Assessment of Municipal Population and Water Demand Projections in Texas Regional and State Water Supply Planning

TWDB Contract Agreement 2500012881
Biannual Progress Report 3

February 23, 2026

Submitted by:
Fitsume Wolkeba and Bridget R Scanlon
University of Texas Bureau of Economic Geology
Pickle Research Campus, Bldg. 130
10100 Burnet Rd., Austin, TX 78758
Ph: 512 471 8241; bridget.scanlon@beg.utexas.edu

For Work Completed 08/2025 to 03/2026
Date Submitted: February 23, 2026

Submitted to: Amanda Covington – Contract Manager, TWDB

Submitted by: Bridget Scanlon, Research Professor, UTBEG

Project Goals

1) Better understand how other states, major water providers, and organizations assess or project future municipal water demands for their long-range water supply planning processes, including for water uses not served by utilities (i.e. exempt domestic and other use), including whether/how they use scenarios and why.

2) Assess the accuracy (quantified error) of Texas regional water supply plans' previous *population and municipal water demand* projections (as negotiated and modified through the planning process via TWDB and the Regional Water Planning Groups (RWPGs)) and compare to the accuracy of the Texas Demographic Center's population projection accuracy over the same time period (2000 – 2020). Accuracy of population projections should also be measured against actual population counts by the U.S. Census Bureau in the decennial Censuses in 2000, 2010, and 2020 at various geographic levels, including the county and municipal WUG level.

3) From the accuracy assessment, determine a reasonable quantified uncertainty of the projections and recommend how to apply that quantified uncertainty to future projections developed using TWDB's current projection methodologies.

4) Assess whether different projection methodologies, such as land use modeling for utility population projections, would increase the accuracy of future projections.

5) The project should result in implementable recommendations including, but not limited to:

   a) determining whether TWDB's current population and municipal water demand projection methodologies are still appropriate for future projections or if the methodologies should be changed.

   b) assessing what type of flexibility, if any, could be directly incorporated into draft TWDB population projections to allow RWPGs and TWDB to more efficiently and quickly assess and arrive at an agreeable but credible set of population projections – especially for the later planning decades.

   c) comparing the margin of error in the population projections versus the municipal water demand projections for same geographic area.

PROJECT PROGRESS DISCUSSION

Summary

- We have made substantial advances in the following areas since the previous progress report:
- Finalization of web and literature review of other state water planning process
- Detailed analysis of draft and final population projections at state, regional water planning area (RWPA), county, and WUG level, with particularly emphasis on the WUG level.
- Completion of five county area analysis, including predominantly rural and urban counties.
- Evaluation of water demand projections, including GPCD and water use, with comparison of draft and final projections at state, RWPA, and county level and WUG level.
- Evaluation of city projections of population and GPCD, based on discussions and report review for City of Austin and San Antonio
- Uncertainty analysis with relative contributions from uncertainties in populations and GPCD

Subtask 1. Projection Methodology Literature and Data Review

Status: We consider Subtask 1 80% complete.

Subtask 1a. Literature Review

We completed preliminary compilation of population projections and gallons per capita per day (GPCD) data to support their water supply planning programs for the western states. The level of detail in each of the state plans is highly variable. It seems that many states rely on population projections and GPCD values estimated developed at subbasin or utility scale and aggregate to state level. Projections have been developed by different states up to a maximum of 2075. Assessment of the accuracy of population projections or GPCD estimates is not evident in the state water planning documents.

We also reviewed other organizations in Texas, including Harris-Galveston and Fort Bend Subsidence Districts Joint Regulator Plan Review, North Texas Municipal Water District, North Central Texas Council of Governments, Capital Area Council of Governments, and Texas Department of Transportation.

Subtask 1b. Data Review

This subtask focuses on analysis of municipal WUGs to be included in the study. However, we have analyzed all the municipal WUGs at this stage.

Next Steps:

We have compiled a first draft of the literature review and are checking with agencies to obtain more details about their planning process. We will finalize the report.

Subtask 2. Accuracy of both the draft and final, where available, population projections in regional water plans, and accuracy of Texas Demographic Center's population projections

Status: We consider Subtask 2 to be 90% complete.

Subtask 2.1:

Comparison of TDC projected population to population counts in 2000, 2010, and 2020 by the U.S. Census Bureau in the decennial Censuses for the whole state and each county.

Completed

TDC project population was compared to population counts in 2000, 2020, and 2020 by the US Census Bureau for the state and each county. Results at the state level show that for 2020 (2017 projection) that TDC population underestimates census population with migration scenarios of 0.0 (6%) and 0.5 (1%) but overestimates for migration scenario of 1.0 (5%). Comparison at the county level shows that the best comparison was found with migration scenario of 0.5 with the recent data. The county populations were ranked, and comparisons were conducted using percentiles. Most of the analyses show divergence between TDC and the census when percentiles exceeded ~60th percentile. Uncertainties in projections increase with the length of the projection time.

Subtask 2.2:

Comparison of draft Regional Water Plan projected population to population counts in 2000, 2010, and 2020 by the U.S. Census Bureau in the decennial Censuses for the whole **state**, **regional water planning area**, **county**, and **select WUGs**.

Completed

Analysis of the **WUG** data indicates that the population projections diverge from the census population with increasing county WUG population with greater divergence about the 90th percentile (with projections above and below actual WUG populations). WUG population projections are generally overestimated relative to the census data.

Subtask 2.3

Status: We consider Subtask 2.3 to be 95% complete.

Comparison of final Regional Water Plan projected population to population counts in 2000, 2010, and 2020 by the U.S. Census Bureau in the decennial Censuses for the whole state, regional water planning area, county, and select WUGs.

The final population projections are similar to the draft projections and census data at the **state** level with differences of ~ 1% for 2020 data.

There is only one set of **RWPG** population projections; therefore, we evaluated the available projection with the census data. Results at the **state** level show low percentage errors between draft and census population data (~ 1%). Differences between projected and census data range from -3 to 14% at the scale of the **RWPGs** for 2017 projection of 2020 census data. Projection uncertainties are similar or increase with the length of the period projected.

