Memorandum

To: OWRD Groundwater Allocation Rulemaking Team

From: Ben Scandella, Groundwater Data Chief

Date: 2/22/2024

Regarding: Analysis of Oregon wells correlated with precipitation

---

Summary

The Department is in the process of updating its rules for issuing new groundwater permits, and the proposed new rules include a new definition for Reasonably Stable Groundwater Levels. This proposed definition attempts to balance multiple policy objectives, including characterizing groundwater levels as Reasonably Stable as long as they remain within the range of observed variability consistent with long-term stability. This memo evaluates two aspects of this variability relevant to the proposed definition: the magnitudes of water level cycles and the rates of decline calculated using the proposed definition. Those two attributes are evaluated using data from wells that are assumed to be Reasonably Stable, with water level records at least 25 years long that (1) are correlated with precipitation and (2) show limited best-fit linear rates of change over the period of record.

The observed attributes vary among wells but do not show obvious dependence on location. This memo characterizes their statewide statistical distributions. These distributions may be combined with policy objectives, such as the percentage of Reasonably Stable wells that should be correctly classified as Reasonably Stable, to determine appropriate thresholds in the proposed definition. The determination of appropriate thresholds may also consider errors of incorrectly classifying declining trends as Reasonably Stable and impacts of additional groundwater level declines, which are not addressed directly in this memo.

Contents

Introduction

The proposed definition of Reasonably Stable Groundwater Levels balances multiple policy objectives, including being both sensitive to the onset of clearly declining trends and robust to water-level fluctuations that are likely still consistent with long-term stability. These objectives exist in tension with each other and with others, including to limit the complexity of the proposed definition and to limit the burden of additional data collection in areas without sufficient data. The proposed definition includes thresholds on the maximum allowable magnitude and rate of decline that are considered consistent with reasonable stability, and the Department seeks to inform the values of those thresholds using relevant data and analysis. This memo primarily addresses the ability of the proposed definition to characterize groundwater levels as Reasonably Stable as long as they remain within the range of variability consistent with long-term stability.

The range of water levels that fluctuate around a value that remains constant over the full period of measurement has been termed the “dynamically stable range” (Gleeson *et al.*, 2020). Groundwater use may be considered renewable if water levels are expected to remain within this range or to be able to recover to within this range over a timescale for human planning (suggested as 100 years in Cuthbert *et al.*, 2023). In the context of allocation of issuing new groundwater permits without end dates, and where regulation of groundwater use has not produced timely relief, long-term groundwater level declines caused by additional groundwater permits do not lead to an expectation of recovery. Therefore, for the purposes of this analysis, the dynamically stable range does not include expected drawdown to a new dynamic equilibrium that may be established following the full capture of surface water by increased groundwater pumping.

This memo characterizes the dynamically stable range to-date using water level data in wells that show limited long-term declines over their full period of record of at least 25 years. The memo further focuses on wells with groundwater-level behavior correlated with regional precipitation. Precipitation trends are a major driver of significant groundwater level fluctuations that have the potential to remain stable over the decades to centuries. Water levels in most groundwater reservoirs in Oregon rise and fall due to natural variability in precipitation and consequent groundwater recharge, and these natural fluctuations are often amplified by anthropogenic activities. In some settings, persistent groundwater pumping causes water levels to decline, but in others, the additional pumping is offset by capture of surface water that allows water levels to remain stable over the long run. Because of the potential for stable fluctuations to be amplified by anthropogenic activity, attributing the causes of declines is less important than observing that the water level record remains stable over the full period of record.

Even if the water-level record remains stable over decades, these sources of variability have the potential to cause the determination of "Reasonably Stable Groundwater Levels" to oscillate between true and false in the intervening years. Such oscillations between Reasonably Stable and not could create substantial uncertainty and conflict between applicants for water rights under the Prior Appropriation system. Based on a Reasonably Stable determination, an applicant could be issued a permit after an earlier applicant was denied because prior water levels led to a determination of not Reasonably Stable. Some instances of this undesirable situation are unavoidable with a rule-based definition of Reasonably Stable. The lack of a mechanistic understanding of *why* groundwater levels are declining limits the ability of any rule to predict whether observed declines are part of a natural cycle that is expected to recover, or whether they are the beginning of a new system state characterized by ongoing declines.

In the absence of such mechanistic understanding, the thresholds in rule should be set so that the proposed definition of Reasonably Stable achieves a tolerable balance of the two types of errors of inference: type I (false positive) and type II (false negative). If the test for stability evaluates a null hypothesis of a stable water level trend, then a type I error is when the test characterizes a well that is truly Reasonably Stable as declining. By the same token, a type II error is when the test characterizes a well that is declining as Reasonably Stable. The fact that the proposed definition offers no middle ground of "Cannot be Determined" means that there is a zero-sum game to be played in setting the thresholds. Higher tolerances on the magnitude and rate of decline will reduce the incidence of type I errors by allowing for larger or faster declines before a well is classified as declining, but they will consequently increase the incidence of type II errors.

The approach pursued in this memo is to select and construct a set of water level trends that are assumed to represent Reasonably Stable, and then test how consistently the components of the proposed definition successfully characterize those trends as Reasonably Stable. The consistency depends on the threshold values on the magnitude and rate tests in the proposed definition. Thus, it directly quantifies type I errors as a function of the threshold values. To the extent that those errors can be limited, the results here will indicate confidence that determinations of Not Reasonably Stable are correct. However, the thresholds in the proposed definition should also be set to limit type II errors, which indicate that the test is insufficiently sensitive and does not meet the stated objective to detect the onset of consistent declining trends. Given the perpetual nature of groundwater permits in Oregon, the consequences of failing to detect a declining trend are significant. This analysis lacks a data set that represents declining water level trends across the state so does not attempt to quantify the rate of type

II errors. Instead, it seeks to limit them by excluding water level trends that may not represent Reasonably Stable behavior from the sample set. The criteria used to include wells and trends in this analysis received significant attention in peer review and are discussed and analyzed below.

A key assumption of the analysis below is that the past is the key to the future: that precipitation and groundwater recharge can be expected to fluctuate around a long-term stable value. While climate forecasts suggest that annual precipitation is expected to remain stable over the foreseeable future, average groundwater recharge may decline due to increasing evapotranspiration with rising temperatures, decreased snowmelt, and increased temporal concentration of precipitation events, which can reduce the portion of precipitation that can infiltrate and recharge aquifers (Albano *et al.*, 2022; Mote *et al.*, 2019).

Methods

Identification Wells with Precipitation-Correlated Groundwater-Level Records

The set of wells analyzed was restricted to those whose best-fit linear rate of change over the period of record are limited to less than 0.5 feet per year. Sensitivity analysis to this threshold is presented below. This restriction, combined with removal of the best-fit linear trend over the period of record, was used to enforce that the water-level trends analyzed represented reasonable stability. 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. So long as the best-fit linear trend over the period of record are within this limit, even water level fluctuations amplified by anthropogenic activity were considered valid for inclusion in this analysis (after detrending, discussed below) as representative of the dynamically stable range. Sensitivity analysis to this threshold is presented below and shows that results are effectively insensitive above 0.5 feet per year.

Water levels were compared against precipitation from the National Oceanic and Atmospheric Administration's "Climate at a Glance" dataset in its Climatic Division (9 in Oregon) containing each well (National Oceanic and Atmospheric Administration (NOAA), 2024). Precipitation was averaged with a backward-looking moving-average window with window spans of 2 through 50 years, reflecting a wide range of response times between recharge and water-level changes. Wells were included in the analysis set if they were sufficiently correlated with precipitation averaged over any of these averaging windows, allowing inclusion of wells in a range of hydrogeologic settings around the state without needing to specify the appropriate groundwater response time in any particular well. Wells were required to be correlated with precipitation using a Pearson's correlation coefficient $R^2 >= 0.2$ with $R > 0$. This threshold is discussed in the following section, "Selection of and Sensitivity to Analysis Parameters". A selection of hydrographs paired with precipitation records among included wells is displayed in Figure 6 through Figure 12.

At least 25 years of annual high groundwater-level measurements were required for wells to be included in this analysis to ensure both the robustness of water-level trend over the period of record, and the correlation with precipitation. This minimum 25-year threshold also increased the likelihood of sampling multiple cycles of precipitation response within a particular record. A sensitivity analysis is shown below.

Detrending

In order to represent the best evaluation of long-term stability, water level records were “detrended” by subtracting the Theil-Sen best-fit slope. This detrending was performed after evaluation of the correlation with water levels (described above) but before evaluation of the characteristic magnitude and rate (described below). An additional sensitivity analysis evaluates the impact of whether precipitation is correlated with raw or detrended water levels below. Examples of the detrending are shown in (Figure 1 through Figure 5).

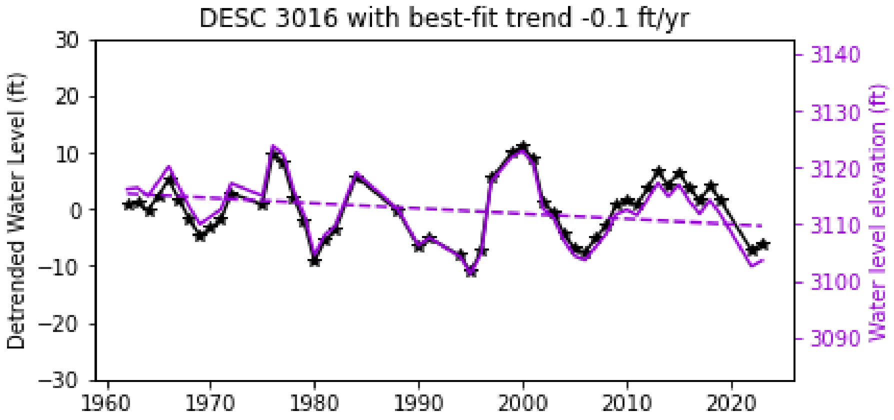

Figure 1: Example of detrending water levels in well DESC 3016. The right y-axis and purple colors indicate raw water level elevations. The right y-axis and black data show the detrended water levels after the best-fit line (dashed purple) was removed.

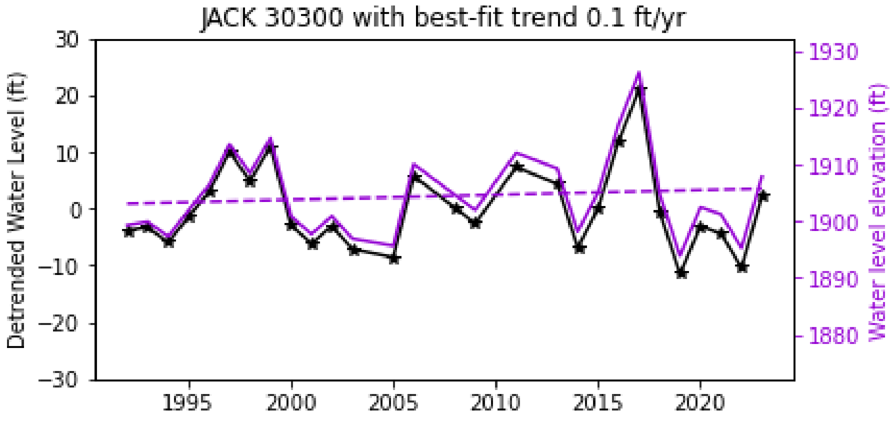

Figure 2: Example of detrending water levels in well JACK 30300. The right y-axis and purple colors indicate raw water level elevations. The right y-axis and black data show the detrended water levels after the best-fit line (dashed purple) was removed.

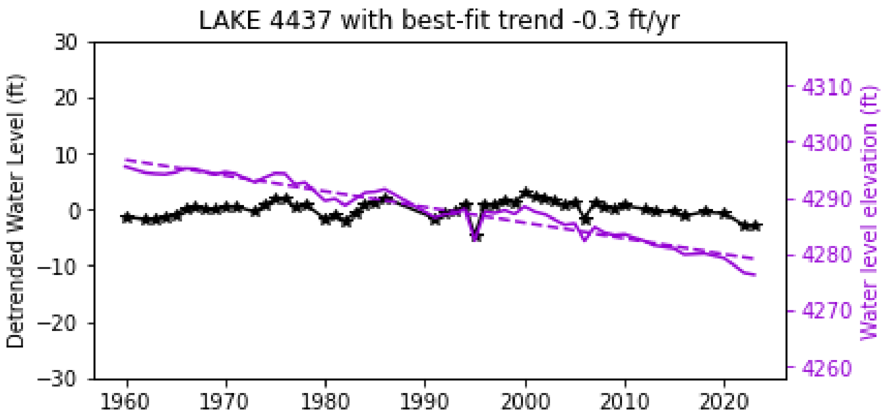

Figure 3: Example of detrending water levels in well LANE 4437. The right y-axis and purple colors indicate raw water level elevations. The right y-axis and black data show the detrended water levels after the best-fit line (dashed purple) was removed.

