SPE 152923

Probabilistic Reserves and Resources Estimation: A Methodology for Aggregating Probabilistic Petroleum Reserves and Resources

M. Galvao, L. Hastenreiter, and J. Molina, Petrobras, A. Quadros, J. Montechiari, and S. Hamacher, PUC-Rio

This paper was prepared for presentation at the SPE/EAGE European Unconventional Resources Conference and Exhibition held in Vienna, Austria, 20-22, March 2012.

Abstract

The oil industry strives to create an international standard for classification and estimation of resources since the 1930s. The goal is to provide investors with information obtained under the same assumptions, as to facilitate the comparison between petroleum companies. In 2007 the four major international organizations, *Society of Petroleum Engineers* (SPE), American *Association of Petroleum Geologists* (AAPG), *World Petroleum Council* (WPC) and *Society of Petroleum Evaluation Engineers* (SPEE), jointly released a single set of guidelines for classification and evaluation of oil and gas resources, the *Petroleum Resources Management System* (PRMS, 2007).

Several methodologies for estimating reserves can be employed within the PRMS's (2007) premises, which can be classified as deterministic or probabilistic. Unconventional resources emerge as a new frontier for the oil industry, thus implying high uncertainty levels in both technical and economic assessments. The main purpose of this paper is to explore this issue and to propose a correlation-based probabilistic methodology for aggregating oil and gas reserves of conventional and unconventional resources. The methodology is in accordance with the guidelines of the PRMS (2007) and with the new rules of the *Securities Exchange Commission* 2009 (SEC). The correlation assessment evaluates technical and operational features, and the probabilistic aggregation is performed by Monte Carlo Simulation (MCS).

Besides the introductory section, this paper comprises four other sections. A literature review presents definitions of conventional and unconventional resources, and an examination of classification, estimation and aggregation of reserves, important for better understanding the following sections. The third section describes the proposed correlation-based probabilistic methodology. Afterward, a case study presents an application of the methodology. Finally, the last section synthesizes the main conclusions.

Introduction

The estimation of petroleum reserves entails complex assumptions and calculations, once there is uncertainty associated with volumes, recovery, development, and marketability of resources. Unconventional resources bring new challenges to the industry, reinforcing that uncertainty analyses are paramount for estimating reserves.

Yet, uncertainty based assessments are not trivial assignments. There must be a thorough understanding of the models to be employed, in order not to overestimate proved reserves.

When reserves are determined by probabilistic methods, the arithmetic addition of individual accumulations within an integrated project will understate the aggregated reserves at the proved (P90) level. On the other hand, if full independency is assumed and the accumulations share common risks, the probabilistic addition of reserves will overstate the proved (P90) reserves (Carter and Morales, 1998).

This paper introduces a practical correlation-based methodology for aggregating probabilistic reserves from both conventional and unconventional resources. As long as production forecasts have passed through an uncertainty analysis and they have been built under the correct premises, the methodology can be applied. Figure 1 presents the basic information flow that summarizes the whole reserves estimation process.

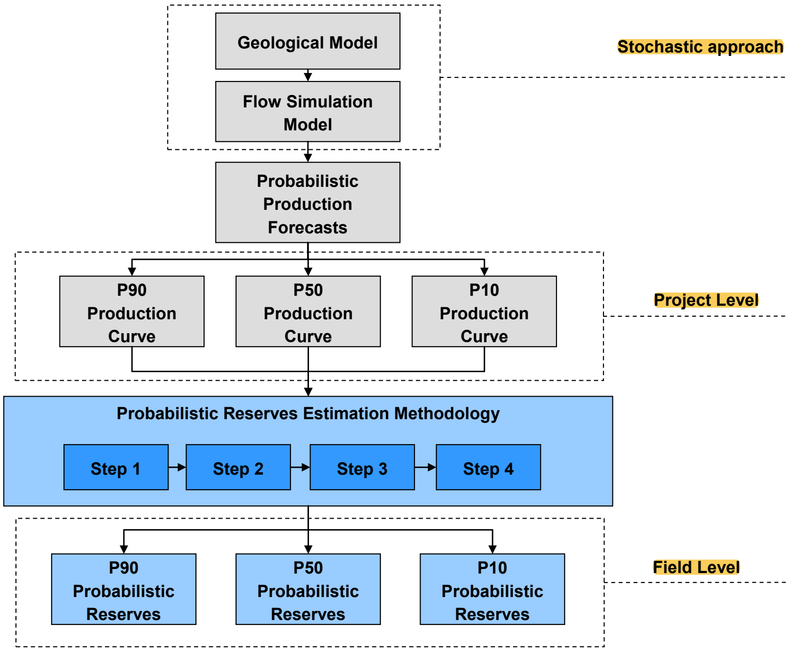

Figure 1: Information flow

It is important to state that the gray boxes are inputs for the methodology. The focus lies on estimating the projects' reserves probability distribution and aggregating these curves.

Literature Review

Conventional and Unconventional Resources

PRMS (2007) defines conventional resources as discrete petroleum accumulations limited to a geological structural feature and/or stratigraphic condition, where the trapping mechanism is dominated by hydrodynamic or buoyancy forces. Additionally, conventional resources are recovered through wellbores and typically require minimal processing prior to sale.

Regarding unconventional resources, PRMS (2007) defines them as hydrocarbon accumulations that are pervasive throughout a large area and that are generally not significantly affected by hydrodynamic influences (also called "continuous-type deposits"). Such accumulations require specialized extraction technology and, in the case of oil, the raw production may require significant processing prior to sale.

According to Elliot (2008), the fundamental difference between conventional and unconventional resources lies in the trapping mechanism and its influence on the production mechanism.

Chan *et al.* (2010) present the Petroleum-Resource Triangle (modified from Holditch, 2002) shown in Figure 2.

