Proceedings World Geothermal Congress 2010
Bali, Indonesia, 25-29 April 2010

Probabilistic Resource Estimation of Stored and Recoverable Thermal Energy for Geothermal Systems by Volumetric Methods

Mustafa Onur, Hülya Sarak, and Ö. İnanç Türeyen

Department of Petroleum and Natural Gas Engineering, Istanbul Tech. U., Maslak, 34469, Istanbul, Turkey

onur@itu.edu.tr, hulya@itu.edu.tr, and inanct@itu.edu.tr

Keywords: Resource Estimation, Volumetric Methods, Probabilistic, Uncertainty Characterization, Adding Field Resources

ABSTRACT

This paper investigates the propagation of uncertainties in the input variables (used in the volumetric method) on to stored and recoverable thermal energy resources calculated from volumetric methods. Both Monte Carlo (MC) and the analytic uncertainty propagation (AUP) methods are considered and compared for uncertainty characterization. Analytic uncertainty propagation equations (AUPEs) are derived based on a Taylor-series expansion around the mean values of the input variables. The AUPEs are general in that correlation among the input variables, if it exists, can also be accounted for on the resulting uncertainty. Monte Carlo methods (MCMs) were used to verify the results obtained from the AUPEs.

A comparative study that we have conducted shows that the AUPM is as accurate as the MCM for the problem of interest. Hence, it can be used as a fast tool eliminating the need for MCM because the resulting distributions of thermal resources always tend to be log-normal. We also discuss the problem of aggregating thermal resources for projects involving more than one field. Should one use an arithmetic or probabilistic addition to determine “proved,” “probable,” and “possible” (which corresponds to P10, P50, and P90 percentiles of the cumulative distribution function, respectively) corresponding a geothermal project or a country which may involve many diverse fields, each with its estimated values of P10, P50, and P90? We show that the arithmetic sum underestimates the P10 value, compared to probabilistic sum which is the statistically proper method for adding resources or reserves. Applications on synthetic and field data cases are presented to demonstrate the methodology considered in this work.

1. INTRODUCTION

Uncertainty is inherent in estimation of any type of resources (oil, gas or heat) from underground energy systems. The thermal energy or power resource (or “reserve”[†]) of a given geothermal reservoir is no exception. Unfortunately, this is also true regardless of any method used for estimation, e.g., volumetric, decline curve, or reservoir simulation methods because the input variables required for the resource estimation problem always contain uncertainties to some degree that propagate into resource estimates. Therefore, to make good decisions, one must be able to accurately assess and manage the uncertainties and risks.

In this work, we limit our study to the assessment of uncertainty in estimated thermal energy resource (in the form of stored or recoverable) by the volumetric methods (for example, see Muffler and Cataldi, 1978).

Volumetric methods are usually used to estimate stored heat and recoverable power resources in the early life of geothermal reservoirs. Estimation of the thermal energy requires geological, geophysical, and petro-physical data including reservoir temperature, reservoir area, thickness, porosity, rock and fluid specific heats, etc. The values of these input variables have usually large uncertainties associated with them, and hence it is very important to propagate these uncertainties on to the estimation of the thermal energy reserves. From the view point of a field investment, an accurate assessment of uncertainty in stored and recoverable heat is a crucial task from which to make decisions that will create value and/or mitigate loss in value (risk).

In the past, various authors have considered assessing uncertainty in estimated stored and recoverable power from the volumetric method by using the Monte Carlo (MC) method (Brook et al., 1979; Serpen, 2001; Lovekin, 2004; Sanyal and Sarmiento, 2005, Arkan and Parlaktuna, 2005; Serpen et al., 2008). However, none of these works provide a deep investigation of and insight into the uncertainty assessment from the volumetric method. These works simply apply the MCM to characterize uncertainty in estimated stored heat or recoverable power for the geothermal reservoirs interest.

It is no doubt that the MCM used in the previous references cited is a general approach for assessing the uncertainty. In this work, however, we show that there is a simple and fast alternative method – which we refer to it as the analytic uncertainty propagation method (AUPM) – to the MCM method for characterizing uncertainty. The validity of the AUPM for accurately characterizing uncertainty results from the fact the distributions of stored heat and power for a zone, well, or field to be computed from the volumetric method always tend to be log-normal. This result simply follows from the fundamental theorem of statistics and probability – the Central Limit Theorem (CLT) (e.g., see Parzen, 1962). The CLT states that the sum of a sufficiently large number of independent random variables each with finite mean and variance will be approximately normally distributed. As a consequence of this theorem and the functional relationship of the volumetric method which involves a product/quotient of several independent random variables for computing thermal energy reserves, one should expect that the resulting distribution of the thermal energy is to be nearly log-normal as the number of input random variables increases. This result is in fact valid no matter what form of uncertainty the input variables assume. The same findings have been reported previously for the

[†] We use the word “resources” rather than “reserves” when commerciality is not proven (SPE, 2006).

assessment of uncertainty in oil or gas reserves computed from the volumetric methods by Capen (1996, 2001).

Other objectives of this work are as follows: We would like to show how the different types of input distributions and correlation among input variables propagate into the uncertainty of the estimated thermal energy or power. Furthermore, we discuss the problem of aggregating individual fields' resources for the geothermal projects involving many diverse fields or a country's total resources. It is shown that the arithmetic addition underestimates the P10 that would be obtained by the probabilistic addition, which is the statistically proper method for adding resources or reserves) unless all fields considered in aggregation are fully (or perfectly) correlated with each other.

2. VOLUMETRIC METHOD

Here, we introduce the volumetric method used to compute the stored heat and recoverable thermal energy (in terms of power) for a given geothermal field. Throughout, we will assume a geothermal reservoir containing hot solid rock and single-phase liquid water. For two-phase systems, appropriate equations are given by Sanyal and Sarmiento (2005).

Based on the volumetric method, we can estimate recoverable power (in MW) by the following equation (Serpen, 2001; Sanyal and Sarmiento, 2005; Arkan and Parlaktuna, 2005, Sarak *et al.*, 2009):

$$PW = \frac{H_t R_F Y}{10^3 L_F t_p}. \qquad (1)$$

where H_t is the total stored heat and is usually referred to as the accessible resource base and is estimated from Eq. 2:

$$H_t = \left[\left(1-\phi\right) c_s \rho_s + \phi c_w \rho_w\right] Ah\left(T - T_r\right). \qquad (2)$$

In Eq. 2, the subscript *r*, *s*, *t*, and *w* denote reference, solid rock, total, and water, respectively. The variables used in Eqs. 1 and 3 and their units are given below:

H_t total stored heat in kJ (see Eq. 2)

L_F load factor (fraction)

PW recoverable power in MW

R_F recovery factor (fraction)

Y transformation yield (fraction)

t_p project life in seconds (s.)

A area in m^2

c specific heat in kJ/(kg ^{o}C)

h net thickness in m

H stored heat in kJ

T temperature in ^{o}C

ϕ porosity (fraction)

ρ density in kg/m^3

In Eq. 1, L_F, the load factor, represents the fraction of the total time in which the direct heating or power generation is in operation, and *Y*, the transformation yield, represents the efficiency of transferring thermal energy from the geothermal fluid to a secondary fluid.

It is also worth noting that Eqs. 1 and 2 are valid for predicting the recoverable power for both direct- heating and power (electricity) generation applications from liquid-dominated systems from which only liquid is produced, provided that the variables such as T, T_r, R_F, Y, L_F, and t_p are "adequately" chosen for the specific application of interest. For example, for most of the direct-heating applications, T_r usually ranges from 15 to 60 ^{o}C, whereas for electricity generation, T_r usually ranges between 70 to 100 ^{o}C. Although it is important how to choose the variables and their ranges for accurately assessing uncertainty in *PW* (Eq. 1) and H_t (Eq. 2), our purpose here is not to get into a detailed discussion of appropriate sources of data and how one could determine the appropriate values of the variables and the associated uncertainties. Assessment of uncertainties in the input variables itself is a notoriously difficult problem, because, to our knowledge, there is no standard rule for characterizing uncertainty in the input variables. On this aspect, we refer the readers to the works of Capen (1976) and Welsh *et al.* (2007).

