

From: Ron Weitzman <ronweitzman@redshift.com>
Sent: Sunday, July 12, 2015 3:46 PM
To: MPWSP-EIR
Cc: Californian; Carmel Pine Cone; Cedar Street Times; Channel 11; Herald City Editor; Jim Johnson; Kera Abraham; KION TV ; KSMS TV; MC Weekly Editor; Sara Rubin; Shanna McCord
Subject: Request for Data & Support
Attachments: Problems with Model Evaluationl.docx

Mary Jo Borak, CPUC
c/o ESA

Dear Ms. Borak:

This is a request for data I made prior to the decision to grant such a request. So I am repeating it.

This is also a “vote” by Water Plus as a party to the proceeding on A.12-04-019 for recirculation of the DEIR as a CEQA-NEPA document.

Here is my data request: The scatterplot in Appendix E-2 of the DEIR provides inadequate and misleading information for the evaluation of the model. As a statistician, I need to see the data used to construct that scatterplot so I can professionally and responsibly evaluate the model. Not only does the scatterplot constitute inadequate information for model evaluation, the reported evaluation statistic computed from the same data used to construct the scatterplot is also very likely inaccurate, and the only way I or any other statistician can determine its accuracy is to have the data used to compute it. Unless you provide access to those data, the DEIR will be inadequate and likely erroneous on a matter crucial to the assessment of the project’s viability and environmental impacts. In the absence of available slant-well data collected on site over a period of years, the soundness if the decision to go forward on the project, costing hundreds of millions of dollars, depends on the validity of the model. Its appropriate evaluation is critical.

Please provide me as soon as possible the spreadsheet data used to construct the scatterplot and particularly provide me these data as an Excel document that I can use to check on the calculations and make further calculations, as necessary.

Also, if currently available, please let me know ASAP the numerical value of the correlation coefficient describing the scatterplot and also be sure to report that correlation coefficient in any revised and recirculated DEIR.

Thank you.

Ron Weitzman

President, Water Ratepayers Association of the Monterey Peninsula (Water Plus)

P. S. To reinforce this request, I have attached a comment on the DEIR I submitted to the PUC Energy Division on June 30. This comment further details the problems I see in the DEIR on model evaluation.

Problems with the Evaluation of the CEMEX Model

In investigating the DEIR evaluation of the CEMEX model, I have found some serious problems. Appendix E-2 shows a graph (scatterplot) of observed and model-predicted measurements of groundwater elevations at different sites and times. To critique the evaluation thoroughly, I would need the Excel spreadsheets showing the data and the calculations, which the DEIR does not supply.

For example, I cannot tell from the graph and its description in the text whether the residuals between predicted and observed measurements were computed by subtracting the first from the second (correct) or the second from the first (incorrect), which Appendix A to Appendix A of Appendix E-2 indicates is being done. Supporting the belief that the determination of residuals might be incorrect is the mean residual reported on the graph. It is -0.7 feet when it should be zero if the residuals had been determined correctly. As presented, the graph shows the estimation of predicted from observed measurements when a prediction line having a slope equal to one, as this one does, would be correct only if the estimation were in the opposite direction, unless the prediction were perfect (all the points falling on the line).

If you look at the graph, the difference between these two methods of residual computation for each data point is the identification of its residual as the horizontal distance (correct) or the vertical distance (incorrect) between the point and the diagonal line running through the dots from lower left to upper right.

Doing the calculation incorrectly could give a substantially better, but incorrect, evaluation of model fit. I would need the Excel spreadsheet showing the data and the calculations to determine which way the residuals were determined and, if determined incorrectly, what the correct model evaluation might be.

Another serious problem with model evaluation is that the data points in the graph represent an area far more extensive than the CEMEX site to which the model is applied, as well as depths not accessed by the slant well there. The depths represented include not only the 180-foot aquifer, which the slant well accesses, but also the 400- and 900-foot aquifers, which the slant well does not access. In fact, the

graph shows no data points representing the Dune Sand aquifer, even though it is accessed by the slant well. To provide a truly apt model evaluation, I would need the Excel spread sheet to evaluate the model separately for the 180-foot aquifer. Examination of the graph suggests that the fit for these data points (yellow-filled circles) may not be very good at all, many of the predicted elevations being uniformly much too high.

Since the slant well draws water only from the Dune Sand and 180-foot aquifers, a graph and residual calculations should be based on data from only these depths within the CEMEX area.

All this may seem overly technical, but the model fit is crucial to the DEIR since its conclusions are based on unconfirmed model predictions of future slant-well effects.

REPORTING REMEDIATION. The EIR should report the correlation coefficient between observed and predicted values for both the entire data set and for the data set restricted to the CEMEX area and the 180-foot aquifer. This statistic does not depend on the method of residual determination. One minus its square is the proportion of observed measurement variation that is not predicted, or is unaccounted for, by the model. This proportion is the statistical standard of model evaluation.

OPERATIONAL REMEDIATION. Since the model is unevaluated for one portion of its intake source region (Dune Sand aquifer) and is inadequately and likely erroneously evaluated for the other (180-foot aquifer) at the CEMEX site, reliance on the model to make predictions of the impact of slant-well pumping on groundwater elevations is out of the question for any envisioned scenario. Data collection over dozens of months of test-well pumping is necessary to evaluate the effects of the proposed project on groundwater elevations before a decision can responsibly be made to go forward with the project. Without such data collection, no EIR would be adequate and no CPCN could be justified.

If you do not take these remediation measures, please explain, Why not?