PROCEEDINGS, Thirty-Fourth Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, February 9-11, 2009
SGP-TR-187

ASSESSMENT OF UNCERTAINTY IN ESTIMATION OF STORED AND RECOVERABLE THERMAL ENERGY IN GEOTHERMAL RESERVOIRS BY VOLUMETRIC METHODS

Hülya Sarak, Ö. İnanç Türeyen and Mustafa Onur

Istanbul Technical University
ITU Petrol ve Dogal Gaz Muhendisligi. Bolumu.
Maslak, Istanbul, 34469 TURKEY
e-mail: hulya@itu.edu.tr, inanct@itu.edu.tr, onur@itu.edu.tr

ABSTRACT

In this paper, we investigate the propagation of uncertainties in the input variables (used in the volumetric method) on to stored and recoverable thermal energy estimates calculated from volumetric methods. Effects of the different types of input distributions, correlation among input variables and cognitive biases are also investigated. 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. We show that AUPM can be used as a fast tool, without resorting to the MCM because the resulting distributions of stored and recoverable heat are always log-normal, which makes it possible for the results of the AUPM to accurately characterize the uncertainty. It is also shown that it is incorrect – a commonly made mistake – to add the "proved" and "probable" (which corresponds to P_{10} and P_{90} percentiles of the cumulative distribution function, respectively) thermal energy "reserves" from individual wells (or fields) to get "proved" field (or country) reserves. Applications on synthetic and field data cases are presented to demonstrate the methodology considered in this work.

INTRODUCTION

Uncertainty is inherent and ubiquitous in estimation of any type of reserves (oil, gas or heat) from underground energy systems. The thermal energy or power "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 reserve estimation problem always contain uncertainties to some degree that propagate into reserve 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 reserve (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 reserves 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 points 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; Arkan and Parlaktuna, 2005; Serpen et al., 2008). However, none of these works provide a deep investigation of and insight to 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, 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 are almost always 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 will tend to be 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 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, correlation among input variables, and cognitive biases regarding the input distributions (Capen, 1976; Capen, 2001; Welsh et al., 2007) propagate into the uncertainty of the estimated thermal energy or power. In addition, we show that it is incorrect – a common mistake – to add the "proved" thermal energy reserves from individual wells (or fields) to get "proved" field (or country) reserves, which has long been established by Capen (1996, 2001).

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.

Based on the volumetric method, the total stored heat in the reservoir is equal to the sum of the stored heat in the (solid) rock and water:

$$H_t = H_s + H_w, \tag{1}$$

where H_s and H_w are given, respectively, by

$$H_s = (1-\phi) c_s \rho_s A h (T_s - T_r), \tag{2}$$

and

$$H_w = \phi c_w \rho_w A h (T_w - T_r). \tag{3}$$

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

A area in m^2
c specific heat in kJ/(kg ^oC)
h net thickness in m
H stored heat in kJ
T temperature in ^oC
ϕ porosity (fraction)
ρ density in kg/m^3

A few remarks are in order for Eqs. 2 and 3: Eqs. 2 and 3 are given by Muffler and Cataldi (1978) who consider different temperatures for the solid rock and fluid components of the reservoir. In Eqs. 2 and 3, however, we will assume that a local thermal equilibrium is always valid in the reservoir so that solid rock and fluid temperatures are identical in Eqs. 2 and 3; i.e., $T_s = T_w = T$. Hence, using Eqs. 2 and 3 with this assumption in Eq. 1 yields:

$$H_t = \left[(1-\phi) c_s \rho_s + \phi c_w \rho_w\right] A h (T - T_r). \tag{4}$$

H_t defined by Eq. 4 is usually referred to as the accessible resource base and it can be converted to recoverable power (in MW) by the following equation (Serpen, 2001; Arkan and Parlaktuna, 2005):

$$PW = \frac{H_t R_F Y}{10^3 L_F t_p}. \tag{5}$$

The variables used in Eq. 5 and their units are given below:

H_t total stored heat in kJ
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.)

Here, 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 worth noting that the stored heat in the solid rock is typically an order of magnitude larger than the stored heat in the water. In other words, the ratio of stored heat in the solid rock to the total stored heat in the solid rock plus water, i.e., H_s/H_t, is nearly 0.80 to 0.90 (or in percentage 80 to 90%). Note that in the case of hot dry rock, this ratio is unity because ϕ is zero for hot dry rock applications. In other words, for most of the applications, $H_t \approx H_s$. This observation indicates that ignoring the stored heat in water in computing H_t and PW from Eqs. 4 and 5 may not introduce significant underestimation in H_t and PW.

It is also worth noting that Eqs. 4 and 5 are valid for predicting the recoverable power for both direct-heating and power (electricity) generation applications 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 H_t (Eq. 4) and PW (Eq. 5), 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 stored heat and the recoverable power) results from our lack of knowledge in most of the input variables in Eqs. 4-5. 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 H_t and PW are log-normal regardless of the types of probability distributions chosen for the input variables.

When computing H_t from Eq. 4, we treat H_t as a function of eight random, input variables; namely, A, h, ϕ, c_s, c_w, ρ_s, ρ_w, (T-T_r). When computing PW from Eq. 5, we treat PW as a function of eleven random input variables; the same eight input variables given previously for the H_t plus three more input variables; namely R_F, L_F, and Y. In our applications, we fix T_r at a constant value, but the variable (T-T_r) in Eq. 4 will be treated as a random variable because T in Eq. 4 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-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. 5.

Some input variables involved in Eqs. 4 and 5 can be expected to 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. 4. 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. 4. 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 H_t and PW. 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.

QUANTIFICATON OF UNCERTAINTY

In this section, we review some basic equations and methods used for quantification of uncertainty in the estimation of H_t and PW by the volumetric method (Eq. 1-5). We first consider the Monte Carlo method (MCM) and then the analytic uncertainty propagation method (AUPM).

However, before getting into details of these methods, we first note that Eq. 4 for estimating H_t involves a product of four random variables; $\left[(1-\phi)c_s\rho_s + \phi c_w\rho_w\right]$, A, h, and (T-T_r), whereas Eq. 5 for estimating PW involves a product of three random variables; H_t, R_F and Y, and a quotient of the random variable L_F. As H_t involves a product of four random variables, then PW is a product of seven random variables and a quotient of one random variable.

If we take the natural logarithm of PW given by Eq. 5 and rearranging the resulting equation so that the right-hand side of the resulting equation can only be expressed in terms of the input random variables, we obtain:

$$\ln PW + \ln(10^3 t_p) = \ln H_t + \ln R_F + \ln Y - \ln L_F \,, \tag{6}$$

where $\ln H_t$, which follows from E. 4, is given by:

$$\ln H_t = \ln\left[(1-\phi)c_s\rho_s + \phi c_w\rho_w\right] + \ln A + \ln h + \ln\left(T - T_r\right) \,. \tag{7}$$

Eqs. 6 and 7 clearly indicate that $\ln PW$ and $\ln H_t$ can be written as a sum of the natural logarithms of all independent random variables. If all these random variables are treated as independent, then it follows from the CLT, discussed previously, that the resulting distribution of $\ln H_t$ or $\ln PW$ will tend to be normal. It is important to note that this is true no matter what type of distribution the input random variables assume.

If $\ln H_t$ and $\ln PW$ will tend to be normal, then the corresponding distributions of H_t and PW will tend to be log-normal. We would expect the distribution of PW to be more like log-normal than that of H_t because PW involves more products/quotients than does H_t.

Note that the CLT promises that $\ln H_t$ and $\ln PW$ be normal if all the random variables are independent. Hence, we may not claim that $\ln H_t$ and $\ln PW$ be still normal if some of the input random variables are not independent. However, our results to be shown later indicate that the resulting distributions of $\ln H_t$ and $\ln PW$ still tend to be normal even if some of the input variables are treated as dependent.

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

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 H_t given by Eq. 4 or PW given by Eq. 5. In the applications to be given, we perform Monte Carlo simulations by using @RISK,™ spreadsheet-based software (2004).

Analytical Uncertainty Propagation Method (AUPM)

Here, we derive an uncertainty propagation equation for a function f (such as H_t in Eq. 4 or PW in Eq. 5 or their natural logarithms given by Eqs. 6 and 7) 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 and modeling errors. 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.

