Assessing Well Depth in California Between 1990 and 2021

Halina Do-Linh

12-02-2021

Question

Are we drilling deeper wells faster over time in California?

1.5 million people in California rely on domestic well water. (Johnson) Majority of the 1.5 million reside in rural areas without alternative drinking-water sources. Furthermore, in the Central Valley, it has been found that one in five wells now run dry due to groundwater level declines (Jasechko 1). While some have responded to this issue by drilling deeper, this practice has also been found to be unsustainable because it is costly, it is impractical in certain hydrogeological conditions, and deep groundwater is often brackish or saline. (Perrone 3) While we are aware that groundwater levels are declining and that deeper drilling is unsustainable, we do not know how fast we are drilling deeper in California.

It is important to study groundwater trends because water is an important resource in California and it is crucial for groundwater management to be sustainable and equitable. By understanding the rate at which we are drilling deeper wells over time, we can

Data

My data has been obtained by Dr. Debra Perrone and Dr. Scott Jasechko for their research on groundwater wells. California domestic well water data was obtained through personal communication with state representatives from the California Department of Water Resources (CADWR) in March 2016 and April 2016. The CADWR maintains the California Groundwater Well Completion Report database. Dr. Perrone and Dr. Jasechko reported that the data may contain biases by region, time, purpose, and compliance.

This data is international, but I subsetted the data to only contain wells from California. I also added region codes based on the California Complete Count Office, which grouped California's 58 counties into 10 regions based on hard-to-count populations, like-mindedness of the counties, capacity of community-based organizations within the counties, and state Census staff workload capabilities.

Data Histograms

After taking the natural log of well depth (m), it appears to have a normal distribution . It also looks like majority of wells are located in Region 1 or Superior California (Figure 1).

Exploratory graphs of total counts of wells both over time and by region suggests that overall we are completing less wells (figure 1). However, we are also seeing that we could be potentially drilling deeper wells over time (figure 2).

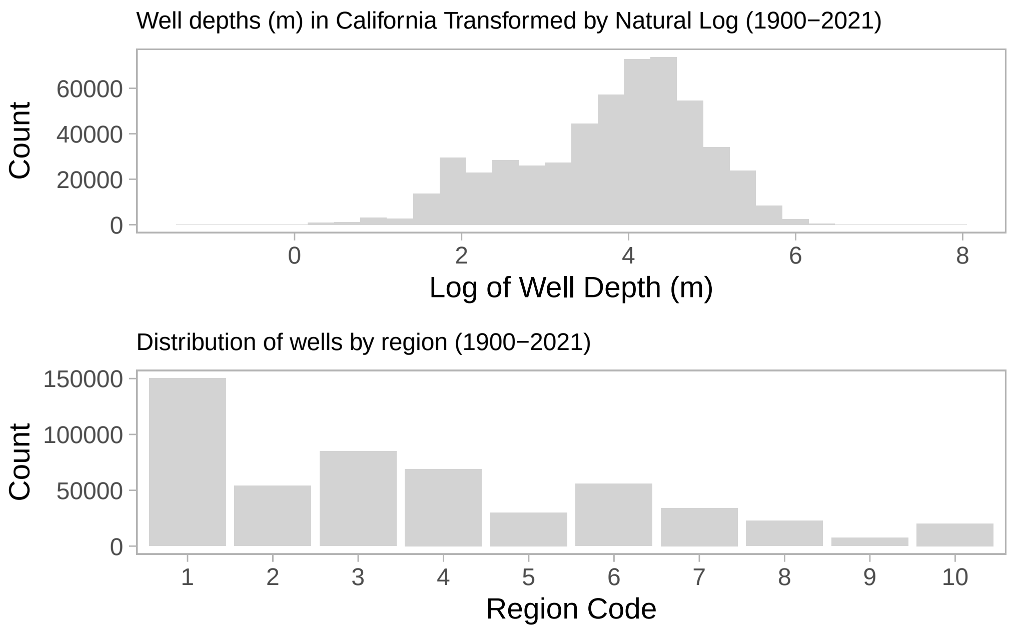

Figure 1: Well depth (m) data transformed by natural log were normally distributed (top). Majority of wells in California can be found in Northern regions of California identified here as region codes 1-5 (bottom).

