfolio, since there is a physical restriction for that.

In order to turn the presented problem a reasonable representation of the reality, the UTCO is decomposed in many components (DELARUE et al, 2011):

$$UTCO_i = \sum_k C_{i,k} = INVe_i + FU_i + FOMe_i + VOM_i$$

Being:

$C_{i,k}$ represents the component cost k of the technology I [$/kWh]

$INVe_i$ represents the invesment cost of the technology i [$/kWh]

$FU_i$represents the fuel cost of the technology i [$/kWh]

$FOMe_i$ represents the fixed cost of Operations and Maintenance of the technology i [$/kWh]

$VOM_i$ represents the variable cost of Operations and Maintenance of the technology i [$/kWh]

LOSEKANN et al. (2013) presents in their study a scatter plot relating the unit cost of energy generation with the standard deviation of this cost, expressed as a percentage of the average cost, for multiples technologies available in the Brazilian market. Such results are presented in Figure 13.

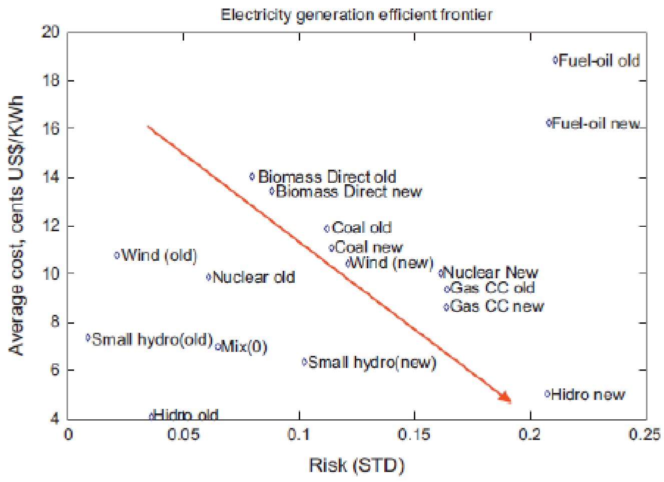



**Figure 13: Average cost and risk (standard deviation) for different technologies**

It is possible to notice that there is a negative correlation between the average unit cost and the standard deviation of this cost. This correlation is analogous to the classic relation between risk and return, which states that the bigger the risk incurred by the investor, the bigger the return required by him. In the case of the energy sector, however, the objective is to minimize the cost variable (and not maximizing the return), thus, this relation is inverted, and in other words, a greater risk (the standard deviation, for example) requires a lower energy generation.

Yet regarding this theme, it is worth saying that the conflict between minimizing the average unit cost of energy generation of a portfolio with multiples technologies and its risk may be represented by the Paretto`s Efficient Frontier, exhibited in Figure 14.

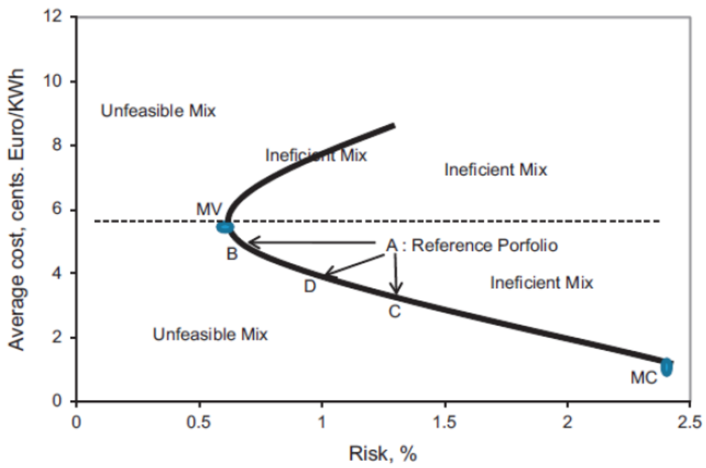



**Figure 14: Paretto`s Efficent Frontier**

Alternatively to the optimization problem above, it is equivalent to present in another form (BARROSA 2015). According to the author, it is convenient to rewrite the problem`s objective function as being a function of the energy matrix portfolio risk, subject to a parameterized constraint which represents the greatest unit cost acceptable. Hence, the problem is formulated as:

Min $RISK$

Subject to (1) $\sum_1^n UTCO_i \leq C$

(2) $\sum_1^n x_i = 1$

(3) $x_n \geq 0, i = 1, \ldots, n$

This new form of representing the energy matrix optimization problem enables the use of different risk measures, and therefore, it will be implemented in this study, since it is proposed here to use not just the standard deviation as risk measure, but also the Value at Risk (VaR) and the Conditional value at Risk (CVaR).

# 3 METHODOLOGY

In this chapter, it is presented, initially, the motivation of the choice of the energy sector analyze in this study and its characteristics. Then, it is described the cost measure used, the Levelized Cost of Energy (LCOE). It is presented also the database chosen for the application of the Kriging Method. Finally, it is demonstrated how Monte Carlo simulations are carried on in order to support the data preparation.

## 3.1 The United States Energy Sector

### 3.1.1 Motivation

The original idea of this study was to analyze the energy sector on the native country of the author, Brazil. But, some limitations and drawbacks emerged on the data collection phase:

1) There is in Brazil a huge lack of public data on energy generation costs. It is worth saying that the major regulatory agency of energy in the country, the ANEEL (National Agency of Electrical Energy) does not disclosure data on the generation plants, since the majority belongs to private non-listed companies, and therefore, are not required to report their results.
2) Yet regarding the private companies responsible for managing the energy generation plants, the author tried to establish contact with them in order to gather information, but, unfortunately, most of these companies refused to disclosure data, stating that these numbers are strictly confidential data.
3) Governmental policies, as federal, as state related, of subsidies and incentives are highly frequent in Brazil. One of the reasons why it takes place is because the State, by many political reasons, has the willingness to keep control on the energy sector. The government does so, for example, through price control over fuel, what, consequently, directly impacts on the costs of a generation plant. This interventionism hinders the market rationale, or in other words, basic "laws" of economy (e.g. the supply and demand law) are put aside. All in, it ends up jeopardizing the implementation of the model proposed in this study, the Kriging Method, since this model is based not in governmental policies, but in market logics, as the risk return relation. As an example of action of this kind by the State, it is possible to verify in Figure 15 the average variable cost of energy generation of four power plants located at the Rio de Janeiro state. Despite of the fact that the four plants have the exact same generation technology, each of them has a different generation cost. It might be explained by the observation that each was constructed in a different date (years apart from each other) and in each date

which the contracts were signed, there was a different governmental subsidy taking place.

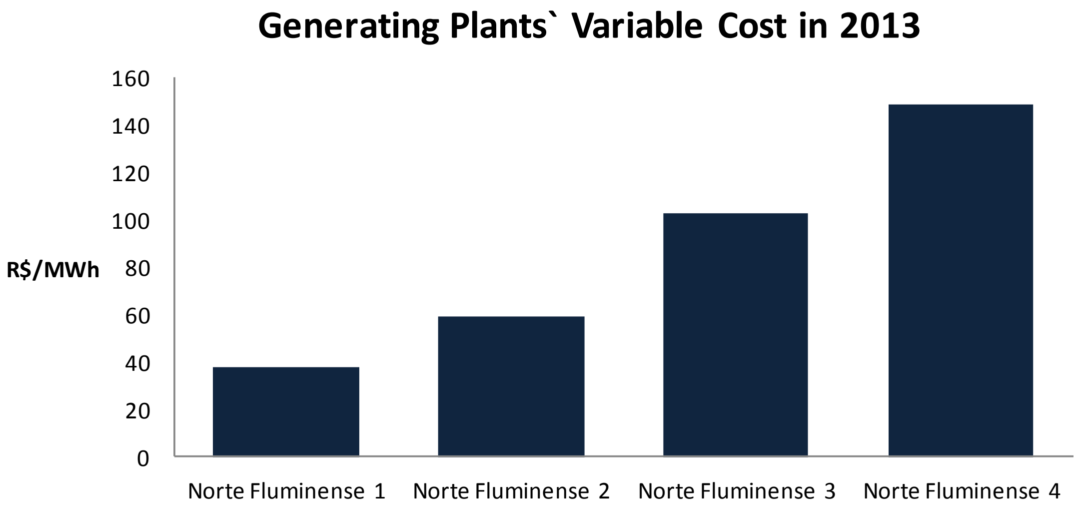



**Figure 15: Variable cost of the power plants**

Hence, this study`s objective is to analyze the United States energy sector. It is so, because there is a great availability of data regarding this market, especially by public agencies that consolidate date, such as the Department of Energy of the United States (DOE) and the Energy Information Agency (EIA). Furthermore, one of the strongest characteristics of this market is the economical liberalism, in which the market is driven by the players` willingness influencing costs and generation prices. Therefore, the proposed method becomes valid in such market, and consequently, a potential tool on the support of energy policies development.

### 3.1.2 Main Characteristics

The United States are second biggest producer and consumer o energy in the world, lagging behind just from China. The country consumes approximately twenty percent of the world production and, notably, has a relevant role in the global market. It stays clear through observation of the Figure 16 that in the past few centuries the per capita demand in the United States increased considerably. However, it is important to notice that also, in the last decades there has been a decreasing on the per capita consumption, which may explained by the surge of more efficient technologies in energy terms, and also explained by the population increase and by environmental policies aiming the reduction of the energy consumption.

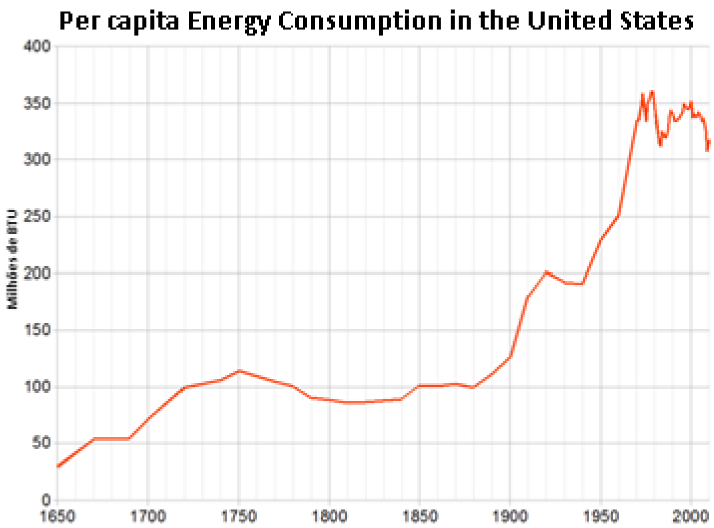



**Figure 16: Energy consumption in the United States**

It is possible to verify in Figure 17 that the sector responsible for the greatest energy consumption is in fact the electrical, which is by the way, the sector analyzed in this study.

