Comments on Trim Line and Model Uncertainty

by
Charles M. Brendecke, PhD, PE
AMEC Earth and Environmental, Inc.
representing Idaho Ground Water Appropriators, Inc.

For consideration by the Eastern Snake Hydrologic Modeling Committee
July 8-9, 2009

Background

In a letter dated February 25, 2009, Director Tuthill addressed the following question from the committee:

As part of the uncertainty analysis, should the ESHMC address the technical aspects (not policy issues) of a trim line as a function of uncertainty?

In his response the Director noted opinions issued by Hearing Officer Schroeder on this topic and concluded that he had "sufficient guidance and a basis for the use of a trim line." Nevertheless he invited committee members to provide write-ups and/or presentations "regarding the <u>technical aspects</u> (emphasis added) of the use of trim line" at a later meeting and indicated that such would become part of a white paper on the topic.

At the ESHMC meeting of March 31- April 1, Koreny, *et.al.*, made a presentation entitled "Technical Analysis of the 'Trim line'." This presentation was followed up with a white paper distributed to the ESHMC on June 5, 2009.

My comments herein are meant to respond to both Director Tuthill's request and to the presentation and white paper by Koreny, *et.al*.

Preface

The present discussion of the trim line intertwines two issues which I believe should be separated. The first issue is how uncertainty in model predictions should be expressed in displaying model results. The second issue is whether and how to define a "zone of exclusion" for administrative curtailment. The first issue is largely a technical one; the second is largely a policy one (though its ramifications can be assessed in a technical analysis).

Apparently Koreny, et.al., at least partially agree with me, for they repeatedly state in their white paper that "The trim line has nothing to do with model uncertainty."

I would note that the use of the trim line was addressed at some length in the hearings for both the spring user and Surface Water Coalition delivery calls. The Director's final orders incorporating the Hearing Officer's opinions and recommendations are under appeal in District Court. In all likelihood they will be further appealed to the Idaho Supreme Court. In my view these appeal schedules and the transition in Department leadership make extended ESHMC discussion of the trim line policy (to be distinguished from the technical issue of model uncertainty) untimely, at best.

The committee's original question and the Director's response plainly seek to exclude the policy aspects of the use of trim line. But the setting of a trim line is essentially a policy matter, and it is nearly impossible to discuss it without venturing into subjective considerations of policy, *viz.* references by Koreny, *et.al.*, as to what is "justified" or "reasonable" in the Director's use of trim line. The Director has the discretion, and is probably compelled, to consider a host of applicable facts, policies, laws and regulations in determining his application of the Conjunctive Management Rules, many of which are not technical in nature.

The remainder of this paper discusses (as separately as I can) the issues of model uncertainty and administrative exclusion, and then responds to a few particular points made by Koreny, et.al., in their white paper.

Expressing Uncertainty in Model Predictions

I submitted a brief write-up to the committee in January of this year in which I outlined the general sources of uncertainty in groundwater modeling. A copy of that document is attached hereto as Appendix A. In it I suggested that there are four fundamental sources of model uncertainty:

- Conceptual uncertainty, aka "structural" or "geological" uncertainty…this arises because we cannot know the true hydro-geologic structure of the aquifer.

- Parameter uncertainty…this arises because, for whatever model structure we adopt, we cannot precisely quantify all the aquifer water budget terms it needs

- Calibration uncertainty (internal)…this arises because, for any given model structure and water budget, there are many combinations of parameter values that can lead to a similarly well-calibrated model.

- Calibration uncertainty (external)…this arises because we are calibrating the model to a set of targets (historical water levels and reach gains) that are themselves uncertain.

There are basically two ways to express quantitatively the uncertainty in model results. One is to determine the probability distribution of the error associated with a model prediction, choose an acceptable confidence limit and state the predicted result along with a range determined from the error distribution and confidence limit. A model prediction would thus be stated as X +/- Y with Z % confidence. These can also be displayed graphically, e.g., as box-whisker plots.

The second basic way to express uncertainty is through sensitivity analysis. Changes in the model predictions associated with alternative (and presumably reasonable) values of model inputs or parameters can be stated as ranges or shown graphically. This method is not as rigorous statistically, but is less intensive computationally. Accordingly, it is more commonly used.