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 variances of H_t and PW given by Eqs. 4 and 5. As can be seen from these equations, H_t is, in general, a nonlinear function of the variables A, h, ϕ, c_s, c_w, ρ_s, ρ_w, $(T\text{-}T_r)$. Because H_t is nonlinear, then PW given by Eq. 5 is also a nonlinear function of the same input variables plus the variables R_F, L_F and Y. As noted before, if we work, however, with the natural logarithms of H_t and PW (Eqs. 7 and 6), then we obtain "nearly linearized" equations for H_t and

PW. We say "nearly linearized" because lnH_t and ln*PW* is still nonlinear, but now are "weakly" nonlinear. For example, lnH_t (Eq. 7) is more linear compared to H_t (Eq. 4) because lnH_t is only nonlinear with respect to the variables; ϕ, c_s, c_w, ρ_s, ρ_w, but is linear with respect to ln$A$, ln$h$, ln(T-T_r), whereas H_t is nonlinear with all of these eight random variables. Similar arguments are valid for *PW* and ln*PW* functions. Therefore, as to be shown later, we obtain more accurate estimates of the variances of lnH_t and ln*PW* if we apply the AUPM directly to lnH_t (Eq. 7) and ln*PW* (Eq. 6) instead of applying it to H_t (Eq. 4) and *PW* (Eq. 5).

One final remark is that if the porosity is nearly zero (e.g., in the case of hot dry rock), then lnH_t given by Eq. 7 can be written as:

$$\ln H_t = \ln \rho_s + \ln c_s + \ln A + \ln h + \ln\left(T - T_r\right). \quad (8)$$

So, in this case, lnH_t is a linear function of the natural logarithms of the input variables; ρ_s, c_s, A, h, and T-T_r. Using Eq. 8 in Eq. 6 yields:

$$\begin{aligned}\ln PW + \ln(10^3 t_p) = {} & \ln \rho_s + \ln c_s + \ln A + \ln h \\ & + \ln\left(T - T_r\right) + \ln R_F + \ln Y \\ & - \ln L_F\end{aligned}, \quad (9)$$

which indicates that ln*PW* is a linear function of the natural logarithms of the input variables; ρ_s, c_s, A, h, T-T_r, R_F, Y, and L_F.

AUP Equations for H_t (Eq. 4) and PW (Eq. 5)

Let's consider a random function f of M variables, X_i, i =1,2,…,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:

$$\begin{aligned}f\left(X_1, X_2, \ldots X_M\right) = {} & f\left(\mu_{X_1}, \mu_{X_2}, \ldots, \mu_{X_M}\right) \\ & + \sum_{i=1}^{M}\left(X_i - \mu_{X_i}\right)\left(\frac{\partial f}{\partial X_i}\right)\Bigg|_{X_i=\mu_{X_i}, i=1,\ldots,M}\end{aligned}. \quad (10)$$

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

$$\mu_f = f\left(\mu_{X_1}, \mu_{X_2}, \ldots, \mu_{X_M}\right), \quad (11)$$

and

$$\sigma_f^2 = \sum_{i=1}^{M}\theta_i^2\sigma_{X_i}^2 + 2\sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\theta_i\theta_j cov(X_i, X_j), \quad (12)$$

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. 12 in terms of the correlation coefficient as:

$$\sigma_f^2 = \sum_{i=1}^{M}\theta_i^2\sigma_{X_i}^2 + 2\sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\theta_i\theta_j\rho_{X_i,X_j}\sigma_{X_i}\sigma_{X_j}. \quad (13)$$

In Eqs. 12 and 13, θ_i is the derivative of f with respect to the variable X_i, i.e.,

$$\theta_i = \left(\frac{\partial f}{\partial X_i}\right)\Bigg|_{X_i=\mu_{X_i}, i=1,\ldots,M}. \quad (14)$$

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. 12-14) that the uncertainty propagation into 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 (Eqs. 4-7).

We can derive several useful forms of AUPEs. For example, if we divide both sides of Eq. 13 by μ_f^2, we obtain:

$$\begin{aligned}\left(\frac{\sigma_f}{\mu_f}\right)^2 = {} & \sum_{i=1}^{M}\left(\frac{\mu_{X_i}\theta_i}{\mu_f}\right)^2\left(\frac{\sigma_{X_i}}{\mu_{X_i}}\right)^2 \\ & + 2\sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\left(\frac{\mu_{X_i}\theta_i}{\mu_f}\right)\left(\frac{\mu_{X_j}\theta_j}{\mu_f}\right)\rho_{X_i,X_j}\left(\frac{\sigma_{X_i}}{\mu_{X_i}}\right)\left(\frac{\sigma_{X_j}}{\mu_{X_j}}\right)\end{aligned}, \quad (15)$$

where σ_f/μ_f is called the relative uncertainty in f and σ_{X_i}/μ_{X_i} represents the relative uncertainty for the variable X_i, and simply denoted by RU_i. It is worth noting that the term $\mu_{X_i}\theta_i/\mu_f$ in Eq. 15 represents the normalized (or dimensionless) sensitivity of f with respect to the variable X_i and is explicitly computed from:

$$\mu_{X_i}\frac{\theta_i}{\mu_f} = \frac{\mu_{X_i}}{\mu_f}\left(\frac{\partial f}{\partial X_i}\right)\Bigg|_{X_i=\mu_{X_i}, i=1,\ldots,M}. \quad (16)$$

Another useful measure of the total uncertainty can be obtained by multiplying both sides of Eq. 15 by $\left(\mu_f/\sigma_f\right)^2$ and then rearranging the resulting equation:

$$\sum_{i=1}^{M}\frac{\left(\frac{\mu_{X_i}\theta_i}{\mu_f}\right)^2}{\left(\frac{\sigma_f}{\mu_f}\right)^2}\left(\frac{\sigma_{X_i}}{\mu_{X_i}}\right)^2 + 2\sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\frac{\left(\frac{\mu_{X_i}\theta_i}{\mu_f}\right)\left(\frac{\mu_{X_j}\theta_j}{\mu_f}\right)}{\left(\frac{\sigma_f}{\mu_f}\right)^2}\left(\frac{\sigma_{X_i}}{\mu_{X_i}}\right)\left(\frac{\sigma_{X_j}}{\mu_{X_j}}\right)\rho_{X_i,X_j} = 1 \quad (17)$$

Eq. 17 can also be simply rewritten as:

$$\sum_{i=1}^{M} UPC_i + \sum_{i=1}^{M-1}\sum_{j=i+1}^{M} UPC_{i,j} = 1, \quad (18)$$

where we define UPC_i as a measure of the individual contribution (in fraction) of each variable's uncertainty to the total uncertainty (or variance) in f:

$$UPC_i = \theta_i^2\left(\frac{\sigma_{X_i}^2}{\sigma_f^2}\right) = \frac{\left(\frac{\mu_{X_i}\theta_i}{\mu_f}\right)^2}{\left(\frac{\sigma_f}{\mu_f}\right)^2}\left(\frac{\sigma_{X_i}}{\mu_{X_i}}\right)^2, \quad (19)$$

and define $UPC_{i,j}$ as a measure of the contribution of the correlated pairs X_i and X_j (in fraction) to the total uncertainty (or variance) in f:

$$UPC_{i,j} = 2\left(\frac{\sigma_{X_i}}{\sigma_f}\theta_i\right)\left(\frac{\sigma_{X_j}}{\sigma_f}\theta_j\right)\rho_{X_i,X_j} = 2\frac{\left(\frac{\mu_{X_i}}{\mu_f}\theta_i\right)\left(\frac{\mu_{X_j}}{\mu_f}\theta_j\right)}{\left(\frac{\sigma_f}{\mu_f}\right)^2}\left(\frac{\sigma_{X_i}}{\mu_{X_i}}\right)\left(\frac{\sigma_{X_j}}{\mu_{X_j}}\right)\rho_{X_i,X_j} \quad (20)$$

Note that UPC_i (Eq. 19) is always positive, but $UPC_{i,j}$ (Eq. 20) can be positive or negative depending on the signs of the normalized sensitivities (Eq. 16) and the correlation coefficient $\rho_{X_iX_j}$. If, and only if, all variables are uncorrelated, all $UPC_{i,j}$ are zero.

