EASTERN IDAHO
WATER RIGHTS COALITION

EASTERN IDAHO WATER RIGHTS COALITION
P.O. Box 50125 • Idaho Falls, ID • 83405-0125

March 15, 2012

Gary Spackman
Interim Director
Idaho Department of Water Resources
322 Front Street
Boise, Idaho 83720-0098

Re: Quantification of Uncertainty.

Dear Director Spackman:

A standard textbook states that "Predictions made with simulation models must be interpreted with caution. The 'aura of correctness' (Bredehoeft and Konikow, 1993) attached to model calculations often exerts much more influence than is reasonable, given the typically uncertain data on which models are built.... This uncertainty does not disappear simply because a model is constructed."[1] In other words, even though the constructed model may be the best scientific tool available, this does not assure 100% correct predictions.

We appreciate the opportunity you have given the Eastern Idaho Water Rights Coalition and all the stakeholders to weigh in on modeling-uncertainty issues. While we respect the ESHMC discussions of the technical difficulty in precisely quantifying uncertainty, we feel that with some common-sense analysis, it is possible to calculate an approximation that is neither capricious nor arbitrary.

The attached memo illustrates how a simple weighted average assessment of water-budget uncertainty can be combined with Darcy's Law and analytical aquifer response analysis to estimate an approximate uncertainty in the representation of the timing and spatial distribution of pumping or recharge effects. By assuming that other sources of uncertainty are small, the following preliminary conclusions can be made:

[1] P.A. Domenico and F.W. Schwartz. 1998. Physical and Chemical Hydrogeology Second Edition. P. 144.

1. Uncertainty in the timing of effects is on the order of 15% to 20% of the estimated time of arrival. This translates to a week or two of uncertainty for nearby locations, or many years uncertainty for distant locations.

2. At scales smaller than a few miles, the spatial uncertainty of partitioning effects to individual model cells can be very large.

3. Uncertainty in the spatial distribution of effects at scales larger than a few miles is on the order of 15% to 20% of the total modeled impact. If the question to be asked relies on the difference between reaches, the uncertainty can be larger than the difference itself.

Please review this methodology with staff and consider presenting it to the ESHMC. We recommend that this procedure be refined with more robust statistical methods. Perhaps IDWR could contract with Dr. Van Kirk of Humboldt State University for an independent analysis. Dr. Van Kirk has also suggested that these concepts could be tested with numerical simulations using parameters randomly selected from appropriate distributions, with a simplified version of the numerical model.

Thank you for this opportunity to suggest an option that may provide a path forward in an approximate assessment of model uncertainty.

Sincerely,

Roger Warner
President

MEMORANDUM

To: Eastern Idaho Water Rights Coalition
Fr: Rocky Mountain Environmental Associates, Inc.
Date: 15 March 2012

Re: Weighted-average evaluation of model uncertainty

This memo presents an outline for two simple procedures to estimate model uncertainty, both based on a weighted-average evaluation of the uncertainty in water budget components. The first procedure estimates uncertainty in temporal output of the model, and the second estimates uncertainty in spatial distribution of effects to model reaches. There are additional sources of uncertainty but the water-budget uncertainty at least provides a starting point for evaluation.

Weighted-Average Water Budget Uncertainty

In 2005 Idaho Water Resources Research Institute estimated the uncertainty of the ESPAM1.1 water budget at approximately plus or minus 17%[1], interpreted to mean that

[1] Snake River Plain Aquifer Model Scenario Update: Hydrologic Effects of Continued 1980-2002 Water Supply and Use Conditions Using Snake Plain Aquifer

482 Constitution, Idaho Falls, Idaho 83402-3537 • (208) 524-2353 • FAX (208) 524-1795 • www.rockymountainenvironmental.com

there is a 95% probability that the true value is somewhere between 83% and 117% of the estimated value. The IWRRI estimate was based on statistical procedures for finding the variance of a sum, based on the variances of the components. This is essentially a weighted average of the uncertainty of the components of the water budget. We recommend that a similar analysis be performed on the ESPAM2.0 water budget, and anticipate that its uncertainty will be of similar magnitude.