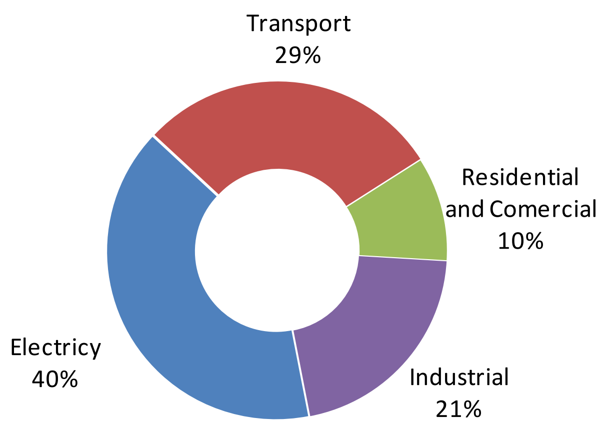



**Figure 17: Participation of each sector in the energy consumption of the United States**

It is important also to define which are the main sources of energy of the electrical sector of the United States, what is exhibited in Figure 18. One may note that almost half of the electrical energy is provided by technologies which use coal as generation source.

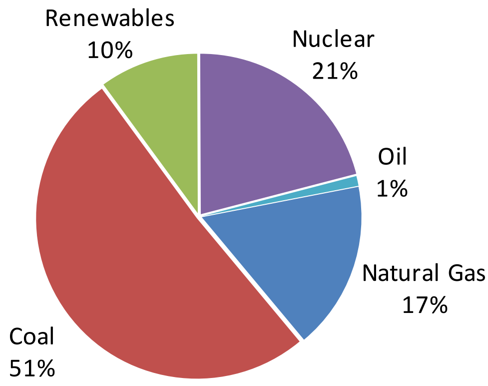



**Figure 18: Energy matrix in the United States**

It is worth noting also the greater importance to renewable energies: according to the Energy Information Agency (EIA), in 2003, the renewable sources represented six percent of the country`s energy matrix, and in 2015, on its turn, represent ten percent of the total electric energy in the country.

Given the recent increase of renewable sources in the country, it is important also to highlight which are the main renewable technologies used, what may be seen in Figure 19.

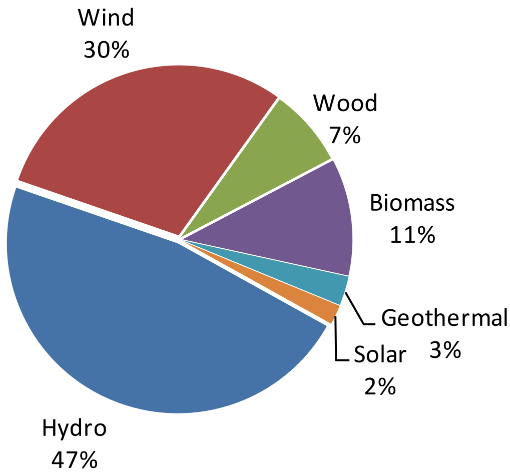



**Figure 19: Participation of renewable technologies**

Hydro, wind and biomass technologies stand out, representing ninety percent of the country`s total renewable production. Because of their relevance, these three technologies, together with solar technologies, were the renewable technologies chosen to further analysis in this study.

## 3.2 LCOE

The cost measure that will be used in this work is the LCOE (Levelized Cost of Energy). The LCOE is a convenient measure to compare the overall competiveness of different energetic generation technologies. It represents the cost per megawatt hour (in real terms) of construction and operation of a generating plant in one financial and one operation cycle previously defined.

The main parameters are necessary to calculate the LCOE the capital cost, the fuel cost, base and variable costs of operations and maintenance, the financing cost, and a certain utilization rate of the generating technology. So, it a general form, it is possible to note that:

$$LCOE = f(Capital\ Cost, O\&Mf, O\&Mv, Fuel) \quad [\frac{R\$}{MWh}]$$

Before demonstrating its general form, it is important to define first the CRF (Capital Recovery Factor), which is a factor uses to annualize the capital cost, or in other words, the incurred capital expenditures to construct the plant. It, depending on the technology adopted, may represent an important part of the total generation cost. The CRF is calculated through the following formula:

$$CRF = \frac{D * (1 + D)^N}{((1 + D)^N - 1)}$$

In which:

D is equal to the discount rate in which the cash flows are discounted to Present Value. This rate is different to each technology.

N represents the activity time of the generating plant.

Therefore, the LCOE`s calculated as:

$$LCOE = \frac{CC * CRF * (1 - I * DPV)}{CF * (1 - I)} + \frac{O\&Mf}{CF} + O\&Mv + Fuel * HR \quad [\frac{R\$}{MWh}]$$

In which:

CC represents the cost of capital to construct a generating plant;

I is the tax rate (%) applied by the government;

DPV represents the present value of the power plant`s depreciation;

CF represents the capacity factor, which is defined as the proportion between the plant`s effective production and the its total capacity in this period;

O&Mf represents the fixed costs of operation and maintenance;

O&Mv represents the variable costs of operation and maintenance;

Fuel represents the costs incurred with fuel acquisition;

HR represents the heat rate, which is the plant`s efficiency in converting fuel into energy.

It is worth noting that the weight of each parameter varies between each technology. For example, in the case of solar and Wind generations, the fuel cost is nil, and the costs with operation and maintenance are low. Yet, in these cases, the capital costs are the costs that represents the biggest weighs on the LCOE, since these power plants require high levels of capital expenditures. On the other hand, technologies like gas, have a different cost structure: the factor that weighs the most on the LCOE is the fuel cost.

## 3.3 Data Collection

### 3.3.1 Database

The data used on this study`s analysis were collected on the Transparent Cost Database, which is a public database that gathers data belonging to different sources, not only academics, but also from governmental agencies, such as the Department of Energy of the United States (DoE), the Energy Information Administration (EIA), the Environmental Protection Agency (EPA), among others.

The data collected are estimates for the LCOE for many technologies in the United States for the next twenty five years, from 2016 to 2040. In this study, it will be analyzed the main generation technologies in the United States, which are: solar, wind, hydro, biomass, nuclear, gas, and coal. Table 4 presents the number of data collected for each technology.

**Table 4: Number of data collected for each technology**

| Technology | Number of Data Collected |
|---|---|
| **Solar** | 692 |
| **Wind** | 795 |
| **Hydro** | 114 |

| | |
|---|---|
| **Biomass** | 350 |
| **Nuclear** | 131 |
| **Gas** | 207 |
| **Coal** | 524 |
| **Total** | **2813** |

In order to compare the estimates from the different sources of this database, which were done in different years, they were adjusted to the actual monetary value (2015 U$ Dollar) through the inflation incurred in the United States. The values of the estimates for the LCOE already adjusted by inflation of the wind technology are presented in Figure 20. For convenience, the others estimates are presented in the Appendix A.

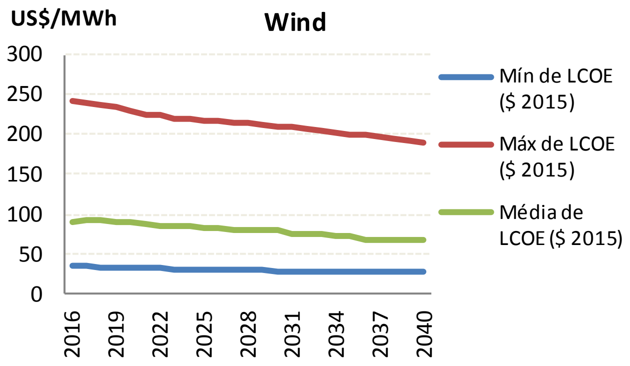



**Figure 20: Estimates of the wind costs until 2040**

Additionally to the total values of the LCOE, it was extracted also from the database its sub-factors, which are the cost of capital, the cost with operations and maintenance and the fuel cost. It is possible to observe in Figure 21 the LCOE composition for each technology studied.

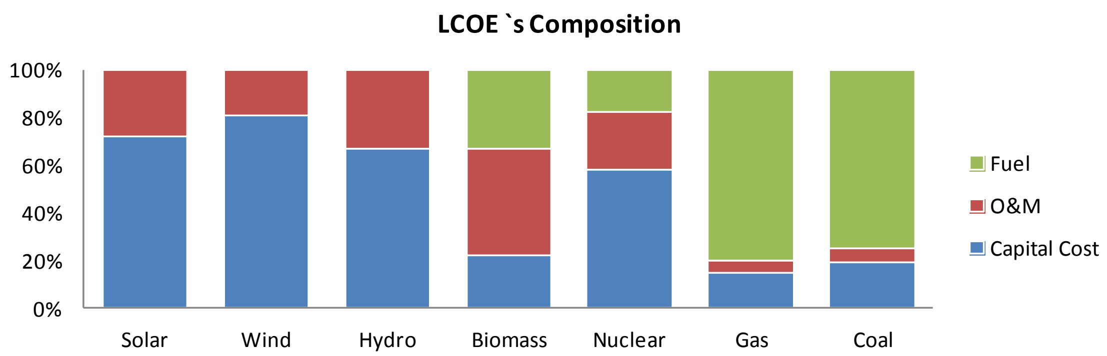



**Figure 21: Composition of the cost of each technology**

Figure 21 proves what was stated in the last chapter, since it is possible to see that in the case of the wind and solar technologies, the greatest sub-factor is in fact the cost of capital. Now, in the case of technologies such gas and coal, the most relevant factor is the fuel. It is worth noting that the biomass generation technology is a special case, in which the cost factor that weighs the most in the cost composition is the Operations and Maintenance factor.

### 3.3.2 Monte Carlo Simulation

It is important to remember that in the application of the proposed method the Kriging Method, as input for the model, it is necessary many scenarios to the chosen variable (i.e. the LCOE). The more scenarios as input, so better is the model`s precision, and therefore, more valid it is. Since data collected is not vast (16 points for each year), it was realized that it was important to generate more data points. In order to do so, the solution was applying Monte Carlo simulations so it could be generated more points based on the existing characteristics.

GLASSERMAN (2003) defines the Monte Carlo method as a statistical method based in a sample base and use heuristic probabilities in order to obtain numeric results. The author defines also a particular class of methods, called Brownian Motion. The idea of such class of methods is to generate random walks from the statistical parameters previously known. In its simplest form, from a data point it is generated a new point supported by a random variable. The generic formulation of the Brownian Motion follows:

$$X(t_{i+1}) = X(t_i) + \sqrt{t_{i+1} + t_i} Z_{i+1} \quad , \quad i = 0, \dots, n-1$$

In which:

X(t) represents the simulated value in the point t;

$Z_i$ represents an independent normally random variable, whose value ranges from 0 to 1.

In the case of a sample with mean μ for each point t and with a standard deviation σ for each point t (the point t represents in this study the year analyzed and the LCOE mean and standard deviation are calculated from the database for each year), it may be formulated (GLASSERMAN, 2003):

$$X(t_{i+1}) = X(t_i) + \mu(t_i)(t_{i+1} - t_i) + \sigma(t_i)\sqrt{t_{i+1} + t_i} Z_{i+1} \quad , \quad i = 0, \dots, n-1$$

From this equation and from the Monte Carlo simulation it is generated the random walks for the LCOR from 2016 to 2040 for each of the seven technologies studied. It was obtained a thousand scenarios, given that each scenario represents one random walk. As example, it is demonstrated in Figure 22 ten random walks for the wind technology.