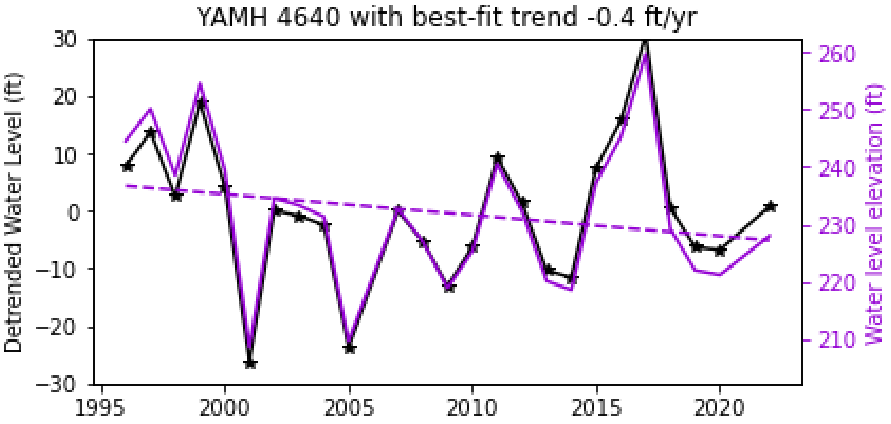

Figure 4: Example of detrending water levels in well YAMH 4640. The right y-axis and purple colors indicate raw water level elevations. The right y-axis and black data show the detrended water levels after the best-fit line (dashed purple) was removed.

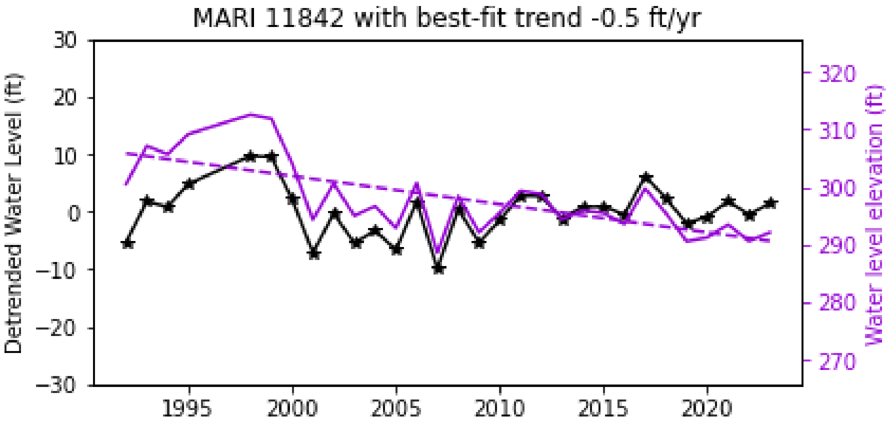

Figure 5: Example of detrending water levels in well MARI 11842. The right y-axis and purple colors indicate raw water level elevations. The right y-axis and black data show the detrended water levels after the best-fit line (dashed purple) was removed.

Detrended water level records has the effect of removing trends over a well's period of record. This may include consistent declines that are correlated with declining precipitation over the same period of record. To the extent that annual precipitation is not expected to change significantly in Oregon in climate projections (and setting aside debates about whether groundwater recharge may decline (Waibel *et al.*, 2013)), wells with long-term declining trends correlated with precipitation may be expected to recover. For those wells, removing the long-term trend would underestimate the magnitude and rate of declines that should be included in Reasonably Stable. On the other hand, correlation between long-term declines in water levels and precipitation does not imply causation, and some of the long-term declines may be attributable to increased groundwater pumping. The portion of a declining trend that is caused by increased pumping should not be expected to recover without intervention, so it would be reasonable to remove that component of the trend. Detailed studies can begin to attribute declines to natural and anthropogenic causes (Gannett and Lite Jr., 2013) but are beyond the scope of this statewide survey. Instead, this work includes below a sensitivity analysis of the results to whether the water level records are detrended or not.

Evaluation of Characteristic Magnitude and Rate of Decline in Each Well

The characteristic magnitude of decline for each well was calculated as the largest decline magnitude among years in the detrended record for that well. Each year's Annual High Water Level was compared against the shallowest preceding Annual High. Representing each well using the maximum decline over the period of record sometimes led to declines that were not broadly representative of the majority of that well's record. For example, well SHER 340 is correlated with precipitation, but the strong dip in the early 1990s is not representative of the bulk of the well's variability (Figure 9). The largest decline magnitude and rate of decline in this well are therefore larger than the typical values over the period of record.

Likewise, the characteristic rate of water-level decline within each well was calculated as the largest "test rate" among all years in the detrended record with preceding data spanning at least 13 years. For each year under evaluation, the "test rate" was the minimum rate of decline among trends using data over immediately preceding averaging periods from 5 to 20 years. While the proposed definition requires only 5 years of data, the requirement for 13 years in this analysis reflects the fact that interannual precipitation variability in Oregon is dominated by decadal fluctuations (Abatzoglou *et al.*, 2014), as well as the sensitivity to this initial span presented in the following section, "Minimum Initial Period of Data Collection for Rate Test Evaluation."

The results of the proposed rate test were compared against the results of the one-sided Mann-Kendall test for monotonic declines (in Figure 27 and Figure 28). While the proposed rate test is applied to data over spans from 5 through 20 years (then with the slowest rate of decline chosen as the "test rate"), the Mann-Kendall test was applied to all data leading up to a given year under evaluation. The Mann-Kendall test has been widely used for evaluating trends in environmental data (Helsel *et al.*, 2020), but it assumes that measurements are independent and not serially correlated (Yue *et al.*, 2002). Annual High Water Level measurements are typically strongly serially correlated, and this violation of a core assumption of the test biases the results toward detecting declines when they may not be present (Yue and Wang, 2002). Corrections are available but require expert judgment to be applied in each case (Yue and Wang, 2004). The proposed form of the rate test was developed because it is relatively easy to implement without advanced statistical software, and because the comparison against a maximum decline rate adds flexibility and facilitates transparent discussion about the limits of detection by the rate test. However, the comparison between it and the Mann-Kendall test provides perspective on the sensitivity and robustness of the proposed definition at different threshold rates. Significance levels of 0.05 (a standard value), 0.01, and 0.001 were tested. The significance level alpha 0.05 for a one-tailed test means that the test required less than a 5% probability that an observed trend could have been generated from a stationary distribution, in order for the test to classify the trend as declining.

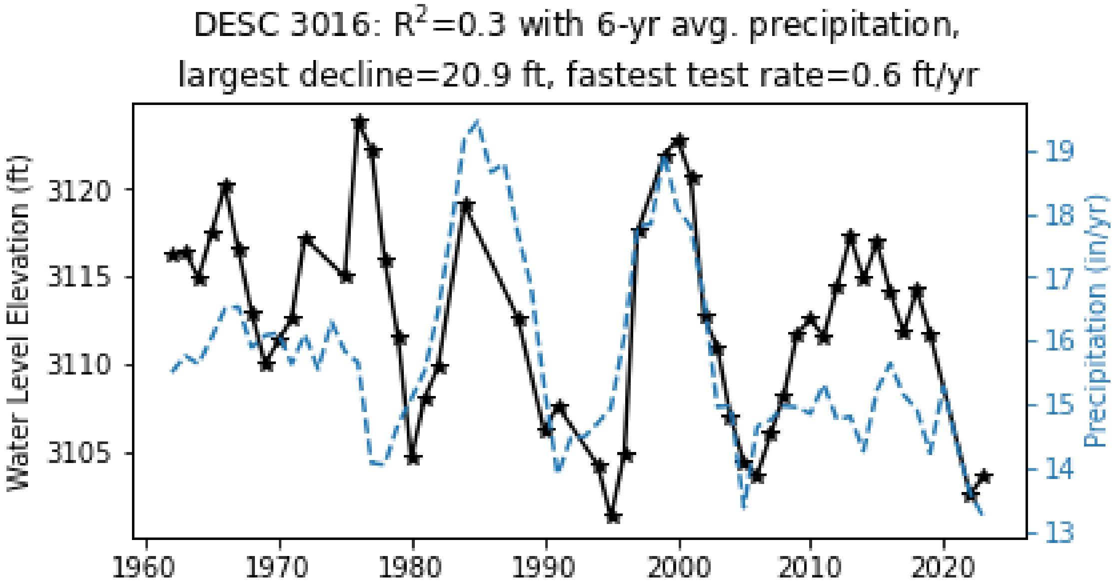

Figure 6: Hydrograph for well DESC 3016 with Annual High Water Levels (solid black line, left axis) correlated (R^2 = 0.30) with 6-year moving-average precipitation (dashed blue line, right y-axis). The largest magnitude of decline = 20.8 feet, and the fastest rate of decline calculated with the proposed definition ("test rate") is 0.6 feet per year in 1970. Among years with preceding data spanning at least 13 years, the steepest decline was 0.3 feet per year in 1995.

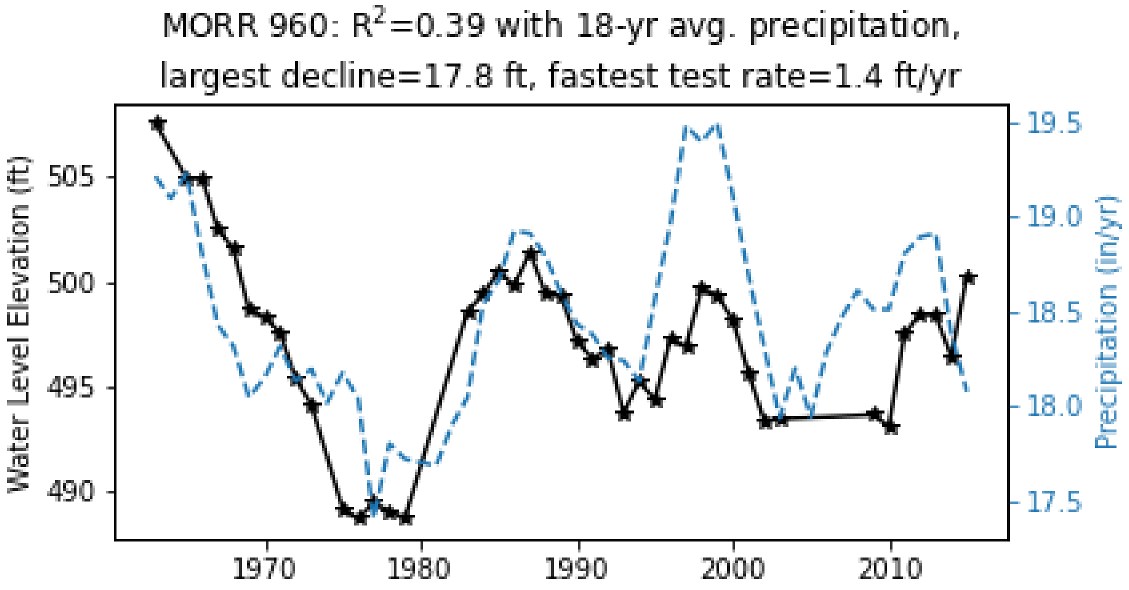

Figure 7: Hydrograph for well MORR 960 with Annual High Water Levels (solid black line, left axis) correlated (R^2 = 0.39) with 18-year moving-average precipitation (dashed blue line, right y-axis). The largest magnitude of decline = 17.8 feet, and the fastest rate of decline calculated with the proposed definition ("test rate") is 1.4 feet per year in 1976, near the bottom of the initial declining trend measured in this well.

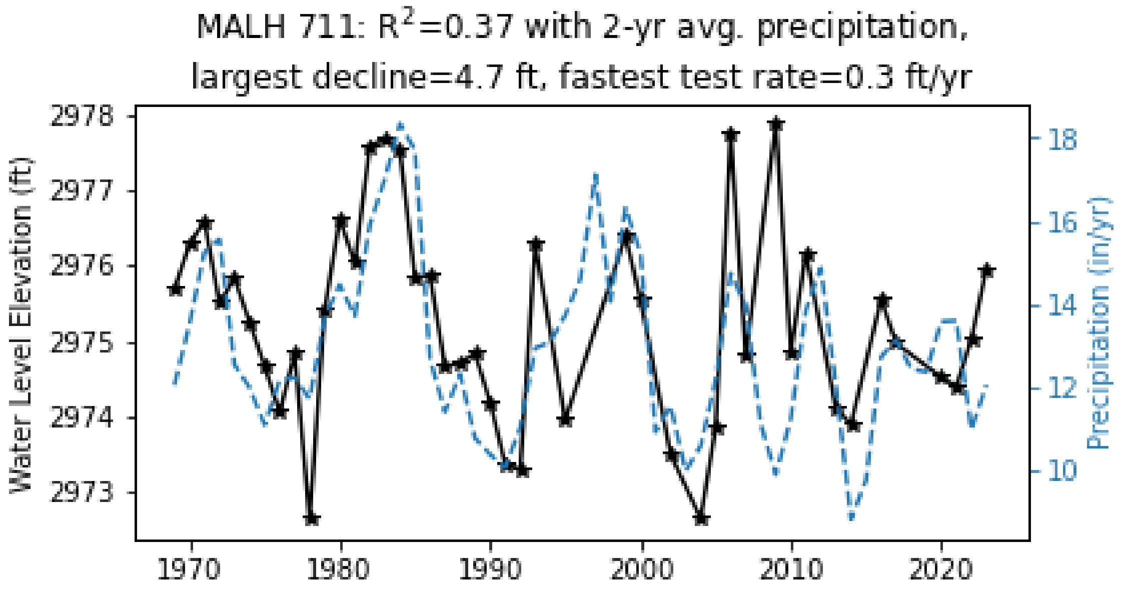

Figure 8: Hydrograph for well MALH 711 with Annual High Water Levels (solid black line, left axis) correlated (R^2 = 0.37) with 2-year moving-average precipitation (dashed blue line, right y-axis). The largest magnitude of decline = 4.7 feet, and the fastest test rate of decline among years was 0.3 feet per year.