Temporal Uncertainty. To propagate the water-budget uncertainty into an indication of the temporal uncertainty of model results, we first applied Darcy's law. It relates the quantity of groundwater flow (in our case, the water budget) and the aquifer transmissivity as follows:

$$Q = T\, w\, dh/dl$$

where	Q =	rate of flow through the aquifer
	T =	aquifer transmissivity
	w =	width of flow tube considered
	dh/dl =	gradient along the length of flow

This can be rearranged to express T as a function of the water budget:

$$T = Q * (1/w\, dh/dl)$$

We can calculate the variance of a product[2] if we know the variances of the factors and the covariance between them. Relative to the water budget (Q), the width and gradient are very well known. If we assume they are perfectly known and their measurement methods are independent of the measurements of the water budget, then the variance of transmissivity is driven entirely by the variance of the water budget. Hence, its uncertainty will also be on the order of plus or minus 17%.

To estimate temporal uncertainty, we must understand how uncertainty in transmissivity affects uncertainty in propagation of effects. A long-established analytical model relying on the same basic flow equations as the MODFLOW model is the Jenkins method. This method allows calculation of a reference time called the "stream depletion factor," which is the time for half of the pumping rate of a continuously-pumping well to be expressed at a hydraulically connected stream. It is defined as:

Model Version 1.1 "Base Case Scenario." http://www.if.uidaho.edu/~johnson/ifiwrri/projects.htm#model

[2] A.J. Clemens and C.J. Burt, 1997. *Accuracy of Irrigation Efficiency Estimates,* Journal of Irrigation and Drainage Engineering. Notation altered.

$sdf = a^2 S/T$

where	sdf =	stream depletion factor
	a =	distance from stream
	S =	storage coefficient
	T =	transmissivity

The estimation process can result in correlation between S and T. Nevertheless, we do not expect that the variance of the ratio will translate into an uncertainty less than that of T alone.[3] Again assuming that uncertainty in distance is very small and that distance measurement is independent of estimation of S and T, the uncertainty in the stream depletion factor is essentially dependent on the ratio S/T and hence can be assumed to be on the order of plus or minus 17%.

Because the stream depletion factor depends on distance, the numerical value of uncertainty will depend on the location of the stress. If the stress is near the river and has a stream depletion factor of 100 days, the temporal uncertainty would be approximately plus or minus 17 days. A stress distant from the river might have a stream depletion factor of 50 years, in which case the temporal uncertainty would be on the order of plus or minus eight years.

<u>Spatial Uncertainty</u>. This analysis relies on the estimated uncertainty in transmissivity discussed in the temporal analysis. If the aquifer transmissivity were absolutely uniform across the plain, the spatial distribution of stresses to reaches would be a function of only geographical distances, which can be measured with great precision. However, if one considers an application of Darcy's law to an aquifer stress located between two surface water bodies, the steady-state partition also depends upon the relative transmissivity in the two flow paths. This is essentially an integration of the heterogeneous transmissivity throughout the flow tube through which effects propagate.

There are two important questions here. The first is whether the scale of analysis is appropriate to the spatial discretization of transmissivity representation, and the second is the effect of uncertainty in transmissivity, when the scale of analysis is appropriately large.

Regarding the scale of analysis, we know that modeled transmissivity in ESPAM1.1 and ESPAM2 spans several orders of magnitude. Because the model smooths and simplifies complex geological features, it is reasonable to expect that even within a small area,

---

[3] This assumption would be a good topic for further investigation.

actual transmissivity may vary by at least an order of magnitude. However, between pilot points the model representation is essentially uniform. Figure 1 illustrates a plausible small-scale heterogeneous aquifer model where the true transmissivity is 100,000 ft^2/day on the east and 10,000 ft^2/day on the west, but the smoothed representation uses the geometric mean of 32,000 ft^2/day across the domain. The smoothed representation would indicate a 50%/50% partition of pumping to the east and west springs, respectively. In reality the partition to the west spring is less than 20% of the indicated value and the partition to the east spring is 180% of indicated.

