23692 Birtcher Drive
Lake Forest CA 92630

949.420.3030 phone
530.756.5991 fax
westyost.com

Date: September 13, 2024

Project No.: 941-80-24-32
SENT VIA: EMAIL

Todd Corbin
General Manager
Chino Basin Watermaster
9641 San Bernadino Road
Rancho Cucamonga, CA 91730

**SUBJECT: Response to Comments from Appropriative Pool following 2025 Safe Yield Reevaluation Calibration Workshop #2**

Dear Todd:

This letter documents our responses to the Appropriative Pool's comments following the August 6, 2024 workshop[1] regarding the calibration and uncertainty analysis of the 2025 Chino Valley Model (2025 CVM) as part of the ongoing 2025 Safe Yield Reevaluation. Thomas Harder and Company (TH&Co; Jim Van De Water, PG, CHG and Thomas Harder, PG, CHG) submitted the comments on behalf of the Appropriative Pool (AP) in a letter dated August 23, 2024.

# TH&CO COMMENTS AND RESPONSES

## Comment 1

It is our understanding from a discussion at the end of the presentation that a covariance matrix is not being used. On Slide 10, could it be the lack of a covariance matrix that is leading to the unrealistic parameter fields as opposed to overfitting?

Parameters that are not spatially correlated (e.g., array multipliers in various input files ['packages']), are defined by their means and standard deviations whereas those that are spatially correlated are defined by their means and covariance matrices. If a covariance matrix is not being used, does it mean that the pilot point parameters are not spatially correlated? That is, are they behaving independently of one another? Could this be the reason for unreasonable parameter fields as opposed to overfitting? From our read of https://help.pesthomepage.org/pestpp-ies.html, it appears that not using a covariance matrix for spatially correlated parameters is an unwise choice.

[1] 8/6/2024 Workshop Agenda; 8/6/2024 Workshop Presentation

**Response:**
We are using a covariance matrix for each aquifer parameter in a model layer. The covariance matrices were generated by using the PEST PPCOV utility. Our results as shown from the August 6th workshop (e.g., slide 28) indicate realistic parameter fields that do not exemplify overfitting. We will include more hydrographs and other exhibits to convey these results in the more detailed documentation that will be included with the 2025 SYR report.

## Comment 2

It is our understanding that [West Yost (WY)] had some difficulties generating a covariance matrix on their system. Which utility did WY use when attempting to generate a covariance matrix? Note that there is a new PEST Suite (Suite 18) and issues encountered by WY may be solved if the updated utility programs are used. We've noted that the new version of "ppcov" performs much faster on our system than previous versions.

**Response:**
Our challenges arose from the initialization of PESTPP-IES, not generating a covariance matrix. We are using PEST Suite 18. We have prepared a parameter uncertainty file that includes covariance matrices for each calibrated aquifer property in each model layer. Those covariance matrices were generated with the PEST PPCOV utility.

The difficulties that we had experienced occurred in the initialization of PESTPP-IES when it attempted to generate an ensemble of initial parameter fields for 160,000 parameters and 400 realizations. Consequently, we instructed PESTPP-IES to load an ensemble of initial parameter fields that was generated by using the PEST RANDPAR3 Utility program. PESTPP-IES was able to start the loading process, but eventually caused a crash of the entire high-performance computing system that we employed for PESTPP-IES runs. Further investigation revealed that loading the initial ensemble of parameters used more than 50 GB of RAM before the system crashed. We are currently attempting to rerun this configuration with increased RAM.