We use the Cumulative Distribution Function (CDF) to summarize the differences across plans and projections. We have provided a CDF interpretation example in Figure 1. A perfect or ideal projection will show as a vertical line (0 difference with census on x axis, purple line). Better projections are shown by steeper curves closer to the 0 vertical line.
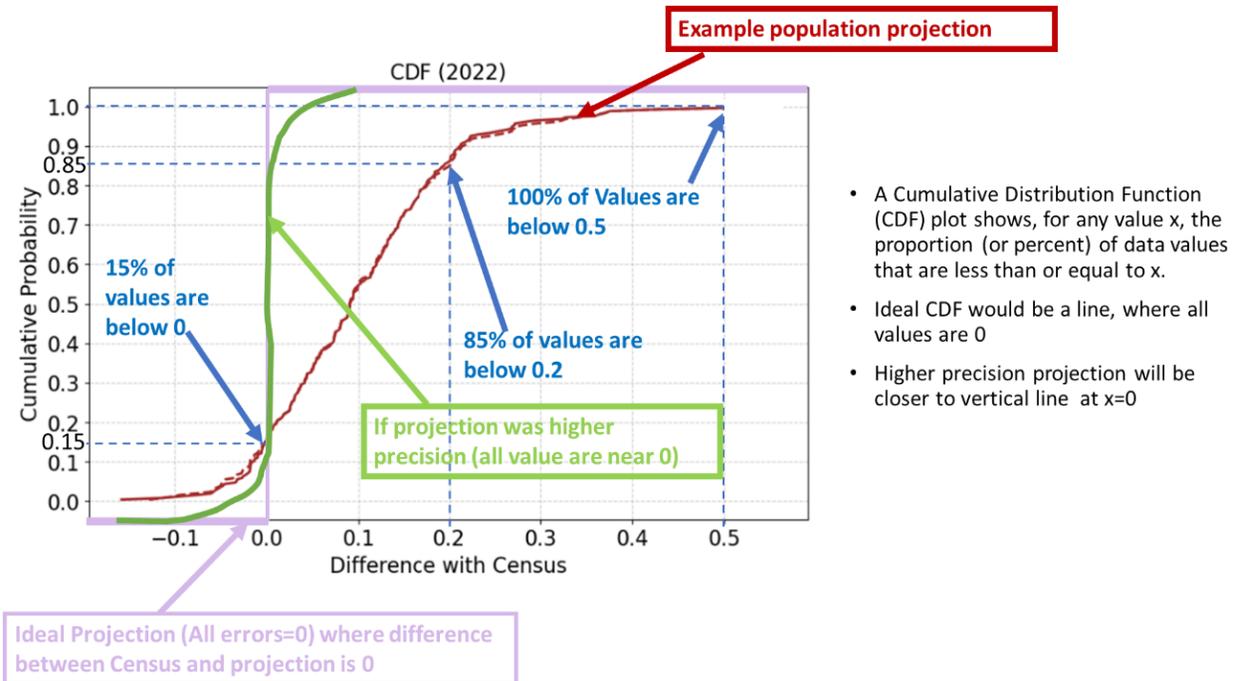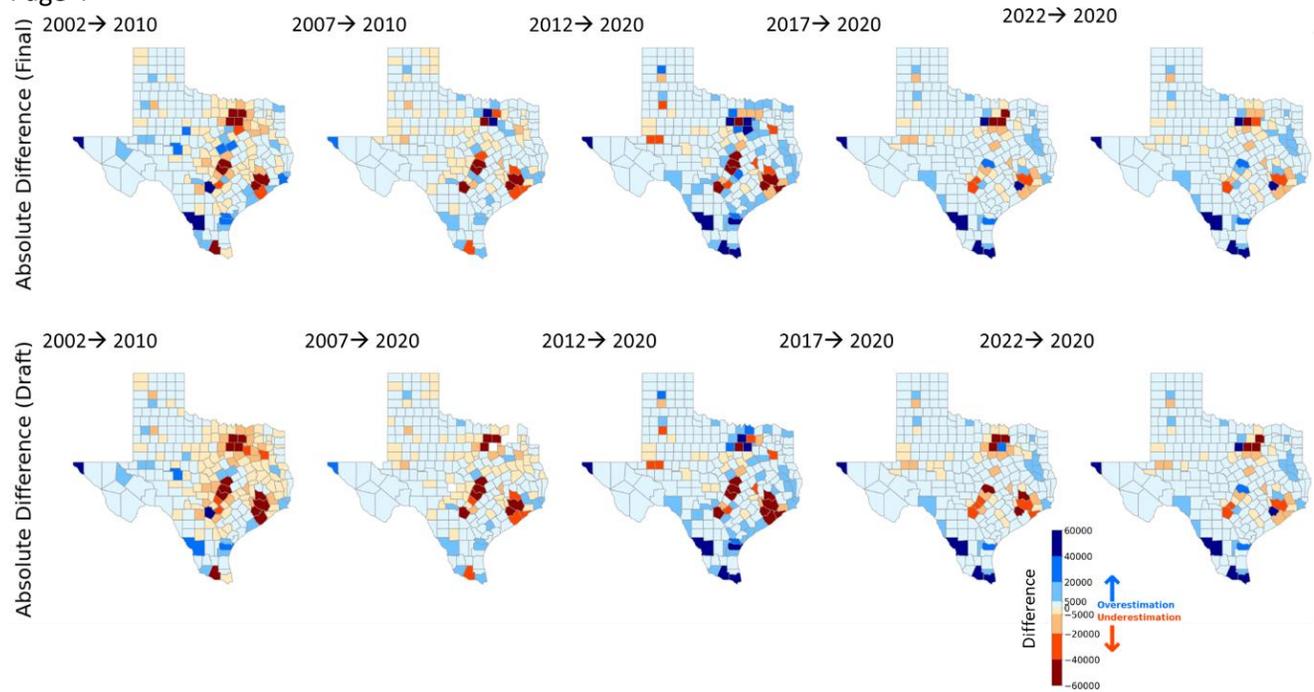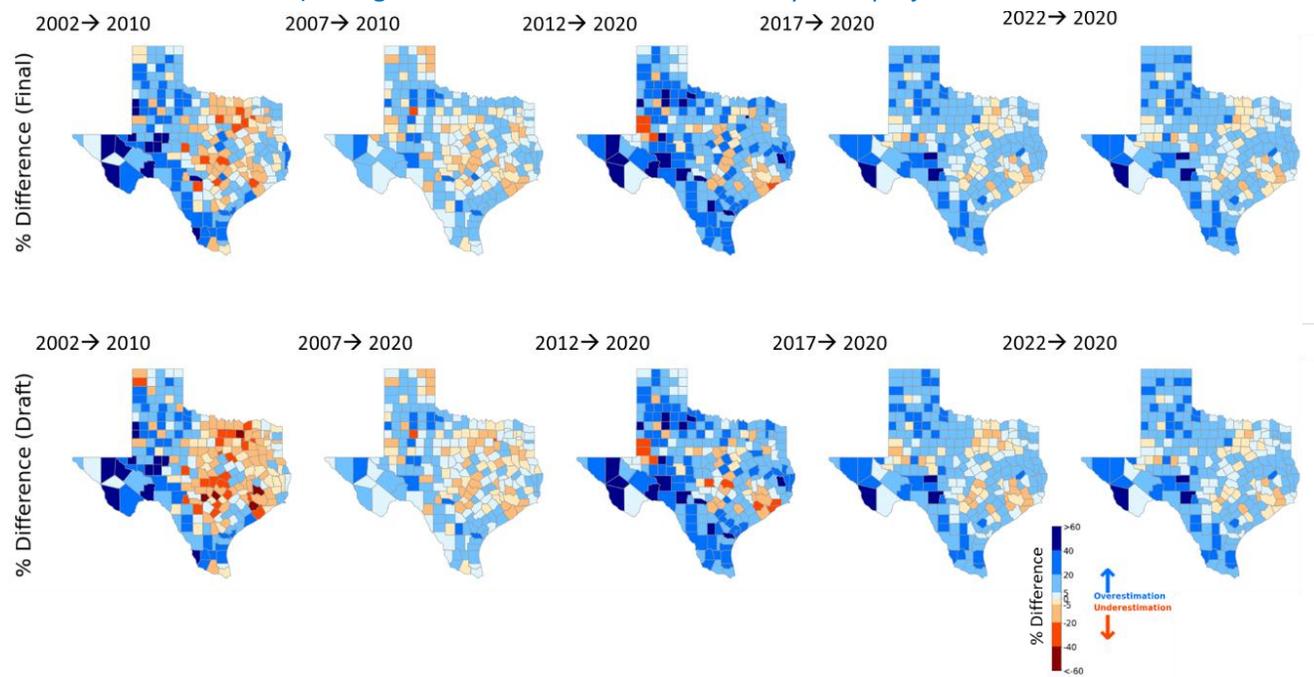