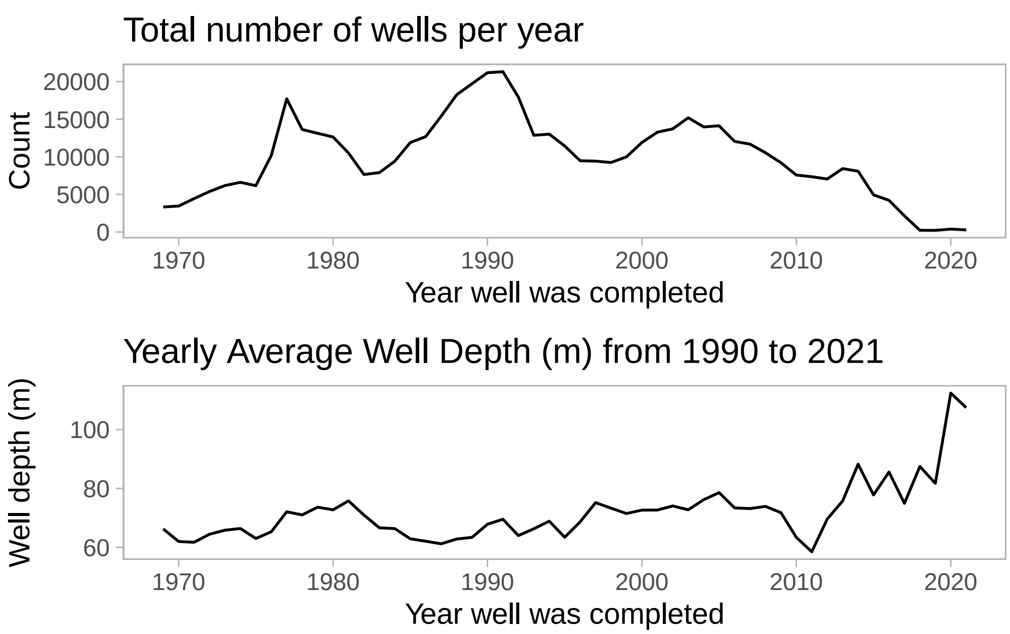

Figure 2: Number of wells are appearing to decrease overtime in California with a significant drop in 2005 (top). Average well depth (m) are appearing to increase overtime in California with a significant increase after 2010 (bottom).

Analysis Plan

1. Regression Analysis

Use a multiple linear regression model to test if there an effect of time and region on well depth (m), and if there is an effect - what is the rate of well depth (m) over time.

$$welldepth(m) = \beta_0 + \beta_1 \cdot year_i + \beta_2 \cdot region_i + \varepsilon_i$$

I added an interaction effect to my multiple linear regression model to see if the effect of year on well depth (m) depends on region. I predict there may be deeper wells in the northern regions than the southern regions since majority of California's drinking water comes from northern regions, and that the rate of well depth will be faster in the northern region than the southern region as well.

$$welldepth(m) = \beta_0 + \beta_1 \cdot year_i + \beta_2 \cdot region_i + \beta_3 \cdot year_i \cdot region_i + \varepsilon_i$$

2. Time Series Analysis

Run decomposition analysis to test for seasonality. From my exploratory graphs I visually saw some evidence of cyclical patterns (see Figure 6 in Appendix) and wanted to test to see if this was significant.