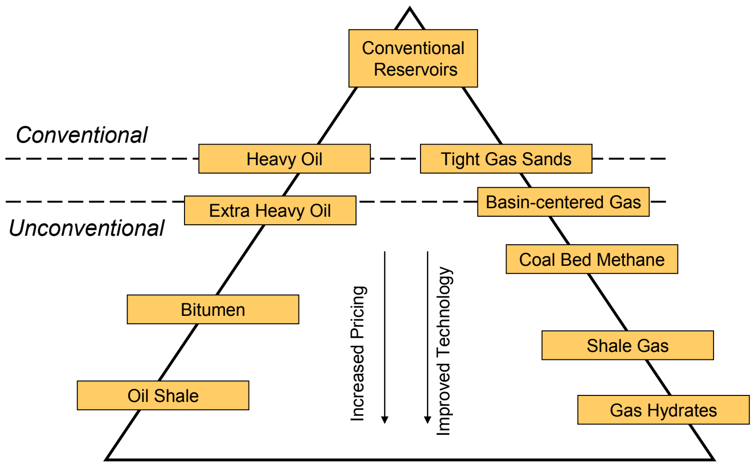

Figure 2: Petroleum-Resource Triangle

The difference of bitumen, heavy oil, and light oil is majorly due to their viscosity and density at standard conditions. As stated by Etherington and McDonald (2004), heavy oil is defined as less than 22° API gravity while extra heavy crude is less than 10°. Bitumen is defined as heavy or extra-heavy crude with viscosities greater than 10,000 centipoises measured at reservoir temperature and atmospheric pressure, on a gas free basis (Canadian Oil and Gas Evaluation Handbook, 2002). Bitumen is either recovered by mining or an in-situ process that reduces its viscosity to the point where it can flow to a wellbore (Elliot, 2008).

Precise quantitative reserves and oil-in place data for natural bitumen and extra-heavy oil on a reservoir basis are seldom available to the public, except in Canada. Bitumen is reported in 598 deposits in 23 countries. No deposits are reported offshore. It occurs both in clastic and carbonate reservoir rocks and commonly in small deposits at, or near, the earth's surface. The three Alberta oil sand areas, Athabasca, Peace River, and Cold Lake, together contain 1.73 trillion barrels of discovered bitumen in place, representing two-thirds of that in the world and at this time are the only bitumen deposits being commercially exploited as sources of synthetic crude oil (World Energy Council, 2010).

Oil Shale is the extraction of liquid hydrocarbons by heating shales containing solid hydrocarbon. The commercial viability of extracting these hydrocarbons is still under study. It is noteworthy that the total world in-place resource of shale oil is estimated at 4.8 trillion barrels (World Energy Council, 2010).

Tight Gas is held in low permeability reservoirs, sometimes "basin-centered" and deep. The recovery depends largely on natural or induced fractures (Elliot, 2008).

Coal Bed Methane is trapped by adsorption in micropores in coal and it is recovered by pressure differential between the coal and fractures (Elliot, 2008).

Shale Gas is trapped by adsorption on the solid hydrocarbon, kerogen, and on clay particles. The recovery process is based on the desorption from the matrix due to a pressure differential and drainage through fractures (Elliot, 2008).

Gas Hydrates is a solid clathrate compound, where methane is trapped within a crystal structure of water. According to Elliot (2008), the viability of extracting these hydrocarbons is still at a very early stage.

Production Forecasts and Reserves Estimation

There are several methods for production forecasting, such as decline curves (Arps, 1945) and material balance. These methods are widely spread in the industry and the reader should be familiar with them. Yet, they may be limited for unconventional reservoirs.

Lee and Sidle (2010) presented a paper entitled "Gas Reserves Estimation in Resource Plays" (SPE 130102), which brings a critique of some methods used to forecast production and estimate reserves in unconventional resource plays. The authors start by the volumetric method, which consists of estimating two pieces of information: the volume of hydrocarbons in place and the recovery factor. In tight gas reservoirs, it can be difficult to estimate the areal extent of the reservoir and its net pay thickness. Besides, in shale gas resources, gas volumes in pore spaces can be difficult to estimate accurately.

The material balance method (p/z *versus* Gp) may not provide reserves estimates and production forecasts of sufficient accuracy for unconventional gas reservoirs. This is due to the assumptions necessary for the applicability of the method, which are an unchanging drainage volume and a stabilized (boundary-dominated) flow.

The analogy method is able to provide reliable results in both conventional and unconventional reservoirs and wells. The difficulty lies on identifying accurate analogies. For unconventional resources, this issue tends to be more sensitive, once production history for specific reservoir properties and technologies are seldom available.

Arps' decline curves method (Arps, 1945) is based on determining the trend in production decline from history and to project the future production. Thus, one can calculate reserves when the economic limit is reached. Its general equation is given by:

$$q = q_i \frac{1}{\left(1 + bD_i t\right)^{(1/b)}} \qquad (1)$$

q: Production rate at time t
q_i: Production rate at time zero
D_i: Arps' initial decline rate
b: Arps' hyperbolic decline constant

For $b = 0$ the decline is called exponential, for $b = 1$ it is called harmonic and for $0 < b < 1$ it is the hyperbolic decline curve.

Lee and Sidle (2010) state that attempts to determine the Arps' constants for tight gas and shale reservoirs result is values of b that exceeds the unity. As a consequence, the reserves estimates present physically unreasonable properties. This result is due to the fact that data are in the transient flow.

History matching with analytical models is normally limited due to the shortage of quality information concerning formation properties. Besides, an incomplete understanding of the gas flow from complex reservoirs and the required effort to adjust each individual well can also impact on the overall accuracy.

Ultimately, type curves methods present limitations regarding newer gas shales plays with limited long-term production and in some layered tight gas formations.

There are some recent empirical models that focus on overcoming the problems listed by Lee and Sidle (2010). A new approach is presented by Duong (2011) to predict the future rate and the estimated ultimate recovery (EUR) for fracture-dominated wells in unconventional reservoirs. Traditional decline methods cannot be applied for wells producing from supertight or shale reservoirs in which fracture flow is dominant. Without the presence of pseudoradial and boundary-dominated flows (BDFs), neither matrix permeability nor drainage area can be established. Thence, the fracture contribution is dominant over the matrix, and the EUR cannot be based on a traditional concept of drainage area.

Duong's work introduces an empirically derived decline model that is based on a long-term linear flow in a large number of wells in tight and shale-gas reservoirs. In addition, this new approach is able to represent any uncertainty in reserves estimation by providing a statistical method to analyze production forecasts of resource plays. It can establish a range of results for these forecasts, including probability distributions of reserves in the form of P90 to P10.

Other studies, such as Ilk *et al*. (2008) and Valkó (2009), also propose decline curve analysis methods different from the Arps' method. Space does not allow a detailed description of these techniques, but the references to sources are provided.