A few important remarks (from Eqs. 13-20) are in order: If the variables are completely independent, the total uncertainty in f will be dominated more by the variables having the larger products of normalized sensitivities and relative uncertainties. So, for this case, one only needs to inspect the product of normalized sensitivity and relative uncertainty, or UPC_i for each of the variables to identify which of the variables contribute more to the total uncertainty in f. If the variables are correlated, then one should compute UPC_{ij}s in addition to UPC_i to identify which of the variables and the correlated pairs contribute more to the total uncertainty in f.

It may not be obvious from Eqs. 13-20 to see how correlation among variables propagates into uncertainty in f because in the case of correlation, the uncertainty in f (Eq. 13) depends on **(i)** the number of correlated pairs, **(ii)** magnitude and sign of the correlation coefficient between the correlated pairs, and **(iii)** magnitude and sign of the sensitivity of f to the correlated variables. So, the uncertainty in f can increase or decrease with correlation. However, how correlations among variables propagate into uncertainty in f can be easily identified by calculating and then inspecting the UPC_i and UPC_{ij} given by Eqs. 19 and 20, derived from the error propagation equation accounting for correlation among the variables.

As mentioned previously, for the problem of interest, f in Eqs. 10-20 could represent H_t given by Eq. 4 or PW given by Eq. 5. The sensitivities of H_t and PW (i.e., θ_is) required in Eqs. 12-20 can be obtained by analytical differentiation of Eqs. 4 and 5. For example, the sensitivity of H_t to a given variable in Eq. 4 is obtained by differentiating Eq. 4 with respect to that variable. These sensitivities are tabulated in Table 1. To compute uncertainty propagation in PW given by Eq. 5 from Eqs. 12-20, then we first compute the sensitivities from the formulas given in Table 2.

Table 1: Sensitivity of H_t (Eq. 4) with respect to a given variable X_i in Eq. 4.

Variable X_i	$\boldsymbol{\theta_i = \partial H_t/\partial X_i^*}$
ϕ	$\left(-\mu_{c_s}\mu_{\rho_s} + \mu_{c_w}\mu_{\rho_w}\right)\mu_A\mu_h\mu_{(T-T_r)}$
c_s	$\left(1-\mu_\phi\right)\mu_{\rho_s}\mu_A\mu_h\mu_{(T-T_r)}$
ρ_s	$\left(1-\mu_\phi\right)\mu_{c_s}\mu_A\mu_h\mu_{(T-T_r)}$
c_w	$\mu_\phi\mu_{\rho_w}\mu_A\mu_h\mu_{(T-T_r)}$
ρ_w	$\mu_\phi\mu_{c_w}\mu_A\mu_h\mu_{(T-T_r)}$
A	$\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s} + \mu_\phi\mu_{c_w}\mu_{\rho_w}\right]\mu_h\mu_{(T-T_r)}$
h	$\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s} + \mu_\phi\mu_{c_w}\mu_{\rho_w}\right]\mu_A\mu_{(T-T_r)}$
T-T_r	$\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s} + \mu_\phi\mu_{c_w}\mu_{\rho_w}\right]\mu_A\mu_h$
*****evaluated at the mean values of the variables X_is	

Table 2: Sensitivity of PW (Eq. 5) with respect to a given variable X_i in Eq. 5.

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	$\dfrac{\left[\left(1-\mu_\phi\right)\mu_{c_s}\mu_{\rho_s}+\mu_\phi\mu_{c_w}\mu_{\rho_w}\right]\mu_A\mu_h\mu_{(T-T_r)}\mu_{R_F}\mu_Y}{-10^3 t_p\mu_{L_F}^2}$
*****evaluated at the mean values of the variables X_is	

AUP Equations for* $\ln H_t$ *(Eq. 7) and* $\ln PW$ *(Eq. 6)

In the case where we consider uncertainty propagation in $\ln f$ (e.g, $\ln f = \ln H_t$ given by Eq. 7 or $\ln f = \ln PW$ given by Eq. 6) instead of f, then we may use two different approaches to derive AUP equations. These approaches yield two different AUP equations for estimating the variance of $\ln H_t$ or $\ln PW$.

Approach 1: The first approach is based on the Taylor series expansion of $\ln H_T$ (or $\ln PW$) around the mean values of the input variables μ_{X_i}s . For this case, the AUP equations are given by Eqs. 10-20 with f replaced by $\ln f$. Hence, the sensitivities θ_is required in Eqs. 12-20 are given by:

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

If we consider uncertainty propagation in $\ln f = \ln H_t$, given by Eq. 7, then the sensitivity of $\ln H_t$ with respect to X_i is simply obtained by using the chain rule as:

$$\frac{\partial \ln H_t}{\partial X_i} = \frac{1}{H_t}\frac{\partial H_t}{\partial X_i}\Bigg|_{X_i=\mu_{X_i},i=1,\ldots,M} , \tag{22}$$

where X_i and $\partial H_t / \partial X_i$ are tabulated in the first and second columns of Table 1.

Similarly, if we consider uncertainty propagation in $f = \ln PW$ given by Eq. 6, then the sensitivity of $\ln PW$ with respect to X_i is obtained by using the chain rule as:

$$\frac{\partial \ln PW}{\partial X_i} = \frac{1}{PW}\frac{\partial PW}{\partial X_i}\Bigg|_{X_i=\mu_{X_i},i=1,\ldots,M} , \tag{23}$$

where X_i and $\partial PW / \partial X_i$ are tabulated in the first and second columns of Table 2. Table 3 and 4 present the formulas for the derivatives for $\ln H_t$ and $\ln PW$ with respect to X_i, respectively.

Table 3: Sensitivity of $\ln H_t$ *(Eq. 7) with respect to a given variable X_i in Eq. 7.*

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

Table 4: Sensitivity of ln*PW (Eq. 6) with respect to a given variable X_i in Eq. 6.*

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

Approach 2: The second approach is based on the Taylor series expansion of lnH_t (or 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. 10-20 with *f* replaced by ln*f*, X_is by lnX_is, and μ_{X_i}s by $\mu_{\ln X_i}$s . Hence, for this approach, the sensitivities θ_is required in Eqs. 12-20 are to be computed from:

$$\theta_i = \left(\frac{\partial \ln f}{\partial \ln X_i} \right)\Bigg|_{X_i=\mu_{X_i},\, i=1,\ldots,M} . \qquad (24)$$

If we consider uncertainty propagation in ln*f* = lnH_t, then the sensitivity of lnH_t with respect to lnX_i is simply obtained by using the chain rule as:

$$\frac{\partial \ln H_t}{\partial \ln X_i} = \frac{X_i}{H_t}\frac{\partial H_t}{\partial X_i}\Bigg|_{X_i=\mu_{X_i},\, i=1,\ldots,M} . \qquad (25)$$

Similarly, if we consider uncertainty propagation in ln*f* = ln*PW* given by Eq. 6, then the sensitivity of ln*PW* with respect to lnX_i is obtained by using the chain rule as:

$$\frac{\partial \ln PW}{\partial \ln X_i} = \frac{X_i}{PW}\frac{\partial PW}{\partial X_i}\Bigg|_{X_i=\mu_{X_i},\, i=1,\ldots,M} . \qquad (26)$$

The derivatives (Eqs. 25 and 26) are simply computed by multiplying the derivatives given in the second columns of Table 3 and 4 by the mean value of the variable interest, μ_{X_i} .

Our numerical results indicate that Approach 2 provides slightly better estimate of the variance than Approach 1. However, unlike the Approach 1, the Approach 2 for estimating the variances of lnH_t and ln*PW* will require 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 $\rho_{\ln X_i,\ln X_j}$. All these require an additional effort to compute the mean $\mu_{\ln X_i}$ and variance $\sigma^2_{\ln X_i}$ from a given distribution of the input variable X_i. If the distribution of X_i is 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 $\sigma^2_{\ln X_i}$, given by:

$$\mu_{\ln X_i} = \ln \mu_{X_i} - \frac{1}{2}\ln\left(1+\frac{\sigma^2_{X_i}}{\mu^2_{X_i}}\right), \qquad (27)$$

and

$$\sigma^2_{\ln X_i} = \ln\left(1+\frac{\sigma^2_{X_i}}{\mu^2_{X_i}}\right). \qquad (28)$$

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 $\sigma^2_{\ln X_i}$.