A partir dessa equação e da simulação de Monte Carlo são gerados caminhos aleatórios para o LCOE entre 2016 e 2040 para cada uma das sete tecnologias estudadas. Foram obtidos mil cenários, sendo que cada cenário representa um caminho aleatório. A título de exemplo, são demonstrados na **Erro! Fonte de referência não encontrada.** dez caminhos aleatórios para a tecnologia eólica. For convenience, the others Monte Carlo simulations are presented in the Appendix B.

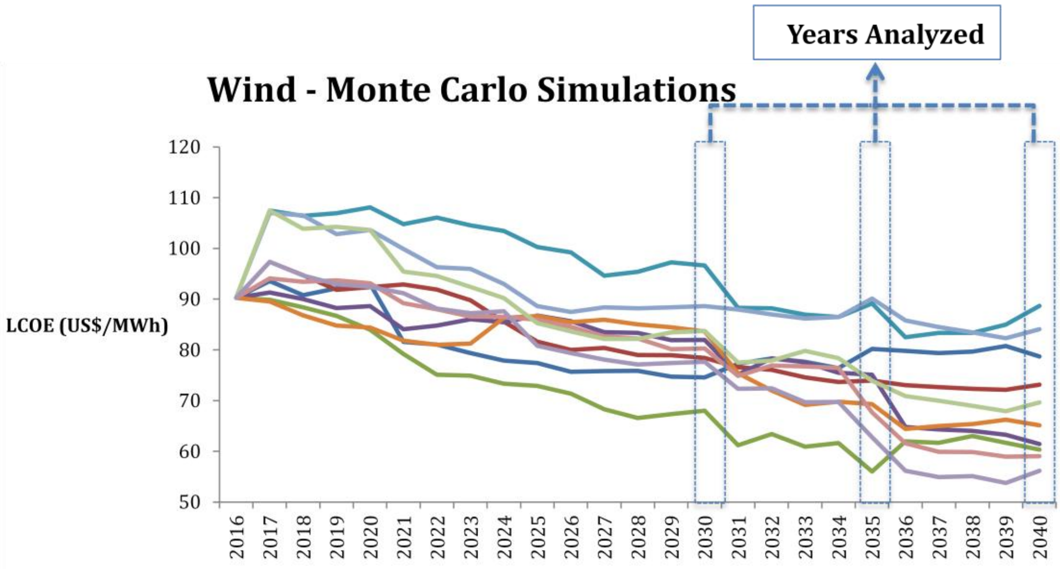



**Figure 22: Ten examples of the Monte Carlo simulations for the wind technology**

It is worth saying that in this study it will be analyzed the energy matrix composition and its costs for three specific years: 2030, 2035 and 2040. The year of 2030 was chosen as the first year analyzed because one of the motivation of this study is to create the decision making process of formulating an optimal energy matrix and for such purpose, a policymaker needs to have a long term perspective, since it takes time to formulate and implement such policies. It is considered, therefore, fifteen years as being a reasonable time period to make such decisions.

The Monte Carlo simulation results are summed up in Figure 23.

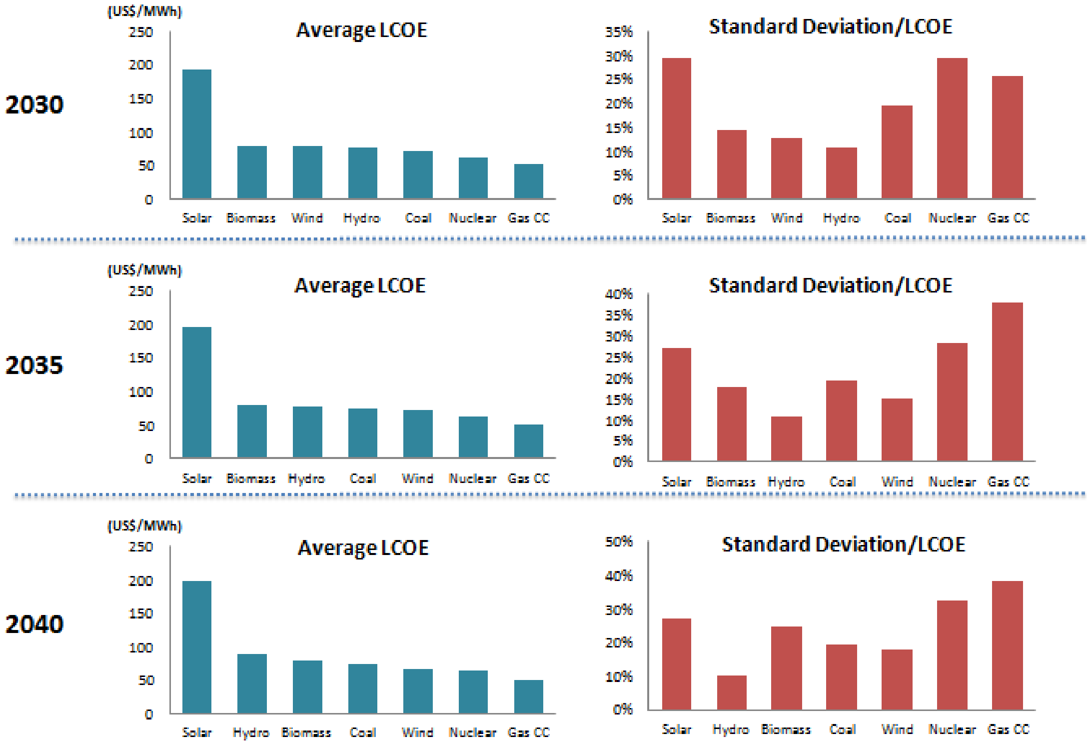



**Figure 23: Average LCOE and risk for the years of 2030, 2035 and 2040**

Furthermore, from the Monte Carlo simulations it is possible to obtain a scatter plot relating the average cost (LCOE) with the standard deviation for each technology, what provides the overall situation for each year. These charts are presented in Figure 24. It is worth noting that the solar technology was omitted in order to facilitate the chart comprehension, since its cost is much greater than the other technologies ones.

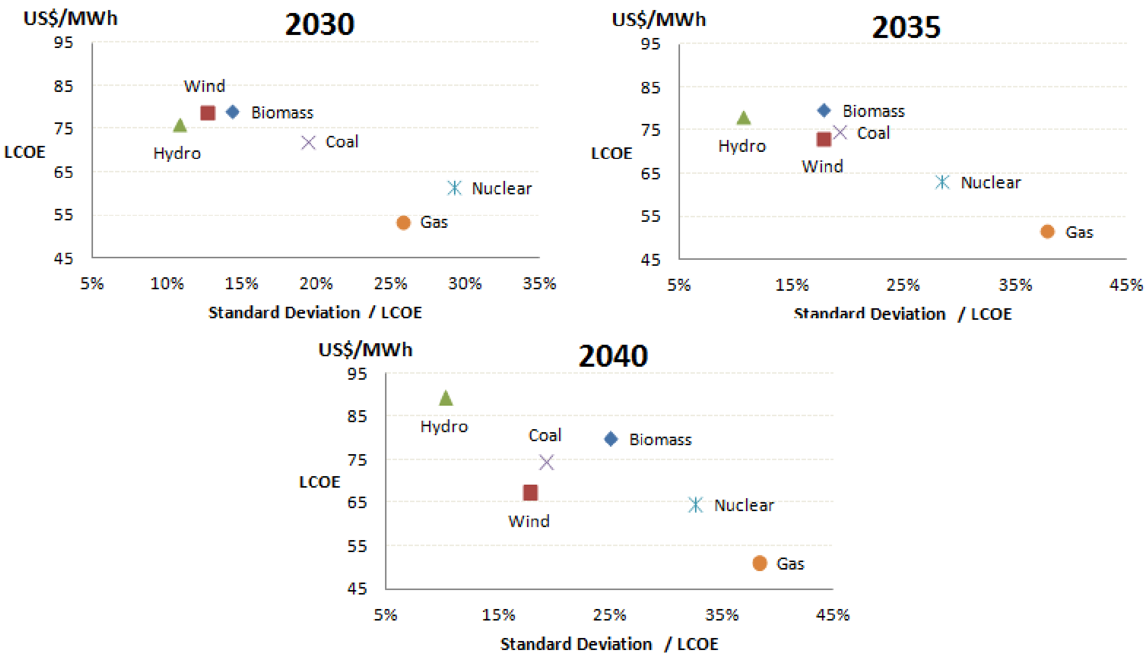



**Figure 24: Scattered plot for the technologies analyzed**

From Figures 23 and 24, is possible to note that the generating technology that uses the gas as input present, in all the years, the lowest cost (lowest LCOE). On the flip side, the technology with the greatest cost (ex-solar) differs from year to year, being in 2030 the wind, in 2035 the biomass, and in 2040 again the wind. By the way, observing Figure 24, it stays clear that the solar technology present much greater cost compared to the others. A preliminary conclusion from this fact is that this technology will not be considered in the optimal matrix, since it does not present a decent risk versus return relation.

It is worth noting also that, as said in the previous chapter, the LCOE present a negative correlation with the standard deviation (risk). For example, the gas technology presents always the lowest total cost, but it is one of the technologies with the greatest risk. The fact that does not exist a dominant technology (i.e. with lower cost and risk), excluding the solar technology, is what turns the proposed method useful in the decision making process of formulating the optimal energy matrix.

# 4 RESULTS

The objective of this study is to apply a new methodology to approximate the solution of the optimal composition of a country`s energy matrix through (i) the simulation to obtain known values of the function Risk(x), and (ii) the interpolation of these data points through the Kriging Method for later optimization.

In this section, it is illustrated initially the sample and correlation function selection, which will be used during the simulations. Follows the proposed application into a energy matrix portfolio, composed by seven technologies (solar, wind, nuclear, gas, coal, biomass and hydro), using three risk measures as objective function: (i) Variance ($\sigma^2$), (ii) Value at Risk (VaR), and (iii) Conditional Value at Risk (CVaR).

## 4.1 Sample and Correlation Function Selection

It was detailed in section 2.3.4 the three main methods to generate a sample of points necessary for the experiment application: random, deterministic and Latin hypercube. Given the similarity of the last two methods, it will be done an analysis between the fit done by the random and deterministic samples. This last analysis aims to assess the influence of the grid selection on the efficiency of the Kriging Method solution.

In the case of the random sample, the weights of each of the seven technologies are generated randomly, such as $0 \leq x_j^{(i)} \leq 1$ and $\sum_{i=1}^{7} x_i = 1$. On the other hand, in the deterministic sample, the values of $x_j^{(i)}$ are obtained maintaining them equally spaced from each other in the function domain, forming hypercube of equal dimension constraining the domain of the function Risk (x).