Reserves Classification

According to PRMS (2007), a quantity of petroleum can be classified as reserves if it satisfies four criteria: it must be discovered, recoverable, commercial, and remaining, based on the development project applied. Reserves are further classified as proved (1P), proved plus probable (2P) or proved plus probable plus possible (3P), in accordance with the level of certainty associated with the estimates. As stated before, the volume needs to be commercial and therefore an economic analysis is required.

Reserves estimates are to be based on "reliable technology" to establish appropriate levels of certainty for the reserves volumes disclosed under SEC (2009) guidelines. "Reliable technology" is described by SEC (2009) as: repeatability and consistency.

Unconventional resources demand different assessment approaches than conventional projects. The main differences are (Chan *et al*., 2010):

- In most cases there is little or no discovery risk of in-place hydrocarbons in unconventional plays;
- The primary challenges in evaluating unconventional resources are in establishing sufficient scale and recovery rates to support commercial projects.

Similarly to improved recovery projects applied to conventional reservoirs, PRMS (2007) requires successful pilots in the subject reservoir or successful projects in analogous reservoirs to establish recovery efficiencies for non-conventional accumulations (Chan *et al*., 2010).

As stated by Ross (2001), two fundamentally different philosophies have been applied to the estimation of recoverable volumes. One is known as "incremental approach", which is a risk-based philosophy, while the other is an uncertainty-based philosophy and it is generally referred to as "cumulative approach". Ross (2001) defines "risk" in the context of the petroleum industry as the probability that a discrete event will or will not occur, and "uncertainty" as the range of possible outcomes in an estimate.

When the range of uncertainty is represented by a probability distribution, a low, a best, and a high estimate should be provided. In this case, proved reserves mean that there should be at least a 90% probability (P90) that the quantities actually recovered will equal or exceed the low estimate (PRMS, 2007). Likewise, proved plus probable reserves represent at least a 50% probability (P50) that the quantities actually recovered will equal or exceed the best estimate, and proved plus probable plus possible reserves represent at least a 10% probability (P10) that the quantities recovered will equal or exceed the high estimate.

Reserves Aggregation

The reserves estimates of each hydrocarbon accumulation are added to arrive at estimates for higher levels, such as fields or properties. This aggregation can be performed deterministically or probabilistically. When performed probabilistically, as a result we have what is known as "Portfolio Effect" (Markowitz, 1952), which is a reduction in the variation of returns on a combination of assets compared with the average of the variations of the individual assets. The larger and more diverse the combination of accumulations, the greater the portfolio effect (Etherington *et al*., 2001). Particularly in the case of reserves represented by probability distributions, for example considering two ventures A and B, it can be verified that:

$$P90_A + P90_B \leq P90_{A+B} \tag{2}$$

$$P50_A + P50_B \approx P50_{A+B} \tag{3}$$

$$P10_A + P10_B \geq P10_{A+B} \tag{4}$$

These results are due to the reduction of the variance of the sum when partially correlated variables are compared to fully dependent variables. So as to illustrate, the variance of the sum of variables A and B are shown in equation 5.

$$VAR(A+B) = VAR(A) + VAR(B) + 2COV(A,B) = \sigma_A{}^2 + \sigma_B{}^2 + 2\rho_{A,B}\sigma_A\sigma_B \tag{5}$$

The parameter $\rho_{A,B}$ is the correlation coefficient between A and B, which is defined in the interval [-1,1]. Fully dependent variables have a correlation coefficient equal to 1, thus resulting in the maximum summation value. As for partially correlated variables, i.e. $\rho_{A,B}$ is less than the unity, the variance of the sum assumes smaller values. Therefore, the aggregated P90 and P10 converge to the mean of the curve, confirming the results presented in equations 2 and 4.

According to Carter and Morales (1998), when probabilistic methods are applied and the resulting reserves are obtained through arithmetic addition, the proved (P90) level is underestimated. If full independency is assumed and the accumulations share common risks, then the resulting proved (P90) reserves are overestimated. Hence, dependencies must be taken into account so that reasonable results can be achieved. In their study, dependencies have been quantified and the probabilistic addition resulted in proved reserves some 9% greater than the deterministic summation.

There is a constraint concerning the order of magnitude for assessing a reserves entity probabilistically. As stated in the PRMS (2007), "*…assessment results should not incorporate statistical aggregation beyond the field, property, or project level. Results reporting beyond this level should use arithmetic summation…*". Regarding this guideline, we define a reserves entity likely for a probabilistic assessment as a venture which scope is a well or a group of wells belonging to a single field, hereinafter referred to as "project".

Sampling Methods

Sampling methods are applied to aggregate the reserves probability distribution curves. The great advantage of sampling methods is to accept any probability distribution for the input data.

Monte Carlo and Latin Hypercube simulations are two examples of sampling methods. Latin Hypercube will accurately recreate the probability distributions specified by distribution functions in fewer iterations, when compared with Monte Carlo sampling (Saliby and Moreira, 2002). Latin Hypercube Sampling stratifies the input probability distributions, i.e. it divides the cumulative curve into equal intervals on the cumulative probability scale (0 to 1). A random sample is compulsorily taken from each interval of the input distribution, thus offering benefits in terms of increased sampling efficiency and faster runtimes.

Correlation Assessment

Different exploitation projects may share similar risks and reservoir properties; hence the correlation among them is a key factor for estimating the aggregated reserves. For instance, if projects A and B are strongly correlated, and there is success in the exploitation of A, so there is a high chance of achieving success in the exploitation of B. The identification of these correlations is very important to avoid the overestimation and underestimation of the P90 and P10 aggregated volumes, respectively.

The direct determination of the correlation coefficient between projects may not be an easy assignment. Carter and Morales (1998) suggest a single assessment for each of the factors that most influence the determination of the correlations. The idea is to compare pairs of projects in each of these factors, instead of comparing the ventures directly.

The main methods to address the correlation between exploitation projects are:

- Direct comparison between each pair of projects made by experts;
- Comparison between each pair of projects on some specific criteria made by experts (Carter and Morales, 1998);
- Binarization: Projects are considered independents and therefore it is assumed that correlations are equal to zero or projects are considered correlated and it is assumed that correlations are equal to one (Delfiner and Barrier, 2004);
- Discretization: A discrete value is assumed for each qualitative level of correlation between projects. For example, "weak – medium – strong" correlations are associated with "0.1 – 0.3 – 0.5" correlation coefficients (van Elk *et al*., 2000);
- Tornado diagrams analysis (van Elk *et al*., 2008).