Estimation of 10th, 50th and 90th percentiles of H_t, PW,* lnH_t *and* ln*PW

As discussed previously and to be shown with numerical results later, the resulting distributions of H_t and *PW* are nearly log-normal due to the CLT. Consequently, the distributions of lnH_t and ln*PW* are nearly normal. Hence, using this observation, we can

compute other statistical measures such as P_{10}, P_{50}, and P_{90} for characterizing uncertainty in H_t or PW in addition to the mean and variance determined from the AUPM. When using the AUPM, we first compute natural logs of P_{10}, P_{50}, and P_{90}; i.e., $\ln P_{10}$, $\ln P_{50}$, and $\ln P_{90}$ from the following formulas given for a normal distribution (also see Capen, 2001):

$$\ln P_{10} = \mu_{\ln f} - 1.28\sigma_{\ln f}\,, \tag{29}$$

$$\ln P_{90} = \mu_{\ln f} + 1.28\sigma_{\ln f}\,, \tag{30}$$

and

$$\ln P_{50} = \frac{\ln P_{10} + \ln P_{90}}{2}\,, \tag{31}$$

where ln$f$ represents either lnH_t or lnPW.

The 1.28 in Eqs. 29 and 30 stems from the fact that for a normal distribution, 10 or 90 percentile lies 1.28 standard deviation units away from the mean.

It is important to note that how we compute the natural log of percentiles from Eqs. 29-31, which require the values of $\mu_{\ln f}$ and $\sigma^2_{\ln f}$ (square root of the variance yields the standard deviation $\sigma_{\ln f}$) from the AUPM. If we directly apply AUPM to H_t (Eq. 4) or PW (Eq. 5), then we can estimate the mean and variance of H_t or PW; i.e., μ_f and σ^2_f, where f represents either H_t or PW. Then, using these values in Eq. 27 and 28 (where X_i is replaced by f) because f is log-normal due to the CLT, we can estimate $\mu_{\ln f}$ and $\sigma^2_{\ln f}$. As mentioned previously, we do not recommend using this approach because applying the AUPM directly to H_t or PW does not provide as accurate estimates of μ_f and σ^2_f as applying it to lnH_t or lnPW due to nonlinearity in H_t or PW.

Therefore, we estimate directly the mean and variance of lnH_t or lnPW; i.e., $\mu_{\ln f}$ and $\sigma^2_{\ln f}$, where f represents either H_t or PW, by using the AUPM method based on either Approach 1 or 2, discussed previously. Then, we compute $\ln P_{10}$, $\ln P_{50}$, and $\ln P_{90}$ by using Eqs. 29-31. We finally exponentiate the computed values of $\ln P_{10}$, $\ln P_{50}$ and $\ln P_{90}$ to obtain P_{10}, P_{50} and P_{90}, respectively, for H_t and PW.

RESULTS

Here, we present our results obtained by considering a few examples to make our points.

Uncorrelated Case

First, we consider a case where all input variables are independent in H_t and PW given by Eqs. 4 and 5. For the purpose of this example, we choose to model all distributions with a triangular distribution. As discussed previously, in fact, there is no reason to insist upon any particular probability distribution for our input variables. The minimum, most likely (mode), and maximum values of the input variables are given in Table 5. The values of mean and variance given in Table 5 were computed from the well-known formulas for a triangular distribution:

$$\mu_{X_i} = \frac{Min + Max + Mode}{3}\,, \tag{32}$$

and

$$\sigma^2_{X_i} = \frac{(Min)^2 + (Max)^2 + (Mode)^2}{18} - \frac{(Min \times Max + Min \times Mode + Max \times Mode)}{18}\,. \tag{33}$$

The data given in Table 5 pertain to Izmir Balçova-Narlıdere geothermal field in Turkey and were taken from Satman et al. (2001). For this application, T_r = 60 °C.

Table 5: Distributions of the input variables.

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. 32 and 33.

Figures 1-4 show histograms of H_s (Eq. 2), H_w (Eq. 3), H_t (Eq. 4), and PW (Eq. 5) generated from the MCM by using the distributions given in Table 5 in @RISK. The statistical variables (e.g., mean, variance, P_{10}, P_{50}, and P_{90}) for each histogram are given in the insets of Figures 1-4.

As is expected from the CLT, all histograms of Figures 1-4 are nearly log-normal. This can be further verified by simply sorting the computed values in ascending order and plotting the resulting values on a log-normal probability paper. If the points fall close to a straight line, then the distribution can be considered as log-normal. Figure 5 shows a plot of PW on a log-normal probability paper. The result of Figure 5 verifies that PW is reasonably log-normal. Although not shown here, plots of H_s, H_w, and H_t on a log-normal probability paper indicate that these variables are also reasonably log-normal.

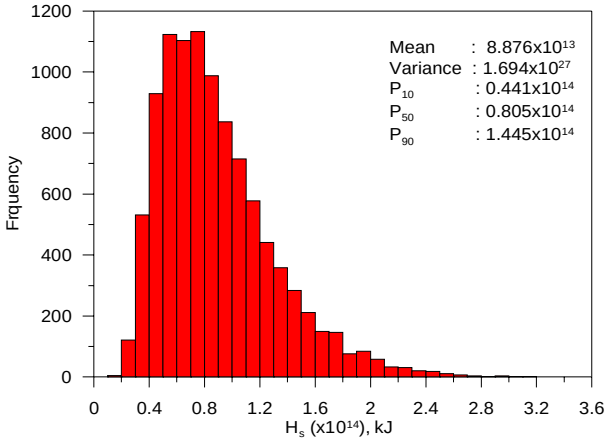

Figure 1: Histogram of stored heat in solid rock, H_s, generated from the MCM.

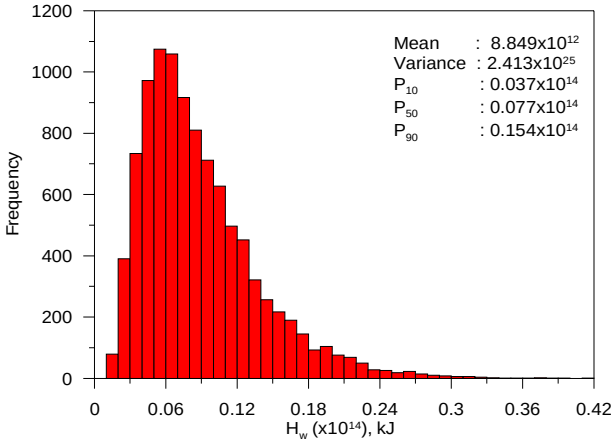

Figure 2: Histogram of stored heat in water, H_w, generated from the MCM.

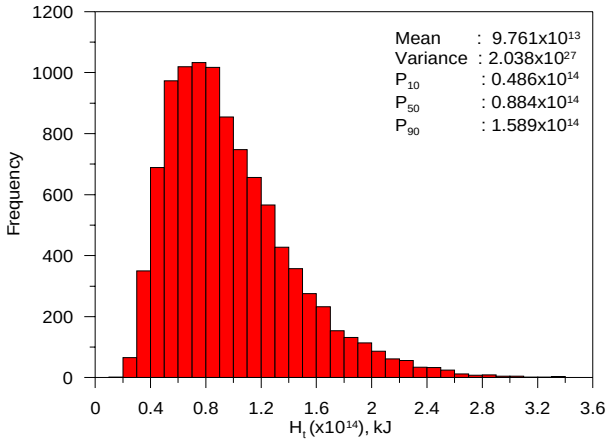

Figure 3: Histogram of total stored heat in solid rock plus water, H_t, generated from the MCM.

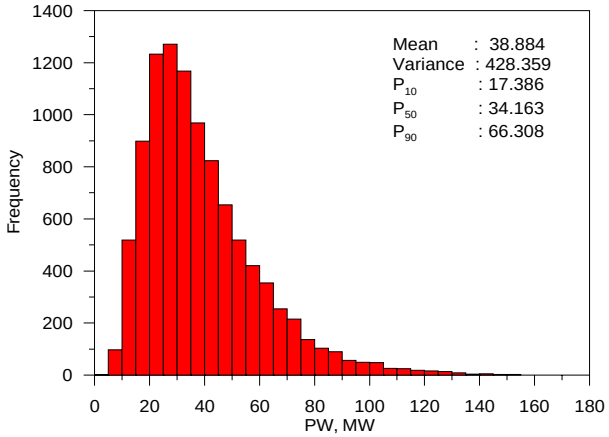

Figure 4: Histogram of recoverable power, PW, generated from the MCM.