Uncertainty in the recoverable power results from our lack of knowledge in most of the input variables in Eqs. 1 and 2. Quantification of uncertainty is inevitably subjective because knowledge about the input variables is dependent on available data and personal experience of the interpreter. As well stated by Welsh *et al.* (2007), it is quite possible for two people to have different probability estimates for the same input variable, based on their differing knowledge. Thus, there is no single "correct" probability distribution, unless all people have identical experience and information, and process it in the same way (Welsh *et al.*, 2007).

Based on the discussion given in the previous paragraph, there is no reason to claim that any particular type of probability distribution (e.g., uniform, normal, log-normal, triangular, etc.) for our input variables be preferable. As we will show later based on the CLT, the resulting distributions of *PW* (or H_t) are almost log-normal regardless of the types of probability distributions chosen for the input variables.

When computing *PW* from Eq. 1, we can treat *PW* in general as a function of eleven random input variables; A, h, ϕ, c_s, c_w, ρ_s, ρ_w, $(T\text{-}T_r)$, R_F, L_F, and Y. In our applications, we fix T_r at a constant value, but the variable $(T\text{-}T_r)$ in Eq. 2 will be treated as a random variable because T in Eq. 2 is treated as a random variable. If the mean and variance of T are μ_T and σ_T^2, then the mean and variance of the $(T\text{-}T_r)$ are $\mu_{(T-T_r)} = \mu_T - T_r$ and σ_T^2, respectively. Furthermore, we assume that there is no uncertainty associated with the variable t_p in Eq. 1.

It is worth noting that some input variables involved in Eqs. 1 and 2 can be statistically correlated. For example, ρ_w, the density of water, is expected to be dependent on the value of T, reservoir temperature in Eq. 2. We would expect that increasing T decreases ρ_w, which indicates that these two variables are negatively correlated. Hence, this indicates that ρ_w and T may not be treated as two independent variables in Eq. 2. We may also expect that c_s, the solid rock specific heat be negatively correlated with ρ_s, the solid rock density, and c_s be positively correlated with temperature. In addition, reservoir area may be positively correlated with the net thickness (Murtha, 1994). Our point is that ignoring existing correlations between input variable pairs may lead to an incorrect characterization of uncertainty in *PW or* H_t. If data and available information permit, one should make scatter plots of input variable pairs to identify the correlation between them, if any, and then include these correlations into the uncertainty assessment procedure.

3. LOG-NORMALITY OF THERMAL RESOURCES

If we take the natural logarithm of *PW* given by Eq. 1, we obtain:

$$\ln PW = \ln H_t + \ln R_F + \ln Y + \ln \frac{1}{L_F} + \ln\left(\frac{1}{10^3 t_p}\right), \quad (3)$$

where $\ln H_t$, which follows from Eq. 2 by taking the natural logarithm of it, is given by:

$$\ln H_t = \ln\left[(1-\phi)c_s\rho_s + \phi c_w \rho_w\right] + \ln A + \ln h + \ln(T - T_r) \quad (4)$$

Eq. 3 clearly indicates that ln*PW* can be written as a sum of the natural logarithms of H_t, R_F, Y and L_F. If all these random variables are treated as independent, then it follows from the CLT, discussed previously, that the resulting distribution of ln*PW* will tend to be normal. Hence, *PW* will tend to be log-normal. It is important to note that this is true no matter what type of distribution the input random variables assume.

Note that the CLT promises that ln*PW* be normal if all the random variables are independent. However, as we will show later [also see, Sarak *et al.* (2009)] the resulting distributions of ln*PW* still tend to be normal even if some of the input variables are treated as dependent.

Sarak *et al.* (2009) have also studied the resource estimation for stored heat, H_t, based on Eq. 2 and 4. They found that the distribution of H_t is also log-normal, even though it may not be apparent from Eq. 4 as the first logarithmic term in the right-hand-side of Eq. 4 cannot be written as the sum of the natural logarithms of the individual input parameters, i.e.,

$$\ln\left[(1-\phi)c_s\rho_s + \phi c_w\rho_w\right] \neq \ln(1-\phi) + \ln c_s + \ln \rho_s + \ln\phi + \ln c_w + \ln\rho_w . \quad (5)$$

Sarak *et al.* (2009) shows that for most of the cases of practical interest, the most of the heat is stored in the solid part (typically 80 to 90 percent of the total heat in rock and fluid) and hence the first term in the right-hand side of Eq. 4 is well approximated by

$$\ln\left[(1-\phi)c_s\rho_s + \phi c_w\rho_w\right] \approx \ln(1-\phi) + \ln c_s + \ln\rho_s . \quad (6)$$

Eq. 6 may justify why H_t (and also *PW* which depends on it) follows closely a log-normal distribution.

4. QUANTIFICATION OF UNCERTAINTY

We first state our definitions to be used for characterizing uncertainty in *PW*. For this characterization, we adopt the convention proposed by Capen (2001). We will refer to P10 as "proved", P50 as "probable", and P90 as "possible", where P10, P50, and P90 correspond to 10th, 50th and 90th percentiles of the cumulative distribution function, respectively, for *PW*.

It is worth noting that in SPE literature and also in some papers in geothermal literature (e.g., Sanyal and Sarmiento, 2005), P10 used throughout in this paper is referred as "P90" (or proved resources) indicating that there is at least a 90% probability that the quantities actually recovered will be equal or exceed the estimate, and P90 used in this paper is referred to as "P10" indicating that there is a 10% probability that the quantities actually recovered will be equal or exceed the estimate. In our definition, P10 (or 50 or 90) refers to 10th (or 50th or 90th) percentile indicating that there is a 10% (or 50% or 90%) probability that the quantities actually recovered will be equal or less than the estimate.

As discussed in the previous section, based on the CLT theorem, the uncertainty on the resource *PW* of a single field will tend to be a log-normal distribution. So, ln*PW* follows a normal distribution characterized by its mean ($\mu_{\ln PW}$) and variance ($\sigma^2_{\ln PW}$), while *PW* is a log-normal by its mean (μ_{PW}) and variance (σ^2_{PW}). The two sets of parameters are related by the equations:

$$\mu_{PW} = \exp\left(\mu_{\ln PW} + \frac{\sigma^2_{\ln PW}}{2}\right), \quad (7)$$

and

$$\sigma^2_{PW} = \mu^2_{PW}\left[\exp\left(\sigma^2_{\ln PW}\right) - 1\right]. \quad (8)$$

We can also derive the following equations from Eqs. 7 and 8:

$$\mu_{\ln PW} = \ln \mu_{PW} - \frac{1}{2}\ln\left(1 + \frac{\sigma^2_{PW}}{\mu^2_{PW}}\right), \quad (9)$$

and

$$\sigma^2_{\ln PW} = \ln\left(1 + \frac{\sigma^2_{PW}}{\mu^2_{PW}}\right). \quad (10)$$

The 10th, 50th and 90th percentiles of *PW* are computed from the following equations:

$$\text{P10} = \exp\left(\mu_{\ln PW} - 1.28\sqrt{\sigma^2_{\ln PW}}\right), \quad (11)$$

$$\text{P50} = \exp\left(\mu_{\ln PW}\right). \quad (12)$$

and

$$\text{P90} = \exp\left(\mu_{\ln PW} + 1.28\sqrt{\sigma^2_{\ln PW}}\right). \quad (13)$$

In following subsections, we review some basic equations and methods used for quantification of uncertainty (i.e., computing mean, variance, P10, P50, and P90 given by Eqs. 7-13) in *PW* (or ln*PW*) by the volumetric method (Eq.

1 or 3). We first consider the Monte Carlo method (MCM) and then the analytic uncertainty propagation method (AUPM).

4.1 Monte Carlo Method (MCM)

The MCM relies on a specified probability distribution of each of the input variables and generates an estimate of the overall uncertainty in the prediction due to all uncertainties in the variables (Kalos and Withlock, 2008). As it does not require a linearization of the function and a continuity of the random variables, it is a more general approach for characterizing the uncertainty for any given nonlinear random function *f*. In our case, *f* represents *PW* given by Eq. 1 or ln*PW* given by Eq. 3. In the applications to be given, we perform Monte Carlo simulations by using @RISK,TM spreadsheet-based software (2004).