Results

Multiple Linear Regression

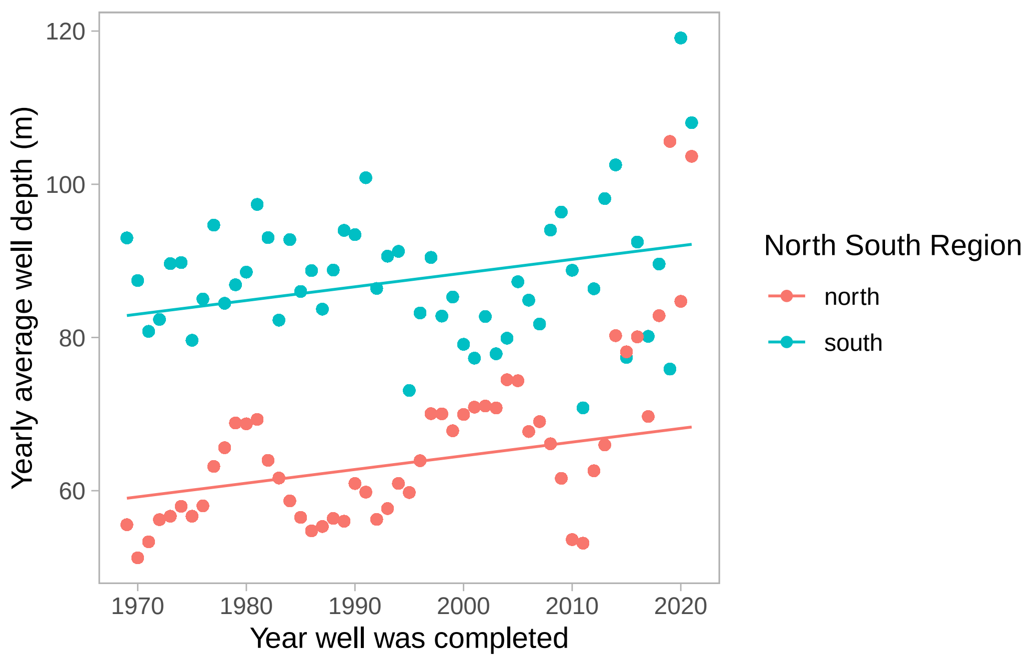

Figure 3: Mulitple linear regression model of year and region on well depth. P-value < 0.00000000000000022. Adjusted R-value 0.7354648.

Reporting & Interpretation of Coefficients: The `intercept` coefficient tells us that we expect to see a well depth of *-292.98 m* at year 0 in the south region (this is nonsensical). The `year_completed` coefficient tells us that we expect well depth to increase by *0.178 m* each year, holding region constant. The `north_south` coefficient tells us that we expect to see, on average, a *23.85 m* increase in well depth in the south region compared to the north region, holding `year` constant.

Overall Interpretation: This parallel slopes model tells us that `year`, regardless of `region`, will have the same impact on yearly average well depth (m). This impact is significant with a p-value of *0.00000000000000022* at the significance level 0.01. The adjusted r-value also reports that *74%*of the data is explained by `year`. So over time we are seeing that `year` is significantly increasing well depth (m) by 0.178 m every year. This is a fairly fast rate.