Given that there are 11,000 model cells and only a few hundred pilot points, it is clear that this inter-pilot-point distance is larger than a single model cell, and hence the uncertainty of partitioning modeled effects to a single cell can be very large.

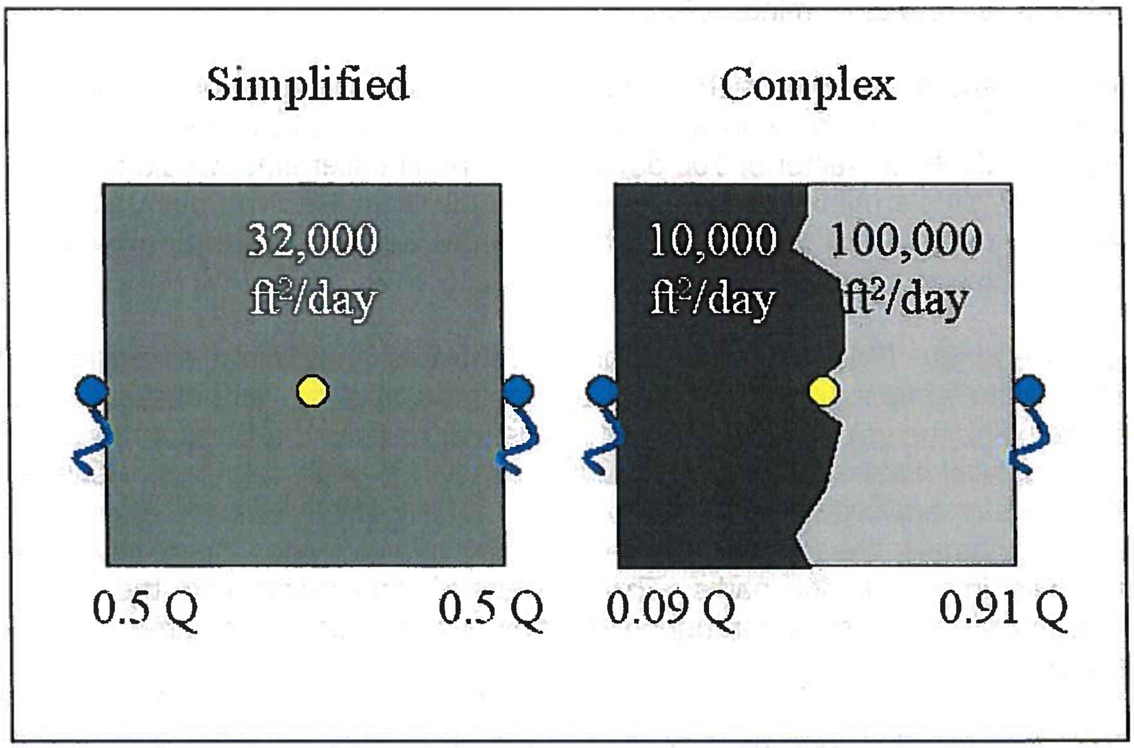

Figure 1. Partition of flux at spatial scales smaller than the inter-pilot-point distance.