Figures 25 represents two sample techniques previously described and used for simulation, relaxing the constraint $\sum_{i=1}^{7} x_i = 1$ and applied for two assets arbitrarily obtained, in order to facilitate its representation in $\mathcal{R}^2$.

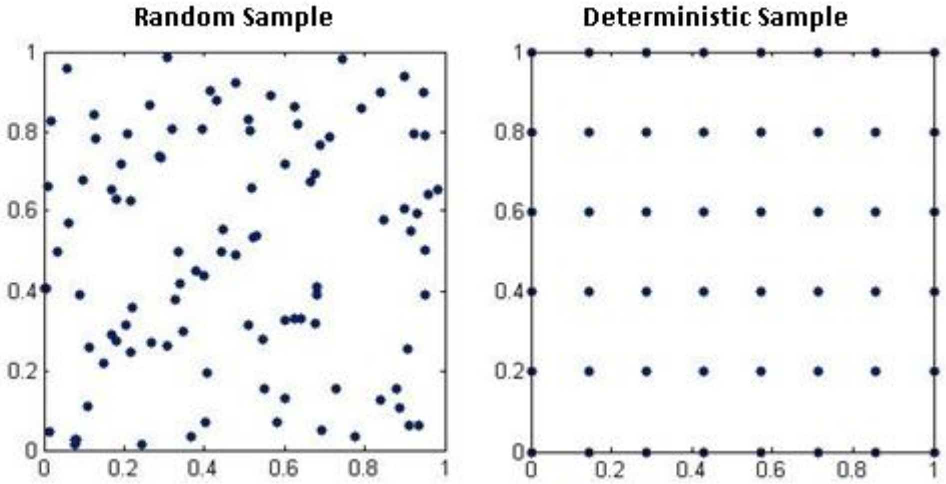



**Figure 25: Examples of random and deterministic sample**

Other important parameter for the Kriging Method application is the quantity of scenarios used in the simulation process. In the case of random sample, the number of points is arbitrarily defined. In the deterministic sample, in turn, the sample size varies according to the number of assets, to the distance defined between the grid points in the simulations. Hence, given $n$ assets composing the portfolio, a distance $d$ between the simulated points, and defining $k = d^{-1}$ the deterministic sample size, respecting the restriction $\sum_{i=1}^{n} x_i = 1$, will be given by (BARROSA, 2015):

$$q = \binom{k+n-1}{k} = \frac{(k+n-1)!}{k!\,(n-1)!}$$

In order to verify which will be the type of sample used in the study, the proposed method was applied for the year of 2030 with the two different sample types, and then, the Mean Squared Errors (i.e. $MSE = \frac{1}{q}\sum_{i=1}^{q}[f(x) - \hat{f}(\mathrm{x})]^2$) were calculated for each risk measur considered. The results are exhibited in Table 5. It is worth noting that in the case of the deterministic sample, a distance of 0.2 (i.e. d=0.2) generates 669 points, a distance of 0.1 generates 3,003 points, and finally, a distance of 0.05 generates a total of 53,130 points.

**Table 5: Mean Squared Error for different samples in 2030**

| | **Mean Squared Error (MSE) - 2030** | | | | | |
|---|---|---|---|---|---|---|
| | **Random Sample** | | | **Deterministic Sample** | | |
| | 669 points | 3.003 points | 53.130 points | d=0,2 | d=0,1 | d=0,05 |
| **Standard Deviation** | 8,4E-02 | 2,7E-02 | 3,1E-03 | 1,1E-02 | 7,8E-04 | 8,2E-05 |
| **VaR** | 4,3E-02 | 1,2E-02 | 8,6E-04 | 2,3E-03 | 5,3E-05 | 1,1E-05 |
| **CVaR** | 3,1E-02 | 9,2E-03 | 6,3E-04 | 4,9E-03 | 8,0E-05 | 7,8E-06 |

It stays clear that the Kriging Method is sensible to sample selection criteria used for the fit. Table 5 reveals that when applied to random points, the method presents greater errors compared to its application using the deterministic sample, which manages to occupy the whole grid. It is clear also that as the number of simulated points increase, the mean squared error decreases.

In this study, therefore, it will be used the deterministic sample with distance of 0.1. This distance was chosen because, it is noticeable that the precision gain with the distance of 0.05 is not that big, and it is important to remember that there is a trade-off between the precision gain the computation efficiency, what justify the choice of distance 0.1.

In the previous phase (i.e. the sample selection) it was used the Gaussian correlation function in order to apply the Kriging Method. However, given the possible correlation functions, presented by LOPHAVEN, NIELSEN & SONDEGAARD (2002) and named in section 2.3.4, it is important to make an analysis to assess which function better fit for the data collected.

For such analysis, it is applied the same method used in the previous phase, in other words, the proposed method is applied for the year of 2030 with a deterministic sample (distance of 0.1) regarding all existing types of correlation functions and the Mean Squared Errors are calculated for each simulation. The results are exhibited on Table 6.

**Tabela 6: Mean Squared Error for different correlation functions in the year of 2030**

| | **Mean Squared Error (MSE) - 2030** | | | | |
|---|---|---|---|---|---|
| **Correlation Function:** | **Exponential** | **Gaussian** | **Linear** | **Spheric** | **Spline** |
| **Standard Deviation** | 1,3E-04 | 7,8E-04 | 1,7E-03 | 3,3E-02 | 7,1E-04 |
| **VaR** | 5,9E-06 | 5,3E-05 | 7,2E-05 | 8,2E-03 | 1,1E-04 |
| **CVaR** | 1,1E-06 | 8,0E-05 | 2,6E-04 | 5,9E-03 | 9,6E-05 |

It is noticeable that the simulation which presented the lowest MSE for the three risk measures was the one that used the Exponential correlation function, and hence, is the one considered the most suitable to be applied in this study.

## 4.2 Kriging Method Application

In the last section it was carried out a study in order to define the best input parameters for the application of the proposed method (i.e. sample and correlation function selections).

It is known that this model is a function of the grid sample, the response values (Risk(x)), the regression and the correlation chosen, and the value of $\theta$. So, the input parameters are defined as:

- The sample composed by seven technologies was generated by a deterministic process which divided the grid in $k$ equidistant intervals, defining what is called the pace, in other words, the distance $d$ between two consecutive intervals for $x_j^{(i)}$;
- The regression method chosen was the linear, which uses a polynomial of first degree to approximate the function;
- The correlation function chosen was the Exponential, which provides the best approximation of the function, according to the MSE analysis;
- The definition of the correlation model implies on the determination of the value of $p_h$, which, in the case of the exponential correlation, is equal to 2;
- In a similar approach to QUEIPO (2002), it will be used $\theta = 1$.

The next step is to apply the proposed method with the input parameters above mentioned for the three risk measures named previously, the Variance, the VaR and the CVaR. It is worth noting that in the case of the variance, it will be done an experiment control, applying also the Markowitz Model, described in section 2.3.1, in order to assess the validation of the results of the Kriging Method.

## 4.2.1 Variance

In this and in the following two sections (i.e. in sections 4.2.2 and 4.2.3), the results obtained in the simulation in the software MATLAB are exhibited in both the forms of efficient frontier (i.e. Paretto`s Optimality Frontier) and of composition of energy matrix for different risk levels.

First of all, the Markowitz and the Kriging Methods are applied with the same input data for the year of 2040, in order to validate the proposed method. It is worth remembering that the Markowitz Method uses as risk measure the variance (here exhibited in the form of standard deviation). Figure 26 shows the two efficient frontiers and Figure 27 demonstrates the optimal compositions for both methods.

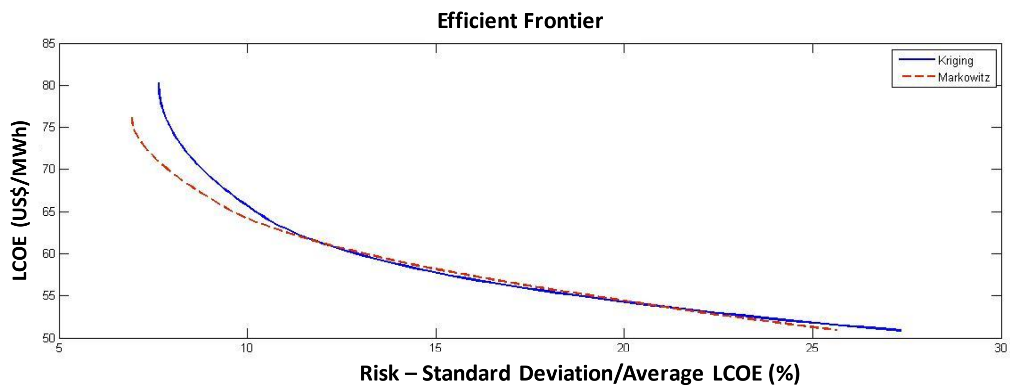



**Figure 26: Efficient frontier for the year of 2040, for the Markowitz and Kriging Methods**

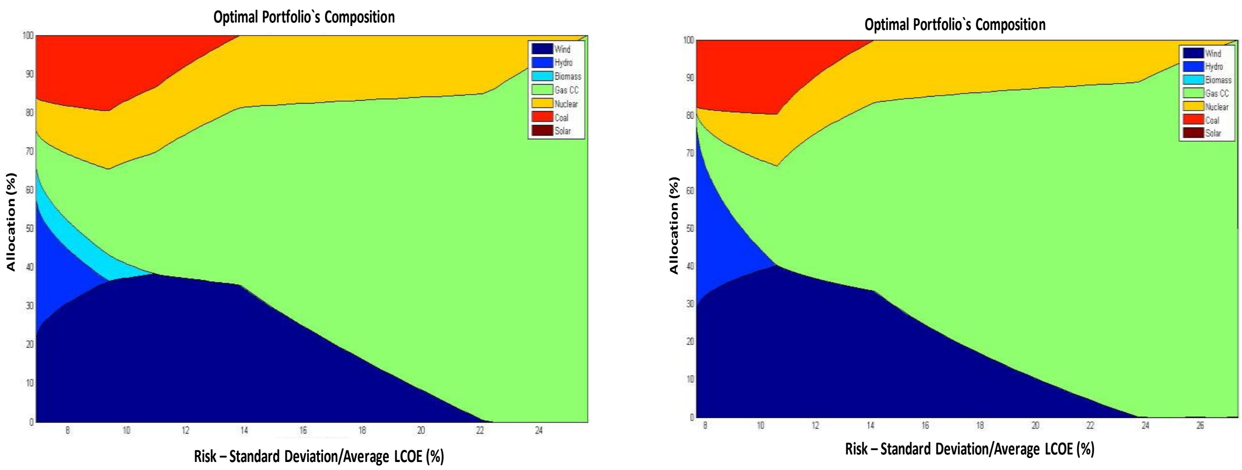



**Figure 27: Optimal composition portfolio for the year of 2040**

As it may be observed, the proposed method revealed being very similar to the original method of Harry Markowitz, not just on the efficient frontier, but also on the optimal composition. It is worth remembering that the Kriging Method is robust, in the sense that it allows considering different risk measures, including those which considers the tail risks, such as the VaR and the CVaR. It also provides a better computational efficiency when compared to the VaR and CVaR models, previously described.