Mulitple Linear Regression Model with Interaction Effect

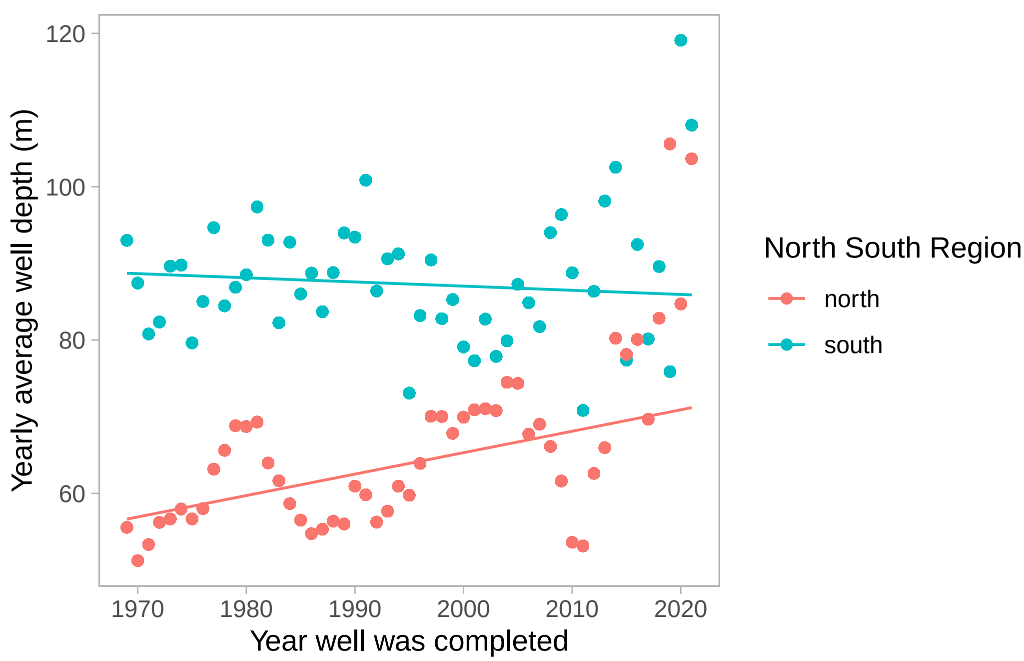

Figure 4: Interaction effect of year and region on well depth. P-value < 0.00000000000000022. Adjusted R-value 0.7563347

Reporting & Interpretation of Coefficients: The `intercept` coefficient tells us that we expect to see a well depth of **-493.84 m** at year 0 in the north region (this is nonsensical). The `year_completed` coefficient tells us that we expect well depth to increase by **0.27 m** each year for a well located in the northern region. The `north_south` coefficient tells us that well depths are, on average, **689.50 m** deeper in the north region than the south region at year 0 (this is also nonsensical). The `year_completed:north_south` coefficient tells us that the change in average well depth (m) after one year for a well in a northern region compared to a well in the southern region is **-0.334 m**.

Overall Interpretation: This interaction effect tells us that the relationship between `year` and `well_depth` varies with `region`. This is a significant effect with a **p-value of 0.00000000000000022** at the the significance level 0.01. The adjust r-value is reports that **75%** of the data is explained by this interaction effect. Since this is a slight increase from the adjusted R^2 value from the model with no interaction, this tells us that the interaction model improves the model fit (at least a little). The rate of the wells being drilled in the north are at a **rate of 0.28 m per year**. This is faster than the rate we found without the interaction effect, so we are seeing that not only are we drilling deeper wells faster over time, but that we are drilling deeper in northern regions than southern regions.