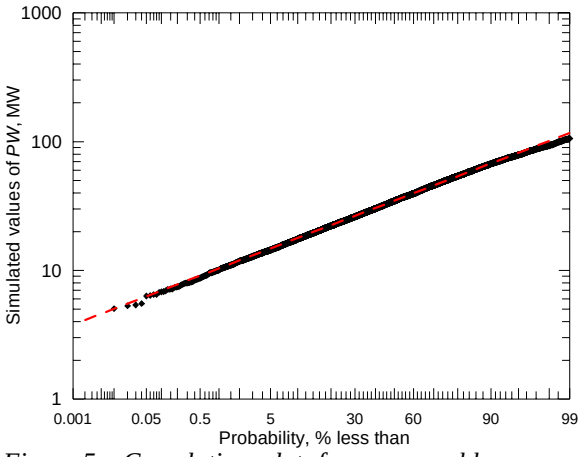

Figure 5: Cumulative plot for recoverable power, PW, generated from the MCM.

It is worth noting that the total stored H_t is the sum of stored heat in the solid rock and the stored heat in water occupying the "pore" spaces. Inspecting the histograms of H_s (Fig. 1), H_w (Fig. 2) and H_t (Fig. 3),

we see that the contribution of H_w to H_t is small, and H_t is dominated mainly by H_s, stored heat in the solid matrix. In fact, as discussed previously, this is not surprising because most of the heat is normally stored in the solid rock in geothermal reservoirs.

It may be also worth noting that as expected from a statistical point of view, the sum of the means of the distributions of H_w and H_s is exactly equal to the mean of the distribution of H_t. However, we see that the sum of individual variances of H_w and H_s distributions, though they are close and they agree within 15%, is not exactly equal to the variance of the distribution of H_t. It is actually less than the variance of the distribution of H_t. The reason for this is due to the fact that H_w and H_s are somewhat (positively) correlated with each other due to the common variables (ϕ, A, h, and T-T_r) involved in sampling the distributions of H_w and H_s. Hence the variance of the distribution is different (greater) from (than) the sum of the individual variances of H_w and H_s distributions. This indicates that we can simply add the means, but not the variances of individual distributions for estimating the mean and variance of a random function involving the sum of two individual random distributions unless they are uncorrelated.

Now, we compare the estimates of means, variances, P_{10}, P_{50}, and P_{90} computed for H_t, PW, lnH_t, lnPW from the MCM and AUPM. As mentioned previously, when we use AUPM, we can consider two different approaches for estimating the variances of lnH_t and lnPW functions. The second approach (Approach 2) requires that we should work with the mean and variances of lnX_is instead of X_is. Table 6 gives $\mu_{\ln X_i}$ and $\sigma^2_{\ln X_i}$ values – computed from descriptive statistics after applying natural log transformation to each X_i – for the input variables listed in Table 5.

Table 6: Mean and Variance of X_i and lnX_i.

Variable X_i	μ_{X_i}	$\sigma^2_{X_i}$	$\mu_{\ln X_i}$	$\sigma^2_{\ln X_i}$
ϕ	0.057	2.722×10^{-4}	-2.919	9.324×10^{-2}
c_s, kJ/(kg °C)	0.883	2.639×10^{-3}	-0.127	3.437×10^{-3}
ρ_s, kg/m^3	2650	1.667×10^{3}	7.882	2.382×10^{-4}
c_w, kJ/(kg °C)	4.130	2.150×10^{-3}	1.418	1.261×10^{-4}
ρ_w, kg/m^3	946.7	2.067×10^{2}	6.853	2.264×10^{-4}
A, m^2	1.1×10^{6}	1.006×10^{11}	13.896	8.104×10^{-2}
h, m	533.3	2.764×10^{4}	6.236	9.630×10^{-2}
T-T_r, °C	66.67	9.306×10^{1}	4.200	2.296×10^{-2}
R_F	0.163	1.239×10^{-3}	-1.837	5.487×10^{-2}
Y	0.817	1.806×10^{-3}	-0.204	2.854×10^{-3}
L_F	0.42	9.500×10^{-4}	-0.871	5.314×10^{-3}
t_p, s.	8×10^{8}	0.0	20.5	0.0

Table 7 compares the values of means and variances computed from MCM and AUPM for H_t, PW, lnH_t and lnPW. The computed values of mean and variance from the MCM and AUPM for H_t, PW, lnH_t, and lnPW functions agree well (Table 7). We also notice that the values of mean and variance for lnH_t and lnPW computed from the AUPM based on Approach 1 and 2 are nearly identical to those estimated from the MCM. The means and variances computed from the MCM and AUPM for H_t and PW functions, however, do not agree as good as those computed for lnH_t and lnPW functions. 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, both H_t (Eq. 4) and PW (Eq. 5) are in fact nonlinear functions of their input variables. On the other hand, lnH_t (Eq. 7) and lnPW (Eq. 6) are almost linear (or weakly nonlinear) functions of the input variables. Consequently, the AUPM (based on either Approach 1 or 2) provides estimates of means and variances for lnH_t and lnPW functions that agree very well with those computed from the MCM. Notice that AUPM based on Approach 2 provides estimates essentially identical to those from the MCM.

Table 7: A comparison of means and variances from the MCM and AUPM for H_t, PW, lnH_t, *and* lnPW *functions.*

	Mean		**Variance**	
	MCM	AUPM	MCM	AUPM
***H_t*, kJ**	9.76×10^{13}	9.50×10^{13}	2.04×10^{27}	1.85×10^{27}
***PW*, MW**	38.9	37.7	428.4	367.4
lnH_t	32.11	32.19† 32.11*	0.2051	0.2044† 0.2045*
lnPW	3.529	3.629† 3.531*	0.2681	0.2590† 0.2680*
†AUPM based on Approach 1 *AUPM based on Approach 2				

Next, we compute the 10th, 50th and 90th percentiles for H_t, PW, lnH_t and lnPW functions. As mentioned previously, to compute these percentiles when using the AUPM, we apply the AUPM (based on either Approach 1 or 2) directly to lnH_t or lnPW to estimate $\mu_{\ln f}$ and $\sigma^2_{\ln f}$, where f represents either H_t or PW. Then we use these values of mean and variance in Eqs. 29-31 to compute lnP_{10}, lnP_{50}, and lnP_{90}. Exponentiating the values of lnP_{10}, lnP_{50}, and lnP_{50} yields P_{10}, P_{50}, and P_{90}, respectively, for the H_t or PW

function. Table 8 compares the values of P_{10}, P_{50}, and P_{90} computed from the MCM and AUPM (based on Approach 1 and 2). The values of P_{10}, P_{50}, and P_{90} computed from the AUPM for H_t and PW (Table 8) were computed with the procedure just mentioned above. The values of P_{10}, P_{50} and P_{90} given in the 5th and 6th rowa of Table 8 for $\ln H_t$ and $\ln PW$ represent the values $\ln P_{10}$, $\ln P_{50}$ and $\ln P_{90}$ computed from Eqs. 29-31.

The values of P_{10}, P_{50}, and P_{90} percentiles computed from the MCM and AUPM agree well (Table 8). The maximum relative difference for the MCM and AUPM based on Approach 1 is around 10%, whereas the maximum relative difference for the MCM and AUPM based on Approach 2 is around 2%.

Another nice feature of the AUPM is that we can identify and quantify which of the input variables contribute more to the total uncertainty in H_t, $\ln H_t$, PW, or $\ln PW$. Because $\ln H_t$ and $\ln PW$ are nearly linear functions of the input variables as discussed previously, we consider $\ln H_t$ and $\ln PW$ to identify which of the input variables contribute more to the total uncertainty in $\ln H_t$ and $\ln PW$. For this purpose, as the input variables are treated as independent random variables in this example, we can simply use UPC_i given by Eq. 19, the relative measure of the individual (fractional) contribution of each variable's uncertainty to the total uncertainty (or variance) in $\ln H_t$ or $\ln PW$. Table 9 tabulates the values of UPC_i along with the relative uncertainty (RU_i) and the products of normalized sensitivity (NS_i) and relative uncertainty for the input variables involved in $\ln H_t$. Table 10 present the results for the input variables involved in $\ln PW$. In Tables 9 and 10, RU_i, NS_i, and UPC_i are defined as:

$$RU_i = \frac{\sigma_{\ln X_i}}{\mu_{\ln X_i}}, \tag{34}$$

$$NS_i = \frac{\mu_{\ln X_i}}{\mu_{\ln f}} \frac{\partial \ln f}{\partial \ln X_i}, \tag{35}$$

and

$$UPC_i = \frac{\left(\frac{\mu_{\ln X_i}}{\mu_{\ln f}} \frac{\partial \ln f}{\partial \ln X_i} \right)^2}{\left(\frac{\sigma_{\ln f}}{\mu_{\ln f}} \right)^2} \left(\frac{\sigma_{\ln X_i}}{\mu_{\ln X_i}} \right)^2. \tag{36}$$

In Eqs. 34-36, f represents H_t or PW.