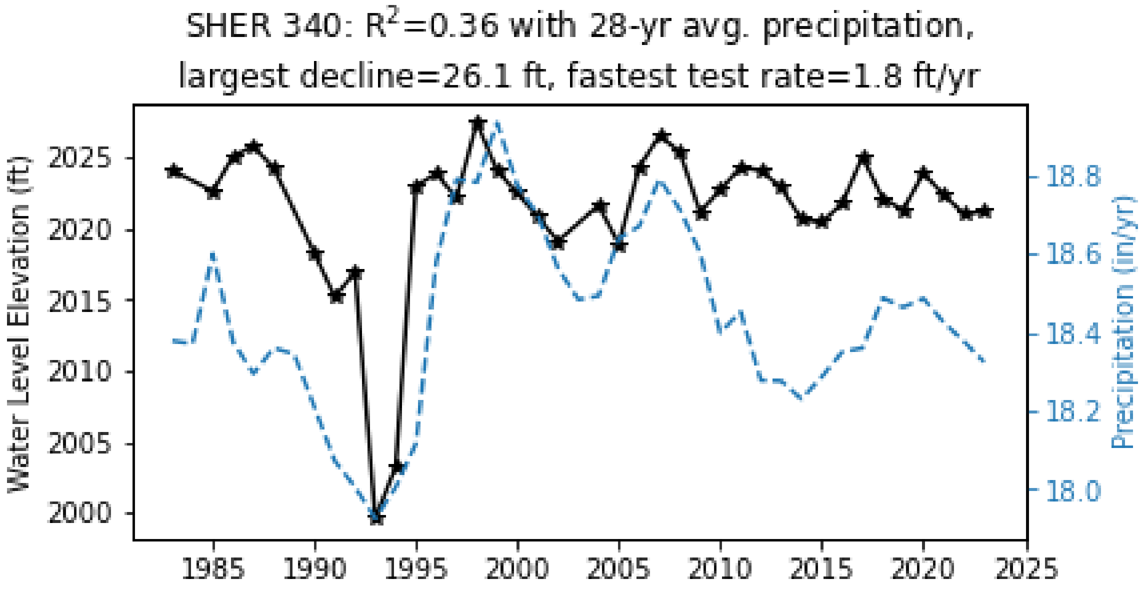

Figure 9: Hydrograph of well SHER 340 with Annual High Water Levels (solid black line, left axis) correlated (R^2 = 0.36) with 28-year moving-average precipitation (dashed blue line, right y-axis). The largest magnitude of decline = 26.1 feet, and the fastest rate of decline calculated with the proposed definition ("test rate") is 1.9 feet per year in 1994.

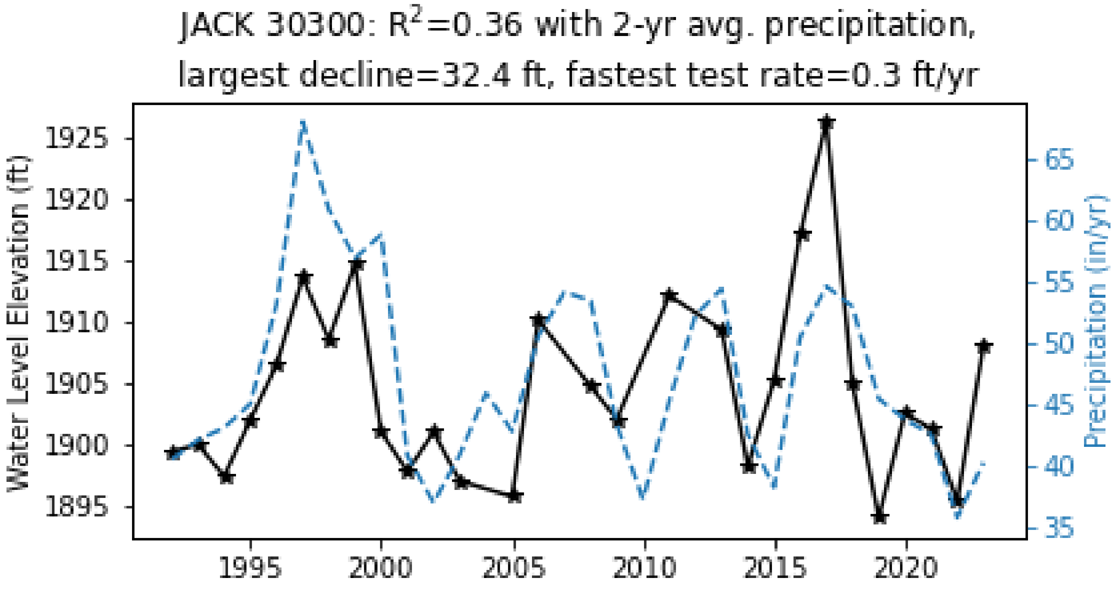

Figure 10: Hydrograph for well JACK 30300 with Annual High Water Levels (solid black line, left axis) correlated (R^2 = 0.36) with 2-year moving-average precipitation (dashed blue line, right y-axis). The largest magnitude of decline = 32.4 feet, and the fastest rate of decline calculated with the proposed definition ("test rate") is 0.3 feet per year in 2005.

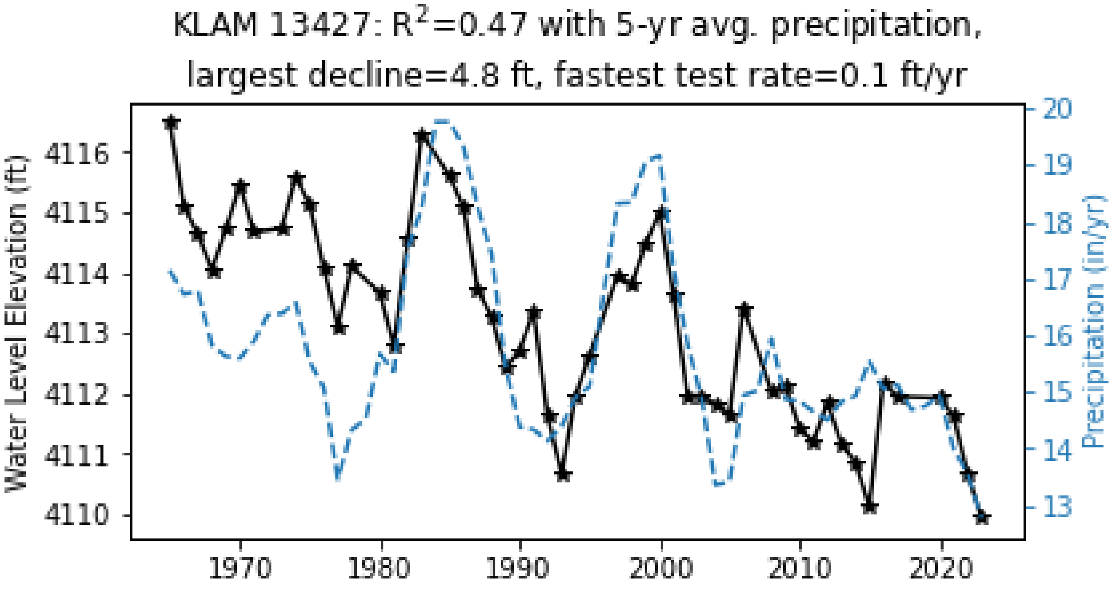

Figure 11: Hydrograph for well KLAM 13427 with Annual High Water Levels (solid black line, left axis) correlated (R^2 = 0.47) with 5-year moving-average precipitation (dashed blue line, right y-axis). The largest magnitude of decline = 4.8 feet, and the fastest rate of decline calculated with the proposed definition ("test rate") is 0.1 feet per year.

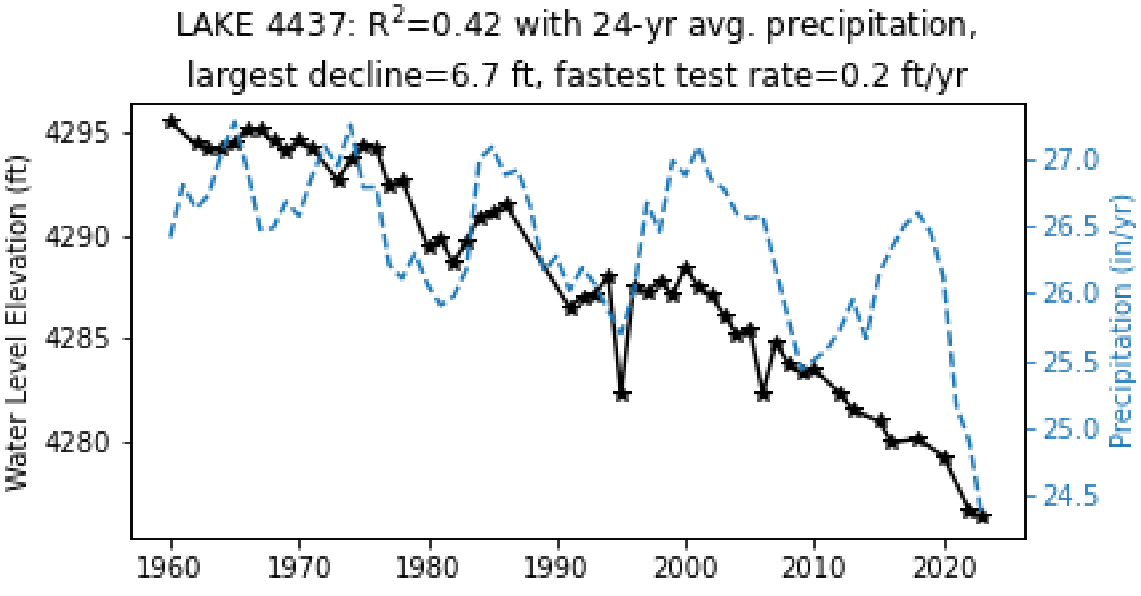

Figure 12: Hydrograph for well LAKE 4437 with Annual High Water Levels (solid black line, left axis) correlated (R^2 = 0.42) with 24-year moving-average precipitation (dashed blue line, right y-axis). The largest magnitude of decline = 6.7 feet, and the fastest rate of decline calculated with the proposed definition ("test rate") is 0.2 feet per year.

Clustering of Similar Wells

Applying the method outlined above to a collection of 9,397 wells with Annual High Water Levels in the Department's Groundwater Information System, 357 wells had the appropriate qualifications (Figure 13). Those wells spanned the state, but they also showed enhanced density in some regions of the state, especially in portions of the Willamette and Klamath basins. Higher density of wells included in this study may be attributed to a high density of observation wells that have been monitored for at least 25 years. This over-representation was in tension with the intent of this analysis: to characterize statewide variability with the distribution of values measured in individual wells.

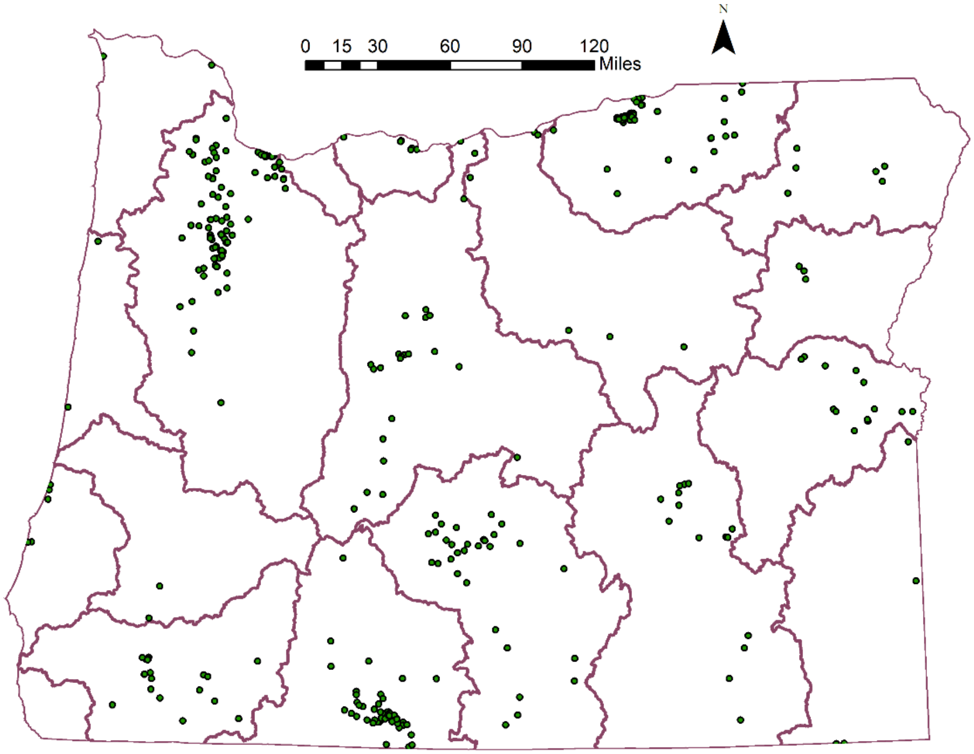

Figure 13: Map of 357 wells with appropriate qualifications for inclusion in the analysis of precipitation-correlated variability in water levels.