Allen Wylie has proposed two approaches for quantification of model uncertainty, a "multiple models" approach and a "bend but don't break" approach. Conceptually, the former lends itself better to a probabilistic depiction of uncertainty while the latter implies more of a sensitivity analysis approach. However, these distinctions may not be bright ones if the "multiple models" are in reality simply a few differently parameterized versions of the same model structure. There has not been significant ESHMC discussion of Allen's proposals since his initial presentation.

Trim Line as Zone of Administrative Exclusion

The trim line defines an area of the ESPA beyond which the Director does not, under the present orders, seek administration against junior priority groundwater rights. It was established by evaluating the steady-state response functions for each connected sub-reach across the entire ESPA model domain. Model developers determined those model cells in which the steady-state response for a sub-reach was greater than 0.1, meaning that more than 10% of the aquifer depletion in that cell would be translated to a river depletion in the sub-reach. The trim line is essentially the envelope of all such model cells. Groundwater users outside the trim line are excluded from administration (curtailment or mitigation) to benefit calling rights in that sub-reach.

Former Director Dreher loosely tied this 10% definition of the trim line to the uncertainty in estimates of sub-reach river gains, more particularly the uncertainty in river flow gages. He stated in hearing testimony that he viewed this as a minimum level of model uncertainty, noting that model uncertainty had not been quantified. He went on to say that:

"…I made the determination it was not appropriate to curtail such junior priority ground water use if, in fact, we didn't know whether curtailment would result in a ***<u>meaningful amount</u>*** (emphasis added) *of water reaching the calling senior right."* (Transcript at 1167: 4-8)

It is important to note that the Director did not assert that pumping outside the trim line had no effect on calling senior water rights.

What is a "meaningful amount" of water plainly involves subjectivity; the ESPA model can provide quantitative estimates of hydrologic effects but it cannot tell us if these effects are "meaningful." What is meaningful to one party may not be to another, and what is meaningful in one context may not be in another. The Director must resolve this by setting a policy that reflects his duties to administer and distribute water under the laws in the state of Idaho.

It is my view that the Director Dreher's response reflected his subjective consideration of, among other things, model uncertainty, the potential futility of administrative curtailment, the larger benefits and costs of curtailment, and the policies, regulations and statutes under which he had to operate. For example, Rule 20.03 of the CMR states the following in describing the purpose of the Rules:

"These rules integrate the administration and use of surface and ground water in a manner consistent with the traditional policy of reasonable use of both surface and ground water. The policy of reasonable use includes the concepts of priority in time and superiority in right being subject to conditions of reasonable use as the legislature may by law prescribe as provided in Article XV, Section 5, Idaho Constitution, optimum development of water resources in the public interest prescribed in Article XV, Section 7, Idaho Constitution, and full economic development as defined by Idaho law. An appropriator is not entitled to command the entirety of large volumes of water in a surface or ground water source to support his appropriation contrary to the public policy of reasonable use of water as described in this rule". (IDAPA 37.03.11.20.03.)

There is much room in this language for argument (and there has been such), but it is clear that a broad view of water resource management is required to implement it.

The Director's application of a trim line may be based solely, in part, or not at all on model uncertainty. He may or may not consider how *de minimus* is defined or exclusion is effected in other states. That he has such discretion has been made clear by the Idaho Supreme Court, by the Hearing Officer and by the District Court. The ESHMC can assist the Director in better defining and describing model uncertainty. It can assist the Director in quantifying the effects of uncertainty and of various forms of administrative exclusion. But it cannot, in my view, tell the Director how to develop and apply administrative policy or opine, at least as a group, about what is reasonable or justified in that policy.

Clarification of Specific Points Made by Koreny, et.al.

In their white paper, Koreny, et.al., ("the authors") make certain statements and present certain data that I believe are either incorrect or incomplete and that may be misconstrued by readers unfamiliar with the details of ESPA modeling or of the matters addressed in the delivery call hearings. In this section I address a few of these.