4.2 Analytical Uncertainty Propagation Method (AUPM)

Here, we derive an uncertainty propagation equation for a function *f* (*PW* in Eq. 1 or its natural logarithm given by Eq. 3) where it is treated as a continuous random function due to uncertainties in the input variables. We assume that all uncertainties are due to the random uncertainties in the input variables and ignore the systematic errors in the input variables. The error propagation equation we present is based on a Taylor series approximation of the function around the mean values of the variables up to its first derivatives with respect to each of the input variables. As a consequence of this approximation, the uncertainty propagation equation provides a linearization of the function in terms of its input random variables (Barlow, 1989; Coleman and Steele, 1999; Zeybek *et al.*, 2009).

The AUPM provides a simple approach for estimating the variance of a function defined by several random variables − particularly so, of a function defined by products and quotients of random variables, whether they are independent or correlated. The method does not assume a specific type of distribution for the input variables and all needed to use the AUPM are the statistical properties of the distribution of each random variable; specifically the mean, variance (or std. dev.), and the covariance (or correlation coefficient) among variable pairs if the random variables are correlated (Sarak *et al.*, 2009).

Before we present the derivation of the AUPM, it is worth noting that the AUPM provides an exact result for the mean and variance of a random function *f* if *f* is linear with respect to the input random variables. Otherwise, i.e., if *f* is nonlinear, then the AUPM provides only approximate estimates of the mean and variance of *f*. The approximation gets better if nonlinear *f* can be well approximated by a linear function near the means of the input random variables.

For the problem of interest in this work, we wish to estimate the mean and variances of *PW* given by Eq. 1. *PW* given by Eq. 1 is, in general, a nonlinear function of the input variables. As noted before, if we work, however, with the natural logarithm of *PW*, we obtain a partially linearized equation for ln*PW* (Eq. 3). We say "partially linearized" because ln*PW* is still nonlinear with respect to the input variables involved in $\ln H_t$ (Eq. 4), but linear with respect to the variables $\ln R_F$, $\ln Y$ and $\ln(1/L_F)$.

As shown by Sarak *et al.* (2009), three different approaches could be considered when the AUPM is used to derive approximations for the mean and variance of *PW*. For example, we can directly apply the AUPM to the *PW* given by Eq. 1 as a function of the input variables (Approach 1) or can directly apply the AUPM to the ln*PW* function given by Eq. 3 by treating it as a function of input variables (Approach 2) or as a function of the natural logarithms of the input variables (Approach 3). Sarak *et al.* (2009) shows that the Approach 3 provides the best accurate estimates of the mean, variance, P10, P50, and P90 of the ln*PW*; i.e., we apply the AUPM directly ln*PW* (Eq. 3) as a function of the natural logarithms of the input variables. In this paper, we present the equations for the AUPM based on Approach 1 and Approach 3. The equations of the AUPM based on Approach 2 can be found in Sarak *et al.* (2009). In this paper, from this point on, the Approach 3 is to be referred to as Approach 2.

4.2.1 AUPM for *PW* (Approach 1)

Let's consider a random function *f* of *M* variables, X_i, $i = 1,2,\ldots,M$, i.e., $f = f\left(X_1, X_2, \ldots, X_M\right)$. Then, expanding *f* around the mean (or true) values of X_is (denoted by $\mu_{X_i}, i = 1,2,\ldots,M$) by using a Taylor series up to first derivatives, we obtain:

$$f\left(X_1, X_2, \ldots X_M\right) = f\left(\mu_{X_1}, \mu_{X_2}, \ldots, \mu_{X_M}\right) + \sum_{i=1}^{M}\left(X_i - \mu_{X_i}\right)\left(\frac{\partial f}{\partial X_i}\right)\Bigg|_{X_i=\mu_{X_i}, i=1,\ldots,M} . \tag{14}$$

It can be shown that the mean (μ_f) and variance of *f* (σ_f^2) are approximated by:

$$\mu_f = f\left(\mu_{X_1}, \mu_{X_2}, \ldots, \mu_{X_M}\right), \tag{15}$$

and

$$\sigma_f^2 = \sum_{i=1}^{M} \theta_i^2 \sigma_{X_i}^2 + 2\sum_{i=1}^{M-1}\sum_{j=i+1}^{M} \theta_i \theta_j cov(X_i, X_j), \tag{16}$$

where $cov(X_i, X_j)$ represents the covariance between the variable pairs X_i and X_j, and if we use the relation between covariance and correlation coefficient, ρ_{X_i,X_j} , then we can express Eq. 16 in terms of the correlation coefficient as:

$$\sigma_f^2 = \sum_{i=1}^{M} \theta_i^2 \sigma_{X_i}^2 + 2\sum_{i=1}^{M-1}\sum_{j=i+1}^{M} \theta_i \theta_j \rho_{X_i,X_j} \sqrt{\sigma_{X_i}^2 \sigma_{X_j}^2} . \tag{17}$$

In Eqs. 16 and 17, θ_i is the derivative of *f* with respect to the variable X_i, i.e.,

$$\theta_i = \left(\frac{\partial f}{\partial X_i}\right)\Bigg|_{X_i=\mu_{X_i}, i=1,\ldots,M} . \tag{18}$$

Note that θ_i represents the sensitivity of *f* to the variable X_i evaluated at the mean values of all the variables. It can be noticed (from Eqs. 16 and 17) that the uncertainty propagation on to *f* is determined not only by the variances of the variables and correlation among them, but also the sensitivity of *f* to each variable in the volumetric method for *PW* (see Eqs. 1).

As mentioned previously, for the problem of interest, *f* in Eqs. 14-17 represents *PW* given by Eq. 1. The sensitivities of *PW* (i.e., θ_is) required in Eqs. 16 and 17 can be obtained by analytical differentiation of Eq. 1 with respect to the input variables in Eq. 1. These sensitivities are presented in Table 1.

Table 1: Sensitivity of PW (Eq. 1) with respect to a given input variable X_i in Eq. 1.

Variable X_i	$\theta_i = \partial PW/\partial X_i^*$
ϕ	$\dfrac{\left(-\mu_{c_s}\mu_{\rho_s}+\mu_{c_w}\mu_{\rho_w}\right)\mu_A\mu_h\mu_{(T-T_r)}\mu_{R_F}\mu_Y}{10^3\mu_{L_F}t_p}$
c_s	$\dfrac{\left(1-\mu_\phi\right)\mu_{\rho_s}\mu_A\mu_h\mu_{(T-T_r)}\mu_{R_F}\mu_Y}{10^3\mu_{L_F}t_p}$
ρ_s	$\dfrac{\left(1-\mu_\phi\right)\mu_{c_s}\mu_A\mu_h\mu_{(T-T_r)}\mu_{R_F}\mu_Y}{10^3\mu_{L_F}t_p}$
c_w	$\dfrac{\mu_\phi\mu_{\rho_w}\mu_A\mu_h\mu_{(T-T_r)}\mu_{R_F}\mu_Y}{10^3\mu_{L_F}t_p}$
ρ_w	$\dfrac{\mu_\phi\mu_{c_w}\mu_A\mu_h\mu_{(T-T_r)}\mu_{R_F}\mu_Y}{10^3\mu L_F t_p}$
A	$\dfrac{\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s}+\mu_\phi\mu_{c_w}\mu_{\rho_w}\right]\mu_h\mu_{(T-T_r)}\mu_{R_F}\mu_Y}{10^3\mu_{L_F}t_p}$
h	$\dfrac{\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s}+\mu_\phi\mu_{c_w}\mu_{\rho_w}\right]\mu_A\mu_{(T-T_r)}\mu_{R_F}\mu_Y}{10^3\mu_{L_F}t_p}$
T-T_r	$\dfrac{\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s}+\mu_\phi\mu_{c_w}\mu_{\rho_w}\right]\mu_A\mu_h\mu_{R_F}\mu_Y}{10^3\mu_{L_F}t_p}$
R_F	$\dfrac{\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s}+\mu_\phi\mu_{c_w}\mu_{\rho_w}\right]\mu_A\mu_h\mu_{(T-T_r)}\mu_Y}{10^3\mu_{L_F}t_p}$
Y	$\dfrac{\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s}+\mu_\phi\mu_{c_w}\mu_{\rho_w}\right]\mu_A\mu_h\mu_{(T-T_r)}\mu_{R_F}}{10^3\mu_{L_F}t_p}$
L_F	$\dfrac{\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s}+\mu_\phi\mu_{c_w}\mu_{\rho_w}\right]\mu_A\mu_h\mu_{(T-T_r)}\mu_{R_F}\mu_Y}{-10^3t_p\mu_{L_F}^2}$
*****evaluated at the mean values of the variables X_is	