Table 8: A comparison of the values of 10th, 50th and 90th percentiles computed from the MCM and AUPM for H_t, PW, $\ln H_t$, and $\ln PW$ functions.

	P_{10}		P_{50}		P_{90}	
	MCM	AUPM	MCM	AUPM	MCM	AUPM
$H_t \times 10^{13}$, kJ	4.86	5.29† 4.95*	8.84	9.64† 8.83*	15.9	17.6† 15.8*
PW, MW	17.4	19.5† 17.6*	34.2	37.7† 34.2*	66.3	72.2† 66.3*
$\ln H_t$	31.5	31.6† 31.5*	32.1	32.2† 32.1*	32.7	32.8† 32.7*
$\ln PW$	2.86	2.97† 2.87*	3.53	3.63† 3.53*	4.19	4.28† 4.19*

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

Table 9: The values of RU, UPC and normalized sensitivities (NS) for each variable in $\ln H_t$.

Variable X_i	$\lvert RU_i \rvert$	$\lvert NS_i \rvert$	$\lvert RU_i \times NS_i \rvert$	UPC_i
ϕ	0.105	3.4×10^{-3}	3.5×10^{-4}	6.1×10^{-4}
c_s, kJ/(kg °C)	0.469	3.6×10^{-3}	1.7×10^{-3}	1.4×10^{-2}
ρ_s, kg/m^3	0.002	2.2×10^{-1}	4.5×10^{-4}	9.8×10^{-4}
c_w, kJ/(kg °C)	0.008	4.2×10^{-3}	3.2×10^{-5}	5.1×10^{-6}
ρ_w, kg/m^3	0.002	2.0×10^{-2}	4.3×10^{-5}	9.2×10^{-6}
A, m^2	0.021	4.3×10^{-1}	8.9×10^{-3}	4.0×10^{-1}
h, m	0.050	1.9×10^{-1}	9.7×10^{-3}	4.8×10^{-1}
T-T_r, °C	0.036	1.3×10^{-1}	4.7×10^{-3}	1.1×10^{-1}

Table 10: The values of RU, UPC and normalized sensitivities (NS) for each variable in $\ln PW$.

Variable X_i	$\lvert RU_i \rvert$	$\lvert NS_i \rvert$	$\lvert RU_i \times NS_i \rvert$	UPC_i
ϕ	0.105	3.0×10^{-2}	3.2×10^{-3}	4.8×10^{-4}
c_s, kJ/(kg °C)	0.462	3.3×10^{-2}	1.5×10^{-2}	1.1×10^{-2}
ρ_s, kg/m^3	0.002	2.0×10^{0}	4.1×10^{-3}	7.5×10^{-4}
c_w, kJ/(kg °C)	0.008	3.7×10^{-2}	3.0×10^{-4}	4.0×10^{-6}
ρ_w, kg/m^3	0.002	1.8×10^{-1}	3.9×10^{-4}	7.0×10^{-6}
A, m^2	0.021	3.9×10^{0}	8.1×10^{-2}	3.0×10^{-1}
h, m	0.050	1.8×10^{0}	8.8×10^{-2}	3.6×10^{-1}
T-T_r, °C	0.036	1.2×10^{0}	4.3×10^{-2}	8.6×10^{-2}
R_F	0.128	5.2×10^{-1}	6.7×10^{-2}	2.1×10^{-1}
Y	0.262	5.8×10^{-2}	1.5×10^{-2}	1.1×10^{-2}
L_F	0.084	2.5×10^{-1}	2.1×10^{-2}	2.0×10^{-2}

The UPC_i values given in Table 9 indicate that for the example considered, mainly three variables are contributing to the total uncertainty in lnH_t. These are, in the order of, thickness (h), area (A), and temperature (T-T_r). Similarly, from Table 10, we see that mainly four variables are contributing to the total uncertainty in lnPW, which are, in the order of, thickness (h), area (A), recovery factor (R_F), and temperature (T-T_r).

Another important point that we wish to make from the results of Tables 9 and 10 is that the variables that are more influential on the total uncertainty in lnH_t or lnPW can be identified by inspecting the absolute value of the products of RU_i and NS_i, but cannot be identified by inspecting solely their individual values. In other words, inspecting RU_i or NS_i alone is not sufficient to identify the variables that are more influential on the total uncertainty in lnH_t or lnPW. Note that such identification can be made before applying the MCM or AUPM by computing the values of normalized sensitivities from the formulae presented (Tables 1-4) and the given values of relative uncertainty for each input parameter.

Correlated Case

Here, we investigate the propagation of correlations among input variables into the total uncertainty in lnH_t and lnPW. 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 5, but assume correlation between the five correlated pairs and the correlation coefficients given in Table 11.

Table 11: Correlated variable pairs and correlation coefficients.

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

Figures 6 and 7 show histograms of H_t (Eq. 4), and PW (Eq. 5) generated from the MCM by using the distributions given in Table 5 and the correlation coefficients given in Table 11 in @RISK.

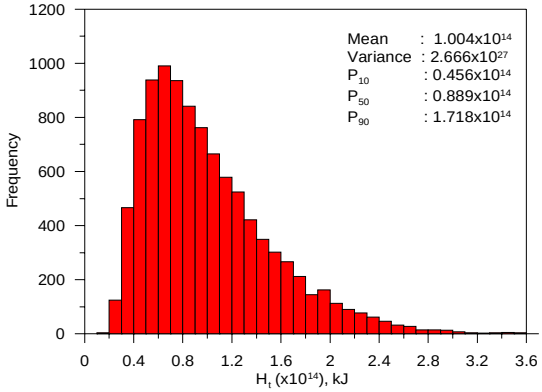

Figure 6: Histogram of total stored heat H_t, generated from the MCM, correlated case.

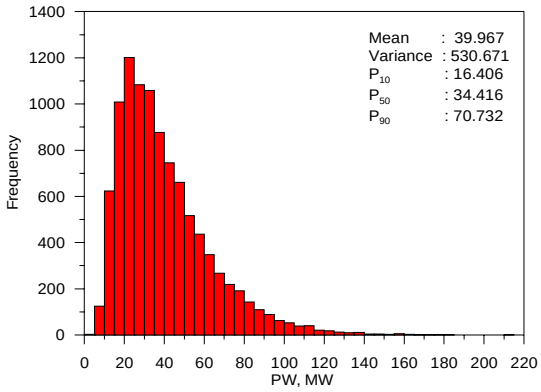

Figure 7: Histogram of recoverable power, PW, generated from the MCM, correlated case.

It can be seen that correlation did not change the shape of the H_t and PW distributions. When the results of Figs. 6 and 7 for the correlated case are compared with the results of Figs. 3 and 4 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 12 compares the values of means and variances computed from MCM and AUPM for H_t, PW, lnH_t and lnPW. We used the AUPM based on Approach 2 to compute means and variances of lnH_t and lnPW 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 lnH_t and lnPW are essentially identical.

Table 13 compares the values of P_{10}, P_{50}, and P_{90} 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 12: A comparison of means and variances from the MCM and AUPM for H_t, PW, lnH_t, and lnPW functions, correlated case.