Figure 1: Cumulative Distribution Function (CDF) interpretation

Comparison at the county level shows more overestimation than underestimation relative to the census (7 out of 9 plans with overestimation in more than 60% of counties). In terms of magnitude difference, considering all planning periods and projections, ≤ 25% of counties overestimate by >5000 and ≤ 20% of counties underestimate by >5000 (Figure 2c). In terms of percent differences, ≤ 30% of counties overestimate by >25%, and ≤ 10% of counties underestimate by >25% (Figure 2c).

Results for the WUGs are similar to the county-level analysis, showing more overestimation than underestimation relative to the census. Plots for WUGs are not included in this report. Up to 70% of WUGs are overestimated relative to the census depending on the plan. There was more agreement with the census at the WUG level than at the county level (steeper CDF plots in WUGs). In terms of magnitude differences, ≤10% of WUGs overestimate (underestimate) by 5000. In terms of percent differences, ≤ 40% of WUGs overestimate by more than 25% and ≤ 15% of WUGs underestimate by more than 25%.

a) Magnitude differences between county level projections and census

b) Percent differences between county level projections and census
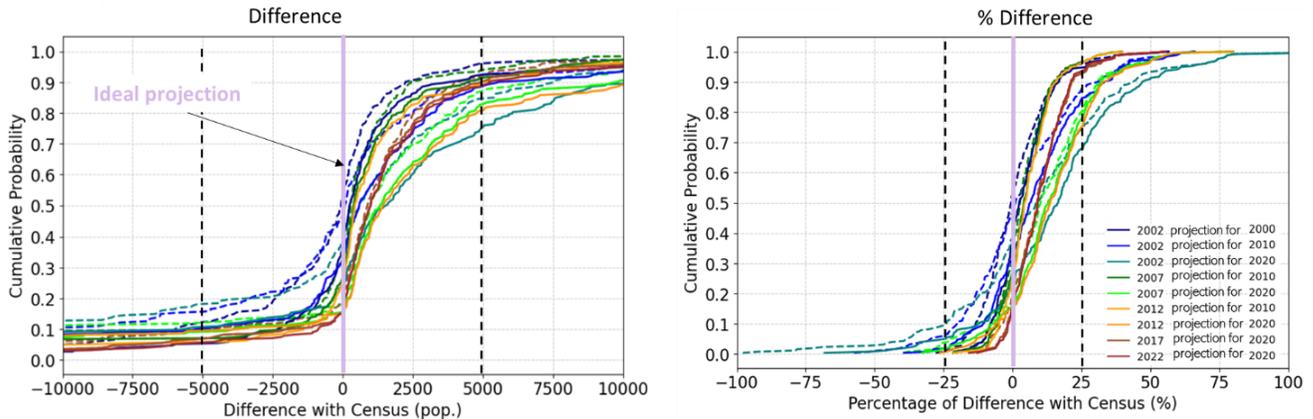
c) Cumulative distribution function (CDF) plots

Figure 2. The spatial maps illustrate the a) magnitude difference between draft and final projection against census for **counties**. b) percent difference between draft and final against census for **counties**.

Subtask 2.5:

Accuracy of urban versus rural populations per TWDB's definition Texas Water Code Ch. 15.001(14):

a) Counties where the population is less than 50,000.

b) WUGs where the population is 10,000 or less and does not overlap an urban area, as defined by the U.S. Census Bureau in 41 CFR 102-83.80.

c) Other utility size thresholds as determined appropriate and in line with uniform reporting requirements across the state (small, medium, medium-large, large, and metropolitan categories[1]).

Urban and rural population analysis is only applied to WUGs. There is less magnitude difference in rural than in urban WUGs and more percentage difference in rural than in urban WUGs (Figure 3).
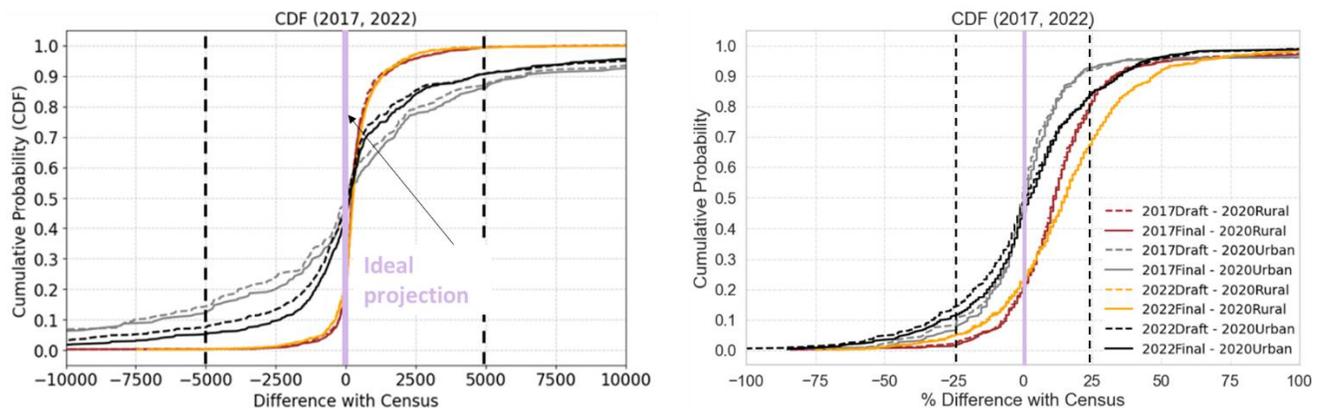

Figure 3: Cumulative distribution function showing the magnitude and percentage differences in urban and rural WUGs across all 2017 and 2022 plans and projections.