$$\mu_{\ln PW}=\ln PW\left(\mu_{\ln X_1},\mu_{\ln X_2},\ldots,\mu_{\ln X_M}\right), \tag{19}$$

and

$$\sigma^2_{\ln PW}=\sum_{i=1}^{M}\theta_i^2\sigma^2_{\ln X_i}+2\sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\theta_i\theta_j\rho_{\ln X_i,\ln X_j}\sigma_{\ln X_i}\sigma_{\ln X_j} \tag{20}$$

where the sensitivities θ_is in Eq. 19 are given by:

$$\theta_i=\left(\frac{\partial\ln PW}{\partial\ln X_i}\right)\bigg|_{X_i=\mu_{X_i},i=1,\ldots,M}, \tag{21}$$

and are tabulated in Table 2.

Table 2: Sensitivity of ln*PW (Eq. 3) with respect to a natural logarithm of a given variable X_i in Eq. 3.*

Variable X_i	$\theta_i = \partial\ln PW/\partial\ln X_i^*$
ϕ	$\dfrac{\mu_\phi\left(-\mu_{c_s}\mu_{\rho_s}+\mu_{c_w}\mu_{\rho_w}\right)}{\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s}+\mu_\phi\mu_{c_w}\mu_{\rho_w}\right]}$
c_s	$\dfrac{\mu_{c_s}\left(1-\mu_\phi\right)\mu_{\rho_s}}{\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s}+\mu_\phi\mu_{c_w}\mu_{\rho_w}\right]}$
ρ_s	$\dfrac{\mu_{\rho_s}\left(1-\mu_\phi\right)\mu_{c_s}}{\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s}+\mu_\phi\mu_{c_w}\mu_{\rho_w}\right]}$
c_w	$\dfrac{\mu_{\rho_w}\mu_\phi\mu_{c_w}}{\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s}+\mu_\phi\mu_{c_w}\mu_{\rho_w}\right]}$
ρ_w	$\dfrac{\mu_\phi\mu_{c_w}\mu_{\rho_w}}{\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s}+\mu_\phi\mu_{c_w}\mu_{\rho_w}\right]}$
A	1
h	1
T-T_r	1
R_F	1
Y	1
L_F	−1
*****evaluated at the mean values of the variables X_is	

Once the mean (μ_{PW}) and variance (σ^2_{PW}) of *PW* are computed by the use of Eqs. 15 and 16 (or 17), we then use these values in Eqs. 9 and 10 to compute the mean ($\mu_{\ln PW}$) and variance ($\sigma^2_{\ln PW}$) of ln*PW*, which is normal based on the CLT theorem. The other uncertainty markers such as P10, P50, and P90 can be computed from Eqs. 11-13.

4.2.2 AUPM for ln*PW* (Approach 2)

The second approach is based on the Taylor series expansion of ln*PW* around the mean values of natural log of the input variables; i.e., $\mu_{\ln X_i}$s. For this case, the AUP equations are given by Eqs. 14-18 with *f* replaced by ln*f*, X_is by lnX_is, and μ_{X_i}s by $\mu_{\ln X_i}$s. So, in this approach, the mean and variance of ln*PW* are computed from:

Once the mean ($\mu_{\ln PW}$) and variance ($\sigma^2_{\ln PW}$) of ln*PW* are computed by the use of Eqs. 19 and 20, we then use these values in Eqs. 7 and 8 to compute the mean (μ_{PW}) and variance (σ^2_{PW}) of *PW*. The other uncertainty markers such as P_{10}, P_{50}, and P_{90} can be computed from Eqs. 11-13 by using the values of $\mu_{\ln PW}$ and $\sigma^2_{\ln PW}$.

Finally, a few remarks are order for the AUP equations based on the Approach 1 or 2. Our numerical results indicate that Approach 2 provides a slightly better estimate of the variance than does Approach 1. However, unlike the Approach 1, the Approach 2 requires us to work with the means and variances of the natural-log of the input model variables, i.e., $\mu_{\ln X_i}$ and $\sigma^2_{\ln X_i}$. In the correlated case, we will also need to convert to correlation coefficient between two pairs, say ρ_{X_i,X_j} to $\rho_{\ln X_i,\ln X_j}$, but our results show that $\rho_{\ln X_i,\ln X_j} = \rho_{X_i,X_j}$ for all practical purposes. If the distribution of X_i is chosen as a log-normal with mean μ_{X_i} and variance $\sigma^2_{X_i}$, then lnX_i is normal with the mean $\mu_{\ln X_i}$ and variance $\sigma^2_{\ln X_i}$ which can be simply computed from:

$$\mu_{\ln X_i} = \ln \mu_{X_i} - \frac{1}{2}\ln\left(1+\frac{\sigma^2_{X_i}}{\mu^2_{X_i}}\right), \tag{22}$$

and

$$\sigma^2_{\ln X_i} = \ln\left(1+\frac{\sigma^2_{X_i}}{\mu^2_{X_i}}\right). \tag{23}$$

If the chosen distribution for the input variable X_i is not log-normal, then we can use descriptive statistics on the available data to compute $\mu_{\ln X_i}$ and $\sigma^2_{\ln X_i}$. If such exhaustive data are not available, then we may generate samples from a known distribution and use descriptive statistics on these samples to compute $\mu_{\ln X_i}$ and $\sigma^2_{\ln X_i}$.

4.3 Example Applications

In this section, we consider some example applications comparing the results obtained from MC and AUP approaches for predicting the uncertainty in thermal resource PW for a single field.

4.3.1 Example Application 1

The first example application pertains to a case where all input variables are independent in *PW* given by Eq. 1. For the purpose of this example, we consider two cases: **Case 1** assumes that the distribution of each input variable in Eq. 1 can be characterized by a triangular distribution, whereas **Case 2** assumes that the distribution of each input variable in Eq. 1 can be characterized by a log-normal distribution. For **Case 1**, the minimum, most likely (mode), and maximum values of the input variables are given in Table 3. The values of mean and variance given in Table 3 were computed from the well-known formulas for a triangular distribution:

$$\mu_{X_i} = \frac{Min + Max + Mode}{3}, \tag{24}$$

and

$$\sigma^2_{X_i} = \frac{(Min)^2 + (Max)^2 + (Mode)^2}{18} - \frac{(Min\times Max + Min\times Mode + Max\times Mode)}{18}. \tag{25}$$

The data given in Table 3 pertain to Izmir Balçova-Narlıdere geothermal field in Turkey and were taken from Satman *et al.* (2001). For this application, T_r = 60 °C.

For Case 2, we assume that the distribution of each input variable X_i is a log-normal with the mean and variances computed from Eqs. 24 and 25 (see 5th and 6th columns of Table 3). Table 4 presents the values of the mean and variances for lnX_i computed (see 4th and 5th columns) by using the values of mean (μ_{X_i}) and variance ($\sigma^2_{X_i}$) for X_i given in the 2nd and 3rd columns in Eqs. 22 and 23.

