Memorandum

To: OWRD Groundwater Allocation Rulemaking Team

From: Ben Scandella, Groundwater Data Chief

Date: 1/16/2024

Regarding: Response to feedback on draft memo, “Analysis of Oregon wells correlated with precipitation”

Summary

After developing a draft memorandum on Analysis of Oregon wells correlated with precipitation, the Department sought technical feedback from both a scientific institution and the Rules Advisory Committee (RAC). The draft was sent to the U.S. Geological Survey (USGS) for peer-review by two scientists not involved with the allocation rulemaking. It was also shared with the RAC ahead of and during the December 11th, 2023 meeting and in two technical information sessions held on January 8th and 9th, 2024. The USGS summary review is attached, and this document summarizes the major points brought up by both the USGS and RAC commentors, along with descriptions of how the draft analysis memo was updated (1/16/2024) to incorporate their feedback.

U.S. Geological Survey Peer Review

Detrending

The USGS peer review expressed a primary concern that detrending the water levels before analysis biased the selection of wells used in this analysis. They noted that detrending during the first step of the analysis would include wells that are insufficiently correlated with precipitation after detrending and exclude wells that would be sufficiently correlated without detrending. This critique was substantially addressed by reversing the order of operations of the assessment process; correlating water level records with precipitation for the sake of well selection, then detrending levels ahead of subsequent analysis. This reordering of the analysis method now includes wells that had a long-term trend that was correlated with temporally-averaged precipitation. However, those long-term trends were still removed from the ensuing analysis in order to create a set of water level fluctuations consistent with remaining stable over their period of record. The purpose and impact of detrending is discussed more below in response to feedback from the Rules Advisory Committee. The updated analysis is described in an revised memo dated 1/16/2024, which includes additional sensitivity analysis to quantify the impact of detrending water levels later in the assessment on the resulting water level trend statistics.

Minimum Span of Record for Evaluation of Rate Test

An additional comment from the USGS, which was also echoed by a RAC member, is that the minimum span of 10 years for evaluating the rate test should be increased to 16 years to reflect the stabilization of the results with minimum time spans of 16 years and greater. This comment was evaluated with an updated sensitivity analysis incorporating the processing parameters, which indicated that the distribution largely stabilizes with minimum initial spans of 13 years or greater. Therefore, the minimum span was increased from 10 to 13 years. The details of this analysis are described in the updated memo dated 1/16/2024.

Rules Advisory Committee Feedback

Natural vs. Stable Records

RAC members noted that the approach does not select for water level trends that are exclusively driven by precipitation fluctuations, but also include fluctuations driven or amplified by anthropogenic activities. We agree and clarified that the approach is designed instead to select and refine water level records that represent stability, rather than natural fluctuations. This point was clarified in the Technical Information Sessions on January 8th and 9th, as well as in the 1/16/2024 version of the memo.

A RAC member asked how the requirement for correlation with precipitation would influence the results. That comment is addressed in the sensitivity analysis of the results to the minimum coefficient of determination (R^2), which showed modest increase in magnitude of decline and a negligible impact on the rate as the minimum R^2 is reduced to 0. The limited increase in number of well clusters as the minimum is set to zero suggests that the vast majority of wells meeting the other selection criteria also have at least some correlation with precipitation averaged over up to 50 years. Wells excluded from the analysis even with a minimum of 0 are those with negative correlation coefficients, R.

Consideration of Type II Errors

RAC members pointed out that the analysis characterizes type I errors and requested that additional analysis be performed to characterize and limit type II errors. Type I errors in the context of this analysis are when stable records are incorrectly classified as not Reasonably Stable, and type II are when declining records are incorrectly classified as Reasonably Stable. RAC members suggested that, given the goal of the rulemaking process and the perpetual nature of groundwater permits, consideration of type II errors is more important than that of type I errors. They pointed out that, given the structure of the proposed definition, increasing the limits on magnitude and rate of decline will reduce type I errors but increase type II errors, and decreasing the limits will increase type I errors but decrease type II errors. A more sophisticated definition of Reasonably Stable could do more to limit both types of errors but would require more complexity and/or professional judgment, in tension with the objective to keep the definition simple. Requiring longer periods of initial data collection could also reduce errors but would be in tension with the goal to limit the burden of initial data collection in areas without sufficient data.