Time Series Analysis - Decomposition

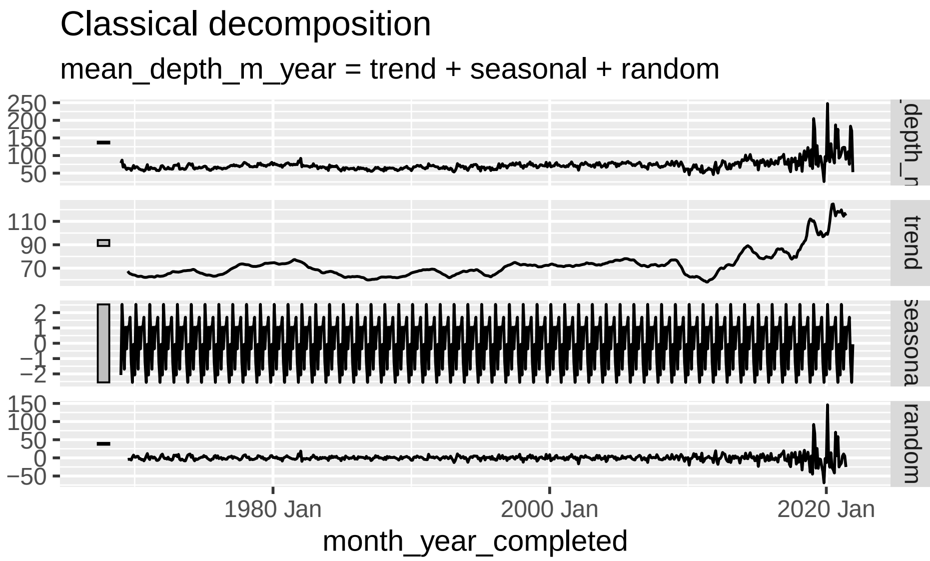

Figure 5: Decomposition of yearly mean well depth (m) with significance with trend, but not seasonal.

Interpretation: We see there is evidence of a long-run upward trend over time and there is no evidence of seasonality. Based on the size of the bars, the `trend` component is more important in driving overall variation in yearly average well depth (m) since the gray bar for `trend` is much smaller than the bar for the `seasonal` value. This confirmed for me that the "cycles" I was seeing in Figure 6 were not significant. It also confirms what we already know: that over time we are drilling deeper wells in California.

Future Research

Future research would check to see if seasonality is significant in only northern regions. There could be a more seasonal affect in northern regions due to geology or climate since this varies from the southern regions.

Other research could include policy into regression analysis or other statistical analysis to see what kind of impact policy has on well drilling in California. This would be important to look at since we know that a large amount of wells are running dry in California (Jasecko 1). And if groundwater levels continue to decline, we need to understand what policies are most effective so that we can better manage water now and in the future.

Appendix

Figures

References

1. Jasechko, Scott, and Debra Perrone. "California's Central Valley Groundwater Wells Run Dry During Recent Drought." Earth's Future, vol. 8, no. 4, Apr. 2020. DOI.org (Crossref), https://doi.org/10.1029/2019EF001339.

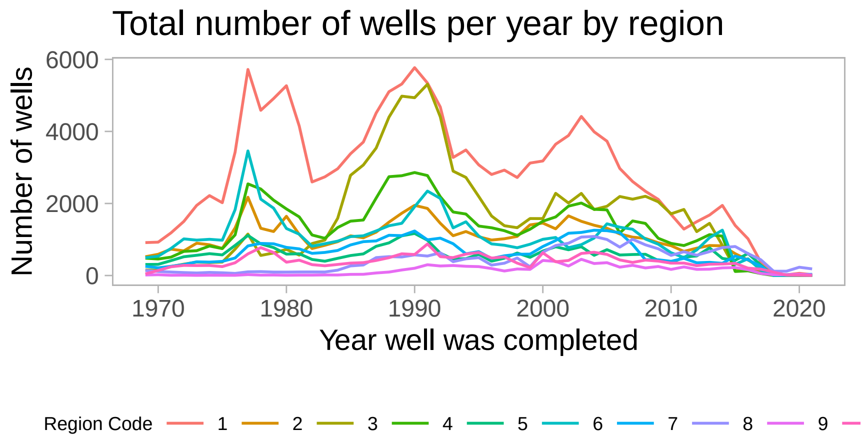

Figure 6: Total number of wells by region with potential for cyclical or seasonal trending.

2. Johnson, Tyler D., and Kenneth Belitz. "Identifying the Location and Population Served by Domestic Wells in California." Journal of Hydrology: Regional Studies, vol. 3, Mar. 2015, pp. 31–86. DOI.org (Crossref), https://doi.org/10.1016/j.ejrh.2014.09.002.

3. Perrone, Debra, and Scott Jasechko. "Deeper Well Drilling an Unsustainable Stopgap to Groundwater Depletion." Nature Sustainability, vol. 2, no. 8, Aug. 2019, pp. 773–82. DOI.org (Crossref), https://doi.org/10.1038/s41893-019-0325-z.