Follows the application of the Kriging Method for the years of 2030 and 2035, always with the input parameters described in the beginning of this section, using the standard deviation as risk measure and for an energy matrix portfolio composed of seven technologies.

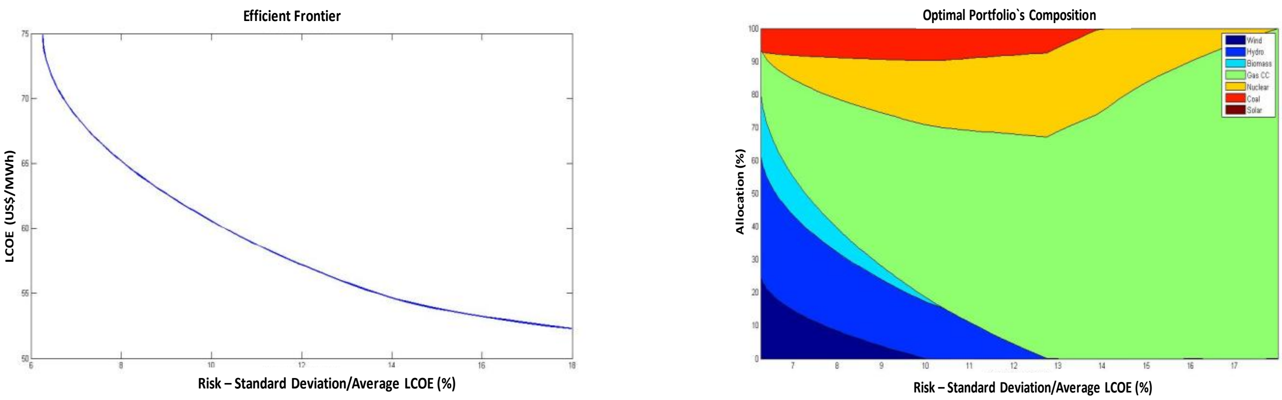



**Figure 28: Efficient frontier and optimal portfolio composition for the year of 2030, taking the standard deviation as risk measure**

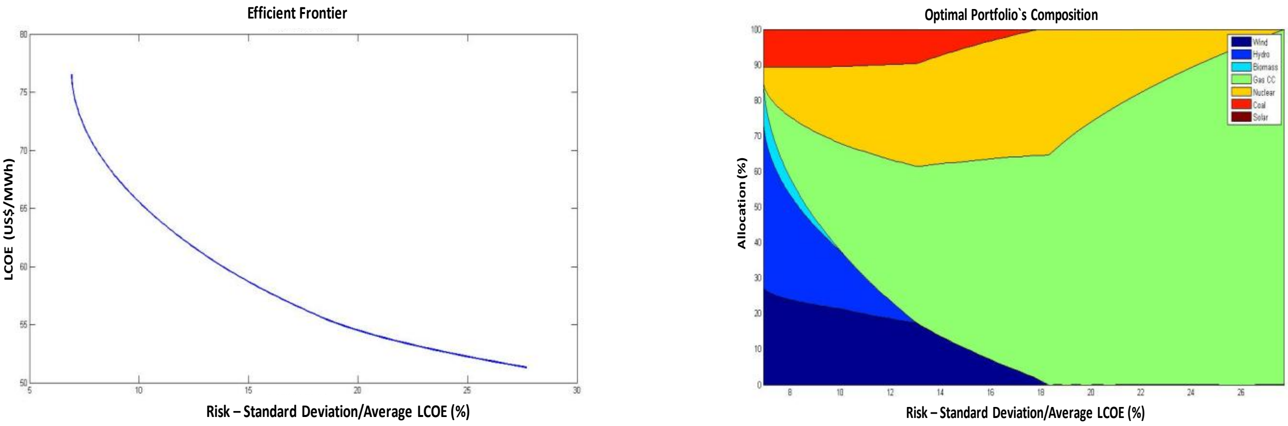



**Figure 29: Efficient frontier and optimal portfolio composition for the year of 2035, taking the standard deviation as risk measure**

As it may be seen, the results for the three years (i.e. 2030, 2035 and 2040) revealed to be very similar, but not exactly the same, as it was already expected, since throughout ten years it is practically impossible to appear technology improvements so innovative that completely change the cost structures of one technology.

Regarding the similarities, it is noticeable in the charts of the efficient frontier that, as observed in section 2.4, the optimal cost (i.e. the LCOE) is negatively correlated with the portfolio`s standard deviation. It is explained by the fact that it is required (by the policymaker, in this case) a lower cost for a greater risk level. Now, in the case of the compositions charts, it may be observed a total allocation on the gas technology for high levels of risk, what is completely reasonable given that this technology presents a lower cost but also a greater risk, preventing its allocation on low risk portfolios. Also, as expected, the solar technology was not allocated in any of the portfolios, since it does not present an adequate risk versus return relation.

It is important to note the impact of the diversification theory in the three years that the simulations were applied. In other words, the MARKOWITZ (1952) theory is valid for this set of data: it is possible to minimize the portfolio`s risk by diversifying the assets allocated.

Moreover, regarding the differences of the simulations, it may be seen that there is a greater allocation in the hydro technology on low risk portfolios through the years. It may be explained by the fact that this technology kept on a low degree of level, despite of the high costs, differently from the other technologies, such as the wind, which has a high cost (high costs of capital) and presented a risk increase through the years. Another rationale to explain such fact is that the hydro technology presents greater risks in its construction phases (e.g. environmental and/or labor related), turning this technology more allocated in long term portfolios. Furthermore, it is observed a grater allocation on nuclear technology, what may be explained by:

(i) lower costs with the acquisition of the uranium ore; and (ii) lower environmental and/or regulatory risks.

### 4.2.2 Value at Risk (VaR)

In this phase of the study, the Kriging Method is applied with the same inputs, but now using the VaR as risk measure. Again, results are exhibited in the form of efficient frontier and optimal portfolio composition.