At scales larger than the distance between pilot points, one can consider that if there were only two reaches, the total calculated flux would be the sum of flux to each of the reaches. Suppose there are two reaches with characteristics as illustrated in Figure 2:

The discharge to the reaches can be calculated by Darcy's Law:

$$Q_A = 100{,}000 \text{ ft}^2\text{/day} * 20{,}000 \text{ ft} * (10 \text{ ft}/ 100{,}000 \text{ ft}) = 200{,}000 \text{ ft}^3\text{/day}$$
$$Q_B = 200{,}000 \text{ ft}^2\text{/day} * 20{,}000 \text{ ft} * (30 \text{ ft}/ 50{,}000 \text{ ft}) = 240{,}000 \text{ ft}^3\text{/day}$$

$Q = Q_A + Q_B = 440{,}000\ ft^3/day$
$Q_D = Q_B - Q_A = 40{,}000\ ft^3/day$

Where Q_A = flux to Reach A (east)
Q_B = flux to Reach B (west)
Q = total flux
Q_D = difference in flux

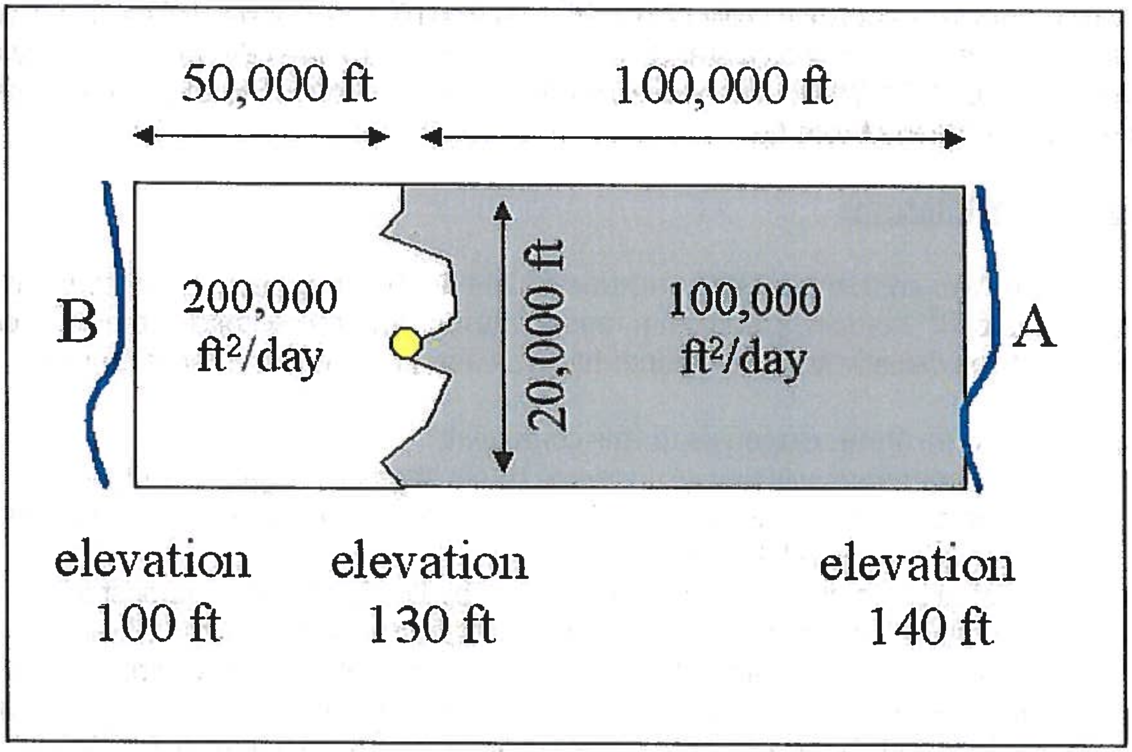

Figure 2. Partition of flux at spatial scales larger than the inter-pilot-point distance.

Since the distances and heads are assumed to be perfectly known, the estimated uncertainty in Q will be approximately the same as in Transmissivity, or about plus or minus 17%. Hence, the uncertainty of Q_A is plus or minus 34,000 ft^3/day or a range of 68,000 ft^3/day. The range of Q_B is similarly 80,000 ft^3/day.