Reserves Distribution Estimation Methods

In order to perform a probabilistic aggregation, all field resources have to be modeled by probabilistic distributions. A possible approach is to apply the multivariate normal distribution analytical method. In this method the distributions of the field reserves are normal, thus the aggregated project reserves is also normal. According to van Elk *et al*. (2008), this represents a disadvantage, once all the distributions have to be symmetrical. The authors state that for fields in the early stage of appraisal and for undiscovered prospects symmetry of the resources distribution is not often a good approximation. Their work introduces a method to model the distributions' skewness through a shape parameter. When this parameter is zero, the resulting distribution is the standard normal distribution. When this parameter reaches very large values, the distribution converges to the half-normal distribution. The half-normal distribution is obtained by left-truncating below zero a standard normal distribution with mean zero (Kemp, 2006). It is noteworthy that the skewness of the half-normal distribution with mean zero is approximately one. For the probabilistic aggregation, the authors use the multivariate form of the skew-normal distribution.

Another approach is to model the reserves curve by a Lognormal distribution. As stated by Delfiner and Barrier (2004), the uncertainty on the resources of a single petroleum field is represented by a Lognormal distribution. In addition, Etherington *et al*. (2001) state that undiscovered and newly discovered exploitation projects tend to have Lognormal distributions heavily skewed right. This means that P90 volumes are small relative to P50 and the P10 upside. In accordance, Demirmen (2007) states that the reserves probability distribution approaches Lognormal. Finally, Gair (2003) states that it is expected the distribution of hydrocarbon discoveries to follow a Lognormal distribution pattern and that the uncertainty in estimating reserves volumes must also be lognormally distributed.

There are other methods and methodologies available in the literature; however they eventually present two major limitations, which are:

1. In some cases the methods for obtaining a project's reserves probability distribution are complex and normally they require an intensive participation of experts, who are expensive and limited resources. In other cases the estimation methods are not robust enough to model particular conditions.
2. The correlation coefficients assessment is either complex or superficial, thus overestimation and underestimation are constantly a present issue.

The proposed methodology focuses on presenting significant advances to overcome these limitations.

The Proposed Methodology

Performing an economic assessment in an entire cumulative production (Np) distribution curve is not feasible, once a single economic scenario cannot describe the reality of each single point of the curve. Moreover, it is not possible to appraise each point individually, since it would require an infinite number of economic scenarios. The best way to provide a probability reserves curve is to assess economically some cumulative percentiles of the probabilistic Np curve and then rebuild it, thus providing an estimate of the probability distribution of reserves.

The proposed methodology is based on the curve assembling technique above-mentioned and it encompasses four major steps. As shown in Figure 1, the basic inputs for the methodology are the probabilistic production curves at the project level.

The first step generates a list of the eligible projects for a probabilistic assessment and the second step performs an economic assessment in order to calculate three discrete cumulative reserves percentiles for each project (P90, P50, and P10). It is important to clarify that the scope of this study does not go deep into economic premises and concepts, and that the idea of this step is to give a general overview of the economic screening.

Emphasis is given to the third and fourth steps. The former consists of defining the projects' probability reserves curves and the latter aggregates these reserves curves. We consider these steps as the great contributions of the methodology.

Step 1 – Selecting the projects for probabilistic assessment

The exploitation areas of an oil company generally hold a great number of projects. Nevertheless, not all of them may be qualified for a probabilistic assessment due to three requirements:

1. There must be significant uncertainty associated with geoengineering parameters and the oil and gas recovery factor;
2. The production forecasts have to consider the correct premises regarding conventional and unconventional resources. Besides, the curves must be generated by probabilistic methods, i.e. different uncertainty-based production curves for each project ought to be built. These production curves have to be associated with a probability of occurrence, therefore representing a cumulative percentile of the Np curve. A thorough uncertainty analysis must be carried out regarding the key uncertainties of a project, which can be reservoir fluid gravity, porosity, permeability, net to gross ratio, and so forth. Commercial softwares and/or uncertainty-based methods are able to provide the corresponding P90, P50, and P10 production curves;
3. The project must be considered relevant to the company.

Besides, after defining the eligible projects, the methodology suggests applying the Pareto principle in order to choose the most representative ventures. The representativeness is calculated by an *uncertainty versus EUR* quantification, which can be expressed by the sample variance of the EUR for each project. Table 1 illustrates an example.

Table 1: *Uncertainty versus EUR* Quantification

Project	EUR (10^6 boe) low estimate scenario	EUR (10^6 boe) most probable scenario	EUR (10^6 boe) high estimate scenario	Sample Variance (Volume)^2
I	10.0	12.0	18.0	17.3
II	5.3	6.8	8.8	3.1
III	3.3	4.5	6.2	2.1
IV	6.7	8.0	9.4	1.8
V	10.7	14.1	18.1	13.7
VI	3.0	3.2	6.0	2.8
VII	8.7	12.3	15.9	13.0
VIII	2.0	2.4	4.9	2.5
IX	5.0	6.4	7.8	2.0
X	9.8	11.2	13.4	3.3

As a rule of thumb, considering total independency, 30% of the projects (I, V, and VII) represent 71% of the total variance among the projects, therefore these projects shall be assessed probabilistically and the others shall be considered deterministic. Nonetheless, all projects could be classified as probabilistic. This assessment may only be necessary if there is a need to prioritize some particular ventures, especially when computational effort is a restriction.

The projects which have not been selected shall contribute to the overall reserves by deterministic summation at the end of the process.

Step 2 – Probabilistic economic assessment

As previously mentioned in this paper, the focus of this step is to simply point out the basic data treatments needed for a probabilistic economic screening.

Production curves ought to be built under uncertainty analysis resulting in the P90, P50, and P10 curves. Then, the revenue for each scenario is calculated by multiplying the curves by the hydrocarbon price, which can be modeled by any probability distribution. Past data should be analyzed in order to give information about the historical uncertainty and the experts' opinion should also be taken into account (Kuhl *et al.*, 2009). However, in many cases it can be difficult to determine the correct probability distribution for an input. In these situations it is suggested that either the Pert or the Triangular distributions should be chosen. Both only need three relatively easy-to-determine parameters, which are worse, most probable, and optimistic scenarios. The other cash flow inputs should also be modeled by probability distributions.

The next step is to determine the correlations among the cash flow inputs. Once again, a historical data analysis should be carried out and the experts' opinion should be taken into account for this procedure.