Special focus on County-Other WUG population projections.

We have compiled County-other WUG data and identified which County-other WUGs have the largest magnitude and percent differences with the census. Similar analysis was conducted for water demand projections.

---

[1] https://www.twdb.texas.gov/publications/reports/special_legislative_reports/doc/Water-Use-of-Texas-Water-Utilities-87th-Legislative.pdf?d=7743.300000071526

Subtask 3. Accuracy of state and regional water plans' municipal demand projections (with consideration of drought supply planning)

Status: We consider subtask 3 to be 80% complete.

We have evaluated the difference between actual drought year (2011) demand and demand projections from all plans. The plans underestimated water demand compared to the 2011 demand in 45%-60% of counties. We have also made similar calculations against decadal maximum (2000-2020) water demand and projected water demand (Figure 5). The 2017 and 2022 projections of 2020 show relatively smaller differences compared to previous plans.
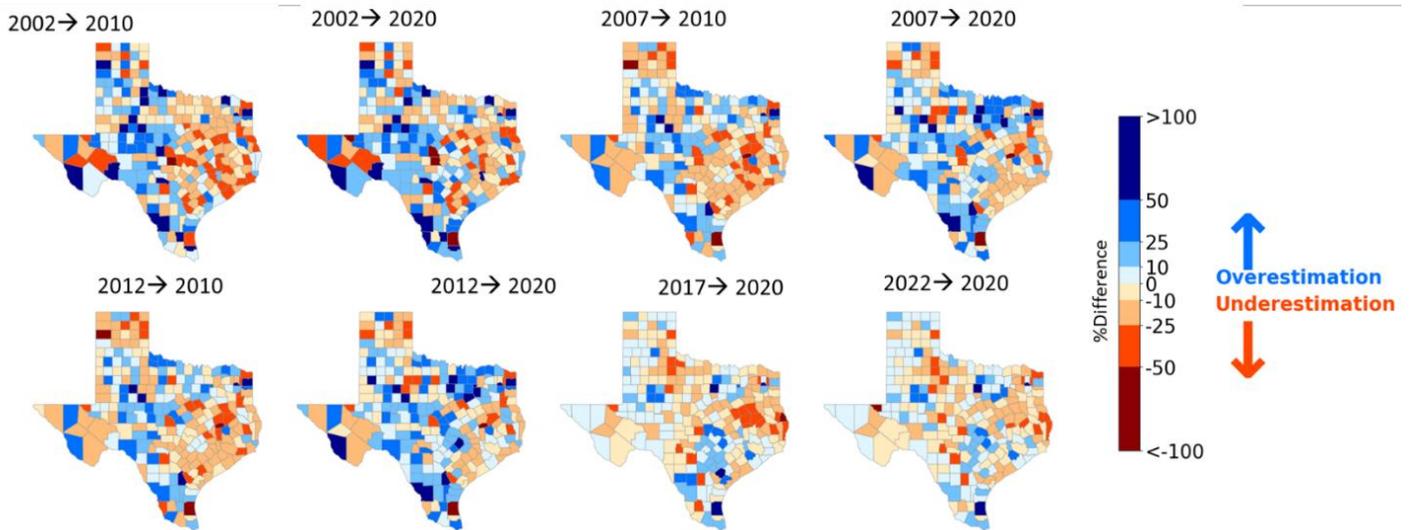

Figure 5: Percentage of differences showing over and underestimation of 2011 drought demand in counties.

Subtask 3.1

Temporal analysis of projected municipal demand considering the various factors of municipal water use:

a) Residential

b) Per capita

We have assessed municipal demand projections at state, RWPG, county, and WUG levels.

Subtask 3.2:

2) Analysis results from various geographic levels: State, regions, counties, select WUGs.

The analysis for the regional water state and planning group has been completed and reported in the previous progress report.

At the county level, 60% to 80% of counties overestimate water use. Populated counties such as Travis, Dallas, and Harris show the highest magnitude differences. The reliability of the water demand projections improved with time as seen from the 2002 versus 2017 projections (Figure 6). Overall, the different planning periods overestimate water demand in ≤30% of counties by more than 1000 af and underestimate water demand in ≤10% of counties by more than 1000 af. In terms of percent difference, ≤40% of counties overestimate by >25% and less than 5% of counties underestimate by >25% (Figure 7).
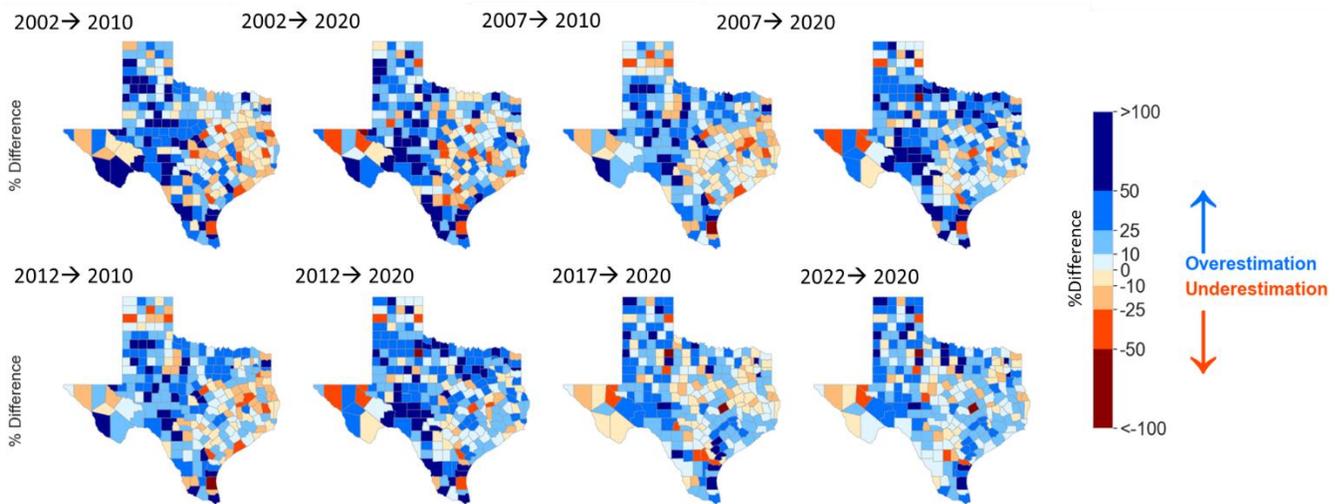
Figure 6: Percentage of over and underestimation of actual water use in counties. (2022 projections for 2020 were finalized before 2020)
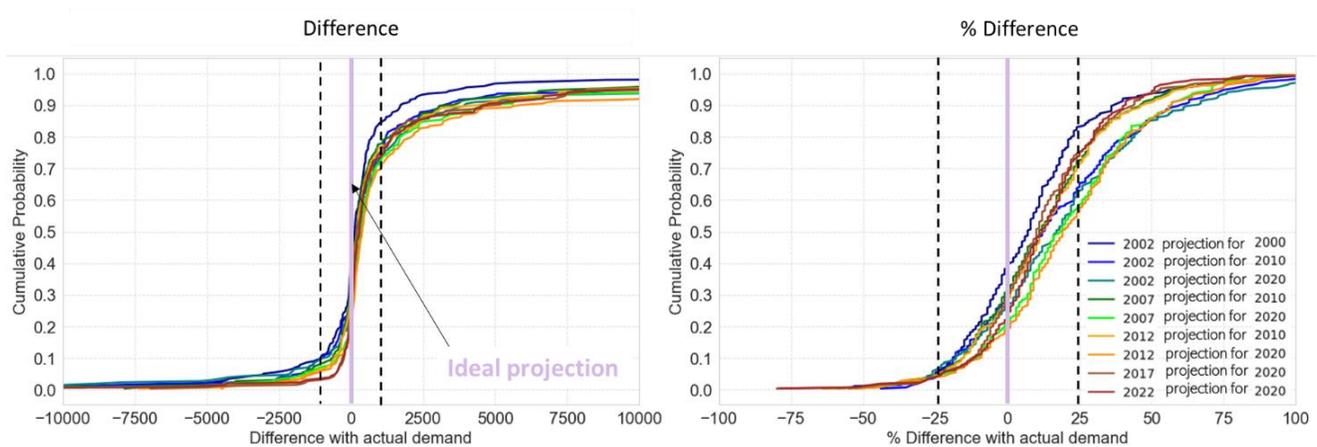