To address this issue, RAC members suggested that type II errors could be evaluated by applying the proposed definition with a range of thresholds to a set of wells that represent declines. However, identifying an analogous set of water level trends that represents declines is more difficult than identifying a set of water level trends that clearly represents Reasonably Stable behavior, as was done here. Wells may experience periods of both stability and declines over their periods of record, and an objective of the proposed test is to detect the onset of declines after an initial period of stable behavior. That objective is partly addressed by limiting the duration of historical averaging in the rate test to 20 years. However, selecting a representative set of water level records that include such a transition would be a challenge that would require either expert judgment in each case, or some set of criteria that would likely have a significant influence on the results. This expanded analysis was not integrated into the 1/16/2024 draft of the memo but is being considered.

Instead, the present analysis sought to limit type II errors by excluding from the sample set water level fluctuations that had lower likelihood of representing stable variations. Those fluctuations were excluded by limiting declines over the period of record to 0.5 feet per year and by removing the best-fit trend before quantifying declines. Both of those restrictions had the impact of limiting the measured magnitude and rate of decline in the final results, as a consequence of offering more confidence that the analyzed records represent stable behavior.

Trends Over the Period of Record

RAC members pointed out that the approach outlined here excludes some declining water level trends that could be expected to recover due to their correlation with precipitation. The exclusion of wells from the precipitation-correlated set was caused by a number of factors, including the limited averaging period used for precipitation (2 through 10 years), the detrending of water levels before correlating with precipitation (as presented in the memo draft from 12/11/2023), and the maximum allowed rate of decline over the period of record (0.5 feet per year). On top of this, the detrending before evaluation of total declines and rates of decline limited both metrics compared with not detrending (see sensitivity analysis of detrending below). The discussion below addresses these critiques through a combination of describing responsive improvements to the analysis and explaining why the restrictions imposed were appropriate.

RAC members noted that the range of groundwater response times (GRTs) in Oregon may extend beyond the range of averaging periods used for smoothing precipitation records before correlating with water level records. The backward-looking averaging of precipitation is intended to reflect the GRT between recharge and observation in a well, in that longer times for a pressure perturbation to propagate from a zone of recharge to an observation well can be captured by smoothing the precipitation signal over a correspondingly longer period. Allowing correlation of water levels with precipitation over longer backward-looking averaging periods enhances the effect described above, because longer averaging periods enhance temporal smoothing. Thus, the fluctuations in longer-averaged precipitation records will be typically extended over longer time periods, increasing the potential for inclusion of trends that span the entire measurement period for water levels in a given well. In response to this critique, the longest backward-looking averaging periods was increased from 10 years to 52 years. Water levels were also allowed to correlate with precipitation before they were detrended. As a result, additional wells were included in the sample set that revealed issues warranting additional requirements on the sample set. Wells with deepenings had resulting water level impacts included due to those deepenings (e.g. UMAT 1311), so deepened wells were excluded from the sample set. In addition, some wells had substantial increasing water level trends over the long run (e.g. JEFF 821 and JEFF 822) but, after detrending, exhibited substantial decreasing trends. To limit this effect, wells were excluded from the analysis set if the magnitude of their long-term trend (either declining or increasing) exceeded 0.5 feet per year. These collective changes had the effect of increasing the resulting sample of wells from 236 to 357 and clusters from 106 to 160.

RAC members asked for justification for the maximum rate of change over the period of record of 0.5 feet per year. The threshold of 0.5 feet per year was set in consideration of both the reliable limits of detection with the measurement method and ability for a plausible record to be maintained after detrending. Sensitivity analysis to this threshold is presented below and shows that results are effectively insensitive above 0.5 feet per year. It is possible in principle that water levels could recover from declines exceeding 0.5 feet per year over a period of record at least 25 years, so the restriction of 0.5 feet per year over the period of record may have exclude some valid stable fluctuations from the test set. However, loosening the limitation on maximum decline rate over the period of record would also include more wells that are not expected recover due to the potential for declines to be driven by non-cyclical effects like additional pumping. Doing so would then increase the frequency of type II errors, incorrectly identifying unstable wells as stable.