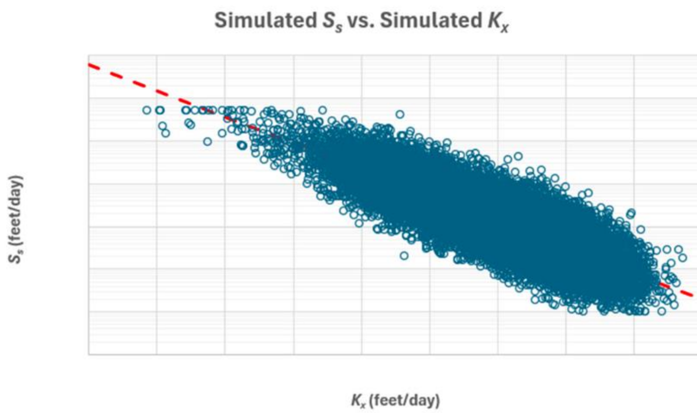
## Comment 3

Regarding overfitting, do individual hydrographs show simulated heads to be underpredicted during periods of low measured groundwater elevations and overpredicted during periods of high groundwater elevations? That is, do the simulated "peaks" and "troughs" exceed those that are measured? If not, is it reasonable to conclude that the model is not overfit?

**Response:**
The hydrographs for a vast majority of the calibrated realizations do not indicate overfitting.

## Comment 4

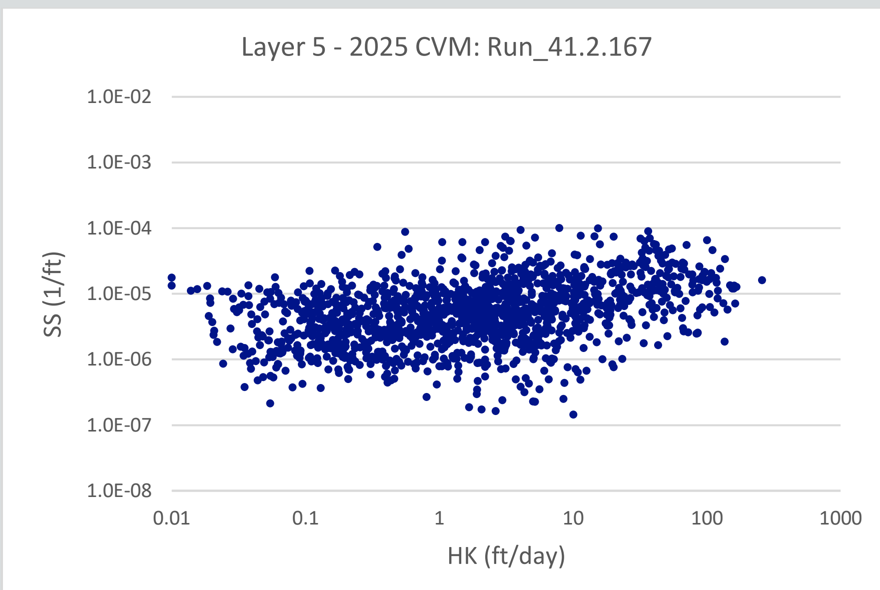
Has WY checked the parameter fields for internal consistency? That is, are areas of relatively high specific yield and high vertical conductivity reasonably co-located with areas of high lateral hydraulic conductivity (Kx)? Or, conversely, are areas of relatively low specific storage (Ss) reasonably co-located with areas of high lateral hydraulic conductivity (Kx and/or Ky)? The figure below shows a model with roughly 28,000 parameters assigned to Ss and another 28,000 parameters to Kx. WY may want to consider developing this type of figure as an internal check for inconsistent parameter associations. Inconsistent parameter associations could lead to the unrealistic parameter field shown on Slide 10.




**Response:**

The parameter field presented in slide 10 was shown to clarify our use of the term "overfitting" from the 5/29 workshop. Our current results do not suggest that parameter fields are overfit. We have developed figures as you suggested, such as the one below. We will include similar charts in the more detailed documentation of the calibration and uncertainty analysis.

This figure shows HK versus SS for model cells in Layer 5 for one of the realizations in the PESTPP-IES run with 25,326 parameters. HK and SS are slightly positively correlated. This observation is in line with studies[2,3] that suggest no clear spatial correlation between K and Ss in heterogeneous aquifers. The relationship between K and Ss is complex and depends on various factors, such as the size and compressibility of the porous media. It's often necessary to consider both properties together to understand the behavior of groundwater systems.