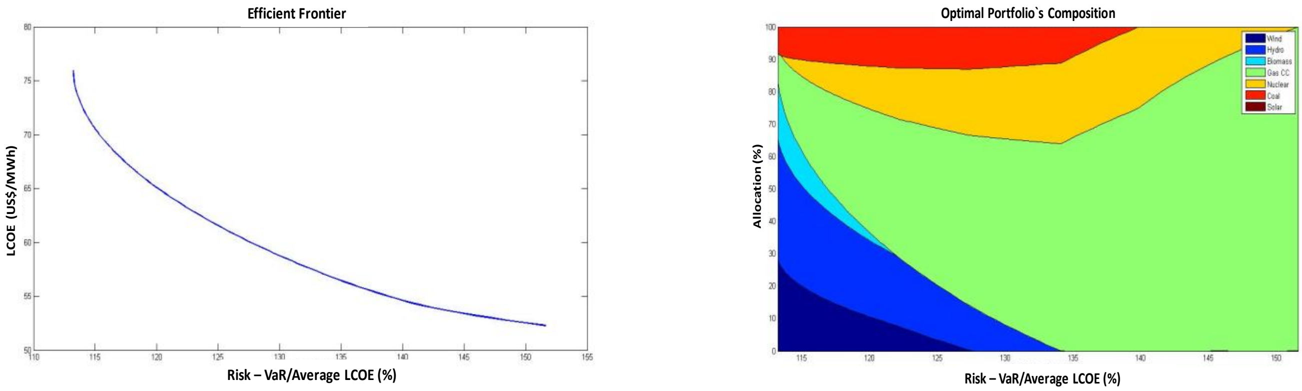



**Figure 30: Efficient frontier and optimal portfolio composition for the year of 2030, taking the VaR as risk measure**

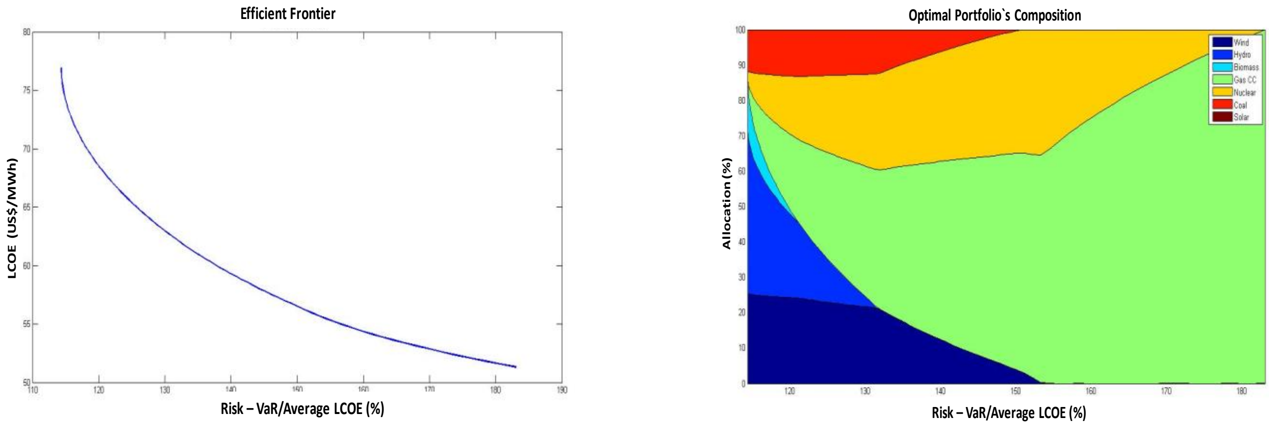



**Figure 31: Efficient frontier and optimal portfolio composition for the year of 2035, taking the VaR as risk measure**

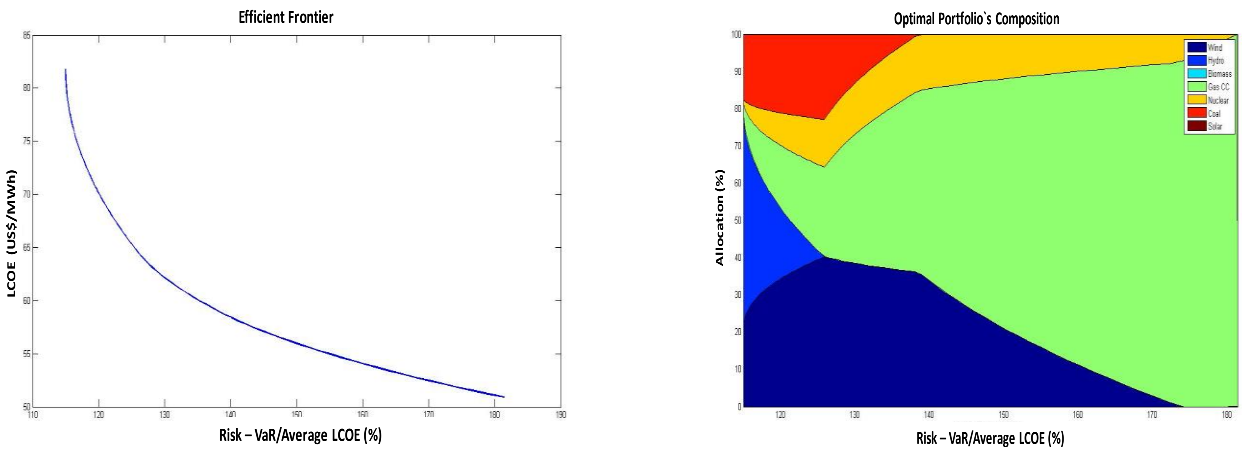



**Figure 32: Efficient frontier and optimal portfolio composition for the year of 2040, taking the VaR as risk measure**

One can notice that despite of the general forms being the same, there are evidently some differences between the application of the VaR and the variance as risk measures. Initially, it is worth high lightening that the risk demonstrated on the efficient frontier, given a certain LCOE, is greater when the VaR is used as risk measure. This is so, because the VaR is a tail risk measure and so captures more efficiently extreme events, which should be taking into account. Now, regarding the composition chart, ideal low risk portfolios in 2030 present a greater allocation in wind technology in detriment of the biomass technology. It may be explained by the fact that the VaR is a tail risk measure and is able to capture some risk that the standard deviation is not able to do so, such as strikes and workers` claims, which is very common within the biomass technology, since it is a labor-intensive business.

### 4.2.3 Conditional Value at Risk (CVaR)

Finally, the Kriging Method is applied using the CVaR as risk measure, with the same input parameters. Results are showed in the following figures:

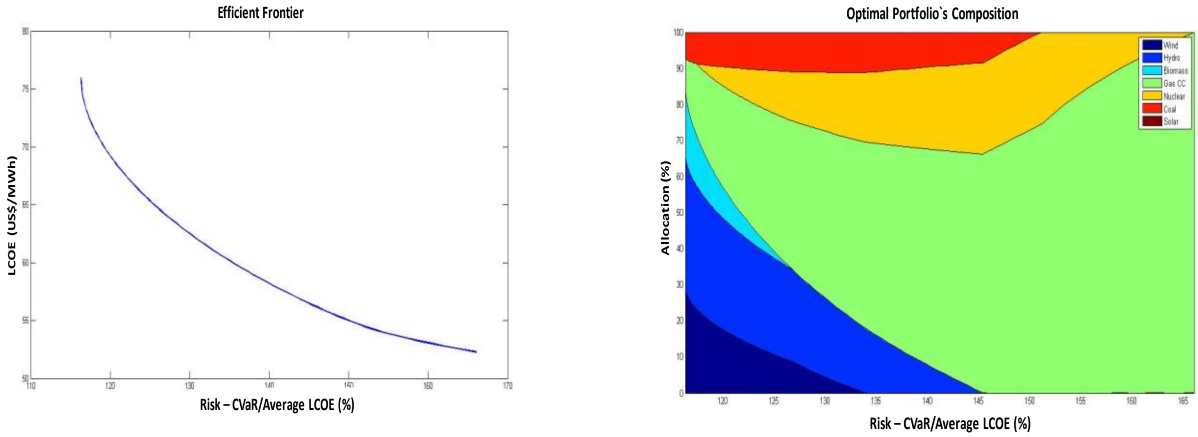



**Figure 33: Efficient frontier and optimal portfolio composition for the year of 2030, taking the CVaR as risk measure**

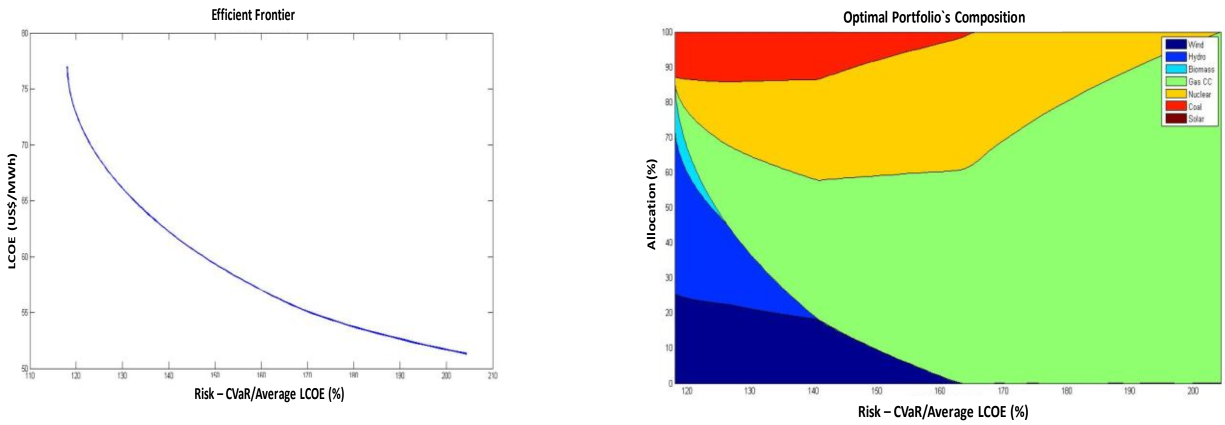



**Figure 34: Efficient frontier and optimal portfolio composition for the year of 2035, taking the CVaR as risk measure**

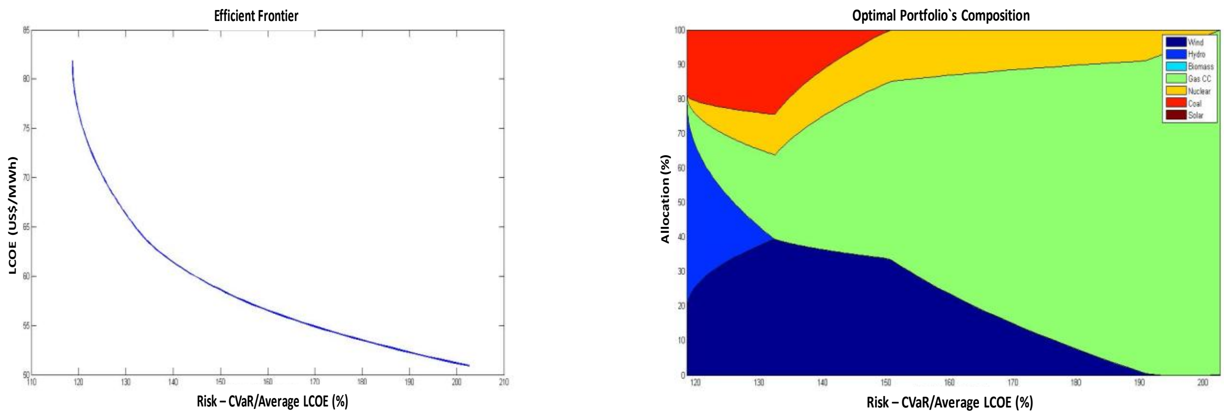



**Figure 35: Efficient frontier and optimal portfolio composition for the year of 2040, taking the CVaR as risk measure**

It is interesting to note that the results obtained in the phase (i.e. using the CVaR as risk measure) are very similar to those obtained using the VaR as risk measure, since both are tail risk measures, and therefore, capture both sporadic events. It is worth noting that in the case of the efficient frontier, for a same level of cost, the CVaR presents a greater risk than the VaR, what may be explained by the fact that the CVaR is defined as the average of the values that exceed the VaR. Finally, it is important to state that these soft changes in the composition charts (compared to these VaR`s charts) are again explained by the fact that the CVaR captures events even more drastic than those captured by the VaR. In this sense, the CVaR may be considered more efficient, since policymakers should take in consideration such events when allocating and selecting a country`s energy matrix portfolio.

### 4.2.4 Minimum Risk Portfolios` Compositions

In this section, it is exhibited the optimal portfolios for each year regarding the minimum possible risk to be reached for such portfolio. This portfolio is important because, many times, it is in fact the portfolio that is intended to achieve.