	Mean		**Variance**	
	MCM	AUPM	MCM	AUPM
H_t, kJ	1.00×10^{14}	9.50×10^{13}	2.67×10^{27}	2.30×10^{27}
PW, MW	39.9	37.7	530.7	438.5
lnH_t	32.12	32.12	0.2553	0.2547
lnPW	3.534	3.536	0.3148	0.3177

Table 14 tabulates the values of UPC_i (Eq. 36) and UPC_{ij} (Eq. 37) values computed from the AUPM method for lnPW function. UPC_{ij}s is computed from:

$$UPC_{i,j} = 2\left(\frac{\sigma_{\ln X_i}}{\sigma_{\ln PW}}\frac{\partial \ln PW}{\partial \ln X_i}\right)\left(\frac{\sigma_{X_j}}{\sigma_{\ln PW}}\frac{\partial \ln PW}{\partial \ln X_j}\right)\rho_{\ln X_i,\ln X_j} \quad (37)$$

The results of Table 14 indicate that correlations in five correlated pairs contribute positively about 16% to the total uncertainty. The remaining 84% in the total uncertainty is due to individual variables contribution (UPC_is) to the total uncertainty. In fact, these results explain why the variance of lnPW (or lnH_t) for the correlated case increased about 17% (or 24%), compared to that of lnPW (or lnH_t) for the uncorrelated case. From Table 14, we also notice that the correlation between area (A) and thickness (h) contribute the most (about 13%), among the other four correlated pairs, to the total of UPC_{ij}s.

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 H_t, PW, lnH_t, and lnPW. The results also indicate that the AUPM works as good as the MCM to estimate uncertainty (variance, 10th, 50th, and 90th percentiles) in lnPW and lnH_t functions even for the correlated input variables.

Table 13: A comparison of the values of 10th, 50th and 90th percentiles computed from the MC and AUP methods for H_t, PW, lnH_t, and lnPW functions, correlated case.

	P_{10}		**P_{50}**		**P_{90}**	
	MCM	AUPM	MCM	AUPM	MCM	AUPM
$H_t\times10^{13}$, kJ	4.56	4.65	8.89	8.87	17.2	16.9
PW, MW	16.4	16.7	34.4	34.3	70.7	70.6
lnH_t	31.5	31.5	32.1	32.1	32.8	32.8
lnPW	2.82	2.81	3.53	3.54	4.25	4.26

Table 14: The values of UPCi, and UPCij for lnPW.

Variable X_i	UPC_i	UPC_{ij}
ϕ	4.0×10^{-4}	-
c_s, kJ/(kg °C)	9.1×10^{-3}	-
ρ_s, kg/m^3	6.3×10^{-4}	-
c_w, kJ/(kg °C)	3.3×10^{-6}	-
ρ_w, kg/m^3	6.6×10^{-6}	-
A, m^2	2.6×10^{-1}	-
h, m	3.0×10^{-1}	-
T-T_r, °C	7.3×10^{-2}	-
R_F	1.7×10^{-1}	-
Y	8.8×10^{-3}	-
L_F	1.7×10^{-2}	-
Correlated Pairs, (X_i, X_j)		
$(c_s, T-T_r)$	-	3.3×10^{-2}
(c_s, ρ_s)	-	-2.1×10^{-3}
$(\rho_w, T-T_r)$	-	-8.2×10^{-4}
(c_w, ρ_w)	-	-3.8×10^{-6}
(A, h)	-	1.3×10^{-1}

Cognitive Biases

It has long been known – e.g., Tversky and Kahneman (1974), Capen (1976), and Welsh et al. (2005, 2007) – that cognitive biases impact decisions taken under conditions of uncertainty. Cognitive biases results from discrepancies between calculated, optimal decisions and those made using intuition (Welsh et al., 2005). There are three different specific cognitive biases of interest; *overconfidence*, *trust heuristic (or anchoring), and availability.*

Overconfidence results from people's tendency to understate uncertainty due to having no good quantitative idea of uncertainty. Thus, they overestimate the precision of their own knowledge (Capen, 1976, Welsh et al., 2005 and 2007). For example, 80% confidence ranges are commonly used when interpreters are asked to estimate variables such as average porosity and thickness of reservoir. Data collected from interpreters however, indicate that such ranges on average include the actual values less than 50% of the time rather than 80% as the confidence range should indicate (Welsh et al., 2007). In other words, commonly stated or used ranges for the variables by interpreters are too narrow. In this work, we limit our investigation only to how cognitive bias due to overconfidence propagates into the total uncertainty in $\ln H_t$ and $\ln PW$. We will not consider cognitive biases due to trust heuristic – results from managers' tendency to rely on the judgments of individuals that they have learnt to trust, rather than incorporating the opinions of everyone working on an interpretation task – and availability – results from people's tendency to ignore the total number of possible events and to use the number of events that can be recalled (Welsh et al., 2005 and 2007). We refer the interested readers to Welsh et al. (2007) for details regarding modeling trust heuristic and availability.

To model cognitive bias due to overconfidence, we use the same approach used by Welsh et al. (2007). We consider the input variable distributions given in Table 5 as the base distributions for the overconfidence model. To model the impact of overconfidence on the total uncertainty, the 10th and 90th percentiles of each of the base input distributions given Table 5 were calculated. Then, for each of the base distributions, a triangular distribution with the same mode and mean was found by adjusting the minimum and maximum values by equal amounts so that the 10th and 90th percentiles of the base distribution corresponds to 20th and 80th percentiles of the "real" distribution. Hence, the base distributions given in Table 5 is 20% overconfident compared to the transformed distributions (referred to as "real"). For example, Figure 8 compares the real (wider) distribution of porosity (ϕ) against the 20% overconfident (narrow) distribution of ϕ.

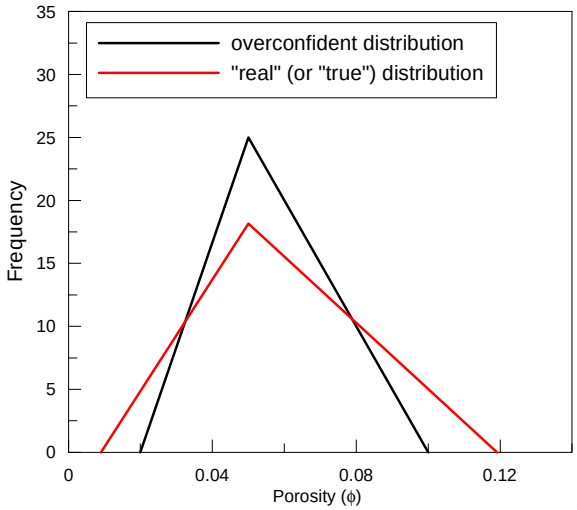

Figure 8: A comparison of overconfident ϕ distribution with "real" ϕ distribution.

To assess the effect of overconfidence, transformations described above were applied to all of the input variables (given in Table 5) used in calculating the H_t, PW, $\ln H_t$ and $\ln PW$. We assume that all input variables are independent. Here, we only present our results obtained for PW. Similar results obtained for H_t.

Table 15 compares the mean, variance and 10th, 50th, and 90th percentiles computed for PW based on the 20% overconfident input variables and the real input variables by using the AUPM (based on Approach 2). The mean and variance given for PW in Table 15 were obtained after computing the mean and variance for the $\ln PW$ and then using the following transformation from normal to log-normal:

$$\mu_{PW} = \exp\left(\mu_{\ln PW} + \frac{\sigma^2_{\ln PW}}{2}\right), \tag{38}$$

and

$$\sigma^2_{PW} = \exp\left(2\mu_{\ln PW} + \sigma^2_{\ln PW}\right)\left[\exp\left(\sigma^2_{\ln PW}\right) - 1\right]. \tag{39}$$

As can be seen from Table 15, although the mean is identical for both the 20% overconfident model and the real model, the uncertainty markers; variance and percentiles, are significantly different, as expected. The variance for the real case is about twice the variance of the 20% overconfidence case. Hence, the 10th and 90th percentiles of the PW pertaining to these two cases are different. As expected, the 10th and 50th percentiles for the 20% overconfidence case are

larger than the corresponding ones for the real case, and 90th percentile for the 20% overconfidence case is smaller than that for the real case.

Hence, 20% overconfidence model predicts significantly too low uncertainty in the estimates of stored heat and recoverable power.

Table 15: The values of mean, variance, 10th, 50th and 90th percentiles for the real and %20 overconfidence models.

	μ_{PW}	σ^2_{PW}	P_{10}	P_{50}	P_{90}
PW for the 20% overconfidence	39.1	468.8	17.6	34.2	66.3
PW for the real	39.1	929.4	10.7	29.4	78.2

Adding Individual Well (or Field) Reserves to Determine Field (or Country) Reserves

Here, we show that it is incorrect – a commonly made mistake – to add the "proved" and "probable" (P_{10} and P_{90} percentiles) thermal energy reserves from individual wells (or fields) to get "proved" field (or country) reserves, which has long been established by Capen (1996, 2001) in the context of stochastic estimation of oil and gas reserves.