This spatial sampling bias was mitigated by clustering wells that were nearby and showed similar relevant properties. Wells were clustered using single-linkage agglomerative clustering (Kawa *et al.*, 2023). That means a well is added to an existing cluster if it meets all of the following criteria with respect to any other member of the cluster:

- Located within 10 miles, or within 25 miles if either well is indicated in the OWRD Groundwater Information System as a basalt well (due to their much larger hydraulic diffusivity (Freeze and Cherry, 1979)).
- The magnitudes of water level decline agree within 30% of the mean value between the pair. A percentage threshold was used rather than an absolute value because the distribution of well-specific maximum decline magnitudes is roughly exponential (Figure 24). Sensitivity analysis to the 30% minimum is presented below in the section, “Maximum Percent Difference in Characteristic Magnitude of Decline Between Similar wells.”
- The wells potentially access the same aquifer. Aquifer compatibility could be established through any of the following methods:
    - If both wells have been identified to access the same aquifer in the Groundwater Information System.
    - If either of the wells only had an aquifer *system* selected and not a specific aquifer, and the wells are in the same aquifer system.
    - The aquifer and aquifer system for either of the wells has not been selected. This means that wells with neither aquifer nor aquifer system selected could be spuriously clustered with wells in different aquifers, but only if they also met the other similarity criteria for clustering.

Single-linkage clustering was used instead of complete-linkage clustering because wells along a gradient of elevation within the same aquifer may respond to hydraulic forcing with a gradient of magnitudes. Other clustering methods may also be useful but were not investigated.

Applying the above method grouped the 357 wells into 160 distinct clusters (Figure 14). Of these clusters, 124 had only a single well, 36 had more than 1 well, 6 had at least 10 wells, 4 had at least 20, and the largest cluster had 41 wells. Values of water level range and characteristic rate of decline were averaged among wells within each cluster using the mean value.

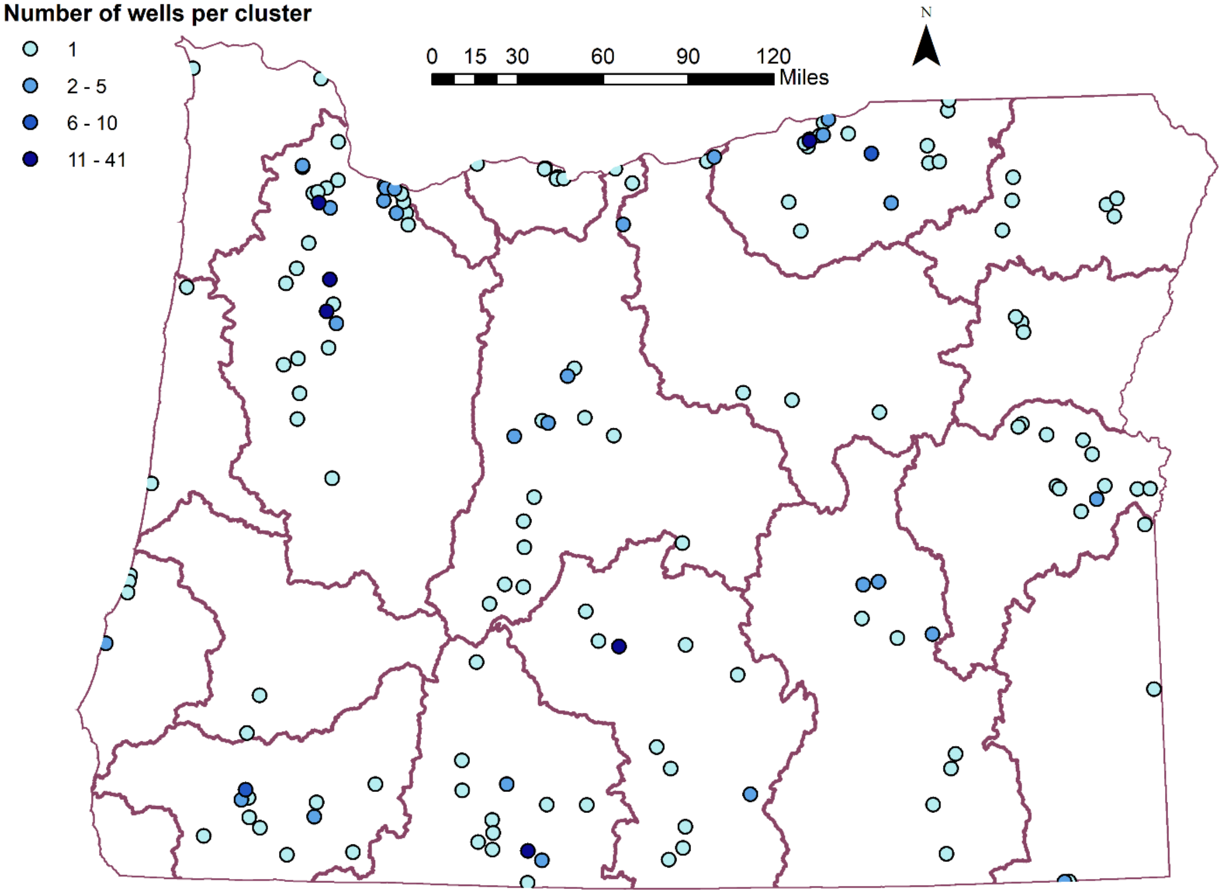

Figure 14: Map of 160 precipitation-correlated clusters, colored by number of wells per cluster.

Selection of and Sensitivity to Analysis Parameters

This analysis depends on empirical parameters to appropriately identify wells correlated with long-term precipitation cycles and with limited influence of pumping. The most sensitive of parameters include the minimum degree of correlation with regional precipitation, the minimum number of Annual High Water Levels, the maximum long-term rate of decline, the minimum percent difference in total decline between cluster members, and the initial span of data required to evaluate the rate test. The particular parameter values were selected manually using professional judgement on review of (1) hydrographs from wells determined to be precipitation-driven or not, and (2) sensitivity of the results and sample size to the parameter values. Below is presented a discussion of these selections, along with a sensitivity analysis of the results to changing each parameter's values (Figure 15 through Figure 23).

Minimum Coefficient of Determination (R^2)

A primary filter on inclusion of wells is the requirement that they be sufficiently correlated with the regional precipitation record from their respective climate divisions. The criterion for "sufficiently correlated" was set considering the coefficient of determination ($R^2 = 0.2$) for wells with known strong climatic influence, including DESC 3016 (Figure 6). Higher minimum values would have ensured more significant correlations, but the number of qualifying well clusters decreased roughly linearly to nearly 0 by $R^2 = 0.5$ (Figure 15, bottom). The sensitivity of the results to the minimum coefficient of determination are shown in Figure 15. The upper subfigure shows how the distribution of water level declines display a rough local minimum around $R^2 = 0.2$. The distribution is represented using the 50th, 67th, 80th, and 90th percentile values, which are offered as round portions of well clusters that the Department could consider as objectives. These percentiles correspond to the appearance of type I errors in ½, 1/3rd, 1/5th, and 1/10th of the well clusters, respectively. In other words, higher percentiles and lower incidence of type I errors lend higher confidence that findings of Not Reasonably Stable are likely to be correct. The 90th percentile decline decreases slowly from about 29 feet to 24 feet between $R^2 = 0$ and $R^2 = 0.35$. Above that, it decreases more rapidly to 15 feet at $R^2 = 0.5$.

The middle subfigure shows how the characteristic rates of decline retain stable distributions for minimum R^2 values between approximately 0.1 and 0.4. Above this, the distribution widens, but the 90th percentile fluctuates erratically between 0.5 and 0.7 feet per year. These fluctuations can be attributed in part to the rapidly shrinking set of well clusters with increasing coefficient of determination.

Minimum Number of Annual Water Level Measurements

The 90th percentile magnitudes of water level decline decreases from approximately 24 and 29 feet as the minimum number of Annual High Water Level measurements is increased from 10 to 40 (Figure 16). Over the same range, the 50th and 67th percentiles increase modestly, collapsing the distribution. Meanwhile, the distribution of rates of decline increase as the minimum number of measurements increase, with the 90th percentile increasing from 0.4 to 0.7 feet per year. The minimum of 25 that was used to generate the final results (Figure 24 through Figure 26), shown in the dashed vertical line, was selected for balance of relatively large sample size (with a lower minimum) and more robust indication that the best-fit linear trend over the period of record indicates stable behavior (with a higher minimum).

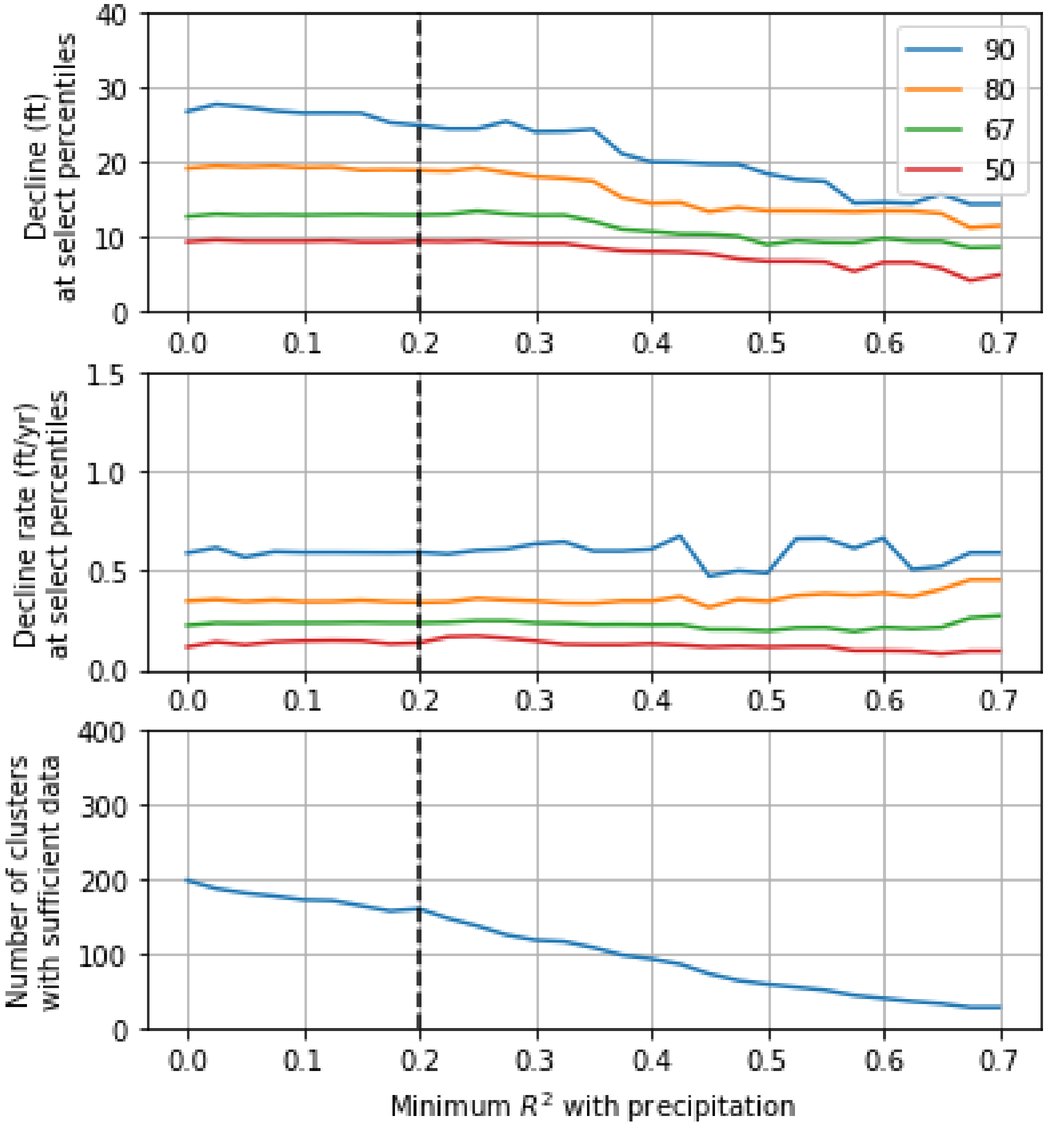

Figure 15: Sensitivity of results to minimum Pearson's correlation coefficient R2. Top: water level range at 50th, 67th, 80th, and 90th percentiles (line colors indicated in legend) among clusters of qualifying wells. Middle: characteristic of decline at 50th, 67th, 80th, and 90th percentiles. Bottom: number of clusters with sufficient data given the threshold displayed in the x-axis. The vertical dashed line in all subfigures represents the final threshold of 0.2.