Considering that standard deviations can be estimated as one fourth the range,[4] the approximate standard deviation of Q_A is 68,000/4 or 17,000 ft^3 day. Its estimated variance is the standard deviation squared, or 2.9 x 10^8 ft^6/day^2. Likewise the approximate standard deviation of Q_B is 20,000 ft^3/day and its variance 4.2 x 10^8 ft^6/day^2.

[4] R. Lyman Ott. 1993. An Introduction to Statistical Methods and Data Analysis.

Since both transmissivity estimates were informed by the same water budget, their correlation coefficient can be assumed to be approximately 0.75. Inverting the equation for correlation coefficient, the covariance can be estimated as the product of correlation coefficient and the two standard deviations.[5] Hence, the covariance is approximately (0.75 * 17,000 * 20,000) or 2.6×10^8 ft^6/day^2.

The variance of a sum or difference is the sum of the variance of the components, plus twice their covariance. In this case, the variance of either the sum or difference of the two reach effects is approximately ($2.9 \times 10^8 + 4.2 \times 10^8 + 2 * 2.6 \times 10^8$) = 1.2×10^9 ft^6/day^2. This is equivalent to a standard deviation of approximately 35,000 ft^3/day or a range of +/- 70,000 ft^3/day. The value 70,000 ft^3/day is about 16% of Q, about one third of Q_B and nearly twice Q_D.

<u>Anticipation of Criticisms</u>

We anticipate two criticisms of this procedure. One is that it appears to be counter to the large overall R^2 statistic produced in model output, and the second is that it appears to contradict the visually stunning match that was obtained at some target springs.

Large R^2. We offer three responses to this complaint:

1. Adding parameters will always increase the R^2 statistic, regardless of improvement in predictive power. There are very many parameters implicitly reflected in the overall R^2 value.
2. The scatter plots of model run 8 performed by Dr. Schreuder and posted by IDWR show R^2 statistics for many individual targets in the 0.4 to 0.7 range. Since the R^2 statistic is an indication of the fraction of variability explained by the predicted values, ($1 - R^2$) is an indication of the fraction of variability that is *not* explained. In this case ($1 - R^2$) is generally within the range of 0.3 to 0.6, which is comparable with the uncertainties estimated with the simplistic procedures shown here.
3. The model R^2 statistic represents the relationship between the model and the underlying data; we are interested in the relationship between the model and the actual underlying physical parameters:
   a. The R^2 statistic describes the modeling procedure effects and assumes the data are perfect;
   b. Our estimates describe the data effects and assume the modeling procedure is perfect;

---

[5] Douglas C. Montgomery and George C. Runger. 1994. <u>Applied Statistics and Probability for Engineers.</u>

c. True uncertainty is actually a combination of data uncertainty, modeling procedure uncertainty, and additional sources. Hence, either of these indications are parts of a larger whole and are not contradictory.

Visual match to targets. The commendable matches to springs like Rangen must be considered in light of the conceptual model limitations and spatial limitations of the underlying data. Getting the right answer with wrong inputs may not be as comforting as one would hope; it does not necessarily indicate we will be able to predict the true physical response to a particular modeled stress. Additionally, these good matches must be considered along with poorer matches to other targets.

Summary

This simplistic analysis considers only two components of uncertainty, the uncertainty in the water budget and the spatial resolution of estimates of transmissivity. Using basic hydrologic and statistical relationships, it suggests that the uncertainty in estimates of timing of stress is on the order of 17% of the characteristic stream depletion factor for each site. Depending on location, this can range from half a month to years.

The analysis indicates that actual partition of effects can be 20% to 180% of modeled values, at spatial scales smaller than the inter-pilot-point distance. At larger scales the can be plus or minus 16% of the total stress. However, this can be more than 100% of the difference in the partition between reaches, which is often the driving factor in transfer decisions.

These are rough-and-ready calculations which can be refined with additional work. They provide an order of magnitude estimate and an outline for refinement.