After the inputs and their correlations have been modeled, a Latin Hypercube simulation has to be performed so as to obtain the yearly cash flow curves. Furthermore, a thorough analysis of these resulting cash flow curves must be carried out, so that the economic limit can be determined for each production scenario. This piece of information can be calculated either by the Net Present Value (NPV) method or another method that may be more suitable for one's reality. Then, the projects' economic volumes are classified as P90, P50, and P10 reserves.

Step 3 – Estimating the probabilistic reserves curve for each project

As the reserves for each scenario have been calculated, a project reserves distribution curve can be estimated. Each production scenario contributes with a discrete cumulative percentile of the reserves curve, e.g. the P90 production curve results in the proved reserves, which provides the P90 cumulative reserves percentile. Figure 3 shows an example of the cumulative percentiles plotted on a cumulative descending chart.

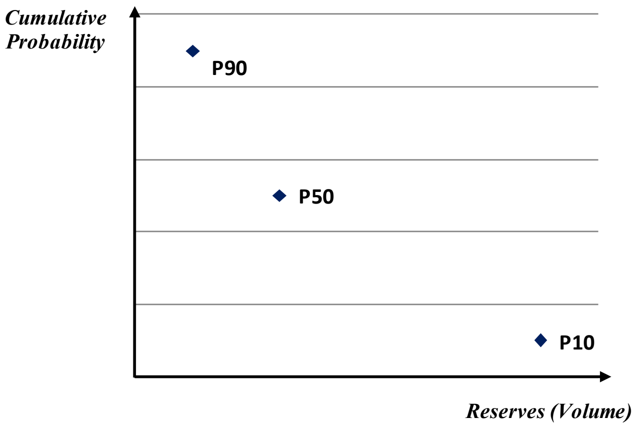

Figure 3: Cumulative percentiles of a single project

Henceforth, we present a procedure to estimate the reserves probability distribution of a project. The distribution model varies depending on the difference between the cumulative percentiles, generating three different conditions.

Condition 1: (P10 – P50) > (P50 – P90)

This is the expected behavior of a reserves curve, once it presents a positive skewness. For this condition we suggest applying the Lognormal distribution, as defended by several authors aforementioned. The mean and the standard deviation are sufficient to define a Lognormal distribution; nonetheless we got three pieces of information: P90, P50, and P10. In order to impose a third parameter to a Lognormal distribution, a location factor is added to the model. Commercial softwares easily incorporate this third parameter and build a Lognormal distribution directly with the P90, P50, and P10 cumulative percentiles.

Condition 2: (P10 – P50) < (P50 – P90)

This result may occur when the premises and concepts of reservoir engineering are supplanted by the influence of economics and other restrictions, which generate a negative skewness. A Lognormal distribution cannot be fitted to this condition. This restriction is demonstrated in Appendix I.

According to Kuhl *et al*. (2009), the identification of the best distribution family for stochastic simulation input models comprises several procedures. These procedures include from informal graphical techniques up to statistical goodness-of-fit tests, such as the Kolmogorov-Smirnov, Chi-Squared, and Anderson-Darling tests. Besides, experts' opinions have to be taken into account in some cases so as to identify the appropriate distribution family.

Gullco and Anderson (2007) presented a study of the application of the Beta distribution to petrophysics. The authors argue that one of the reasons for the use of this distribution is that it allows the calculation of parameters from the data, without having to introduce them arbitrarily. In addition, numerical characteristics and empirical observations justify the introduction of this distribution.

In our study we arrived at two stochastic models for Condition 2, which are the Beta General distribution and a model derived from the concept that the difference between reserves volumes shall present an exponential pattern with respect to the cumulative probability of recovery, even for negative skewness values. This behavior has been confirmed through past data studies. The first model was validated under a set of empirical and statistical analyses, as suggested by Kuhl *et al.* (2009), and it is presented in Appendix II. The latter was semi-analytically determined and named after *Hyperbolic Reserves Distribution* (HRD).

Hyperbolic Reserves Distribution (HRD)

Equation 6 is a first-order linear ordinary differential equation which defines the expected behavior aforementioned. Through this equation we modeled the probability distribution.

$$\frac{dx}{dy} = C\lambda^{y} \tag{6}$$

The parameter x represents the cumulative reserves, y is the cumulative probability of recovery, and λ and C are constants. This differential equation can be easily solved, and by rearranging the terms we can build the distribution. The PDF and the CDF are calculated and they are shown in equations 7 and 8, respectively. The parameter φ is a constant of integration.

$$PDF = \frac{1}{(x-\varphi)\ln(\lambda)} \tag{7}$$

$$CDF = \frac{\ln\left(\left|\frac{x-\varphi}{\alpha}\right|\right)}{\ln(\lambda)} \tag{8}$$

$$\alpha = \frac{C}{\ln(\lambda)} \tag{9}$$

$$\alpha \in \mathrm{R}^{*};\ \varphi \in \mathrm{R};\ \lambda \in \mathrm{R}_{+}^{*} - \{1\} \tag{10}$$

The distribution's name is derived from the PDF's hyperbolic shape. The mean μ and the standard deviation σ are given by, respectively:

$$\mu = \frac{1}{\ln\lambda}\left[\alpha(\lambda-1) + \varphi\ln\lambda\right] \tag{11}$$

$$\sigma = \frac{\alpha}{\sqrt{2}\ln\lambda}\sqrt{(\lambda^2-1)\ln\lambda - (2\lambda^2 - 4\lambda + 2)} \tag{12}$$

The HRD distribution can be entirely defined with three parameters. These parameters can be written in terms of the P90, P50, and P10 cumulative percentiles, as shown in equations 13, 14, and 15.

$$\varphi = \frac{P_{90}P_{10} - {P_{50}}^{2}}{P_{90} + P_{10} - 2P_{50}} \tag{13}$$

$$\lambda = \left(\frac{P_{90}P_{50} + P_{10}P_{50} - P_{50}^{\ 2} - P_{90}P_{10}}{P_{90}^{\ 2} - 2P_{50}P_{90} + P_{50}^{\ 2}} \right)^{5/2} = \sqrt{\left(\frac{P_{50} - \varphi}{P_{90} - \varphi} \right)^5} \tag{14}$$

$$\alpha = \frac{\left[P_{90} - \left(\frac{P_{90}P_{10} - P_{50}^{\ 2}}{P_{90} + P_{10} - 2P_{50}} \right) \right]^{5/4}}{\left[P_{50} - \left(\frac{P_{90}P_{10} - P_{50}^{\ 2}}{P_{90} + P_{10} - 2P_{50}} \right) \right]^{1/4}} = \left[\frac{(P_{90} - \varphi)^5}{(P_{50} - \varphi)} \right]^{1/4} = \frac{P_{90} - \varphi}{\lambda^{0.1}} \tag{15}$$

Lower and upper limits are:

$$LL = \alpha + \varphi \tag{16}$$

$$UL = \alpha\lambda + \varphi \tag{17}$$

The skewness of the HRD function can be either negative or positive, but not zero. As for Condition 2, it is always negative.