Figure 7: The graphs display the cumulative distribution function (CDF) between projected water demand (af) and actual net water demand (af) across counties. A similar set of analyses and visualizations has been developed for WUGs to identify projection accuracy and usage trends at WUG level. Underestimating actual demands is a concern as the plans are expected to include drought conditions.

Subtask 3.3: Evaluation of urbanization and impacts on GPCD using USGS land use change data.

This subtask has not been initiated.

Subtask 3.4: Accuracy of urban vs rural water use.

a) Counties where the population is less than 50,000.
b) WUGs where the population is 10,000 or less and does not overlap an urban area, as defined by
the U.S. Census Bureau in 41 CFR 102-83.80.
c) Other utility size thresholds as determined appropriate and in line with uniform reporting
requirements across the state (small, medium, medium-large, large, and metropolitan categories).

We used the Texas Water Code for defining rural and urban WUGs (Texas Water Code 15.001(14)). Urban and rural population analysis only applied to WUGs. There is less magnitude difference in rural than urban WUGs and more percentage difference in rural than urban WUGs (Figure 8). It is important to note that 2022 population projections were finalized before 2020 in 2018 and did not include the 2020 census population during development.

We selected six major cities and conducted a time-series analysis of actual population, GPCD, and total water use. This approach allowed us to explore the relationships among these three components and identify trends and patterns in urban water demand (Figure 9).
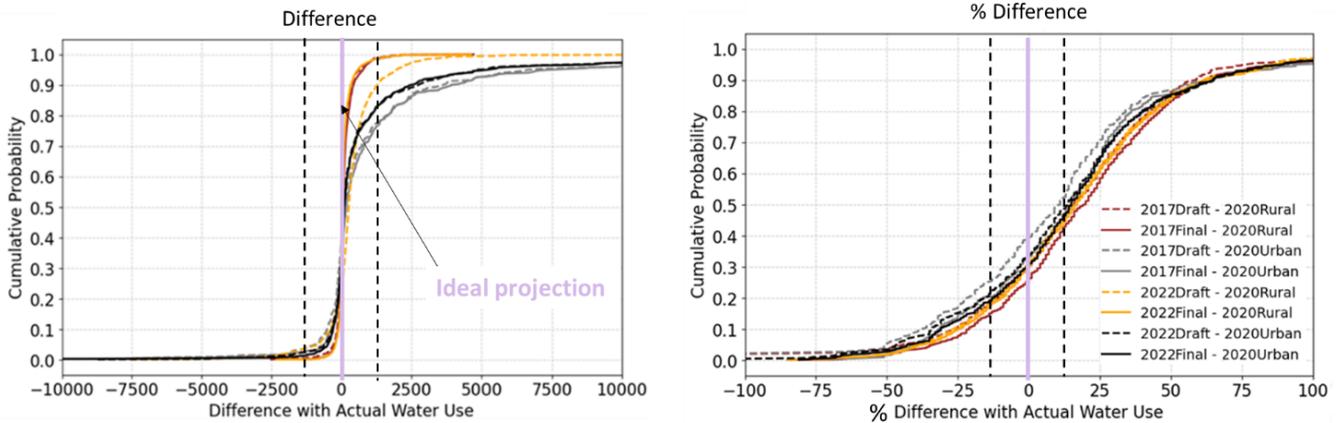