Table 3: Distributions of the input variables; triangular distribution; Case 1

Variable X_i	*Min*	*Mode*	*Max*	*Mean*⁺ μ_{X_i}	*Variance*⁺ $\sigma^2_{X_i}$
ϕ	0.02	0.05	0.1	0.057	2.722×10^{-4}
c_s, kJ/(kg °C)	0.75	0.9	1.0	0.883	2.639×10^{-3}
ρ_s, kg/m^3	2550	2650	2750	2650	1.667×10^{3}
c_w, kJ/(kg °C)	4.00	4.18	4.21	4.130	2.150×10^{-3}
ρ_w, kg/m^3	922	931	987	946.7	2.067×10^{2}
A, m^2	5×10^5	9×10^5	2×10^6	1.1×10^6	1.006×10^{11}
h, m	250	350	1000	533.3	2.764×10^{4}
T-T_r, °C	40	75	85	66.67	9.306×10^{1}
R_F	0.07	0.18	0.24	0.163	1.239×10^{-3}
Y	0.7	0.85	0.9	0.817	1.806×10^{-3}
L_F	0.35	0.41	0.5	0.42	9.500×10^{-4}
t_p, s.	8×10^8	8×10^8	8×10^8	8×10^8	0.0

⁺mean and variance were computed from the known formulas given for a triangular distribution, see Eqs. 21 and 22.

Table 4: Distributions of the input variables; log-normal distribution; Case 2.

Variable X_i	*Mean* μ_{X_i}	*Variance* $\sigma^2_{X_i}$	*Mean* $\mu_{\ln X_i}$	*Variance* $\sigma^2_{\ln X_i}$
ϕ	0.057	2.722×10^{-4}	-2.905	8.045×10^{-2}
c_s, kJ/(kg °C)	0.883	2.639×10^{-3}	-0.1261	3.379×10^{-3}
ρ_s, kg/m^3	2650	1.667×10^{3}	7.882	2.374×10^{-4}
c_w, kJ/(kg °C)	4.130	2.150×10^{-3}	1.418	1.260×10^{-4}
ρ_w, kg/m^3	946.7	2.067×10^{2}	6.853	2.306×10^{-4}
A, m^2	1.1×10^6	1.006×10^{11}	13.871	7.986×10^{-2}
h, m	533.3	2.764×10^{4}	6.233	9.275×10^{-2}
T-T_r, °C	66.67	9.306×10^{1}	4.189	2.072×10^{-2}
R_F	0.163	1.239×10^{-3}	-1.837	4.558×10^{-2}
Y	0.817	1.806×10^{-3}	-0.2035	2.702×10^{-3}
L_F	0.42	9.500×10^{-4}	-0.8702	5.371×10^{-3}
t_p, s.	8×10^8	0.0	2.05×10^1	0.0

Figures 1 and 2 show histograms of *PW* generated from the MCM by using the distributions given in Table 3 and 4, respectively, in @RISK. The statistical variables (e.g., mean, variance, P10, P50, and P90) for each histogram are given in the insets of Figures 1 and 2.

As is expected from the CLT, both histograms shown in Figures 1 and 2 are log-normal and also the statistical parameters obtained for both cases are very similar. So, as discussed previously, in fact, there is no reason to insist upon any particular probability distribution for our input variables provided that means and variances are the same for the chosen distributions for the input variables unless the data we have suggests otherwise.

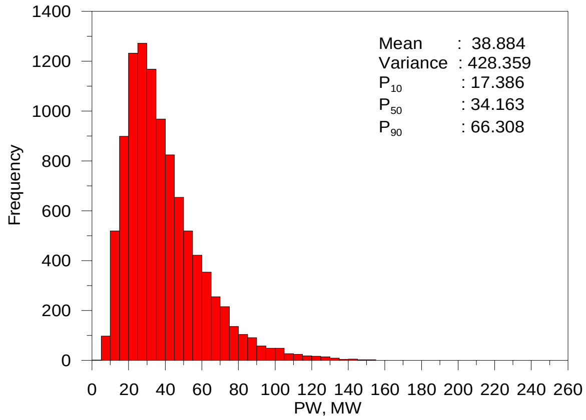

Figure 1: Histogram of recoverable power, PW, generated from the MCM, for the case where each input variable is based on a triangular distribution (see Table 3); Case 1 .

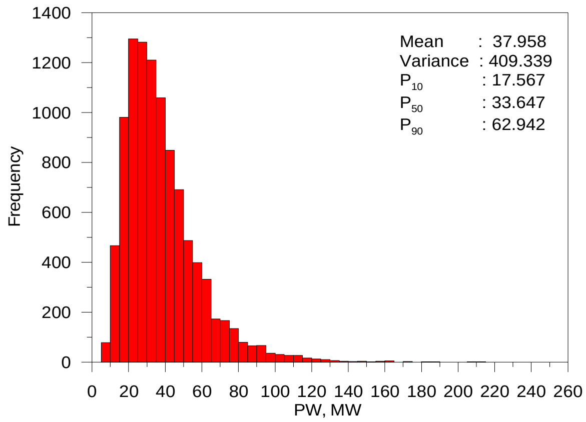

Figure 2: Histogram of recoverable power, PW, generated from the MCM, for the case where each input variable is based on a log-normal distribution (see Table 4; Case 2.

Now, we compare the estimates of means, variances, P_{10}, P_{50}, and P_{90} computed for *PW* from the MCM and AUPM based on Approaches 1 and 2 as described earlier. Table 5 compares the values of means and variances computed from MCM and AUPM for *PW and* ln*PW for* Case 2.As can be seen from Table 5, the computed values of means and variances from the MCM and AUPM for *PW* and ln*PW* functions agree well. We also notice that the values of means and variances for *PW* and ln*PW* computed from the AUPM based on Approach 2 better agrees with corresponding ones from the MCM. This is not surprising though, and is an expected result because as mentioned previously, the AUPM provides a linear approximation to a nonlinear random function around the mean values of the input variables and can provide exact results for the mean and variance for a function *f* if that function is linear with the input random variables. In our case, *PW* (Eq. 1) are in fact nonlinear functions of their input variables. On the other hand, ln*PW* (Eq. 3) are almost linear (or weakly nonlinear) functions of the input variables. Consequently, the AUPM (based on either Approach 2) provides estimates of means and variances for *PW* and ln*PW* functions that agree very well with those computed from the MCM.

Table 5: A comparison of means and variances from the MCM and AUPM for PW (Eq. 1) and ln*PW (Eq. 3) function; Case 2.*

	Mean		**Variance**	
	MCM	AUPM	MCM	AUPM
***PW*, MW**	37.96	37.66† 37.91*	409.3	367.4† 408.3*
ln*PW*	3.511	3.513† 3.510*	0.2509	0.2303† 0.2501*

†AUPM based on Approach 1
*AUPM based on Approach 2

Next, we compare the 10th, 50th and 90th percentiles computed from the MCM and AUPM for *PW*, and ln*PW* functions for Case 2. Table 6 presents the results. As can be seen from Table 6, the values of P_{10}, P_{50}, and P_{90} percentiles computed from the MCM and AUPM (based on Approaches 1 and 2) agree quite well. They are essentially identical.

Table 6: A comparison of the values of 10*th*, 50*th and* 90*th percentiles computed from the MCM and AUPM for PW and* ln*PW functions; Case 2.*

	P10		**P50**		**P90**	
	MCM	AUPM	MCM	AUPM	MCM	AUPM
PW, MW	17.6	18.1† 17.6*	33.6	33.5† 33.4*	62.9	62.0† 63.4*
ln*PW*	2.87	2.90† 2.87*	3.52	3.51† 3.51*	4.14	4.12† 4.15*

†AUPM based on Approach 1
*AUPM based on Approach 2

4.3.1 Example Application 1

Our next example application pertains to a case where some of the input variables in Eq. 1 are correlated. As mentioned previously, it is possible that various input variables in stored heat and recoverable power can be correlated with each other. For example, the solid rock specific heat may be negatively correlated with the density of the solid rock, the density of water may be negatively correlated with temperature, and the solid rock specific heat can be positively correlated with temperature. In addition, we may expect that area (*A*) and thickness (*h*) are positively correlated (Murtha, 1994). For this investigation, we use the same input distributions given in Table 3, but assume correlation between the five correlated pairs and the correlation coefficients given in Table 7.

Figure 3 shows the histogram *PW* (Eq. 1) generated from the MCM by using the distributions given in Table 3 and the correlation coefficients given in Table 7 in @RISK. As can be seen, correlation did not change the shape of the *PW* distribution. It is still nearly log-normal. When the result of Fig. 3 for the correlated case is compared with the results of Figure 1 for the uncorrelated case, we see that correlation increased the variance significantly (about 30%). Correlation increased the 10th, 50th and 90th percentiles of H_t and *PW* slightly (about 6%) compared to the corresponding results for the uncorrelated case.