The distribution may return negative values when P50 is relatively high, i.e. P50 is closed to P10 and distant from P90. By definition, the CDF is a monotone non-decreasing function, so in order to guarantee positive values for reserves estimates, it is sufficient to impose the following restriction on the lowest value of the function:

$$x \geq 0, \forall y \in [0,1] \tag{18}$$

$$\alpha\lambda^0 + \varphi \geq 0 \tag{19}$$

The next test synthesizes the restriction.

$$\alpha \geq -\varphi \tag{20}$$

If the test above is not satisfied, the HRD must be left-truncated at zero, which can be easily done with commercial softwares.

The application of the HRD depends on cumulative function builders, which are tools commonly provided in commercial softwares. At least one hundred points should be considered to model the distribution.

Empirical tests show that the Beta distribution is more conservative than the HRD. Hence, for low risk projects we recommend the application of the latter. The concept of "low risk" can be vague and depends on one's experience; nevertheless projects associated with conventional reservoirs with good analogies can be considered as low risk ventures when compared to unconventional resource plays with poor analogies, lack of quality data, and so forth.

Condition 3: (P10 – P50) = (P50 – P90)

The reasons for the occurrence of this condition are the same listed for Condition 2. Due to this equality, the resulting distribution has to be symmetrical. Thence, we propose the use of either the Normal or the Beta general distribution. Empirical tests show that the Normal distribution is more conservative, as a result we recommend the use of the Beta distribution for low risk projects.

Step 4 – Performing the probabilistic reserves aggregation

As previously mentioned in the literature review, the probabilistic aggregation can be applied to calculate reserves. The main reason for doing so is that the deterministic addition tends to underestimate proved reserves, once it considers total dependency among the ventures. So as to perform the probabilistic addition, we suggest the Latin Hypercube sampling method due to its better accuracy in comparison with Monte Carlo.

In order to avoid an overestimation of proved reserves, the correlation coefficients among the projects have to be identified and quantified. Yet, the direct determination of the correlation coefficients may be a difficult procedure. Based on the method developed by Saaty (1980) – Analytic Hierarchy Process, our methodology suggests the decomposition of the general goal correlation in a hierarchy. Experts determine which factors most influence the dependencies and instead of comparing two projects directly, they compare each of these factors of the projects individually, and later consolidate the comparisons. These factors can be either quantitative or qualitative; hence evaluation can be carried out based on system data and/or on experts' opinions.

Quantitative factors shall address both reservoir engineering and development features. Moreover, it is recommended that experts validate the resulting correlation coefficients, even if only quantitative parameters are used.

The methodology accepts correlation coefficients within the interval [0, 1]. The value 1 means total correlation, which would resemble the deterministic aggregation. The value 0 represents complete independence among projects.

This is just a general overview of the proposed correlation method. It will be fully described in a forthcoming paper.

Case Study

This case study focuses on steps three and four of the methodology, which are considered to be the great contributions of this study. So as to illustrate the procedures, a final example will be modeled in a spreadsheet software jointly with a commercial Monte Carlo Simulation software. It consists of the probabilistic aggregation of ten projects plus the addition of other twenty deterministic projects. All thirty projects belong to a single exploitation field and it is noteworthy that the values' representativeness and complexity have been fully retained, albeit they have been adapted for the example.

Reservoir engineers have built three production scenarios (low, best, and high estimate) for the deterministic projects and the P90, P50, and P10 production curves for the probabilistic projects.

The deterministic projects have been economically assessed with no stochastic inputs. The projects' 1P, 2P, and 3P reserves volumes have been determined and their contribution will be added at the end of the process.

All probabilistic projects have been economically assessed in accordance with the proposed methodology, i.e. the inputs' distribution probabilities and their correlation coefficients have been determined by a group of experts. Then, a Latin Hypercube Simulation has been performed and, by assessing the cash flow curves, the economic limits have been established. Thus, the P90, P50, and P10 reserves volumes for each project have been determined.

Now, focusing on steps three and four of the methodology, the reserves distributions for each project will be modeled, correlated, and aggregated.

The ventures have been categorized according to the distribution estimate procedure previously described and the results are listed in Table 2.

Table 2: Projects' Reserves Probability Distributions (10^6 boe)

Project	P90	P50	P10	Category (Condition)	Distribution
A	4.56	6.02	10.02	1	Lognormal
B	5.52	9.36	11.09	2	Beta
C	2.78	4.87	7.94	1	Lognormal
D	9.99	11.40	15.52	1	Lognormal
E	3.42	6.02	6.87	2	HRD
F	3.74	5.84	9.05	1	Lognormal
G	3.59	5.32	9.26	1	Lognormal
H	2.84	4.80	5.28	2	HRD
I	8.58	11.07	13.06	2	Beta
J	2.97	4.59	8.02	1	Lognormal

Table 2 provides the necessary information to assemble the reserves curve for each probabilistic project. The next step is to determine the projects' correlation coefficients. It has been applied a hierarchical decomposition regarding reservoir engineering and development features. At the end of the process experts have endorsed the final results, which are shown in Table 3.

Table 3: Correlation coefficients

	A	B	C	D	E	F	G	H	I	J
A	1									
B	0.65	1								
C	0.82	0.64	1							
D	0.55	0.50	0.55	1						
E	0.61	0.63	0.63	0.67	1					
F	0.63	0.62	0.66	0.74	0.61	1				
G	0.55	0.65	0.71	0.67	0.52	0.60	1			
H	0.64	0.64	0.65	0.66	0.67	0.62	0.73	1		
I	0.57	0.63	0.67	0.68	0.64	0.59	0.56	0.65	1	
J	0.58	0.69	0.69	0.72	0.63	0.59	0.51	0.62	0.71	1

Finally, a Latin Hypercube simulation can be performed. The resulting aggregated reserves curve is shown in Figure 4.