For the purpose of demonstration, we assume a country *Z* consisting of 10 pseudo geothermal fields, each having a log-normal distribution for *PW* with means, variances and P_{10}, P_{50}, P_{90} percentiles as given in Table 16.

If we let PW_j to denote the recoverable power for field *j*, j=1,2,…,N_f, where N_f denotes the total number of fields, the total recoverable power (denoted by PW_C) for the country *Z* is simply given by

$$PW_C = \sum_{j=1}^{N_f} PW_j \,. \tag{40}$$

Because we treat each field's PW_j as independently distributed random functions, the mean and variance of the PW_C given by Eq. 40 are simply the sums of each field's mean and variance, respectively. These values are given in the last row of the second and third columns in Table 16; i.e., $\mu_{PW_C} = 1612.15$ MW and $\sigma^2_{PW_C} = 110,420$ MW^2 (or $\sigma_{PW_C} = 332.295$ MW).

Based on the CLT, we expect that adding PW_js (Eq. 40) leads asymptotically to the normal distribution as N_f becomes large. Then, assuming that PW_C is normal, we can compute the P_{10}, P_{50} and P_{90} percentiles of the PW_C distribution by simply using the formulas given by:

$$P_{10} = \mu_{PW_C} - 1.28\sigma_{PW_C} \,, \tag{41}$$

$$P_{90} = \mu_{PW_C} + 1.28\sigma_{PW_C} \,, \tag{42}$$

and

$$P_{50} = \frac{P_{10} + P_{90}}{2} = \mu_{PW_C} \,. \tag{43}$$

The values of 10th, 50th and 90th percentiles computed from Eqs. 41-43 for the PW_C for the 10-field example case considered are given in Table 17.

Table 16: 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	856.2	8.07×10^4	538.1	813.2	1224.8
Field 2	408.0	2.57×10^4	234.4	377.0	618.0
Field 3	96.59	1.07×10^3	60.10	91.45	138.4
Field 4	50.25	7.17×10^2	23.93	44.22	84.07
Field 5	49.41	5.57×10^2	24.55	44.74	79.62
Field 6	41.54	5.36×10^2	18.57	36.28	70.69
Field 7	40.39	5.29×10^2	17.65	34.83	69.41
Field 8	26.37	2.02×10^2	12.26	53.19	44.03
Field 9	26.36	2.62×10^2	10.75	22.31	46.58
Field 10	17.04	1.47×10^2	6.10	13.88	31.59
Total	**1612.15**	**110420**	**946.4**	**1531.1**	**2407.2**

Table 17: Comparison of 10th, 50th and 90th percentiles computed from different approaches for the Country Z.

	P_{10} MW	P_{50} MW	P_{90} MW
PW_C [1]	1186.8	1612.2	2037.5
PW_C [2]	1227.1	1574.5	2044.5
PW_C [3]	1216.1	1579.0	2050.0
PW_C [4]	946.4	1531.1	2407.2

[1]percentiles computed based on normal assumption
[2] percentiles obtained from the MC simulation
[3]percentiles computed based on log-normal assumption
[4]percentiles calculated as the simple sum of the individual field's corresponding percentiles.

To check the validity of the percentiles computed (see 2nd row of Table 17) for PW_C based on the normality assumption, we performed Monte Carlo simulation based on Eq. 40 with each field's PW_j,

j=1,2,…10, distribution specified in Table 16. Figure 10 shows the histogram of PW_C generated from the MCM. The percentiles of the PW_C distribution are recorded in the inset of Fig. 9, as well as in the 3rd row of Table 17. Although the agreement between the percentiles computed by assuming normality (Eqs. 40-41) and considering the MC is good (all percentiles agree within less than 3.5%), we notice from Fig. 9 that the histogram of PW_C is more like a log-normal than normal because it is slightly skewed right. Therefore, we check whether assuming a log-normal distribution, instead of normal distribution, for PW_C improves the percentiles. The values of the percentiles computed by assuming a log-normal PW_C are recorded in the 4th row of Table 17. Interestingly, the results indicate that for this example, assuming log-normal distribution for PW_C does provide better estimates of the percentiles than does assuming a normal distribution for PW_C. [The agreement between the percentiles computed by assuming log-normality and the MCM are very good (all percentiles agree within less than 1.0%)]. Although, this result may be viewed as contradictory to the CLT, it seems, however, that adding log-normal field PW_js does not quickly approximate a normal, which is consistent with the findings of Capen (2001).

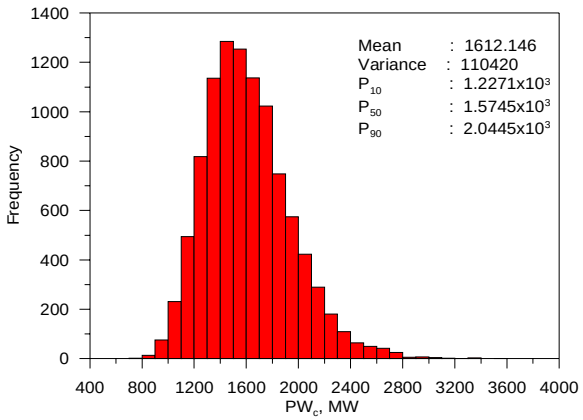

Figure 9: Histogram of the PW_C generated from the MCM.

Finally, as an important point, we note that if we just add the P_{10}s, P_{50}s, and P_{90}s of individual fields (i.e., simple sum), we obtain 946.4, 1531.1, and 2407.2 MW, respectively (see the last row of Table 16, and also the 5th row of Table 17). However, if we combine each field's distribution properly and back-calculate P_{10}s, P_{50}s, and P_{90}s, we obtain 1227.1, 1574.5, and 2044.5 MW, respectively. From a comparison of these numbers, we find that the simple arithmetic sum of each field's P_{10}s is *1.3 times smaller* than the correct P_{10} for the country's total "proved" "reserves" for the recoverable power, whereas, the simple arithmetic sum of each field's P_{90}s is about *1.2 times larger* than the correct P_{90} for the country's total "possible" "reserves." These are sizable errors.

Based on the results of the specific example considered above as well as our other results not shown, it is clear that one *should not add* individual zone, well, or field estimates of P_{10}s, P_{50}s, and P_{90}s to obtain well, field, or country estimates of P_{10}, P_{50}, and P_{90}, respectively. The simple arithmetic sum significantly underestimates the true P_{10}, while it significantly overestimates the true P_{90}. This is also consistent with the findings of Capen (1996, 2001). In addition, like Capen (2001) who looked at the stochastic oil and gas reserve calculations, we found that the distribution resulting from properly combining each zone, well or field "reserve" is more like a log-normal distribution, and hence quick and accurate estimates of P_{10}, P_{50} and P_{90} for a well, field or country could be calculated based on the log-normal assumption without resorting to the MC simulation.

CONCLUSIONS

On the basis of this work, we state that:

1. The distribution of stored heat energy and recoverable power for an individual zone, well, or field for a geothermal reservoir, 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 stored heat and recoverable power for a zone, well, or field. The AUPM method, when combined with the assumption of log-normality for the stored heat and recoverable power, 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 stored and recoverable power.

4. We have shown that cognitive bias (in the form of overconfidence resulting from choosing too narrow variances for the input variables) can significantly underestimate the uncertainty

(variance, percentiles of the cumulative distribution).

5. Finally, we showed that it is incorrect to use a simple arithmetic sum of the "proved" and "probable" (P_{10} and P_{90} percentiles, respectively) thermal energy reserves from individual wells (or fields) to obtain "proved" and "probable" field (or country) reserves. It is shown that the simple arithmetic sum significantly underestimates the true P_{10} and significantly overestimates the true P_{90}, obtained from the probabilistic sum.

REFERENCES

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Tversky, A. and Kahneman, D. (1974), "Judgment under Uncertainty," *Science* **185**, 1124-1131.

Welsh, M.B., Bratvold, R.B., and Begg, S.H., (2005), "Cognitive Biases in the Petroleum Industry: Impact and Remediation," paper SPE 96423 presented at the 2005 SPE Annual Technical Conference and Exhibition, Dallas, TX, USA, 09-12 October.

Welsh, M.B., Begg, S.H., and Bratvold, R.B. (2007), "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.

Zeybek, A.D., Onur, M., Türeyen, Ö.İ., Ma, M.S., Al-Shahri, A.M., and Kuchuk, F.J. (2009), "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.