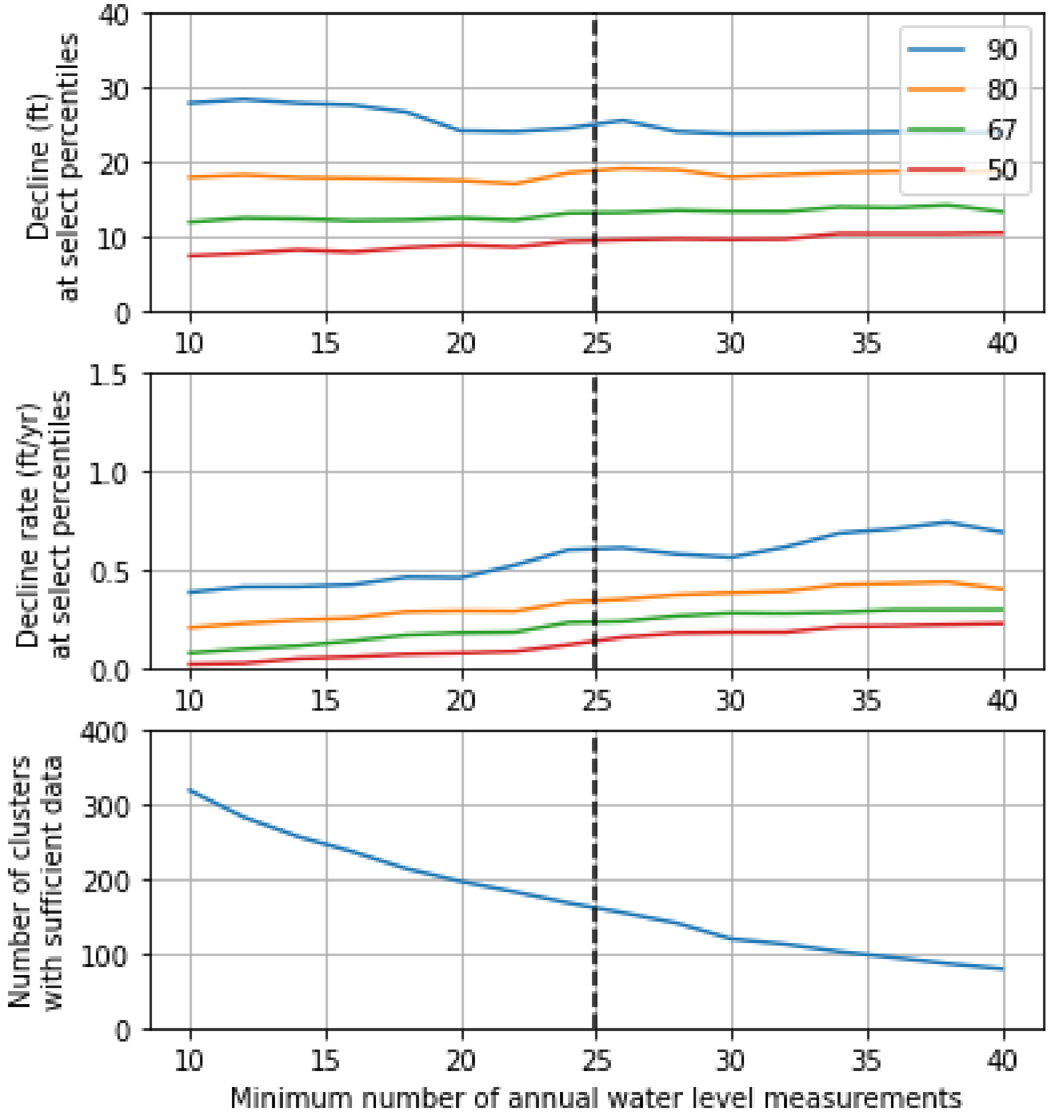

Figure 16: Sensitivity of results to minimum number of Annual High Water Levels. Top: water level range 50th, 67th, 80th, and 90th percentiles among clusters of qualifying wells. Middle: characteristic rate of decline at 50th, 67th, 80th, and 90th percentiles. Bottom: number of clusters with sufficient data given the threshold displayed in the x-axis. The vertical dashed line in all subfigures represents the final threshold of 25 measurements.

Maximum Best-Fit Rate of Change over the Period of Record

The 90th percentile magnitude of water level decline varies between approximately 8 and 30 feet in response to maximum rates of change (decline or rise) over the period of record between 0 and 1 feet per year (Figure 17, top subfigure). The distribution is relatively insensitive for limits between 0.3 and 1 feet per year. The test rates of decline (middle subfigure) also show a distribution that rises as the maximum allowed rates of change rises from 0 to 0.4 feet per year, above which the distributions become generally stable. As the maximum allowed rate of change over the period of record is reduced, fewer wells meet the limitation.

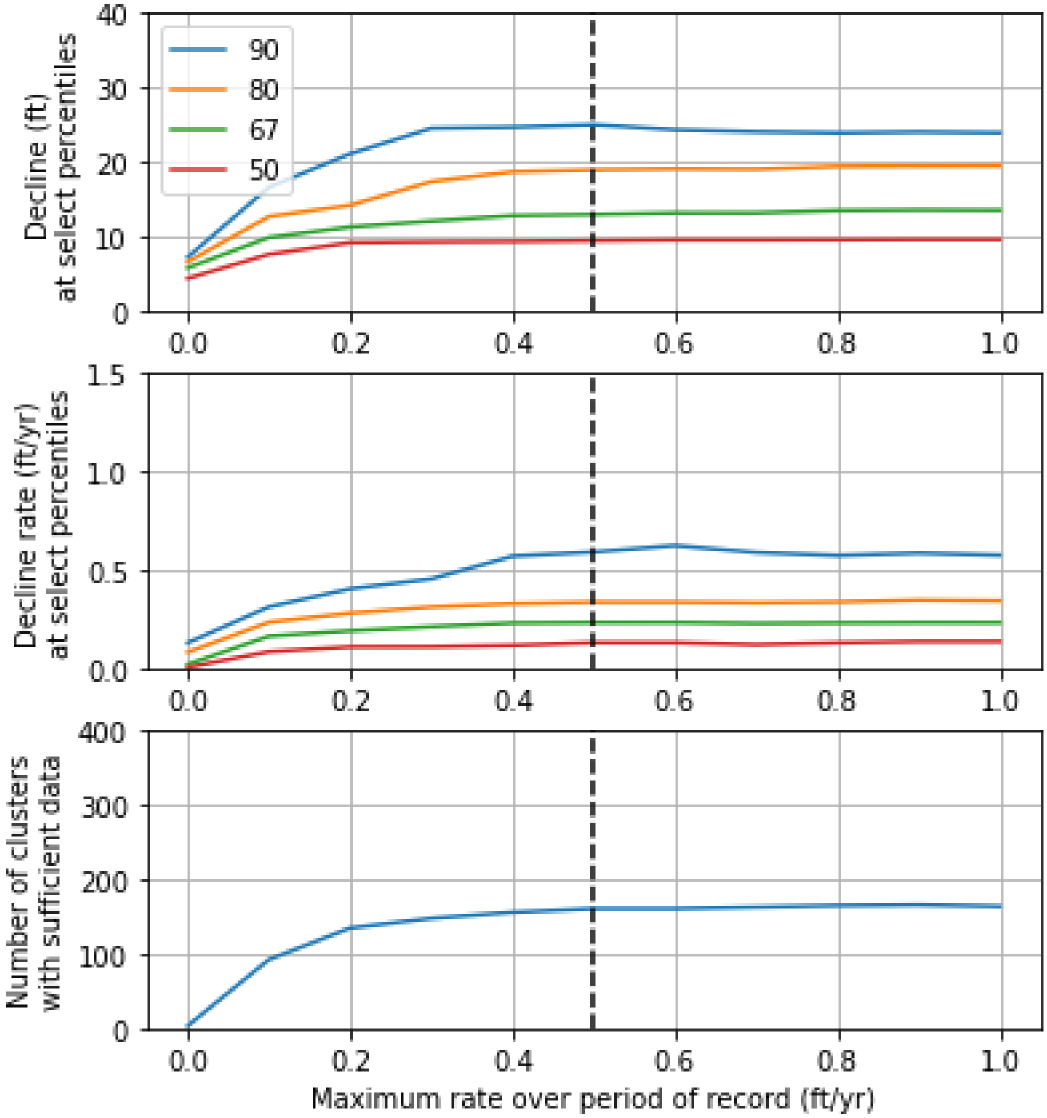

Figure 17: Sensitivity of results to maximum rate of decline over the period of record (feet per year, x-axis). Top: water level range at 50th, 67th, 80th, and 90th percentiles among clusters of qualifying wells. Middle: characteristic rate of decline at 50th, 67th, 80th, and 90th percentiles. Bottom: number of clusters with sufficient data given the threshold displayed in the x-axis. The vertical dashed line in all subfigures represents the final threshold of 0.5 feet per year.

Correlating Precipitation to Raw or Detrended Water Levels

Water levels may be correlated with precipitation either before or after the water level record is detrended. The impact of correlating either raw or detrended water level records with precipitation on decline magnitudes is shown in Figure 18, and the impact on rates of decline is shown in Figure 19. The impact appears fairly limited in both cases; correlating to raw water levels decreased the 90th percentile magnitude from 27 feet to 25 feet but increased the rate from 0.54 feet per year to 0.59 feet per year.

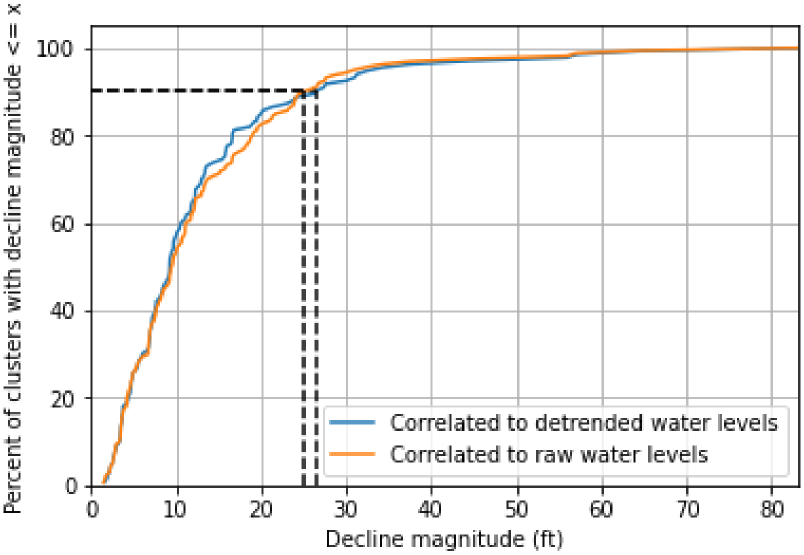

Figure 18: Sensitivity of maximum total decline to ordering of correlation and detrending. Correlating to raw water levels (orange line) reduced the 90th percentile total decline 27 to 25 feet.

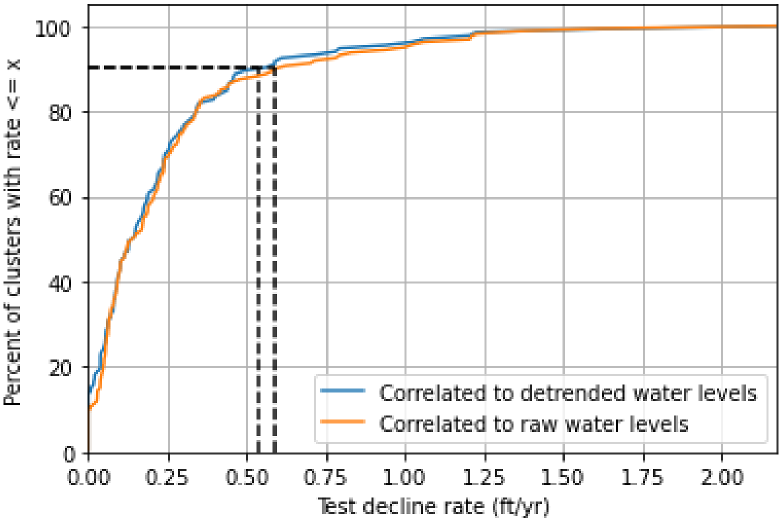

Figure 19: Sensitivity of maximum total decline to ordering of correlation and detrending. Correlating to raw water levels (orange line) increased the 90th percentile rate of decline from 0.54 to 0.59 feet per year.

Detrending

The distribution of maximum total declines per well cluster was calculated both with and without detrending, as shown in Figure 20. It shows that detrending reduces the 90th percentile maximum magnitude by 5 feet from 29 feet to 24 feet, likely reflecting the fact that most wells in the sample show long-term declines over their period of record. The distribution of maximum rates of decline per well cluster is shown in Figure 21. Detrending reduces the maximum rates from 0.75 feet per year to 0.64 feet per year.