Figure 8: The graphs display the cumulative distribution function (CDF) between projected water demand and actual water use across urban and rural WUGs showing the % underestimating and overestimating urban and rural WUGs across plans. Underestimating actual demands is a concern as the plans are expected to include drought conditions.
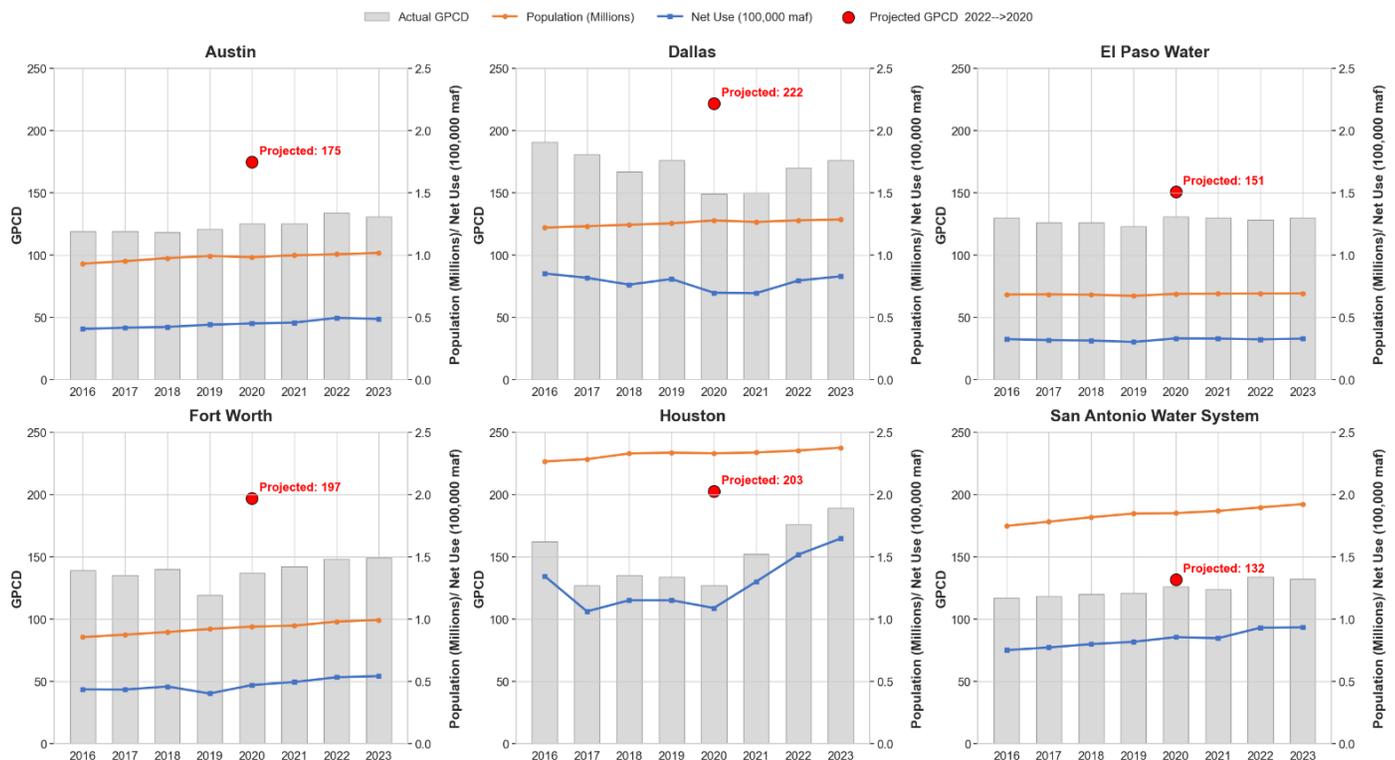

Figure 9: Population (millions, solid line), water use (maf, block), and GPCD (dashed line).
Data source: (https://www3.twdb.texas.gov/apps/reports/WU_REP/SumFinal_UtilityWUGSum). There is no systematic pattern in GPCD over time in these cities. Houston shows a large increase from 2000 – 2023 and also a population increase. In all compared cities 2022 projection was more than actual.

Subtask 3.5: Special focus on the accuracy of the municipal water demand projections compared to the historical estimates for County-Other WUGs.
We have complied with and analyzed County-other WUG water demand and identified County-other WUGS with the largest absolute and percentage differences.

Subtask 3.6: Statistical analysis documentation and outcome: Degree/margin of error, confidence level in the accuracy of the projections at each geographic level.

The degree of uncertainty, margin of error, and confidence level in the accuracy of projections at RWPG, county, and WUG level have been analyzed and documented (Figure 10). The Mean Absolute Error (MAE) for counties and WUGs increases with the projection period (2002→2000 vs 2002→2020).
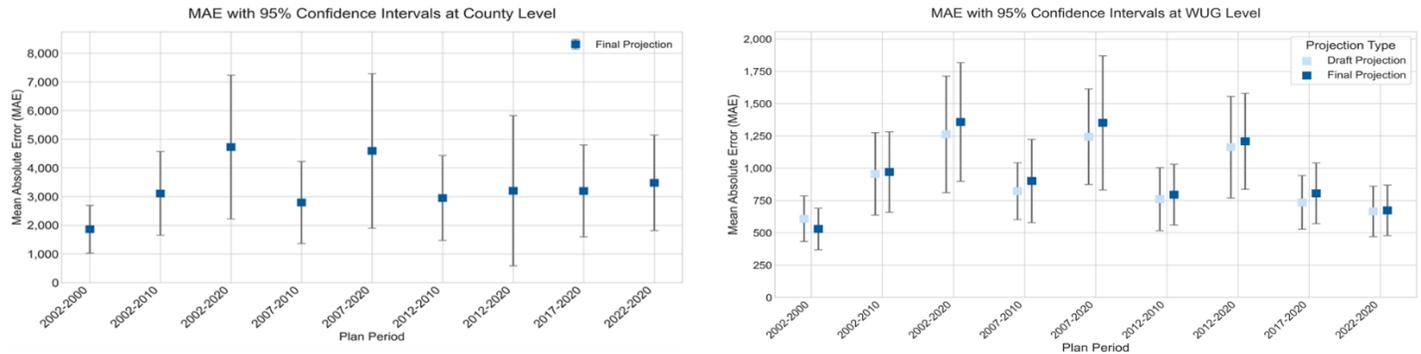

Figure 10: Margin Error and Confidence Interval at 95% for water demand projections

Subtask 3.7: Comparison to the margin of error identified in Subtask 2 for same geographic levels (i.e., compare margin of error in population projections to margin of error in municipal demand projections).

Status: We consider Subtask 3.7 to be 80% complete.

We have compared over- and underestimations by both population and water demand. Plan 2017 for 2020: Found 87 counties (34%) with double overestimation (where both population and water demand were overestimated relative to 2011 drought year). A total of 12 (5%) counties agree in underestimation of population and water demand when compared to drought year 2011 water demand. These results suggest that overestimation in water demand in 34% of counties can be attributed to overestimation of population, but in the remaining majority of counties (61%), the discrepancies between projected and 2011 water demand must be related to GPCD uncertainty. Plan 2022 for 2020: Found 92 (36%) counties with double overestimation (where both population and water demand are overestimated). 13 (5%) counties agree on underestimation in 2022 (Figure 11).

We used Refined Laspeyres Decomposition Error Attribution (Figure 12) and Logarithmic Mean Divisia Index (LMDI) Decomposition Analysis (Figure 13) to assess the source of errors in water demand projections (population or GPCD). In all approaches, GPCD is the largest contributor to the water demand projection uncertainty.

Status: We have initiated this subtask and about 70% completed. We will expand this evaluation over WUGs and RWPG.