<u>Effect of Groundwater Use Outside the Trim Line</u>

In Section 1.1 of their paper, the authors state that pumping outside the trim line *"…is incorrectly assumed to have no effect on spring flow."*

It is not clear who is alleged to be making this assumption. I do not believe that Director Dreher was assuming this when he first developed the trim line policy. His hearing testimony spoke to "meaningful amounts of water"; nowhere did he (to the best of my knowledge) state an opinion that groundwater pumping outside the trim line has no effect on spring flow.

<u>Proportion of Supply</u>

On page 2 of their paper, in conclusion No. 2, they state the following:

“A reduction of the senior’s supply by one-half to one-third is obviously significant and is well above a de minimus impact.”

Clear Springs Foods’ Snake River Farm facility normally diverts in the neighborhood of 100 cfs. By expanding the trim line definition to 1% rather than the present 10%, the flow to Snake River Farm would be increased (according to the authors’ analysis) by 1.81 cfs at steady-state. This about 2% of the Snake River Farm supply, not one-half to one-third of their supply as stated by the authors.

Transient Aspects of Curtailment

The authors present tables showing the flow increases to sub-reaches and spring users from different curtailment dates and trim line definitions. It is important to note that these flow increases are those that would occur under steady-state conditions.

In the documentation for the original curtailment scenarios run by IWRRI is a table that shows the time needed after initiation of curtailment for sub-reach discharge conditions to reach 90% of steady-state values. This table is attached as Appendix B. While the curtailment dates in the IWRRI analysis were not exactly the same as the priority dates of the calling rights at Snake River Farm, the conclusion easily can be reached that it will take in excess of 50 years of curtailment for the spring flow benefits shown in the authors’ analysis to be realized.

Water Use Foregone by Curtailment

The authors present data on the total groundwater irrigated area included under various trim line definitions and show, in a separate table, the benefits to calling rights from curtailment. This presentation makes it somewhat more difficult to understand the relative amounts of water involved.

In the case of the 1964 priority of Snake River Farm, the additional amount of groundwater consumption foregone under the authors’ 1% trim line versus the present trim line used by the Department is 300,000 acre-feet. This translates to a constant year-around flow rate of 414 cfs. This curtailment of 414 cfs would result in an increased flow to Snake River Farm, after 50-plus years, of 1.81 cfs. The ultimate gain in this case to Snake River Farm is 0.4% of the loss to groundwater users. Similar fractions are obtained by analysis of the other data contained in the authors’ tables.

Trim Line as an Assumption

The authors state that the 10% trim line is an “assumption” made by the Director. The use of this word implies that it was a frivolous determination. I believe the hearing record shows that it was deliberately and carefully thought out and derived (numerically) from the ESPA model.

Definition of *de minimus*

In several place in their paper (e.g., the heading of Section 2.2) the authors state that the trim line does not delineate *de minimus* impacts. *De minimus* is a legal term whose definition, if it exists at all, varies between jurisdictions. The authors cite a Colorado definition which is used to determine whether a groundwater source is legally considered to be hydraulically connected (i.e., "tributary") to a surface water source. It is not a definition of injury. I do not believe there is significant disagreement that the ESPA is hydraulically connected to the Snake River.

Mitigation Activities Outside the Trim Line

In Section 2.2 of their paper, the authors describe the trim line as discounting depletions from groundwater pumping and state that "*…contributions to the aquifer* [outside the trim line] *should also be discounted*." The implication of this statement is that groundwater users get credit for mitigation activities (e.g., conversion or retirement of lands) undertaken outside the trim line but are not held accountable for depletions outside the trim line.

The fact is that the Department does not give groundwater users credit for mitigation activities undertaken outside the trim line. It has consistently applied the trim line to both sides of the equation.

Sources of Uncertainty beyond calibration

Authors cite small uncertainty in calibration targets to argue model uncertainty is much less than 10%. Uncertainty arises from many aspects of model development beyond the uncertainty of calibration targets. Model uncertainty has not been quantified. It will differ between scenarios. I fully expect that it increases as the model is employed to answer ever more localized questions.

Accuracy of Calibration Data

The authors cite accuracy of certain spring flow measurements to argue that model is calibrated to very accurate data. The fact is that only a few springs are measured with this level of accuracy. Many springs are not measured and some are not accessible for measurement. However, the model must account for all aquifer discharge in maintaining a balanced water budget, whether it is precisely measured or not.