[2] https://doi.org/10.1016/j.jhydrol.2019.124383

[3] https://doi.org/10.1016/j.jhydrol.2022.127921

## Comment 5

With respect to Slide 17, we have not experienced a notable slowing of the process by simply adding parameters. Perhaps the input files take a bit longer to write but the most notable effect on runtimes we have experienced is associated with increasing the ensemble size (i.e., the number of realizations). This is not to say that doing so (as WY has done in increasing their initial ensemble size from 100 to 400) is a bad thing – we have noticed that the calibration is generally improved with an increased number of realizations.

That said, it is impossible for us to comment without being provided the CVM and the IES setup. Are there plans to release the CVM and the IES setup to the stakeholders? We strongly recommend that this be done.

**Response:**
We recorded the required computing time and resources for each of the simulated model ensembles using the same configuration within our High-Performance Computing environment. The required computing time clearly increased with the increasing number of model parameters. The increased computing time can be attributed by many factors, including the increased time for reading and writing data from/to storage devices, for interpolating from pilot points to model cells, and for consumption of the information stored in the parameters by PESTPP-IES.

Watermaster will not release the CVM pursuant to the 2020 Safe Yield Court Order and Watermaster's Model Information Guidelines.[4] However, through the peer review process, including sharing an earlier version of the PESTPP-IES setup with you, we aim to provide the requisite information to allow peer reviewers to evaluate the model's validity and robustness. If you have specific recommendations for outputs that you would like to see, we will incorporate your requests in future reporting as appropriate.

## Comment 6

With respect to Slide 18, if HK (i.e., Kx in Comment 2 above) is varied, why isn't Ky being varied? Put another way, how can one "not know" the former yet "know" the latter?

**Response:**
We assumed that the K values in a model cell are isotropic, i.e., Ky = Kx. Therefore, Ky varies with Kx.

## Comment 7

Is it possible that parameters not being varied in the IES configuration (i.e., "fixed parameters") are causing those parameters that are varied to assume surrogate roles to account for incorrect assumptions regarding the values assigned to the fixed parameters? Could this lead to problems in calculating the safe yield?

---

[4] July 31, 2020 Orders re Chino Basin Watermaster Motion Regarding 2020 Safe Yield Reset, Amendment of Restated Judgment, Paragraph 6, pp. 6–8; Watermaster Board Resolution 20-05 Regarding Procedure and Fee Schedule for Requesting Information and Documents Related to the Chino Valley Model; Watermaster Board Resolution 01-03 Adopting Procedures, Guidelines and Fee Schedule for Release of Information and Documents.

**Response:**
All aquifer parameters were varied during the calibration.

## Comment 8

Has WY checked the literature to assess what is considered "lots of parameters"? If so, does the literature suggest that using, say, more than 100,000 parameters would be considered too many and could lead to overfitting for large and complex basin-scale models such as the CVM?

**Response:**
Most literature suggests that the number of parameters for an inverse problem should be determined based on the model resolution, problem complexity, data availability and quality, and computational cost, etc. A definition of required or preferred number of parameters does not exist based on our literature review and discussions with other groundwater modelers. The source code of PESTPP-IES and the PESTPP-IES user's manual[5] description of the *ies_updgrades_in_memory_flag* suggest that 100,000 parameters is the threshold for "very high dimensional problems."

We have compared the simulated net recharge of an ensemble of 335 models (each with 2,701 parameters) with an ensemble of 316 models (each with 25,326 parameters). The range of the simulated net recharge values of these two ensembles remained practically unchanged.

It is our professional opinion that a further increase of the number of parameters will not significantly affect the results of the model estimates. However, we are currently attempting to rerun the configuration with 160,000 parameters with increased RAM. If this run is successful, we will describe the results in future documentation.