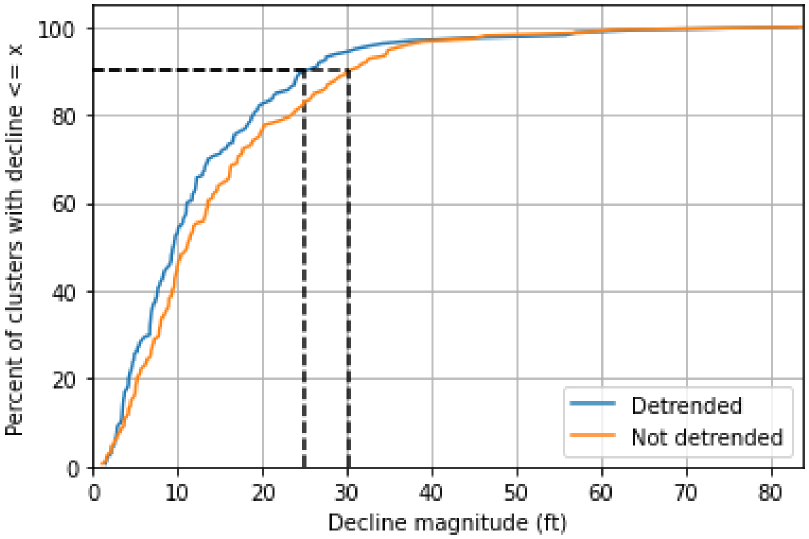

Figure 20: Sensitivity of distribution of maximum total declines per cluster to detrending. Detrending reduces the maximum total decline from 30 to 25 feet.

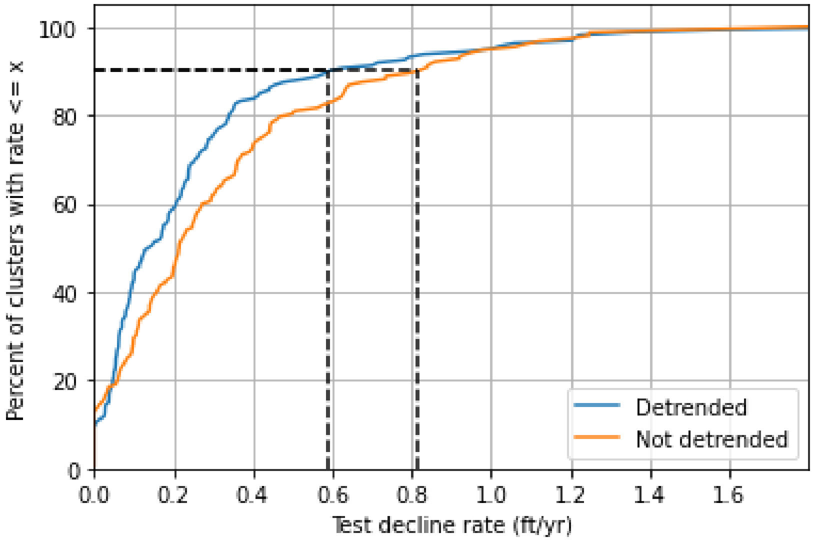

Figure 21: Sensitivity of distribution of decline rates per cluster to detrending. Detrending reduces the 90th percentile rate from 0.81 to 0.59 feet per year.

Maximum Percent Difference in Characteristic Magnitude of Decline Between Similar wells

Decline magnitudes are roughly exponentially distributed (Figure 24), so any sampling from that distribution should be by percent instead of by absolute value in order to avoid distorting it. Thus, when clustering wells to limit spatial sampling bias, similarity was determined using a percentage agreement rather than an absolute value (feet). The results were relatively insensitive to the maximum percent difference, except that the number of clusters decreased strongly with larger values (Figure 22).

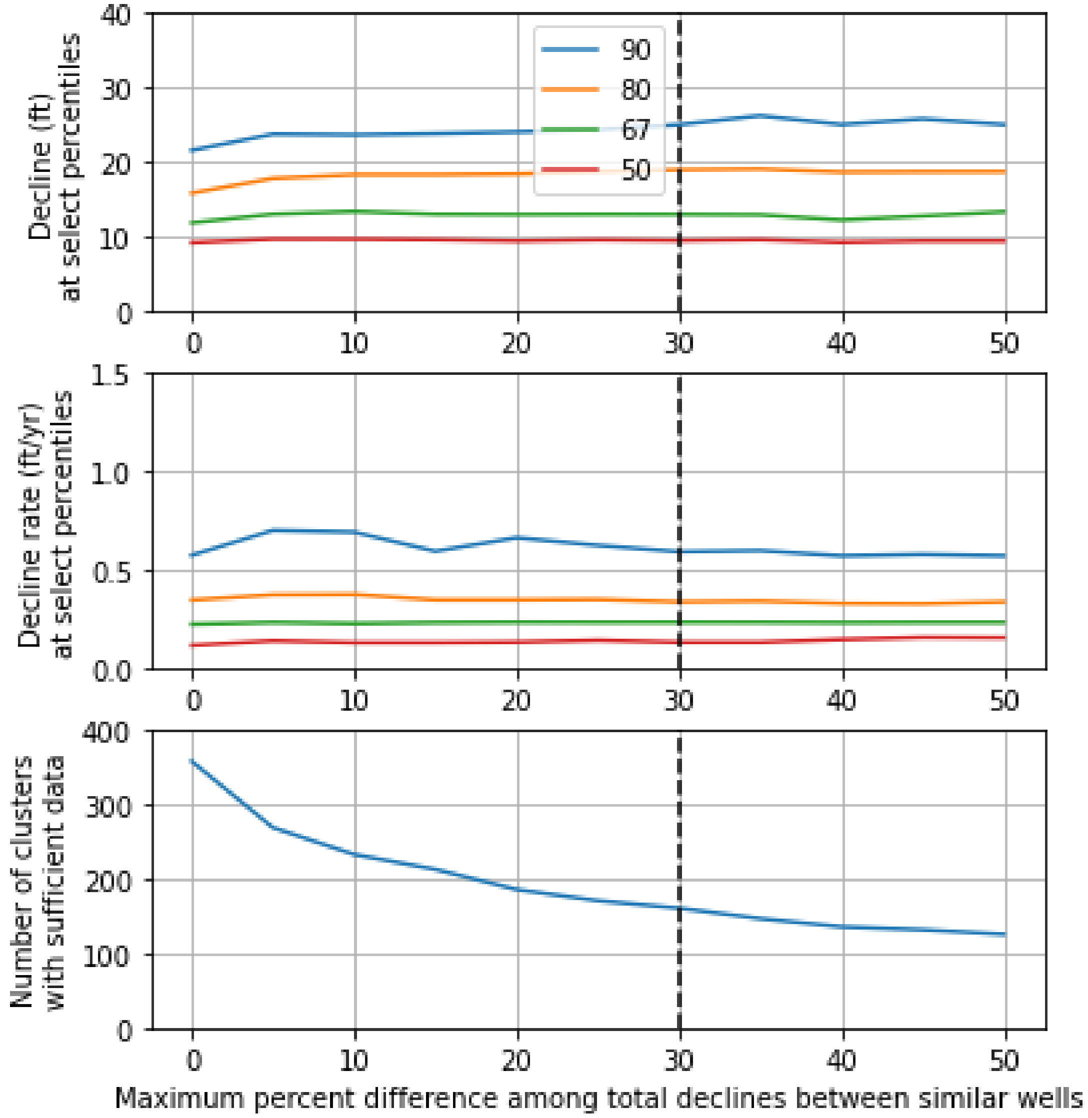

Figure 22: Sensitivity to the maximum percent difference between total declines allowed for wells characterized as similar (eligible to join the same cluster). This parameter had a large influence on the resulting number of clusters but not other results. The vertical dashed line in all subfigures represents the final threshold of 30%.

Minimum Initial Period of Data Collection for Rate Test Evaluation

The proposed definition of the rate test is intended to allow wells to remain Reasonably Stable across roughly decadal periods of the primary interannual variability in precipitation observed in Oregon (Gannett *et al.*, 2001; Abatzoglou *et al.*, 2014). The rate test selects the slowest rate of decline among those measured over any averaging window between 5 and 20 years leading up to the year being evaluated. This definition is therefore most able to remain Reasonably Stable when a full 20 years of data are available for comparison with a given measurement. However, in an area without sufficient data, applicants for water rights may wish to receive a determination of stability with fewer than 20 years of data collection. Breaks in data collection also mean that data are rarely available in all consecutive preceding years. The proposed rule accommodates these interests by requiring a span of 5 years before the rate test to be evaluated. However, in the context of this statewide evaluation of rates of decline in wells correlated with precipitation, it is not obvious that 5 years is the minimum span of data that should exist before the rate test is evaluated.

Instead, the analysis presented here requires that preceding data collection span a period of 13 years before the rate test may be applied. Following are rationale for this choice and a sensitivity analysis of the results. The intent of the threshold of the rate test is to balance sensitivity to the onset of rapid declines against robustness to the roughly decadal precipitation cycles in Oregon. Each well in this analysis is a sample of the statewide variability, and the test should be able to be applied to a full cycle's worth of data. Because the cycles typically last a decade, the test requires data over a similar period operate as intended in Oregon. With fewer than a decade of data, wells may coincidentally begin being measured during a declining limb of a cycle, and after only 5 years, the slowest average rate may be evaluated only over declining data without the benefit of any preceding deeper measurements. This situation is exemplified in well MALH 711 (Figure 8), where measurement began in the late-1960s, near the peak of cycles in precipitation and water levels. The rate according to the proposed rule can be evaluated beginning after data span 5 years, and the fastest rate of decline according to the proposed definition was 0.3 feet per year, measured in 1978. Requiring preceding data to span 13 years instead of 5 reduces the steepest decline rate to 0.04 feet per year.

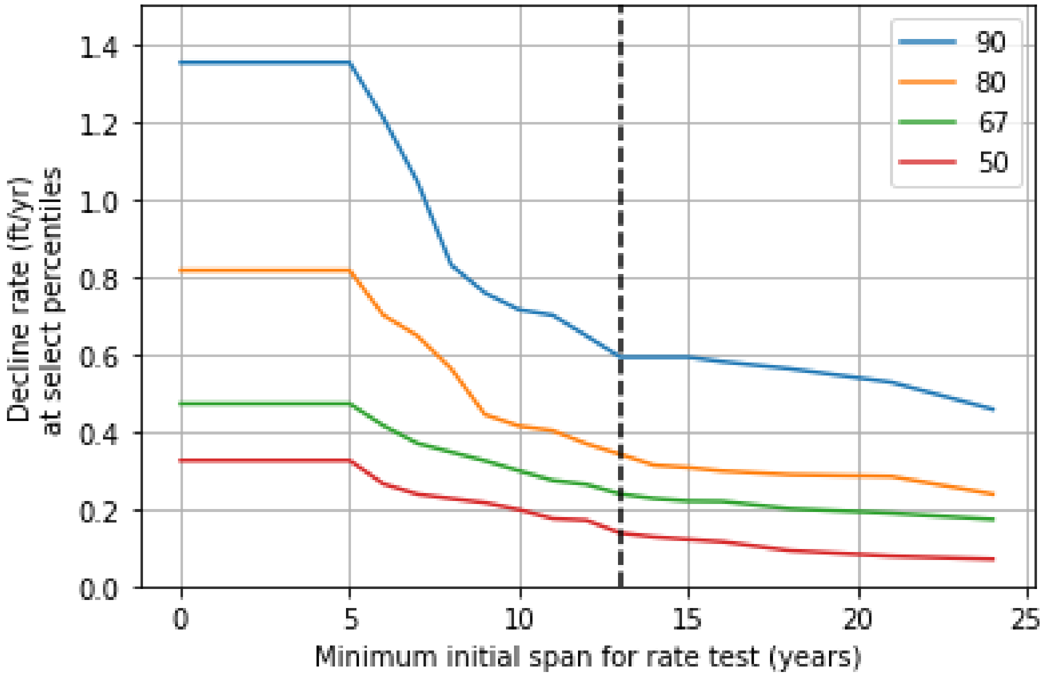

Figure 23: Sensitivity of rate of decline to minimum initial span of data before the rate test is applied.

One could go further to argue that this statewide analysis should be carried out only when preceding data span a full 20 years, in order to fully leverage the smoothing enabled by the form of the rate test. Instead, a sensitivity analysis was performed to identify the minimum period above which results stabilize (Figure 23). Results suggest that 13 years is a suitable compromise between the 5 years required in rule and the 20-year maximum allowed. The sensitivity of the distribution of measured rates according to the rate test (Figure 23) suggests the distribution contracts rapidly as the minimum span is increased from 5 to 13 years. Above that, the distribution continues to contract but at a much slower rate, indicating that initial declines have the largest impact on the resulting decline rate metrics over the first 13 years. Consequently, a minimum initial span of 13 years was adopted for the remainder of this analysis.

Results and Discussion

Magnitudes of Decline

The statewide distribution of maximum total water level declines in well clusters varies from 2 feet to 41 feet (Figure 24). The distribution is roughly concave between the 50th and 100th percentiles, such that the portion of wells represented by an additional increment of range decreases with increasing range. Among this upper half of observed ranges, select percentiles are indicated as dashed lines in Figure 24: the 50th percentile (10 feet), 67th (12 feet), 80th (17 feet) and 90th (24 feet).