Detrending water level records before evaluating the magnitude and rate of decline has the effect of reshaping the sample set to better represent long-term stability. However, given that the majority of wells exhibited average declines over their period of record, this detrending typically had the impact of reducing the largest magnitude and rate of decline measured in wells (see sensitivity analysis in the 1/16/24 memo). Detrending water levels before analyzing the magnitudes and rates of decline therefore removes declining trends over the period of record that are correlated with precipitation, and therefore have the potential to recover if both precipitation and groundwater recharge increase Detrending therefore underestimates the thresholds appropriate to the dynamically stable range to some extent. This means that findings of Not Reasonably Stable will have slightly lower confidence of being correct (not making a type I error of classifying a stable well as unstable). At the same time, detrending water levels respects the fact that correlation between water levels and precipitation does not, by itself, imply that water levels will recover following declines. For one, there is some debate about whether recovery of precipitation will lead to equivalent recovery in groundwater recharge due to increasing atmospheric evaporative demand. In addition, long-term declining trends in groundwater level may be driven by anthropogenic activities that are not expected to be reversed, such as additional groundwater pumping, in addition to declining precipitation. Neither this analysis nor the proposed definition of Reasonably Stable are suited to attribute the causes of impacts on a statewide basis. Instead, both rely on the evidence apparent in water level records to demonstrate stability. This analysis does so by requiring at least 25 annual high water level measurements to show a best-fit rate of change slower than 0.5 feet per year, and by removing that best-fit trend to characterize residual fluctuations (after detrending). The assumption is that the long-term trend is driven by anthropogenic changes and that the residual fluctuations represent the influence of precipitation fluctuations. The resultant benefit of making this assumption and detrending the water levels is that it reduces type II errors (incorrectly classifying an unstable well as stable).

Tuning a Model Using Early Data in Each Well

RAC members suggested an alternative approach of tuning a regression model for precipitation-correlated changes within each well, and then removing this model from the water level record in order to discern water level behavior inconsistent with precipitation variations. As the RAC member succinctly described it, such an approach would help to distinguish the “signal” of non-cyclical water level declines from the “noise” of precipitation-correlated fluctuations. Such an approach could yield substantial increases in sensitivity (reduction in type II errors) while retaining selectivity (avoiding type I errors). However, implementing it would require a significantly more complex definition, in tension with the objective to keep the definition simple to describe and apply. Important questions that would need resolution include (1) what constraints would be applied to such a model to enforce its consistency with physical principles (e.g. requiring the regression coefficients to be positive)? and (2) how far back in time is precipitation eligible to be correlated? Including annual precipitation lagged over decades introduces enough free regression parameters that the regression model may be over-fit when water-level data are available over less than a decade (a minimum of 5-years of data are required for analysis under the current proposed rule language). This approach would require extensive further assessment to implement.

Detrending Method and Illustration

A RAC member asked why the detrending was performed using the least-squares best-fit trend line. A more robust method for identifying a representative trend is the Theil-Sen slope, which is the median slope among all pairwise slopes. The method was updated to use the Theil-Sen slope, with negligible impact on the results. A RAC member asked for a visual illustration of how the detrending process was applied to raw water level records. Example figures were added to the 1/16/2024 memo to address this request.

Spatial Extrapolation

A RAC member asked about whether nearby wells can be used to evaluate Reasonably Stable, and if so, what limitations might exist. While this question did not refer to the analysis presented in the subject memo, staff affirmed in the Technical Information Session on January 9^{th} that nearby wells in the same aquifer could be eligible for determining stability in a proposed point of appropriation without sufficient water level data. The distance over which that extrapolation or interpolation could be applied has not been specified in rule because the geometry and properties of an aquifer and hydraulic barriers will vary in each case.

Clustering over Different Aquifers

A RAC member asked if the clustering method potentially allowed wells in different aquifers to be clustered together. This is possible if either well did not have an aquifer or aquifer system selected and if they also met the other similarity criteria for clustering. This was made explicit in the 1/16/2024 memo.