Table 7: Correlated variable pairs and correlation coefficients.

Correlated Variable Pairs $\left(X_i, X_j\right)$	***Correlation Coefficient*** ρ_{X_i,X_j}
$\left(c_s, T-T_r\right)$	+0.63
$\left(c_s, \rho_s\right)$	−0.44
$\left(\rho_w, T-T_r\right)$	−0.62
$\left(c_w, \rho_w\right)$	−0.42
$\left(A, h\right)$	+0.24

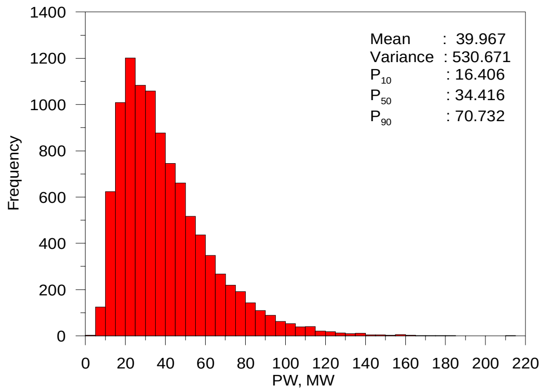

Figure 3: Histogram of recoverable power, PW, generated from the MCM, correlated case.

Table 8 compares the values of means and variances computed from MCM and AUPM for *PW* and ln*PW*. We used the AUPM based on Approach 2 to compute means and variances of *PW* and ln*PW* functions. This approach requires that we work with the correlation coefficient between the pairs in terms of the natural-log of input random variables; i.e., $\rho_{\ln X_i, \ln X_j}$. Our results indicate $\rho_{\ln X_i, \ln X_j} = \rho_{X_i, X_j}$. That is; one can use the correlation coefficients based ρ_{X_i, X_j} when using the AUPM method based on Approach 2. Note that the means and variances obtained from MCM and AUPM based on Approach 2 for ln*PW* are essentially identical.

Table 9 compares the values of P10, P50, and P90 computed from the MCM and AUPM (based on Approach 2). Again, there is a very good agreement in the 10th, 50th and 90th percentiles computed from both methods.

Table 8: A comparison of means and variances from the MCM and AUPM for PW and ln*PW functions, correlated case.*

	Mean		**Variance**	
	MCM	AUPM	MCM	AUPM
***PW*, MW**	39.9	37.7	530.7	438.5
ln*PW*	3.534	3.536	0.3148	0.3177

Table 9: A comparison of the values of 10*th,* 50*th and* 90*th percentiles computed from the MC and AUP methods for PW and* ln*PW functions, correlated case.*

	P10		**P50**		**P90**	
	MCM	AUPM	MCM	AUPM	MCM	AUPM
PW, MW	16.4	16.7	34.4	34.3	70.7	70.6
ln*PW*	2.82	2.81	3.53	3.54	4.25	4.26

In summary, our results show that correlation among variables, particularly between *A* and *h*, if they exist and that data available permits one to identify correlation among variables, should be accounted for accurate characterization of uncertainty in *PW*. The results also indicate that the AUPM works as good as the MCM to estimate uncertainty (variance, 10th, 50th, and 90th percentiles) in *PW* even for the case where the input variables are correlated.

5. AGGREGATION OF THERMAL RESOURCES

Here, we consider the problem of aggregating (or adding) thermal resources of diverse fields. Should one use an arithmetic or probabilistic addition to determine "proved" and "probable" (which corresponds to P10 and P90 percentiles of the cumulative distribution function, respectively) corresponding a geothermal project or a country which may involve many diverse fields, each with its estimated values of mean, variance, P10, P50, and P90?

The aggregation problem has been studied in the petroleum engineering literature by a number of authors; see for example, Capen (1996, 2001), Carter and Morales (1998), van Elk *et al*. (2000), Demirmen (2007), and Delfiner and Barrier (2008). However, this problem has not yet received much attention in the geothermal literature. In our previous work (Sarak *et al.*, 2009), it was shown that the simple arithmetic sum may significantly underestimate the P_{10} and significantly overestimate the P90, relative to the corresponding ones estimated by probabilistic sum and that the statistically proper method of aggregating divers fields' resources is by probabilistic sum. However, it should be note that our previous work assumed that the all fields considered in aggregation process are independent or uncorrelated. Here, we further discuss the aggregation of diverse field reserves by considering the fact that some of the fields may be correlated.

As is well known [for example see, Capen (1996, 2001), Carter and Morales (1998), and Delfiner and Barrier (2008)], the correct procedure when aggregating the resources of many diverse fields is a probabilistic addition whether the fields to be aggregated are independent or not. As to be shown mathematically, the arithmetic sum assumes that all fields considered in aggregation are fully correlated (i.e., pairwise correlation coefficients for all fields is equal to unity). In other words, probabilistic sum will be equal to the arithmetic sum if all fields are fully correlated. On the other hand, if we assume that all fields are independent and apply the probabilistic addition, then the probabilistic sum overestimates the P_{10} and P_{90} of the all fields used in aggregation.

As well stated by Delfiner and Barrier (2008), in reality fields are neither perfectly dependent nor perfectly independent, but instead are correlated. So, although the probabilistic sum is the most general approach that one should use whether the fields are independent or not, however it requires the knowledge of pairwise correlation coefficients for the fields. Suppose that we aggregate thermal resources of n geothermal fields, then one has to consider a total of $n(n$-1)/2 pairwise correlations. For example, if n = 10, then we would need 45 correlation coefficients, if n = 20, then we would need 190 correlation coefficients. In practice, the estimation of all such correlations between field-resource estimates may not be possible and feasible. Therefore, some authors have proposed simplified and pragmatic approaches for aggregating field resources [Carter and Morales (1998), van Elk *et al.* (2000), and Delfiner and Barrier (2008)]. For example, Delfiner and Barrier (2008) propose a partial probabilistic addition in which group of fields presumed to be relatively dependent or relatively are defined, and the summation is performed using arithmetic or probabilistic addition depending on the assumed values of correlation coefficients between pairwise fields. We refer the readers to the work of Carter and Morales (2007) and Delfiner and Barrier (2008) for further details regarding such simplified aggregation procedures.

5.1 Aggregation of Means, Variances, P10, P50, and P90

In the following, we provide a general formulation for estimating the values of mean, variance, P_{10}, P_{50} and P_{90} for a total of n geothermal PW resources, each following a log-normal distribution characterized by its mean (μ_{PW_j}, $j=1,2,\cdots,n$) and variance ($\sigma^2_{PW_j}$, $j=1,2,\cdots,n$). Note that if each PW_j, j = 1,2,...,n, is a log-normal distribution, then lnPW_j, j = 1,2,…,n, is a normal distribution with a mean equal to $\mu_{\ln PW_j}$, $j=1,2,\cdots,n$ and a variance equal to $\sigma^2_{\ln PW_j}$, $j=1,2,\cdots,n$. The two sets of parameters are related by Eqs. 7-10 for each field's resource.

Now suppose that we are interested in the uncertainty of the total resources (denoted by PW_S) which is equal to the sum of the resources of all PW_j, j = 1,2,…,n, given by:

$$PW_S = \sum_{j=1}^{n} PW_j \ , \tag{26}$$

It is not difficult to show that the mean (or expected value denoted by μ_{PW_S}) and the variance of PW_S (denoted by $\sigma^2_{PW_S}$) are given by the following equations, respectively:

$$\mu_{PW_S} = \sum_{j=1}^{n} \mu_{PW_j} \ , \tag{27}$$

and

$$\sigma^2_{PW_S} = \sum_{j=1}^{n} \sigma^2_{PW_j} + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \rho_{PW_i,PW_j}\sqrt{\sigma^2_{PW_i}\sigma^2_{PW_j}} \ , \tag{28}$$

where ρ_{PW_i,PW_j} represents the pairwise correlation coefficient between field resources i and j.