Comments on Model Uncertainty
by
Charles M. Brendecke, PhD, PE

For consideration by the Eastern Snake Hydrologic Modeling Committee
January 13-14, 2009

At the ESHMC meeting in August of last year we began discussion on the issue of ESPA model uncertainty. Allen Wylie presented two potential approaches to quantifying model uncertainty, a “multiple models” approach and a “bend-but-don’t-break” approach, and asked for comments so as to further the discussion of this matter. This memo conveys my thoughts on the matter of model uncertainty and on Allen’s suggested approaches.

Sources of Model Uncertainty

While there are semantic differences among experts and practitioners, I find it helpful to consider four fundamental types of model uncertainty.

- Conceptual uncertainty, aka “structural” or “geological” uncertainty…this arises because we can’t know and fully represent the actual hydrogeologic structure, stratigraphy and boundaries of the aquifer.
- Parameter uncertainty…this arises because, for whatever model structure we adopt, we can’t precisely quantify all the aquifer water budget terms we need as inputs.
- Calibration uncertainty (internal)…this arises because, for any given model structure and water budget, there are many combinations of calibration parameter (e.g., transmissivity, conductance) values that can lead to a similarly well-calibrated model.
- Calibration uncertainty (external)…this arises because we are calibrating the model to a set of targets (water levels and reach gains) that are themselves uncertain.

Overall model uncertainty is a function of all of these and model uncertainty may be different for different scenarios, depending on how different the water budget terms (e.g., aquifer stresses) are between scenarios. Conceptual uncertainty is the most difficult to quantify in a systematic way, partly because it does not lend itself to repetitive, incrementally changing calculations that can be programmed and executed in batch processes.

I think the Committee should discuss and be clear about which of these sources of uncertainty we are attempting to address in our overall quantification of model uncertainty.

Quantifying Model Uncertainty

Our goal should be to express model predictions in probabilistic terms, ideally as probability distributions. The probability distribution of a model prediction is a joint distribution, transformed through the model algorithm, of the probability distributions of all model inputs. This joint distribution is, in turn, one input distribution to the joint distribution derived from all possible models.

The most rigorous approach to generating these distributions in practice is Monte Carlo ("MC") simulation. This approach involves making large numbers of model runs (or "realizations") based on input values drawn randomly from their respective probability distributions (or their joint distribution). The results of this large number of runs are then displayed as the probability distribution of model predictions. Care must be taken in the set up of the random sampling process so as to account for statistical independence (or lack thereof) among the inputs. Unconditional MC simulation allows parameters to vary regardless of their effect on calibration, while conditional MC simulation filters input parameter vectors to keep simulation results within a pre-defined uncertainty range. Needless to say, this is a computationally intensive and time-consuming process.

A less intensive approach is a parametric sensitivity analysis, in which incremental changes are made in model parameters (one at a time) to cover their acknowledged uncertainty range. Effects on model results and calibration "fit" are assessed with each adjustment. While it is more tractable computationally, this approach does not lend itself as well to a probabilistic definition of uncertainty.

In both unconditional MC simulation and sensitivity analysis, there is the risk that the variation in input parameters will lead the model out of calibration. It is my understanding that PEST2000 offers a predictive analysis capability that supports calculation of predictive uncertainty associated with input parameter uncertainties while ensuring that the model remains in calibration. This would appear to avoid the risk in sensitivity analysis and in unconditioned MC simulation that the model is operating outside the range of calibration.

Allen's proposed approaches appear to come closest to the parameter sensitivity method. The "multiple models" approach could accommodate different conceptual models, though it is not clear we are considered these, while the "bend-but-don't-break" approach cannot. The latter would seem to require that the Committee agree *a priori* on a single conceptual model that would be subjected to "bending." Both of Allen's approaches could address parameter and calibration uncertainty. In all cases we will need a definition of the term "calibrated."

Some Suggestions for Ground Work and Discussion

To advance the ball on model uncertainty, I think the following foundational efforts should be considered. Many of these bear on the issue of conceptual uncertainty.