## Comment 9

With respect to Slide 20, which shows that, of the 400 realizations in the initial ensemble, only 316 survived. That is, 21% of the realizations were lost during the IES run using the default setting of 2 times the average runtime to that point. This seems to be a lot for a model as mature as the CVM and may indicate some model defects/stability issues. We commonly see this sort of attrition rate for new or vastly revised models but not for older models.

That said, and as noted above in Comment 5, it is impossible for us to comment further without having the CVM and IES setup.

**Response:**
The run time of a realization depends on its parameter values and sometimes the average runtime of the few models that are completed first. Models with "bad" combinations of parameter values will take much longer to run (or result in a poor fit with the observed data) and eventually be dropped when its runtime exceeded twice the average runtime. The likelihood of "bad" combination of parameter values increases with the allowed range of model parameters.

---

[5] pestpp/documentation/pestpp_users_manual.md at master · usgs/pestpp (github.com)

We have noticed that the calibrated realizations ran reasonably fast when we ran those realizations in the post-process phase to generate net recharge time series.

## Comment 10

As a general observation, we have sensed some frustration among the parties during this workshop and others that the development of the alternative future scenarios has been overly time-consuming and complex. We provided comments to the previous workshop (Scenarios Workshop #3) on July 19th. Those comments provided what we believe to be a simplified way to capture future uncertainty and, if necessary, express multiple futures in the IES setup. While we realize this issue may be addressed at the upcoming Scenarios Design Workshop #4 on August 27th, we feel compelled to ask for the reason why our approach has been apparently rejected.

**Response:**
While we acknowledge the feedback regarding the perceived complexity and time required for developing the alternative future scenarios, Watermaster is required to conduct the 2025 Safe Yield Reevaluation (2025 SYR) in accordance with the Court-ordered 2022 Safe Yield Reset Methodology (2022 SYRM).[6] The 2022 SYRM also better accounts for future uncertainties, responding to the Appropriative Pool's concerns that arose during the 2020 Safe Yield Reset process. Further, the more detailed analysis required by the 2022 SYRM enables more cost-effective development of decision-making tools to optimize Basin management, such as application of the recently released water rights forecasting tool.

Throughout the development of the 2022 SYRM and since initiating the 2025 SYR, parties have been given multiple opportunities, both verbally and in writing, to express any concerns or objections to Watermaster's approach, including regarding the scope and budget for identifying projected cultural conditions and developing scenarios.[7] No written comments were received opposing the 2022 SYRM. We remain committed to transparency and collaboration.

We appreciate your detailed suggestion for an alternative approach; we have responded to your approach in detail in an attachment to Scenario Design TM #3.[8] We plan to incorporate some of your suggestions into our development of the projection scenarios, particularly the use of proxy years for determining multipliers for projected groundwater model inputs. However, portions of the suggested approach do not meet the requirements of the 2022 Safe Yield Reset Methodology.

---

[6] December 16, 2022 Order Granting Chino Basin Watermaster's Motion Regarding the Update to Watermaster's Safe Yield Reset Methodology. No party opposed the Court's approval of the 2022 SYRM.

[7] Draft and final TMs for the 2022 Safe Yield Reset Methodology, including workshop materials, can be found here. We have followed the approach outlined in Section 4.3.2 of the 2022 Safe Yield Reset Methodology TM.

[8] 8/27/2024 Draft Scenario Design TM #3

## CONCLUSION

We look forward to a continued dialog with the peer review committee, and we will include all comments and responses in the final 2025 SYR report. Please let us know if you have any questions or concerns regarding the above comments and responses.

Sincerely,
WEST YOST

Garrett Rapp, PE
Senior Engineer I
RCE #86007

Eric Chiang, PhD
Principal Engineer II

cc: Edgar Tellez-Foster, PhD; Andy Malone, PG