A few remarks are in order for Eqs. 27 and 28: Eq. 27 indicates that the mean of the sum of the resources is equal to the arithmetic sum of the mean of each resource PW_j, whether field resources used in aggregation are correlated or not. So, this result indicates that we can add the mean of each field resources to find the mean of the sum of all fields' resources. Eq. 28 indicates that the variance of the sum of the resources will not be equal to the sum of the variances of individual resources unless all field resources are independent, i.e., $\rho_{PW_i,PW_j}=0$ for all field resource pairs i and j. So if all fields are independent Eq. 28 reduces to:

$$\sigma^2_{PW_S} = \sum_{j=1}^{n} \sigma^2_{PW_j} \ . \tag{29}$$

On the other hand, if we assume that all pairwise correlations in Eq. 28 are equal to unity, i.e., $\rho_{PW_i,PW_j}=1$, for all i and j such that $i \neq j$, then it is not difficult to show that Eq. 28 reduces to:

$$\sigma^2_{PW_S} = \left(\sum_{i=1}^{n} \sigma_{PW_j}\right)^2 . \tag{30}$$

In theory, the sum of log-normal distributions (i.e., the sum given by Eq. 26) is not log-normal. Rather, by the CLT, it is expected to be normal. As discussed by Capen (2006) and shown numerically on an example application by Sarak *et al.* (2009), Monte Carlo simulations show that the convergence to normal distribution is slow because of the skewed shape of the log-normal distribution and that log-normal model is a better approximation for the aggregation of field resources than is a normal distribution. As it is much easier to derive some of the limiting forms we would like to show here, for now we will assume the sum of log-normal distribution of thermal field PW resources is normal so that the P10, P50, and P90 of the PW_S are given by

$$\text{P10}_S = \mu_{PW_S} - 1.28\sqrt{\sigma^2_{PW_S}} \ , \tag{31}$$

$$\text{P50}_S = \mu_{PW_S} \ , \tag{32}$$

and

$$\text{P90}_S = \mu_{PW_S} + 1.28\sqrt{\sigma^2_{PW_S}} \ . \tag{33}$$

Now, we would like to show that the sum of P10, P50, and P90 based on probabilistic addition given by Eqs. 31-33, respectively, will be equal to P10, P50, and P90 that would be obtained by using an arithmetic sum if all fields involved in aggregation are perfectly correlated, i.e., the variance of the sum of n log-normal field resources are given by Eq. 30. Using Eqs. 27 and 30 in Eqs. 31-33 gives:

$$\mathrm{P10}_S = \sum_{j=1}^{n} \mu_{PW_j} - 1.28\left(\sum_{i=1}^{n} \sigma_{PW_j}\right) = \sum_{j=1}^{n}\left(\mu_{PW_j} - 1.28\sigma_{PW_j}\right) = \sum_{j=1}^{n} \mathrm{P10}_j \quad (34)$$

$$\mathrm{P50}_S = \sum_{j=1}^{n} \mu_{PW_j} = \sum_{j=1}^{n} \mathrm{P50}_j , \quad (35)$$

and

$$\mathrm{P90}_S = \sum_{j=1}^{n} \mu_{PW_j} + 1.28\left(\sum_{i=1}^{n} \sigma_{PW_j}\right) = \sum_{j=1}^{n}\left(\mu_{PW_j} + 1.28\sigma_{PW_j}\right) = \sum_{j=1}^{n} \mathrm{P90}_j . \quad (36)$$

So, Eqs. 34-36 clearly show that the probabilistic sum is identical to the arithmetic sum if all fields are perfectly correlated so that Eq. 30 is valid.

In the case where we assume that the sum of n field log-normal resources is still log-normal, i.e., PW_S given by Eq. 26 will have a log-normal distribution. As discussed before, in practice with a finite number of fields, this assumption provides better results for $\mathrm{P10}_S$, $\mathrm{P50}_S$, and $\mathrm{P90}_S$. For this case, PW_S given by Eq. 26 follows a log-normal distribution

With mean equal to μ_{PW_S} (Eq. 27) and variance equal to $\sigma^2_{PW_S}$ (Eq. 28). These parameters are related to the mean $\mu_{\ln PW_S}$ and variance ($\sigma^2_{\ln PW_S}$) of lnPW_S by:

$$\mu_{\ln PW_S} = \ln \mu_{PW_S} - \frac{1}{2}\ln\left(1 + \frac{\sigma^2_{PW_S}}{\mu^2_{PW_S}}\right), \quad (37)$$

and

$$\sigma^2_{\ln PW_S} = \ln\left(1 + \frac{\sigma^2_{PW_S}}{\mu^2_{PW_S}}\right). \quad (38)$$

The P10, P50, and P90 of the PW_S, based on the assumption that PW_S is log-normal, are given by

$$\mathrm{P10}_S = \exp\left(\mu_{\ln PW_S} - 1.28\sqrt{\sigma^2_{\ln PW_S}}\right), \quad (39)$$

$$\mathrm{P50}_S = \exp\left(\mu_{\ln PW_S}\right). \quad (40)$$

and

$$\mathrm{P90}_S = \exp\left(\mu_{\ln PW_S} + 1.28\sqrt{\sigma^2_{\ln PW_S}}\right). \quad (41)$$

5.2 Example Applications

In this section, we consider a few example applications to verify our theoretical findings given in the previous section for aggregation of field resources of PW.

We consider a geothermal project consisting of 5 fields, each having its log-normal distribution of PW with its values of mean, variance, P10, P50, and P90 as given in Table 10.

Table 10: The values of mean, variance, 10th, 50th and 90th percentiles for each pseudo field's PW.

	μ_{PW} MW	σ^2_{PW} MW^2	P_{10} MW	P_{50} MW	P_{90} MW
Field 1	853	80656	538.2	808.7	1229
Field 2	404	25600	230.9	375.9	606.9
Field 3	97	1089	60.11	91.93	141.0
Field 4	51	729	21.67	42.92	82.85
Field 5	41	529	18.42	35.83	70.67

In Tables 11 and 13, we compare the values of mean, variance, P10, P50, P90 for the sum of field thermal resources of the 5 fields, obtained from arithmetic and probabilistic sum (based on MC sampling using @RISK, Eqs. 27, 28, 31-33, and Eqs. 37-41). The results given in Table 11 are for the case assuming that all 5 fields are independent. Table 12 presents the results for the case assuming that all 5 fields are perfectly correlated. Table 13 presents the results for the case assuming that all 5 fields are correlated with all pairwise correlation coefficients that are identical and equal to 0.05.

As expected, the results given in Tables 11-13 indicate that **(i)** arithmetic provides almost identical values of P10, P50, and P90 for the sum of 5 fields' resources if all fields are fully or perfectly correlated with each other, **(ii)** if all fields are independent or correlated with pairwise positive correlation coefficients different from unity, then the arithmetic sum will underestimate the value of P10 and P50, but overestimate the value of P90, and **(iii)** the probabilistic addition assuming normality for the sum of all fields' resources (Eqs. 28-30) does not provide as accurate estimates of P10, P50, and P90 as the probabilistic addition based on the assumption of log-normality for PW_S (Eqs. 36-38).

In summary, we can state that the arithmetic addition assumes that all fields resources considered in aggregation are perfectly correlated and provides a "pessimistic" estimate of P10. On the other hand, the probabilistic addition based on the assumption that all fields considered in aggregation are independent provides a "optimistic" estimate of P10. In reality, the correct P10 value should be between the P10 values estimated from the arithmetic sum which always assumes that all fields are perfectly correlated and the probabilistic addition based on the assumption that all fields are independent because fields are neither perfectly dependent nor perfectly independent, but instead are correlated. Of course, then, the issue is how to estimation the pairwise correlation coefficients of the fields involved in aggregation, though some authors have proposed simplified and pragmatic approaches for aggregating field resources [Carter and Morales (1998), van Elk *et al.* (2000), and Delfiner and Barrier (2008)], as discussed previously.

Table 11: The values of mean, variance, 10th, 50th and 90th percentiles for the sum of all 5 fields for the case all 5 fields are treated as independent.