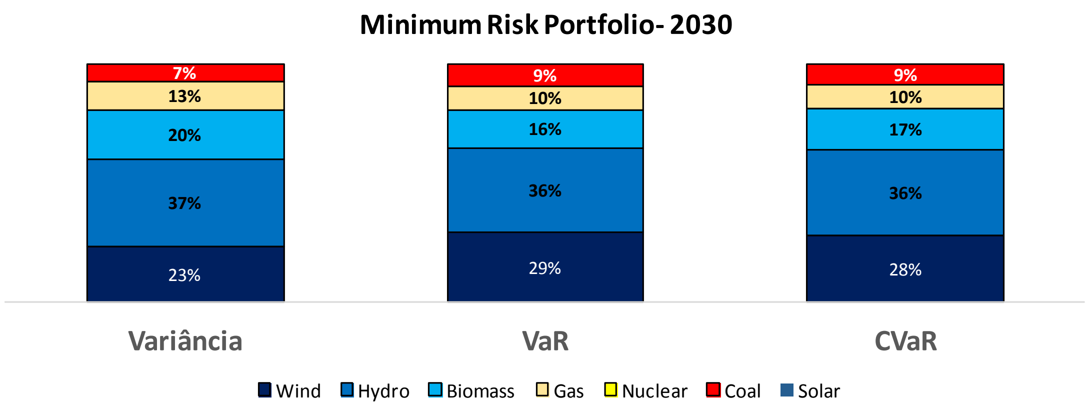



**Figure 36: Minimum risk portfolio for the year of 2030**

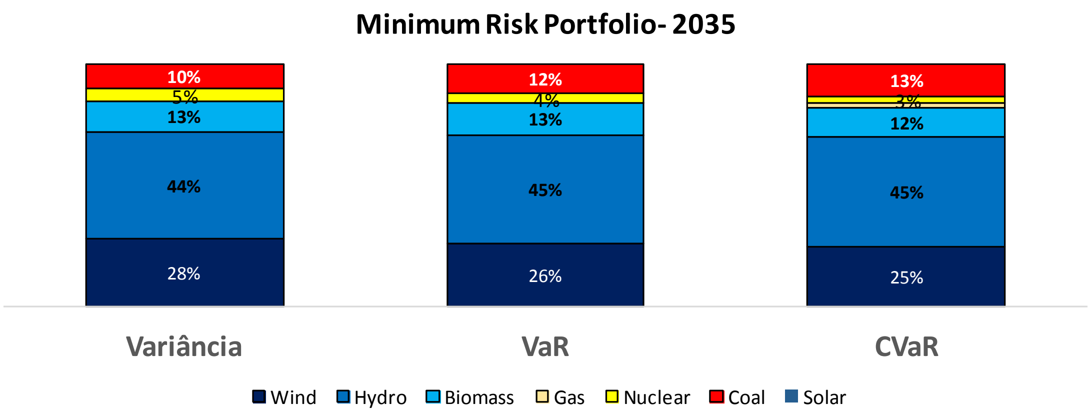



**Figure 37: Minimum risk portfolio for the year of 2035**

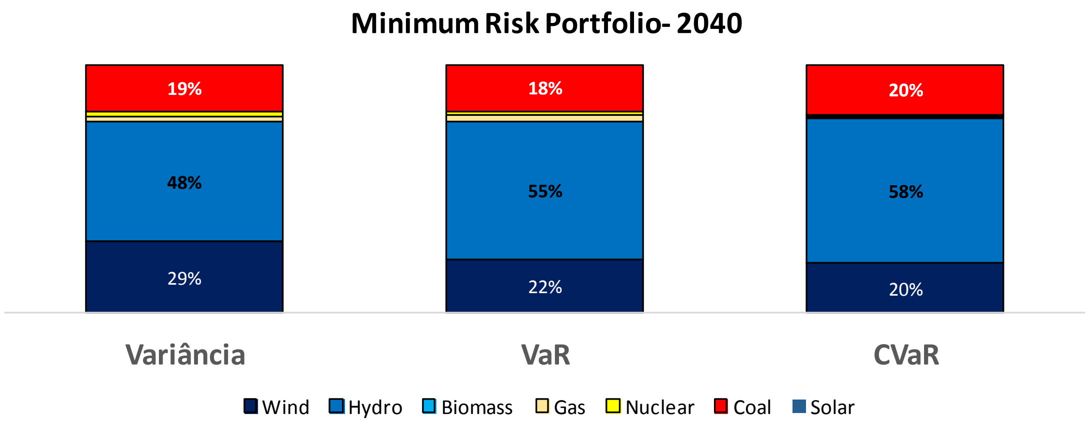



**Figure 38: Minimum risk portfolio for the year of 2040**

Again, it is noticeable that the portfolios have similar allocation, with soft alterations. First, it is possible to observe that there is a trend throughout the years, for the three measures used: a greater allocation in hydro technology in detriment of wind technology. As explained before, this trend is given by the fact that it is estimated that occurs a reduction of the risks associated with the hydro technology, since the greatest risk incurs in the project`s initial phase. Secondly, the decrease of the allocation in gas technology is due to the fact that the risk associated with this technology increases through the years. Finally, it is worth saying that there are differences between the CVaR and the VaR portfolios and it is due to the fact that the former manages to capture event more adverse which negatively impact on the generation cost.

# 5 CONCLUSIONS

This study aimed to use a portfolio selection method capable of optimizing the risk-return relation of an energy matrix portfolio. It was chosen seven generation technologies to compose the portfolio, and the data analyzed was from the United States energy market.

Initially, it was presented the Modern Portfolio Theory, which was created by Harry Markowitz in 1952, standing as the first formulation considered efficient on maximizing a portfolio utility function. Based on it, it was possible to further deepen the analysis, describing other models which use other risk measures than the variance used by Markowitz.

The VaR and the CVaR surged in the literature as risk measures that overcome the drawbacks presented on the variance, given that the CVaR is considered more complete as it is coherent according to Artzner et al (1999), and also as it analyze the probability distribution tail (ROCKAFELLAR e URYASEV, 2002).

Therefore, the Variance, the VaR and the CVaR were studied as risk measures used to evaluate the risk-return relation in the energy matrix optimization problem. In the case of the VaR and the CVaR, theirs traditional models are very complex to be optimized, since it has a great number of variable and constraints, what implies in a non-convenient problem regarding the computational stand point.

In order to overcome these drawbacks, it was proposed the application of the Kriging Method, also known as DACE fit, which is a tool frequently used to solve engineering problems (QUEIPO, et al., 2002). This method creates a response surface which is smoothed based on a sample previously defined. It makes the problem a simple form of resolving the optimization, without losing its validation, what implies in great decrease of the necessity of computation capacity to solve the problem.

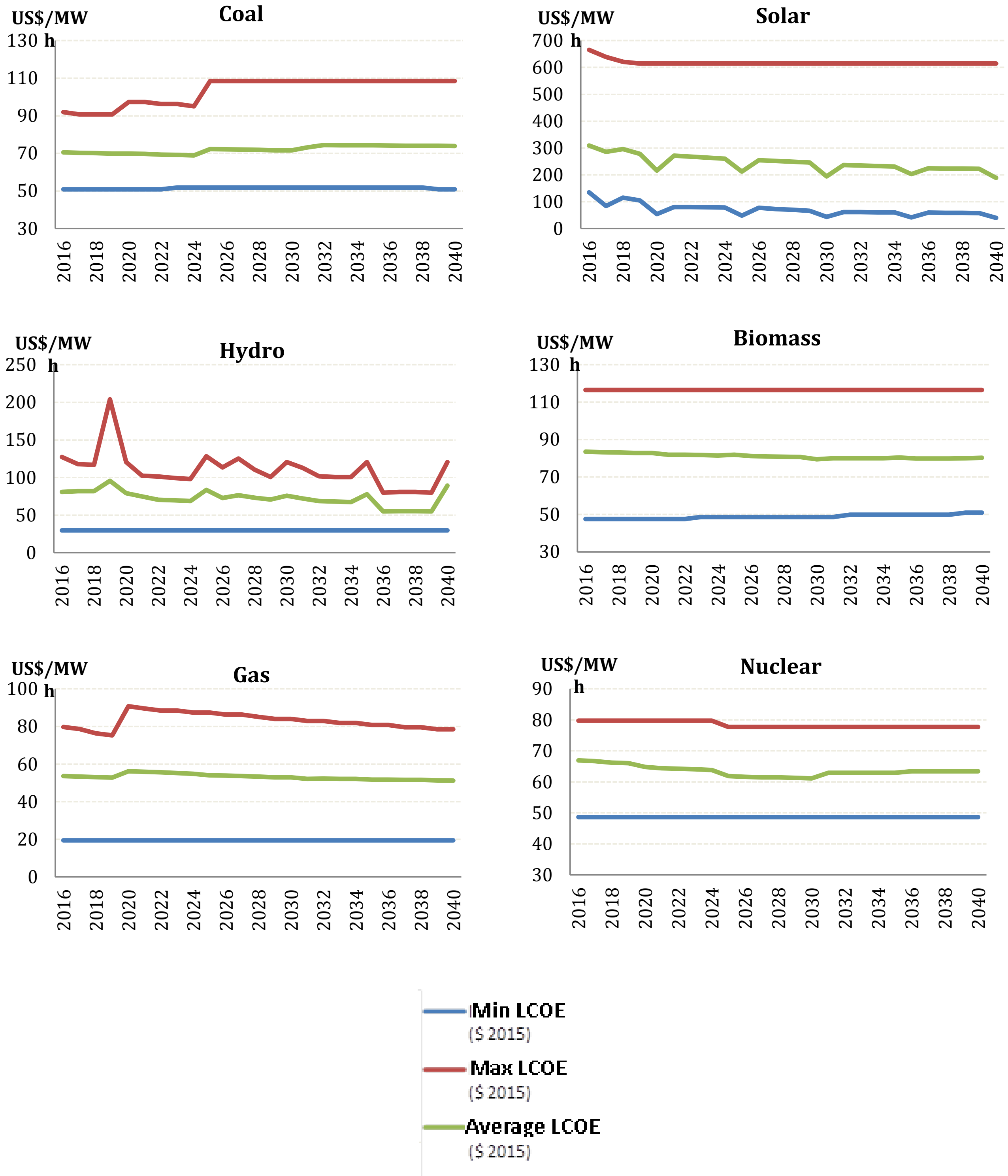
Therefore, based on estimates done by north-American public agencies, Monte Carlo simulations were carried out and the Kriging Method was applied through the use of the software Matlab for the years of 2030, 2035, and 2040, using as risks measures the Variance, the VaR, and the CVaR. In a preliminary observation of the results, it is possible to observe that the solar technology was not allocated in any of the portfolios generated, indicating that this technology still needs improvements in order to reduce and stabilize its costs.

The results suggest that, in the future, there will be a higher concentration in renewable energies if compared to the country`s actual portfolio, mainly the hydro and wind technologies. Moreover, the model also reveals that the technology which uses the biomass as input will present a greater participation in the country`s energy matrix, given that, nowadays, it represents only one percent of the total allocation. Hence, it is possible to assume that environment aggressive technologies (namely coal, gas and nuclear) should play a minimal role in future energy matrix.

It is important to highlight that the model used in the study is a simplified representation of the reality, which may be more robust as new constraints are added to the problem. A possible additional constraint to the problem is one that considers certain environmental issues, such as one related to the abusive use of nuclear energy. Another future extension to this study would be analyzing the probability distributions of each technology cost and incorporate it on the simulation process.

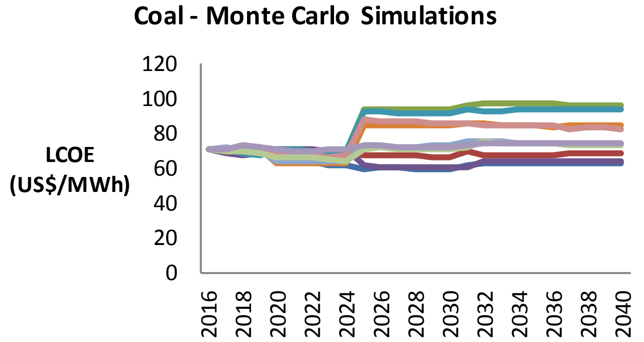
Finally, even with possible extensions, one may consider the initially proposed objective accomplished, since: it was created an energy matrix portfolio optimization problem which allowed (i) considering the tail risks, using risk measures as the VaR and the CVaR; (ii)the increase of the tool`s computational efficiency through the use of the proposed method, the Kriging Method.