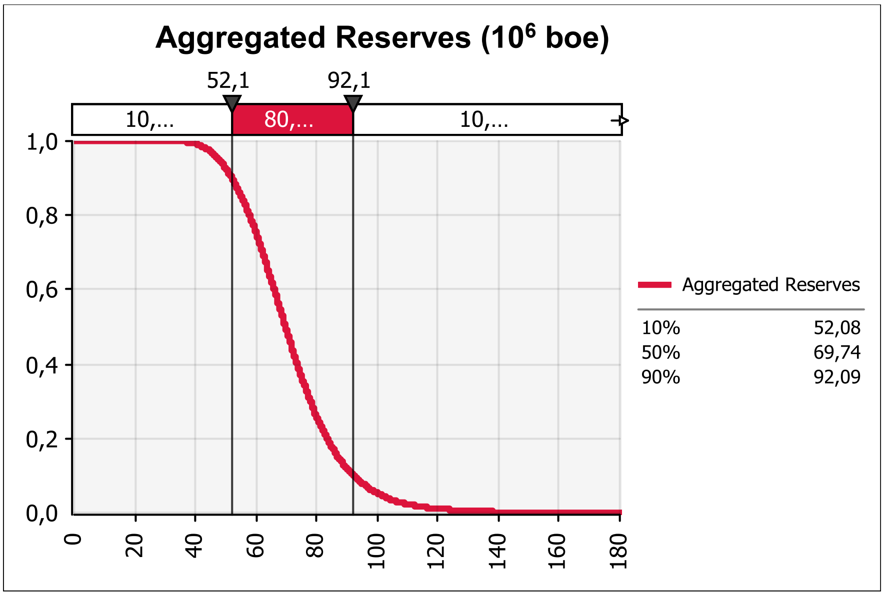

Figure 4: Aggregated Reserves Curve

The deterministic summation of the proved reserves presented in Table 2 is 47.99 10^6 boe. As for the probabilistic approach, proved (P90) reserves are 52.08 10^6 boe. As expected, the deterministic approach understates proved reserves, once it corresponds to the 95% confidence level of the probabilistic addition, i.e. it represents the P95 reserves. The proved reserves resulting from probabilistic addition are 8.52% greater than that resulting from arithmetic summation.

As mentioned before, an overestimation may occur if correlations among projects are not identified. In order to illustrate this situation, the same model presented in the case study has been performed not considering the correlation coefficients in Table 3, i.e. the projects have been considered fully independent. The resulting proved reserves are 63.14 10^6 boe, representing a 31.57% gain over the deterministic summation, which is clearly an overestimation.

The portfolio effect is an outcome of the probabilistic addition and it is notable to state that volumes have not been created, once P90, P50, and P10 are simply shifted towards the mean of the curve. Table 4 shows the absolute variation on each cumulative percentile.

Table 4: Reserves Variation due to the Portfolio Effect (10^6 boe)

Reserves Volumes	Probabilistic Summation	Deterministic Summation	Variation
P90	52.08	47.99	4.09
P50	69.74	69.29	0.45
P10	92.09	96.11	-4.02
Total			0.52

If we had a deterministic scenario for each point of the aggregated probabilistic curve, the total difference would be zero. The 0.52 simply represents the variation of the not considered cumulative percentiles.

In order to calculate the reserves for the entire field, the contribution of the deterministic projects has to be added to the probabilistic reserves. Table 5 shows this summation.

Table 5: Field Reserves (10^6 boe)

Reserves Volumes	Probabilistic Projects	Deterministic Projects	Field Reserves
Proved	52.08	24.56	76.64
Proved plus Probable	69.74	43.18	112.92
Total	92.09	76.45	168.54

Conclusions

This paper provides an auditable, repeatable, and practical usable methodology to aggregate reserves probabilistically. The main advantage of the probabilistic approach is to more accurately assess the proved (P90) reserves, once the arithmetic addition understates this volume. The case study confirmed this underestimation, since the deterministic summation resulted in the P95 reserves. It is noteworthy that the methodology is in accordance with the PRMS (2007) guidelines.

Different projects may share similarities; hence the definition and quantification of the correlations among them cannot be bypassed. If projects are considered fully independent, as a consequence there will be an overestimation of proved reserves and an underestimation of total reserves.

The methodology can be applied to both conventional and unconventional resources, as long as the production forecasts have been built under uncertainty analyses, the correct premises have been taken into account and an appropriate forecasting method has been applied.

This methodology was successfully applied to the probabilistic evaluation and aggregation of ten probabilistic projects, and other twenty deterministic projects. In the case study discussed, proved reserves resulting from the correlation-based probabilistic assessment were 8.52% greater than the deterministic assessment, which is in accordance with the study of Carter and Morales (1998). Furthermore, it has been shown the importance of a systematic appraisal of the dependencies, once the independent aggregation resulted in a 31.57% overestimation in proved reserves when compared to the deterministic assessment.

A system has been developed to support the methodology calculations. The methodology is being tested in several fields and results are being compared with deterministic methods.

Acknowledgements

The authors thank Petróleo Brasileiro S.A. (PETROBRAS) for both assistance and permission for publishing this paper. Special thanks to PUC-Rio for the support on the academic research.

References

Arps, J.J.: “Analysis of Decline Curves”, 1945, SPE 945228-G

Canadian Oil and Gas Evaluation Handbook – Volume 1 prepared by Society of Petroleum Evaluation Engineers and Canadian Institute of Mining, Metallurgy & Petroleum – First edition June 30, 2002

Carter, P.J., Morales, E.: “Probabilistic Addition of Gas Reserves within a Major Gas Project”, 1998, SPE 50113

Chan, P., Etherington, J.R., Aguilera, R.: “A Process to Evaluate Unconventional Resources”, 2010, SPE 134602

Delfiner, P., Barrier, R.: “Partial Probabilistic Addition: A Practical Approach for Aggregating Gas Resources”, 2004, SPE 90129

Demirmen, F.: “Reserves Estimation: the Challenge for the Industry”, 2007, SPE 103434

Duong, A. N.: “Rate-Decline Analysis for Fracture-Dominated Shale Reservoirs”, 2011, SPE 137748

Elliot, D. C.: “The Evaluation, Classification and Reporting of Unconventional Resources”, 2008, SPE 114160

Etherington, J.R., Hunt, E.J., Adewusi, A.: “Aggregating Reserves and Resources for Portfolio Management”, 2001, SPE 71424