1. **Refined Laspeyres Decomposition Error Attribution (Figure 12)**

Population Attribution

$E_{Pop}=(Pop_p-Pop_a) \times GPCD_a$

GPCD Attribution

$E_{GPCD}=Pop_a \times (GPCD_p-GPCD_a)$

Interaction Attribution

$E_{Pop\_GPCD}=(Pop_p-Pop_a) \times (GPCD_p-GGPCD_a)$

Sum to the total error:

$E=E_{Pop}+E_{GPCD}+E_{Pop\_GPCD}$

where,

- Pop_a = Actual population

- Pop_p = Projected population

- GPCD$_a$ = Actual GPCD

- GPCD$_p$ = Projected GPCD

- Total Error in Water demand = (Pop$_p$×GPCD$_p$)−(Pop$_a$×GPCD$_a$)

2. **Logarithmic Mean Divisia Index (LMDI) Calculation (Figure 13)**

Contributing factors

This analysis assumes the total error in demand can be perfectly attributed to errors in population and/or GPCD projections. Given the population projection and drought year, GPCD is used for demand projection calculation, and the refined Laspeyres decomposition error attribution analysis shows relatively smaller interaction error contribution. The perfect decomposition analysis is a reasonable assumption in our context.

D=P x GPCD where D=Demand, P=Population and GPCD (Galons per capita per day) =D/P

Logarithmic Mean Weight (w)

$$w = weight = \frac{D_{proj} - D_{act}}{\ln(D_{proj}) - \ln(D_{act})}$$

Population Contribution

$$E_{pop} = w \; x \; \ln\left(\frac{P_{proj}}{P_{act}}\right)$$

GPCD Contribution

$$E_{GPCD} = w \; x \; \ln\left(\frac{GPCD_{proj}}{GPCD_{act}}\right)$$

LMDI is a perfect decomposition method
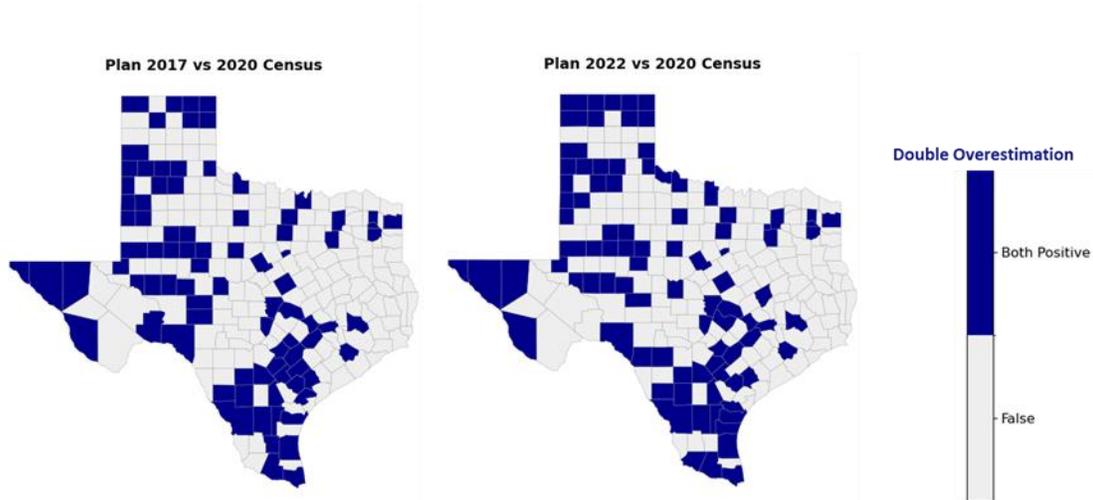
$$E_{pop} + E_{GPCD} = D_{proj} - D_{act}$$

Figure 11: Counties where both population and water demand projections show overestimation (double overestimation) compared to drought year 2011 water use. The same calculation was done for the difference with maximum water demand and also with actual 2020 demand.
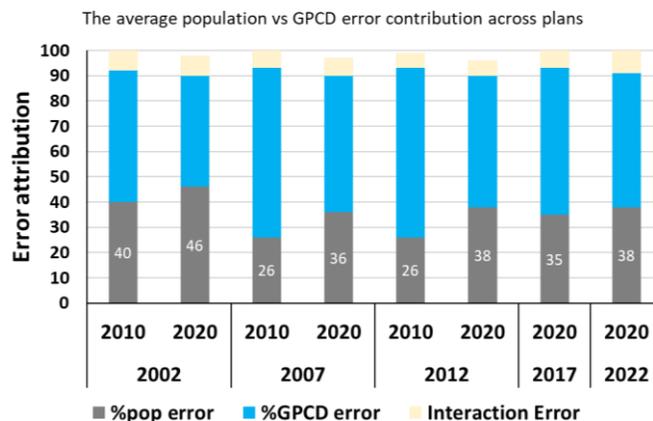
Figure 12: Population and GPCD error attribution across plans. We used refined Laspeyres decomposition error attribution. GPCD has the largest contribution and interaction errors are smaller than individual errors.
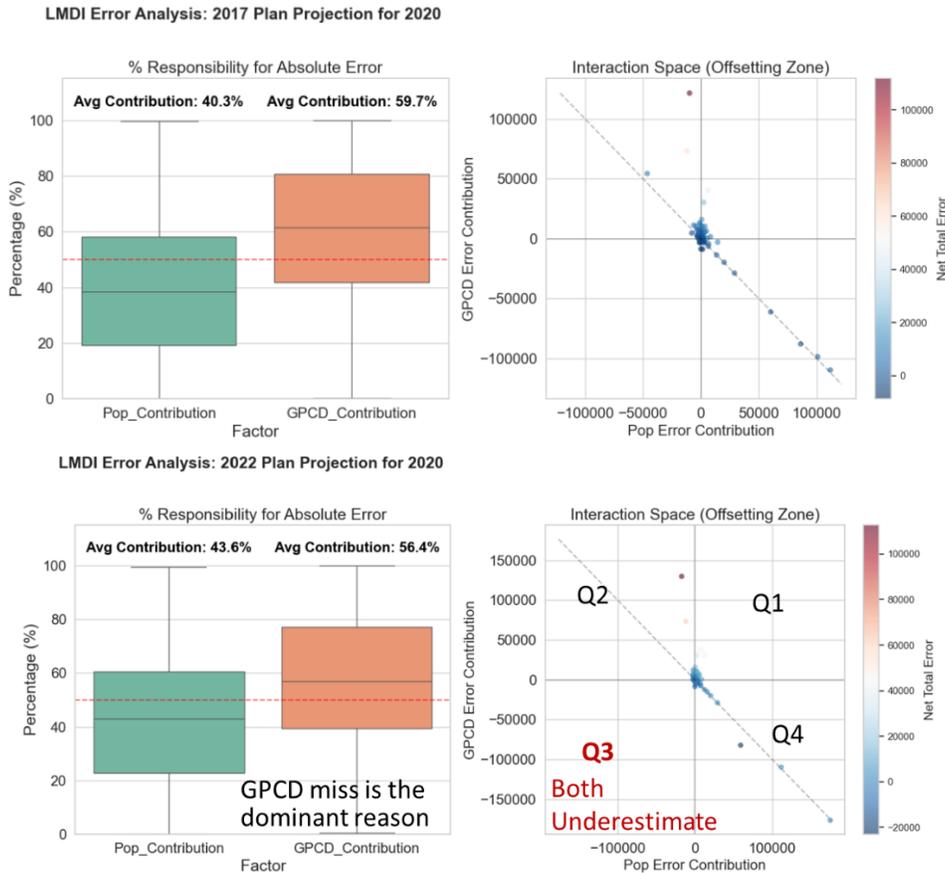