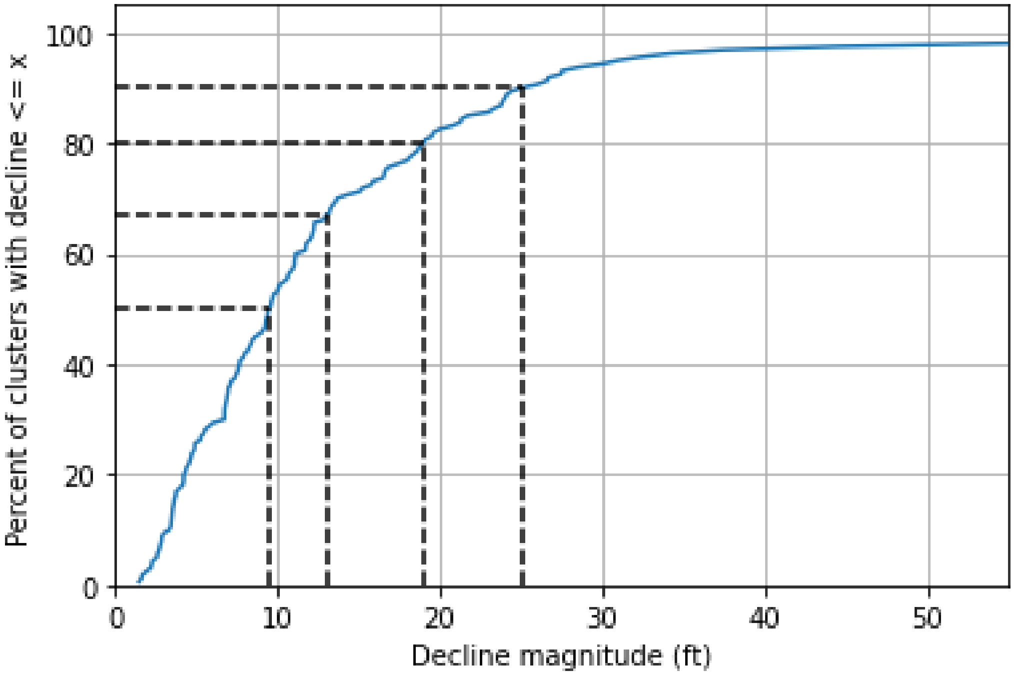

Figure 24: Cumulative distribution of well-specific maximum water level declines in detrended water level records correlated with regional precipitation. The total number of wells analyzed is 357 wells, which were grouped into 160 distinct clusters. Dashed lines indicate the rates corresponding to the 50th, 67th, 80th, and 90th percentiles, which are approximately 10, 12, 17, and 24 feet, respectively.

The well clusters included in Figure 24 are distributed across the state according to the map shown in Figure 25. This map shows that well clusters exceed a threshold of 25 feet in many basins, and there is no discernable spatial trend. To address the variability that does exist between basins, the proposed definition includes an opportunity to modify the definition of Reasonably Stable Groundwater Levels in basin program rules.

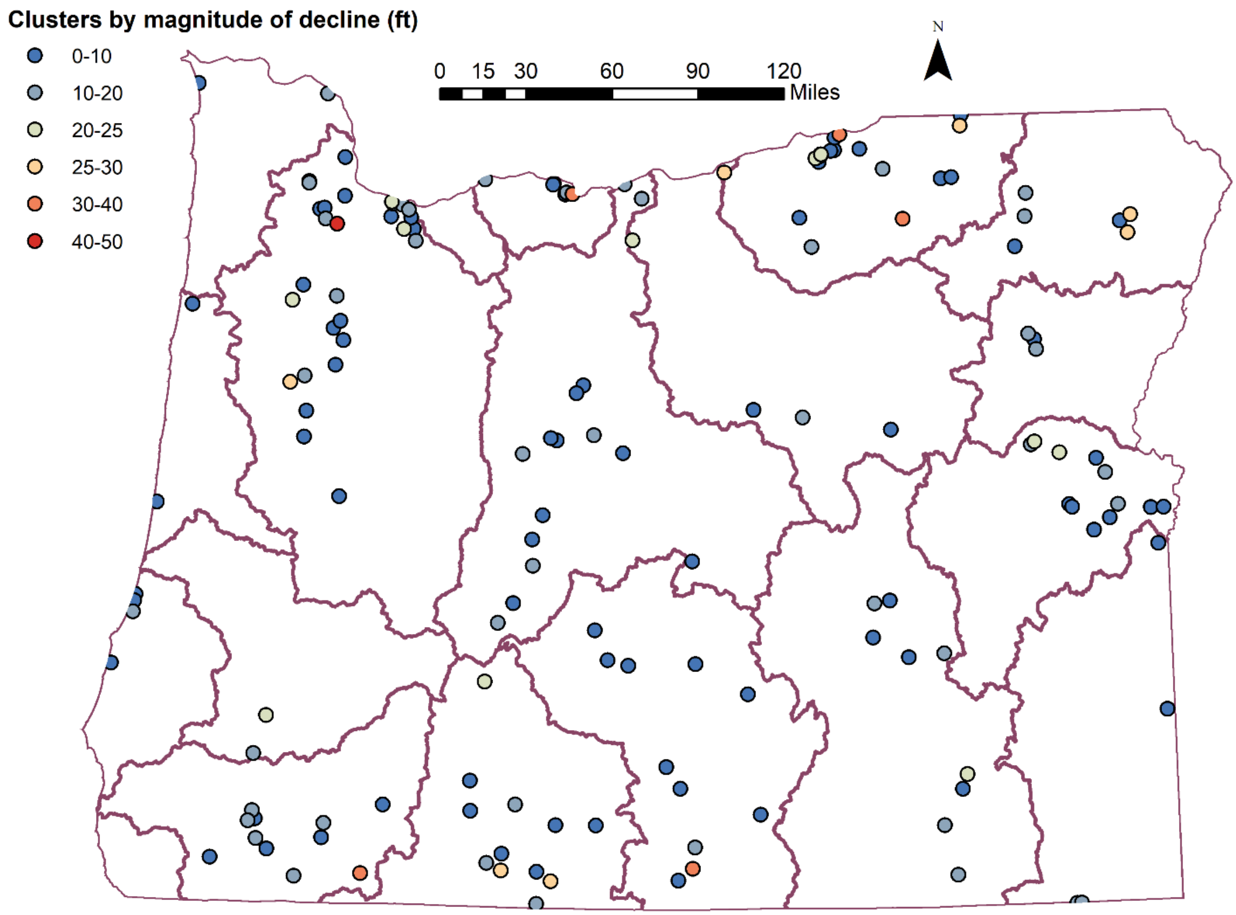

Figure 25: Map of 160 precipitation-correlated clusters colored by magnitude of water level decline (feet). The color scale is warm (yellow to red) for ranges greater than 25 feet and cool (green to blue) for ranges less than 25 feet.

The results shown in Figure 24 are relevant because they indicate the probability that a well influenced primarily by precipitation (and with limited long-term declines) would always remain Reasonably Stable, avoiding any oscillations in reasonable stability. Such oscillations are undesirable for the Department and for water rights applicants, but some instances of oscillation are an inevitable feature of any rule-based definition of Reasonably Stable Groundwater Levels in the context of both climatic and anthropogenic influences. Given that inevitability, it is also worth evaluating the proposed definition from another relevant perspective: the portion of time that a precipitation-correlated well would remain reasonably stable. This portion of time is typically higher than the portion of precipitation-correlated wells that are *always* Reasonably Stable, because even wells that decline to deeper than the allowable maximum total decline may rebound back into the realm of Reasonable Stability.

A sensitivity analysis of the portion of time is presented in Figure 26. That figure shows how the percent of time that the total decline test would be passed (characteristic decline magnitude is less than the allowed decline) depends on the maximum allowed decline (x-axis). The upper, blue line shows the mean percentage of time among all well clusters, including those that pass the total decline test over their entire period of record. This result is analogous to the one presented in Figure 24, and the percentage of time (Figure 26) is larger than the percent of wells always within the maximum (Figure 24) for all thresholds. With a total decline of less than 25 feet, the clusters pass the total decline test an average of 99% of the time. Even among the wells that do exceed the maximum allowed decline at some point in their period of record (orange line in Figure 26), the total decline is less than 25 feet over 86% of the time.

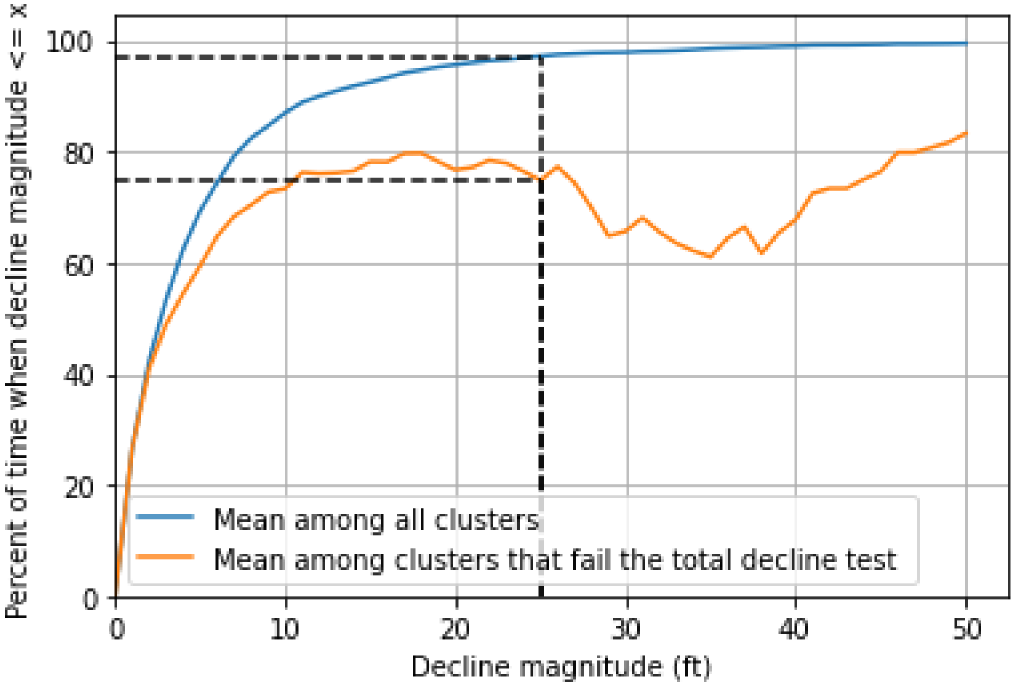

Figure 26: Sensitivity of the percent of time that the total decline test would be passed (decline from highest known less than the maximum allowed) as a function of the maximum characteristic total decline (x-axis) for a cluster. The blue line shows the mean percentage of time within the limit among all well clusters. The orange line shows the percent of time within the limit among well clusters that exceed the limit at some point. The black dashed lines indicate the percentages of time associated with a threshold of 25 feet: 99% over all well clusters, and 75% over those that fail the decline test at some point (have a maximum decline over 25 feet).

Distribution of Rates of Decline

The rate test in the proposed definition attempts to detect consistent declines while retaining Reasonable Stability across cycles up to 20 years in duration. To do so, it evaluates a "test decline rate" as the slowest rate of decline among averaging windows of 5 to 20 years leading up to the year being evaluated. It then compares that test rate against a threshold defining an unreasonable rate of decline. The following section characterizes the distribution of rates of decline observed in the same detrended, precipitation-correlated water level records used in the preceding analysis. For each well, the "test rate" is evaluated in each year with sufficient data. The well's characteristic test rate is then taken as the fastest rate of decline among all valid years. As noted above, valid years are restricted in this analysis to those with a data point measured at least 13 years before. The distribution of well-specific test rates is then collected statewide in order to understand the percentage of precipitation-correlated wells that would remain reasonably stable under a given maximum allowed (reasonable) rate of decline (Figure 27).