Etherington, J.R., McDonald, I. R.: “Is Bitumen a Petroleum Reserve?”, 2004, SPE 90242

Gair, R.: “Integrating Probabilistic and Deterministic Reserves”, 2003, SPE 82000

Gullco, R., Anderson M.: “The Beta Distribution and Its Application to Petrophysics”, 2007, SPE 106746

Holditch, S.A.: “The increasing Role of Unconventional Reservoirs in the Future of the Oil and Gas Business”, SPE Journal of Petroleum Technology, November 2002 pages 34-38

Ilk, D., Rushing, J.A., Perego, A.D. and Blasingame, T.A.: “Exponential vs. Hyperbolic Decline in Tight Gas Sands: Understanding the Origin and Implications for Reserve Estimates Using Arps' Decline Curves”, 2008, SPE 116731

Kemp, A. W.: “The Discrete Half-Normal Distribution”, 2006, International Conference on Mathematical and Statistical Modeling

Kuhl, M. E., Ivy, J., Lada, E. K., S., Steiger, N. M., Wagner, M. A., Wilson, J. R.: “Introduction to Modeling and Generating Probabilistic Input Processes for Simulation”, 2009, Proceedings of the 2009 Winter Simulation Conference

Lee, W.J., and Sidle, R.E.: “Gas Reserves Estimation in Resource Plays”, 2010, SPE 130102

Markowitz, H.: “Portfolio Selection”, 1952, The Journal of Finance, Vol. VII, n.1

Ross, J.G.: “Petroleum Resources Classification and Definitions. Guidelines for the Evaluation of Petroleum Reserves and Resources - A Supplement to the SPE/WPC Petroleum Reserves Definitions and the SPE/WPC/AAPG Petroleum Resources Definitions, 2001.

Saaty, T. L.: “The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation”, 1980, ISBN 0-07-054371-2, McGraw-Hill

Saliby, E., Moreira F.F.P.: “An Empirical Evaluation of Sampling Methods in Risk Analysis Simulation: Quasi-Monte Carlo, Descriptive Sampling and Latin Hypercube Sampling”, 2002, Proceedings of the 2002 Winter Simulation Conference

SPE/WPC/AAPG/SPEE Petroleum Resources Management System, 2007.

Valkó, P.P.: “Assigning value to stimulation in the Barnett Shale: a simultaneous analysis of 7000 plus production hystories and well completion records”, 2009, SPE 119369

Van Elk, J.F., Vijayan, D.K., Gupta, D.R.: “Probabilistic Addition of Reserves: A New Approach, 2000, SPE 64454

Van Elk, J.F., Gupta, D.R., Wann, D.: “Probabilistic Aggregation of Oil and Gas Field Resource Estimates and Project Portfolio Analysis”, 2008, SPE 116395

World Energy Council – 2010 Survey of Energy Resources, 2010, ISBN: 978 0 946121 021

Appendix I

Lognormal restriction: (P10 – P50) > (P50 – P90)

A Lognormal distribution is a probability distribution of a random variable whose logarithm is normally distributed. The P90, P50, and P10 cumulative percentiles are given by:

$$P_{90} = e^{\mu - \left[\sqrt{2}\ erf^{-1}\left(\frac{4}{5}\right)\sigma\right]} \tag{21}$$

$$P_{50} = e^{\mu} \tag{22}$$

$$P_{10} = e^{\mu + \left[\sqrt{2}\ erf^{-1}\left(\frac{4}{5}\right)\sigma\right]} \tag{23}$$

The parameters μ and σ are the mean and standard deviation, respectively, of the variable's natural logarithm and erf^{-1} is the inverse error function. The intention is to prove the following inequation:

$$P_{10} - P_{50} > P_{50} - P_{90} \tag{24}$$

The following notations will be adopted to simplify the demonstration:

$$t = \sqrt{2}\ erf^{-1}\left(\frac{4}{5}\right) \tag{25}$$

$$P_{50} - P_{90} = K_1 \tag{26}$$

$$P_{10} - P_{50} = K_2 \tag{27}$$

Thus, we can write that:

$$K_1 > 0;\ K_2 > 0 \tag{28}$$

$$K_1 = e^{\mu}(1 - e^{-t\sigma}) \tag{29}$$

$$K_2 = e^{\mu}(e^{t\sigma} - 1) \tag{30}$$

$$\frac{K_2}{K_1} = \frac{(e^{t\sigma} - 1)}{(1 - e^{-t\sigma})} = \frac{(e^{t\sigma} - 1)}{e^{-t\sigma}(e^{t\sigma} - 1)} \tag{31}$$

$$\frac{K_2}{K_1} = e^{t\sigma} = (e^{t})^{\sigma} \tag{32}$$

$$e^{t} \cong 3{,}60222 \tag{33}$$

By definition $\sigma > 0$. So, we conclude that:

$$e^{t\sigma} > 1 \tag{34}$$

$$\frac{K_2}{K_1} > 1 \tag{35}$$

As a result, we can state for a Lognormal distribution that:

$$P_{10} - P_{50} > P_{50} - P_{90} \tag{36}$$

Appendix II

The Beta General Distribution

The Beta general is directly derived from the Beta distribution by scaling the [0, 1] range with the use of a minimum and maximum value to define a new range. This distribution satisfies Condition 2 because it admits negative skewness values. The distribution's probability density function is given by:

$$PDF = \frac{1}{B(\alpha,\beta)}\frac{(x-a)^{\alpha-1}(b-x)^{\beta-1}}{(b-a)^{\alpha+\beta-1}} \tag{37}$$

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{38}$$

$$\alpha > 0; \ \ \beta > 0 \tag{39}$$

$B(\alpha,\beta)$ is the Beta function, $\Gamma(\alpha)$ is the Gamma function, and α and β are shape parameters. The lower and the upper limits are a and b, respectively.

The mean μ and the standard deviation σ are shown in equations 40 and 41, respectively.

$$\mu = \frac{\alpha b + \beta a}{\alpha + \beta} \tag{40}$$

$$\sigma = \frac{(b-a)}{(\alpha+\beta)}\sqrt{\frac{\alpha\beta}{(\alpha+\beta+1)}} \tag{41}$$

This distribution requires four parameters, thence we recommend the P90, P50, and P10 cumulative percentiles and the lower limit to serve as inputs. As rule of thumb, the lower limit is considered to be zero. It is important to note that commercial softwares accept these parameters as inputs for this distribution.