Definition of Calibration. In the past we simply looked at various calibration results and picked the model we thought "looked best." This time around we need to find a more objective approach, since we may be trying to compare calibrations from conceptually different models. We might, for example, prepare a map showing the kriged standard error of observed water levels across the ESPA. Any calibration run that produced heads within this standard error would be considered "calibrated." It is a bit harder to see how this would be applied to reach gains; perhaps reach gain error could be derived from the errors in the various components of the relatively simple water balance equation used for gains calculation.

Layering. In the past we adopted the concept of a single layer model. Over the course of the intervening years, arguments have been advanced that there may be multiple operative layers in some portions of the ESPA. Recent MODFLOW versions (HFB package) now accommodate representation of multiple layers in portions of the model domain. I would suggest a review of well logs and development of fence diagrams in certain areas of the aquifer (where such data is reasonably available), as well as water level maps stratified by well completion depth, to review whether there is reason to adopt multiple layers in portions of the aquifer.

Local Distribution of T and S. In the model, transmissivity and storativity are assumed uniform throughout each model cell. In some areas of the aquifer there is probably enough pump test data to examine how much variability exists in observed T and S values within a square mile cell. Understanding this variability would help inform the certainty that should be placed on model simulations that ask ever more detailed (spatially) questions, and might suggest where a finer grid should be considered.

Confined v. Unconfined. Most model simulations have been run in superposition mode. It is my understanding that this approach was adopted at least partly for convenience. There are some areas of the aquifer (e.g., the edges) where this assumption may lead to erroneous conclusions. Since computational power is now much more readily available, we should reconsider whether we ought to be making unconfined model runs and generating differences rather than doing everything with single runs.

Anisotropy. The model in its present form assumes isotropic conditions throughout the aquifer. Although it could readily be done, it does not appear that the MODFLOW isotropy parameter was subjected to PEST manipulation in the original calibration. Since then, arguments have been put forward that the aquifer may be anisotropic in some areas, and MODFLOW now allows isotropy to be set on a cell by cell basis (LPF package). I would suggest a more thorough evaluation of aquifer anisotropy, starting with a statistical analysis of observed T and S values in various parts of the aquifer to see if they contain any directional distribution.

Table 1. Predicted steady state reach accruals for each curtailment period (using ESPAM v1.1).

Cutoff Date	All Pumping		Post January 1, 1949		Post January 1, 1961		Post January 1, 1973		Post January 1, 1985	
Reach Name	Steady State Gain (cfs)	Time (yrs) to realize 90% of gain	Steady State Gain (cfs)	Time (yrs) to realize 90% of gain	Steady State Gain (cfs)	Time (yrs) to realize 90% of gain	Steady State Gain (cfs)	Time (yrs) to realize 90% of gain	Steady State Gain (cfs)	Time (yrs) to realize 90% of gain
Ashton to Rexburg	316	20	290	20	222	19	128	19	22	19
Heise to Shelley	211	18	195	18	141	17	78	17	15	17
Sum of Reaches Above Shelley	**526**		**484**		**363**		**206**		**37**	
Shelley to Near Blackfoot	406	30	368	30	240	28	132	25	26	28
Near Blackfoot to Neeley	1035	35	925	36	593	33	331	29	69	32
Neeley to Minidoka	158	62	134	65	84	61	46	55	10	60
Sum of Reaches Shelley to Milner	**1598**		**1427**		**917**		**509**		**105**	
Devil's Washbowl to Buhl	298	59	257	61	160	57	88	51	18	60
Buhl to Thousand Springs	137	49	122	50	81	43	47	37	8.0	51
Thousand Springs	89	48	79	48	54	42	31	35	5.6	47
Thousand Springs to Malad	10	45	9	46	6.4	38	3.7	32	0.74	42
Malad	77	50	68	51	47	43	29	35	5.5	44
Malad to Bancroft	5.8	38	5.2	38	3.8	34	2.4	31	0.59	36
Sum of Thousand Springs Reaches	**617**		**541**		**352**		**201**		**38**	
Total (cfs)	**2741**		**2453**		**1633**		**916**		**180**	
Total (acre-feet/year)	**1,984,659**		**1,775,637**		**1,182,086**		**663,362**		**130,506**	