	μ_{PW_S} MW	$\sigma^2_{PW_S}$ MW^2	P_{10} MW	P_{50} MW	P_{90} MW
Arithmetic sum	1446	108603	869	1355	2130
Probabilistic sum (MC)	1445	108931	1062	1404	1887
Probabilistic sum (Eqs. 27, 28, 31-33)	1446	108603	1024	1446	1868
Probabilistic sum (Eqs. 37-41)	1446	108603	1057	1410	1880

Table 12: The values of mean, variance, 10th, 50th and 90th percentiles for the sum of all 5 fields for the case all 5 fields are treated as perfectly correlated.

	μ_{PW_S} MW	$\sigma^2_{PW_S}$ MW^2	P_{10} MW	P_{50} MW	P_{90} MW
Arithmetic sum	1446	108603	869	1355	2130
Probabilistic sum (MC)	1446	278959	868	1352	2139
Probabilistic sum (Eqs. 27, 28, 31-33)	1446	280676	768	1446	2124
Probabilistic sum (Eqs. 37-41)	1446	280676	862	1358	2138

Table 13: The values of mean, variance, 10th, 50th and 90th percentiles for the sum of all 5 fields for the case all 5 fields are treated as correlated with identical pairwise correlation coefficients equal to 0.05.

	μ_{PW_S} MW	$\sigma^2_{PW_S}$ MW^2	P_{10} MW	P_{50} MW	P_{90} MW
Arithmetic sum	1446	108603	869	1355	2130
Probabilistic sum (MC)	1446	117696	1056	1406	1909
Probabilistic sum (Eqs. 27, 28, 31-33)	1446	117059	1008	1446	1884
Probabilistic sum (Eqs. 37-41)	1446	117059	1044	1407	1897

5. CONCLUSIONS

On the basis of this work, we conclude that:

1. The distribution of thermal resource power for a single geothermal field, based on a volumetric method, is log-normal, regardless of the types of probability distributions chosen for the input variables in the volumetric equation. This result follows directly from the fundamental theorem of statistics and probability – Central Limit Theorem (CLT).

2. Analytic uncertainty propagation equations (AUPEs) – based on a Taylor-series expansion around the mean values of the input variables – were presented for computing the mean and variance of the recoverable power resource for a field. The AUPM method, when combined with the assumption of log-normality for the recoverable power resource, provides a fast alternative to the Monte Carlo simulation for accurately characterizing uncertainty markers such as variance, P_{10}, P_{50} and P_{90}.

3. The derived AUPEs are quite general in that it can account for correlation among the input variables used in the volumetric equation. It was shown that ignoring correlation, if it exists, may underestimate or overestimate the uncertainty in recoverable power.

4. Finally, we showed that a simple arithmetic sum of the “proved” and “probable” (P_{10} and P_{90} percentiles, respectively) thermal power resources from individual fields assumes that all fields considered in aggregation are perfectly correlated and may significantly underestimate the true P_{10} and significantly overestimate the true P_{90}, obtained from the probabilistic sum accounting for pairwise correlations existing between the fields’ power resources. On the other hand, it was show that using a probabilistic sum based on the assumption that all fields involved in aggregation process are independent may overestimate the true P10 and underestimate the true P90 if some or all of the fields’ power resources are correlated.

REFERENCES

Arkan, S. and Parlaktuna, M.: “Resource Assessment of Balçova Geothermal Field,” proceedings World Geothermal Congress, Antalya, Turkey, 24-29 April (2005).

Barlow, R.J.: Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences, John Wiley & Sons, 204 (1989).

Brook, C.A., Mariner, R.H., Mabey, D.R., Swanson, J.r., Guffanti, M., and Muffler, L.J.P: “Hydrotermal Convection Systems with Reservoir Temperatures ≥90 oC,” Ed., Muffler, L.P.J., Assessment of Geothermal Resources of US, USGS Circular 790, 18-85 (1978).

Capen, E.C.: “The Difficulty of Assessing Uncertainty,” *Journal of Petroleum and Technology*, (August 1976) 843-850.

Capen, E.C.: “A Consistent Probabilistic Definition of Reserves,” *SPE Reservoir Engineering*, (February 1996) 23-28.

Capen, E.C: "Probabilistic Reserves! Here at Last?," *SPE Reservoir Evaluation & Engineering*, (October 2001) 387-394.

Carter, P.J. and Morales, E.: "Probabilistic Addition of Gas Reserves Within a Major Gas Project," paper SPE 50113, *proceedings 1998 SPE Asia Pacific Oil & Gas Conference and Exhibition*, Perth, Australia, 12-14 October (1998).

Coleman, H.W. and Steele, W.G.: *Experimentation and Uncertainty Analysis for Engineers*, 2nd edition, Wiley & Sons, 275 (1999).

Delfiner, P. and Barrier, R.: "Partial Probabilistic Addition: A Practical Approach for Aggregating Gas Resources," SPE Reservoir Evaluation and Engineering, (April 2008), 379-385.

Demirmen, F.: "Reserves Estimation: The Challenge for the Industry," *Journal of Petroleum and Technology*, (May 2007) 80-89.

Kalos, M.V. and Whitlock, P.A.: *Monte Carlo Methods*, Willey-Blackwell, 203 (2008).

Lovekin, J.: "Geothermal Inventory, US Geothermal Development," *GRC Bulletin*, November-December (2004), 242-244.

Muffler, L.J.P. and Cataldi, R.: "Methods for Regional Assessment of Geothermal Resources," *Geothermics*, **7**, 2-4 (1978), 53-89.

Murtha, J.A.: "Incorporating Historical Data Into Monte Carlo Simulation," *SPE Computer Applications*, (April 1994), 11-17.

Parzen, E.: Modern Probability Theory and Its Applications, John Wiley & Sons, 430 (1962).

@RISK Ver.4.5.5 software.: Palisade Corporation, NY, USA (2004).

Sanyal, S.K. and Sarmiento, Z.: "Booking Geothermal Energy Reserves," GRC *Transactions*, 29 (2005), 467-474.

Sarak, H., Türeyen, Ö.İ., and Onur, M.: Assessment of Uncertainty in Estimation of Stored and Recoverable Thermal Energy in Geothermal Reservoirs by Volumetric Methods., *proceedings 34th Workshop on Geothermal Reservoir Engineering*, Stanford University, USA, 9-11 February (2009).

Satman, A., Serpen, U., Onur, M.: *İzmir Balçova-Narlıdere Jeotermal Sahasının Rezervuar ve Üretim Performansı Projesi*, report prepared for Balçova Geothermal Inc., Izmir, Turkey (in Turkish) (2001).

Serpen, U.: "Estimation of Geothermal Field Potential by Stochastically Evaluating Stored Heat Model," Turkish Journal of Oil and Gas, 7 (2001), 37-43.

Serpen, U., Korkmaz, B.E.D., Satman, A.: "Power Generation Potentials of Major Geothermal Fields in Turkey," *proceedings 33nd Workshop on Geothermal Reservoir Engineering*, Stanford University, USA, 28-30 January (2008).

SPE (Society of Petroleum Engineers). "Petroleum Reserves and Resources Classification, Definitions, and Guidelines (draft), http:/www.spe.org/spe/jsp/basic/0,,1104_5806693,00.html (2006), Richardson, TX, USA.

van Elk, J. F., Vijayan, K., and Gupta, R.: "Probabilistic Addition of Reserves," paper SPE 64454 presented at the 2000 SPE Asia Pacific Oil & Gas Conference and Exhibition, Brisbane, Australia, 16-18 October (2000).

Welsh, M.B., Begg, S.H., and Bratvold, R.B.: "Modeling the Economic Impact of Cognitive Biases on Oil and Gas Decisions," paper SPE 110765 presented at the 2007 SPE Annual Technical Conference and Exhibition, Anaheim, CA, USA, 11-14 November (2007).

Zeybek, A.D., Onur, M., Türeyen, Ö.İ., Ma, M.S., Al-Shahri, A.M., and Kuchuk, F.J.: "Assessment of Uncertainty in Saturation Estimated from Archie's Equation," paper SPE 120517 to be presented at the 2009 SPE Middle East Oil Show and Conference, Bahrain, Kingdom of Bahrain, 15-18 March (2009).