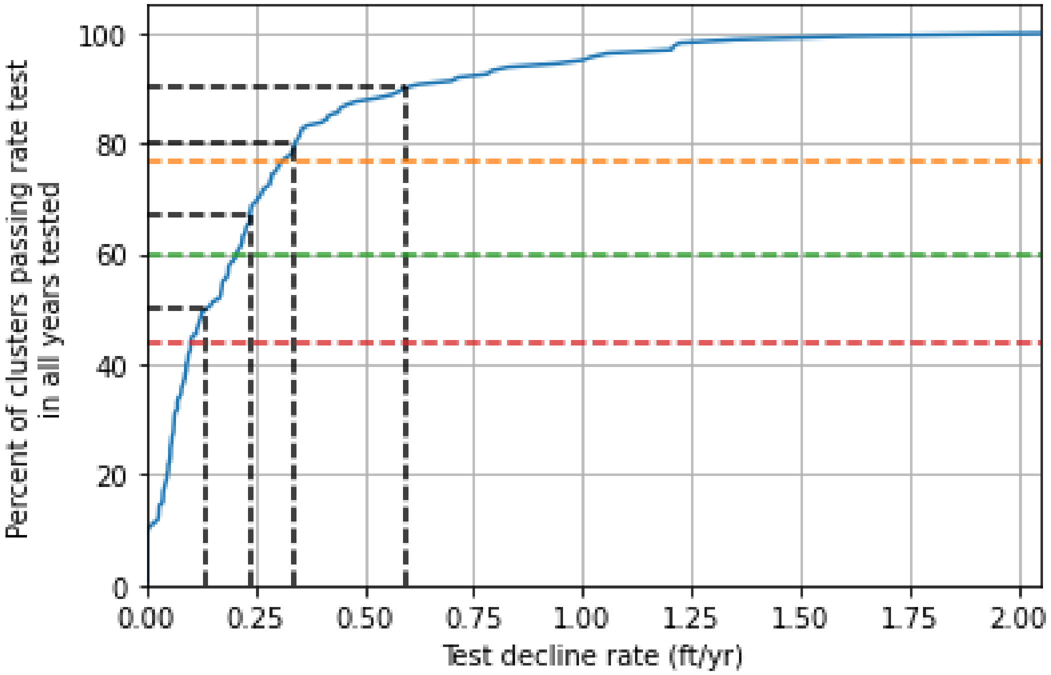

Figure 27: Characteristic rates of decline in detrended water level records among wells statewide. Blue line: cumulative distribution of characteristic rates according to the proposed definition. The characteristic rate for each well was calculated as the maximum rate among all years that could be tested using the proposed definition with a minimum period of 13 years since the first measurement on that well. Black dotted lines indicate the rates corresponding to the 50th, 67th, 80th, and 90th percentiles, which are approximately 0.1, 0.2, 0.3, and 0.6 feet per year, respectively. Orange, green, and red dashed lines: percent of precipitation-correlated clusters that remained "not declining" over the same portion of each well's record, according to the Mann-Kendall test with the values of one-tailed significance levels alpha. Percent of clusters for alpha = 0.05, 0.01, and 0.001 are 43, 60, and 77 percent, respectively.

The percentage of clusters that pass the rate test (i.e. has a test decline rate below the maximum allowed rate) in each year over the entire tested period of record increases with larger rates (Figure 27).

The median rate (among clusters) that encompasses the tested rates in each cluster is 0.14 feet per year, and the 90th percentile is 0.59 feet per year. Because of the restriction requiring a span of at least 13 years for evaluating the rate test, a larger percent of wells would exceed these thresholds during the first 5 years that the rate was eligible to be tested according to the proposed rule. With only 5 years of data collection, the 90th percentile among clusters increases to over 1.3 feet per year (Figure 23). However, these failures of the rate test would be expected to resolve with a additional data collection in order to sample over a full precipitation cycle.

In addition, the 90th percentile rate of 0.6 feet per year is more tolerant of the decline in this analysis than the more common Mann-Kendall test. That test found only 44% of clusters passing the rate test at the standard significance level of alpha = 0.05. Even decreasing alpha to an extremely low value of 0.001 (making the test much less sensitive to declines) only increases the percent of clusters that remain “not declining” over the tested period of record to 77% of clusters. Thus, the proposed rate test offers transparency, simplicity, and robustness to cyclical declines compared with the Mann-Kendall test.

The Department is interested not only in the percent of wells that remain Reasonably Stable over a tested period of record, but also the percent of time that wells tend to remain reasonably stable. The former question is stricter and addresses the potential for any oscillations in the finding of Reasonably Stable, while the latter is practically important for establishing the overall consistency of a test with an expectation of finding water level fluctuations Reasonably Stable among the set evaluated here. The percent of time that water levels pass the proposed rate test as a function of the rate threshold is shown in Figure 28. In contrast with Figure 27, Figure 28 was generated without the 13-year minimum data span restriction, in order to more fully evaluate the impact of the rate limit on the period of record. The percentage of time increases from 79% to 100% as the rate maximum increases from 0 to 2 feet per year. At the rate of 0.6 feet per year, clusters pass the test over 97% of the time. This is larger than the percent passage (86%) by the Mann-Kendall test at standard alpha = 0.05. Among wells that fail the test at some point, the percent of time passing is smaller than the average including all clusters but still high. At the rate of 0.6 feet per year, these clusters still pass the test 90% of the time. By comparison, the Mann-Kendall test passes only 78% of the time. Thus, both in terms of the percent of well clusters and the percent of time over period of record, the proposed rate test with a thresholds 0.6 feet per year is more robust to the fluctuations evaluated in this study. In other words, the proposed definition with a threshold of 0.6 feet per year would offer more flexibility to accommodate additional groundwater development than a standard rate test while being substantially more protective than existing rules.

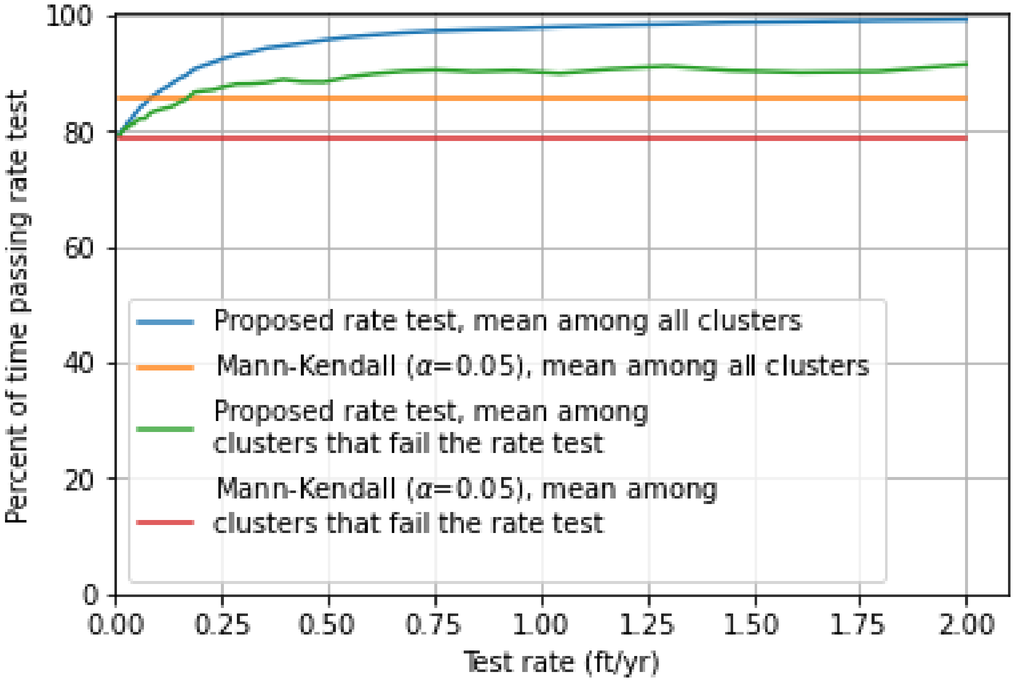

Figure 28: Percent of time that water levels pass the proposed rate test (blue and green) and are not declining according to the Mann-Kendall test on the entire preceding record (orange and yellow). Both tests are applied to detrended data. For the proposed rate test, the percent of time depends on the maximum allowed rate of decline (x-axis), while the nonparametric Mann-Kendall test does not. The blue and orange lines show the mean percent of time among all wells, which addresses the overall robustness of these tests to the decline rates observed in the sample wells. Meanwhile, the green and red lines show the percent of time among wells that fail their respective rate tests during at least part of their records. These lines focus on wells that would fail their respective tests for stability at some point and indicate the portion of time that they nonetheless indicate a lack of declines. For generating this figure, the 13-year minimum span of data was dropped so that the 5-year minimum written into the proposed rule was the active restriction.

Text of Proposed Definitions

The following two definitions are proposed to be added to OAR 690-008:

(1) “**Annual High Water Level**” in a groundwater reservoir or part thereof means the highest elevation (shallowest depth) static groundwater level that exists in a year.

(9) “**Reasonably Stable Groundwater Levels**” means:

(a) The Annual High Water Levels as measured at one or more representative wells in a groundwater reservoir or part thereof:

(A) indicate no decline or an average rate of decline of less than 0.6 feet per year over any immediately preceding averaging period with duration between 5 and 20 years. Four Annual High Water Levels are required to calculate the rate of change; one must have been measured in the year to which the evaluation of reasonably stable applies, and at least one must have been measured between 5 and 20 years prior; and

(B) have not declined by more than 25 feet from a reference level to the level in the year to which the evaluation of reasonably stable applies. The reference level shall be the highest known water level unless Annual High Water Levels have been increased measurably by human activity, in which case the Department may set a different reference level using best available information.

(b) If water level data are insufficient to perform either test in (a) for a given year, then the Department will presume that groundwater levels are not reasonably stable unless:

(A) the most recent evaluation of reasonably stable applies to a year within 5 years of the given year, in which case the Department may presume that the recent evaluation still applies; or

(B) groundwater has not yet been extracted or authorized for extraction from the groundwater reservoir, in which case the Department may presume that groundwater levels are reasonably stable.

(c) The Department may evaluate Reasonably Stable Groundwater Levels for the year of the priority date of a groundwater right application or for a later year if more recent data are available.

(d) The limits in part (a) of this definition may be superseded by limits defined in a basin program rule adopted pursuant to the Commission's authority in ORS 536.300 and 536.310. Any proposed superseding basin program definition must consider, at a minimum, the anticipated impacts of the new definition on:

(A) the number of wells that may go dry; and

(B) the character and function of springs and groundwater dependent ecosystems; and

(C) the long term, efficient, and sustainable use of groundwater for multiple beneficial purposes.

(e) This definition does not apply to Critical Groundwater Areas designated under OAR 690-010.

Works Cited

Abatzoglou, J.T., D.E. Rupp, and P.W. Mote, 2014. Seasonal Climate Variability and Change in the Pacific Northwest of the United States. Journal of Climate 27:2125–2142.

Albano, C.M., J.T. Abatzoglou, D.J. McEvoy, J.L. Huntington, C.G. Morton, M.D. Dettinger, and T.J. Ott, 2022. A Multidataset Assessment of Climatic Drivers and Uncertainties of Recent Trends in Evaporative Demand across the Continental United States. Journal of Hydrometeorology 23:505–519.

Cuthbert, M.O., T. Gleeson, M.F.P. Bierkens, G. Ferguson, and R.G. Taylor, 2023. Defining Renewable Groundwater Use and Its Relevance to Sustainable Groundwater Management. Water Resources Research 59:e2022WR032831.

Freeze, R.A. and J.A. Cherry, 1979. Groundwater. Prentice-Hall, Englewood Cliffs, N.J.

Gannett, M.G. and K.E. Lite Jr., 2013. Analysis of 1997–2008 Groundwater Level Changes in the Upper Deschutes Basin, Central Oregon. USGS Numbered Series, U.S. Geological Survey, Reston, VA.

Gannett, M.W., K.E. Lite Jr., D.S. Morgan, and C.A. Collins, 2001. Ground-Water Hydrology of the Upper Deschutes Basin, Oregon. USGS Numbered Series, U.S. Geological Survey, Portland, OR.

Gleeson, T., M. Cuthbert, G. Ferguson, and D. Perrone, 2020. Global Groundwater Sustainability, Resources, and Systems in the Anthropocene. Annual Review of Earth and Planetary Sciences 48:431–463.

Helsel, D.R., R.M. Hirsch, K.R. Ryberg, S.A. Archfield, and E.J. Gilroy, 2020. Statistical Methods in Water Resources. USGS Numbered Series, U.S. Geological Survey, Reston, VA.

Kawa, N., K. Cucchi, Y. Rubin, S. Attinger, and F. Heße, 2023. Defining Hydrogeological Site Similarity with Hierarchical Agglomerative Clustering. Groundwater 61:563–573.

Mote, P.W., J.T. Abatzoglou, K.D. Dello, K. Hegewisch, and D.E. Rupp, 2019. OCAR4 - Oregon State Oregon Climate Change Research Institute. Oregon Climate Change Research Institute. https://www.occri.net/ocar4/. Accessed 16 Jan 2024.

National Oceanic and Atmospheric Administration (NOAA), 2024. Climate at a Glance: Divisional Time Series. https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/divisional/time-series/3501/pcp/all/11/1895-2023. Accessed 5 Jan 2024.

Waibel, M.S., M.W. Gannett, H. Chang, and C.L. Hulbe, 2013. Spatial Variability of the Response to Climate Change in Regional Groundwater Systems – Examples from Simulations in the Deschutes Basin, Oregon. Journal of Hydrology 486:187–201.

Yue, S., P. Pilon, B. Phinney, and G. Cavadias, 2002. The Influence of Autocorrelation on the Ability to Detect Trend in Hydrological Series. Hydrological Processes 16:1807–1829.

Yue, S. and C.Y. Wang, 2002. Applicability of Prewhitening to Eliminate the Influence of Serial Correlation on the Mann-Kendall Test. Water Resources Research 38:4-1-4–7.

Yue, S. and C. Wang, 2004. The Mann-Kendall Test Modified by Effective Sample Size to Detect Trend in Serially Correlated Hydrological Series. Water Resources Management 18:201–218.