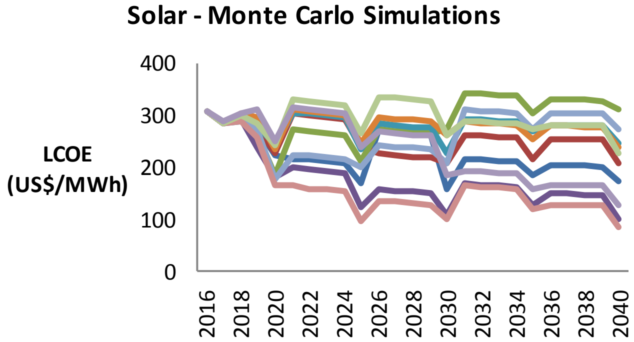
# Appendix A

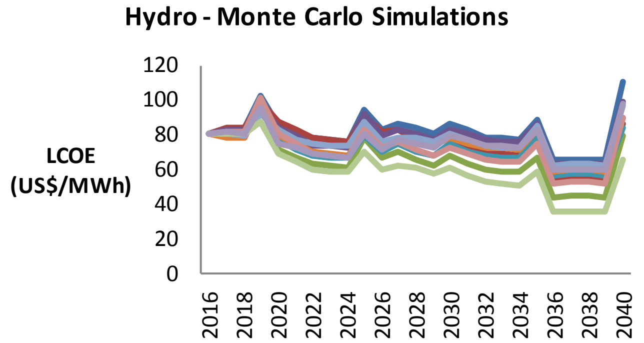


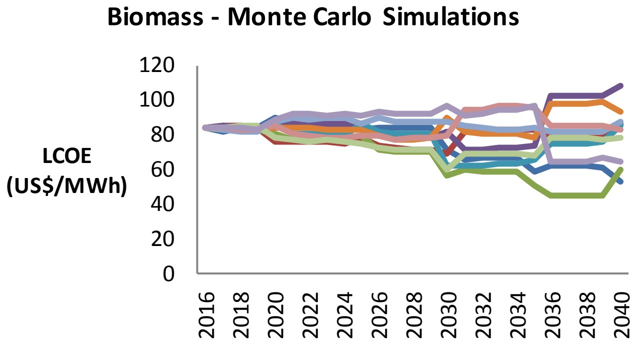
Coal
US$/MWh
130
110
90
70
50
30
Solar
US$/MWh
700
600
500
400
300
200
100
0
Hydro
US$/MWh
250
200
150
100
50
0
Biomass
US$/MWh
130
110
90
70
50
30
Gas
US$/MWh
100
80
60
40
20
0
Nuclear
US$/MWh
90
80
70
60
50
40
30
2016 2018 2020 2022 2024 2026 2028 2030 2032 2034 2036 2038 2040
Min LCOE ($ 2015)
Max LCOE ($ 2015)
Average LCOE ($ 2015)

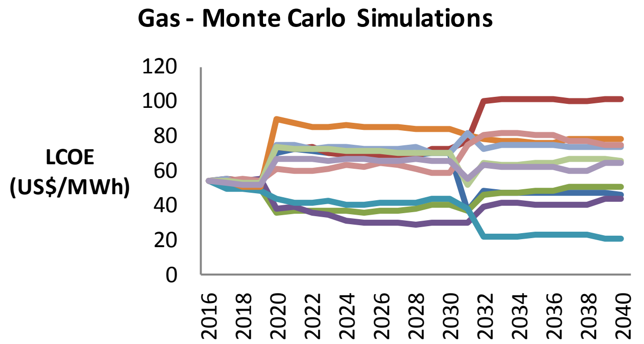
# Appendix B

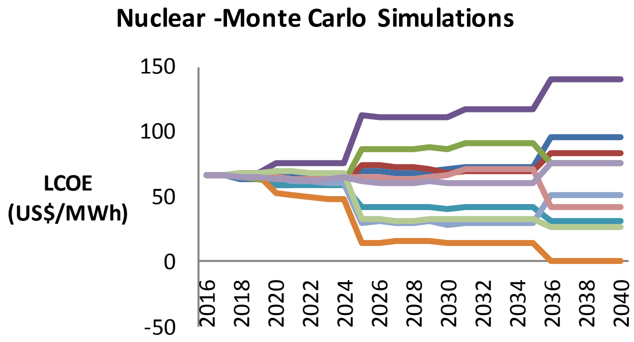

Coal - Monte Carlo Simulations
LCOE (US$/MWh)
0
20
40
60
80
100
120
2016
2018
2020
2022
2024
2026
2028
2030
2032
2034
2036
2038
2040


Solar - Monte Carlo Simulations
LCOE (US$/MWh)
0
100
200
300
400
2016
2018
2020
2022
2024
2026
2028
2030
2032
2034
2036
2038
2040


Hydro - Monte Carlo Simulations
LCOE (US$/MWh)
0
20
40
60
80
100
120
2016
2018
2020
2022
2024
2026
2028
2030
2032
2034
2036
2038
2040


Biomass - Monte Carlo Simulations
LCOE (US$/MWh)
0
20
40
60
80
100
120
2016
2018
2020
2022
2024
2026
2028
2030
2032
2034
2036
2038
2040


Gas - Monte Carlo Simulations
LCOE (US$/MWh)
0
20
40
60
80
100
120
2016
2018
2020
2022
2024
2026
2028
2030
2032
2034
2036
2038
2040


Nuclear -Monte Carlo Simulations
LCOE (US$/MWh)
-50
0
50
100
150
2016
2018
2020
2022
2024
2026
2028
2030
2032
2034
2036
2038
2040

# 6 REFERENCES

ANDRADE, F.W.M., Alocação de Ativos no Mercado Acionário Brasileiro Segundo o Conceito de Downside Risk. Revista de Gestão USP, v.13.,n.2, p.27-36, 2006 ROMAN, 2008.

ARTZNER, P. et al. Coherent mesures of risk. **Mathematical Finance**, v. 9, p. 203-228, July 1999.

AWERBUCH, S., BERGER, M. (2003). Applying portfolio theory to EU Electricity planning and policy making. IEA/EET Working Paper, EET/2003/03.

AWERBUCH, S. (2006). Portfolio-Based Electricity Generation Planning: Policy Implications for Renewables and Energy Security, Mitigation and Adaptation Strategies for Global Change, Vol. 11, n. 3, 2006.

BARROSA, M. R. da, Aplicação de métodos computacionais multidisciplinares de engenharia para otimização de carteiras de investimentos / M.R.. da Barrosa -- São Paulo, 2015.

BAR-LEV, D., KATZ, S. (1976). A Portfolio Approach to fossil fuel procurement in the Electric Utility Industry, *The Journal of Finance*, vo. 31, Issue 3, p. 933-947, 1976.

BERNOULLI, D., Exposition of a new theory on the measurement of risk. Econometrica 22, p. 23-36. Tradução de artigo original publicado em Latim em São Petersburgo em 1738.

BREALEY, R. and S. MEYERS, Principles of Corporate Finance, Seventh Edition, McGraw Hills, 2003.

COSTA, P. e T. BAIDYA, Propriedades Estatísticas das Séries de Retornos das Principais Ações Brasileiras. Pesquisa Operacional, v. 21, n.1, p. 61-88, 2001.COSTA NETO, P. L. O. **Estatística**. 2ª. ed. São Paulo: Editora Blücher, 2002.

DELARUE, E., DE JONGHE, E., BELMANS, R., D'HAESELEER, W., (2011). Applying portfolio theory to the electricity sector: Energy versus power. Energy Economics ,v.33, p.12–23, 2011.

FISHBURN, P.C., Mean-risk analysis with risk associated with below-target returns, American Economic Review 67, no 2, p. 116-126, 1977.

GIUDICI, G., Finanziare le Risorse dell'Impresa, Edizioni CLUP (quarta edizione), 2010.

GLASSERMAN, P., Monte Carlo Methods in Financial Engineering, Springer, New York, 2003

HULL, J. C. **Options, Futures and Other Derivatives**. Fourth. ed. Upper Saddle River: Prentice Hall, 1999.

JONES, D. R.; SCHONLAU, M.; WELCH, W. J. Efficient global optimization of expensive black-box functions. **Journal of Global Optimization**, v.13, n.4, 1998. p.455-492.

JORION, P. **Value at Risk:** the new benchmark of managing financial risk. 3th. ed. New York: McGraw-Hill, 1997.

KONNO, H. and H. YAMAZAKI, Mean-absolute deviation portfolio optimization model and its application to Tokyo stock market, Management Science 37, p. 519-531, 1991.

KREY, B., ZWEIFEL, P. (2006), Efficient Electricity Portfolios for Switzerland and the United States. *Socioeconomic Institute of the University of Zurich*, 2006.

LIM, A., SHANTHIKUMAR G. and GAH-YI VAHN, Conditional Value at Risk in portfolio optimization: coherent but fragile. Operations Research Letters, Volume 39, Issue 3. p. 163-171, 2011.

LOPHAVEN, S. N.; NIELSEN, H. B.; SONDERGAARD, J. **DACE: a matlab kriging tollbox**. Denmark: Informatics and Mathematical Modelling. Technical University of Denmark, p. 28 (Technical Report, IMM-TR-2002-12). 2002.

LOSEKANN, L., MARRERO, G. A.,RAMOS-REAL, F. J., ALMEIDA, E. L. F. (2013). Efficient power generating portfolio in Brazil: Conciliating cost, emissions and risk. Energy Policy ,v.62, p.301–314, 2013.

MARKOWITZ, H. Portfolio Selection. **The Journal of Finance**, v.7, 1952. p.77-91.

MARKOWITZ, H. Portfolio Selection: Efficient Diversification of Investments, Yale University Press, New Haven, 1970.

O mercado de valores mobiliários brasileiro / Comissão de Valores Mobiliários. 3. ed. Rio de Janeiro: Comissão de Valores Mobiliários, 2014.

PAPALAMBROS (2000), P.Y. e D.J. WILDE, Principles of Optimal Design, *Cambridge University Press*, 2000.

PEROLD, A. F., Large scale portfolio optimization, Management Science 30, p. 1143-1160, 1984.

PFLUG, G. C. Some remarks on the Value-at-Risk and Conditional Value-at-Risk. In: LUENBERGER, D. G. **Investment Science**. New York: Oxford University Press, 1998.

QUARANTA, A. G.; ZAFFARONI, A. Robust optimization of condition value at risk and portfolio selection. **Journal of Banking and Finance**, v.32, 2008. 2046-2056.

QUEIPO, V. N. et al. Surrogate Modeling-based optimization for the integration of statistic and dynamic data into a reservoir description. **Journal of Petroleum Science and Engineering**, v.37, n.3-4, 2002. p.167-181.

RIBEIRO, C. O.; FERREIRA, L. A. S. Uma contribuição ao problema de composição de carteiras de mínimo valor em risco. **Gestão e Produção**, v.12, 2004. p.295-304.

ROCKAFELLAR, R. T.; URYASEV, S. Conditional Value-at-Risk for general loss distributions. **Journal of Banking and Finance**, v.26, 2002. 1443-1471.

ROCKAFELLAR, R. T.; URYASEV, S. Optimization of Conditional Value-at-Risk. **Journal of Risk**, v. 2, 2000. 21-41.

ROY, A. D., Safety First and the Holding of Assets, Econometrica 20, p. 431-449, 1952.

RUBINSTEIN, M. Markowitz's "Portfolio Selection": A Fifty-Year Retrospective. **The Journal of Finance**, v.57, 2002. p.1041-1045.

SZEGÖ, G. Measures of risk. **Journal of Banking & Finance**, v. 26, p. 1253-1272, June 2002.

YIN, J.; NG, S. H.; NG, K. M. Kriging metamodel with modified nugget-effect: The heteroscedastic variance case. **Computers & Industrial Engineering**, Singapore, 61, 2011. 760–777.