Figure 13: Population and GPCD error attribution using Logarithmic Mean Divisia Index (LMDI) Decomposition Analysis (2017 and 2022 projection for 2020). It is important to note that 2022 projections were finalized before 2020 in 2018 and did not include the 2020 census population. Percentage of responsibility (box plot): The GPCD holds a higher average responsibility for the absolute error at 56% (in 2022) indicating that in many specific WUGs, the "GPCD miss" was the dominant reason for the error, even if the "population miss" was larger when averaged across all WUGs. The wide boxes show high variability—some cities are strictly "population errors," while others are strictly "GPCD errors." Interaction Space (Right Plot): The Offsetting Zone: Most of the data points are clustered near the center, but several outliers follow that dashed diagonal line. Counteracting Errors: Points on that dashed line represent cases where a population over-projection was partially cancelled by a GPCD under-projection (or vice versa). High-Error Outliers: Q3 is the worst scenario, where both population and GPCD projections were underestimated. We have plotted these figures for each plan and projection year.

Subtask 4. Quantifiable margin of error translated to reasonable degree of uncertainty:

We calculated the margin of error (MoE) using mean absolute error (MAE) for the population for under- and overestimating RWPGs, counties, and WUGS (Figure 14). The MoE is smaller in overestimating RWPGs, counties, and WUGs compared to underestimating RWPGs, counties, and WUGs. There is a larger Mean Absolute Error (MAE) in underestimation (the scale of visualization is a clear indicator of this difference). The final projection resulted in reduced MAE compared to the draft in underestimating counties and WUGs. This is encouraging as we want to avoid underestimation in the final version of the projections. In overestimating WUGs and counties, the final results in a larger MAE. MAEs are larger in more distant (in time) projections. Recent plans show a reduction in MAE.
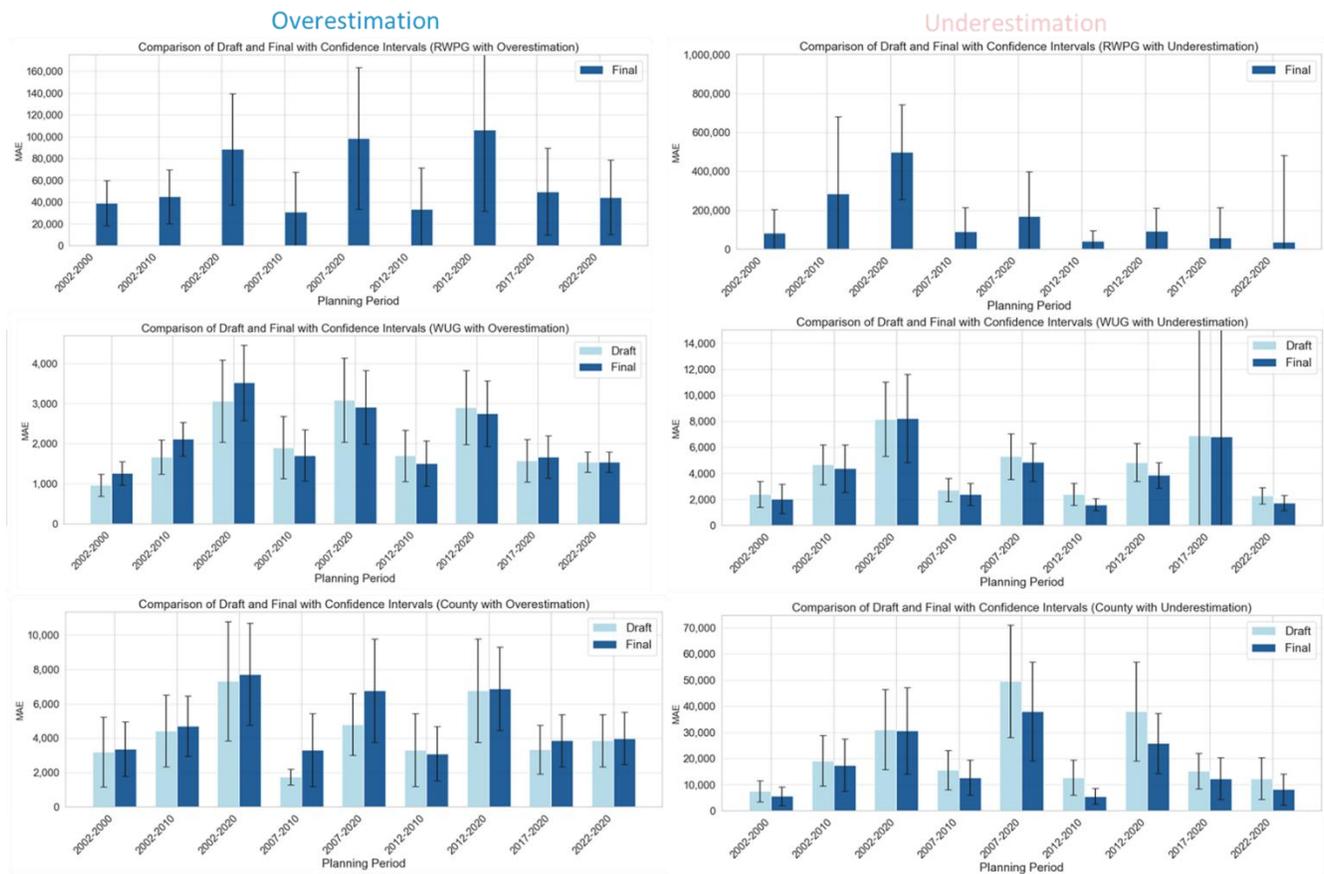
Figure 14: Margin of error and confidence interval at 95% for draft and final population projections against census (in overestimation and underestimation) across RWPG, county, and WUGs. The range lines in the MAE bars show the margin of errors. Note differences in Y scales.

Subtask 4. Recommendations to improve future accuracy of population projections.

Subtask 4 is 60% complete.

We have proposed a reasonable allowable difference in population projections based on nine population clusters (Figure 15 and Table 1). The analysis suggests 10%-30% differences would be permissible, accounting for the 75th percentile of differences across plans for the predefined nine clusters. We have also indicated the percentage of counties that fall within this threshold and if we consider the mean or median of the differences across plans and projections. mean/median/75th percentile of % differences are calculated from % differences with census over all planning periods (2002→2010, 2002→2020, 2007→2010, 2007→2020, 2012→2010, 2012→2020, 2017→2020, 2022→2020).
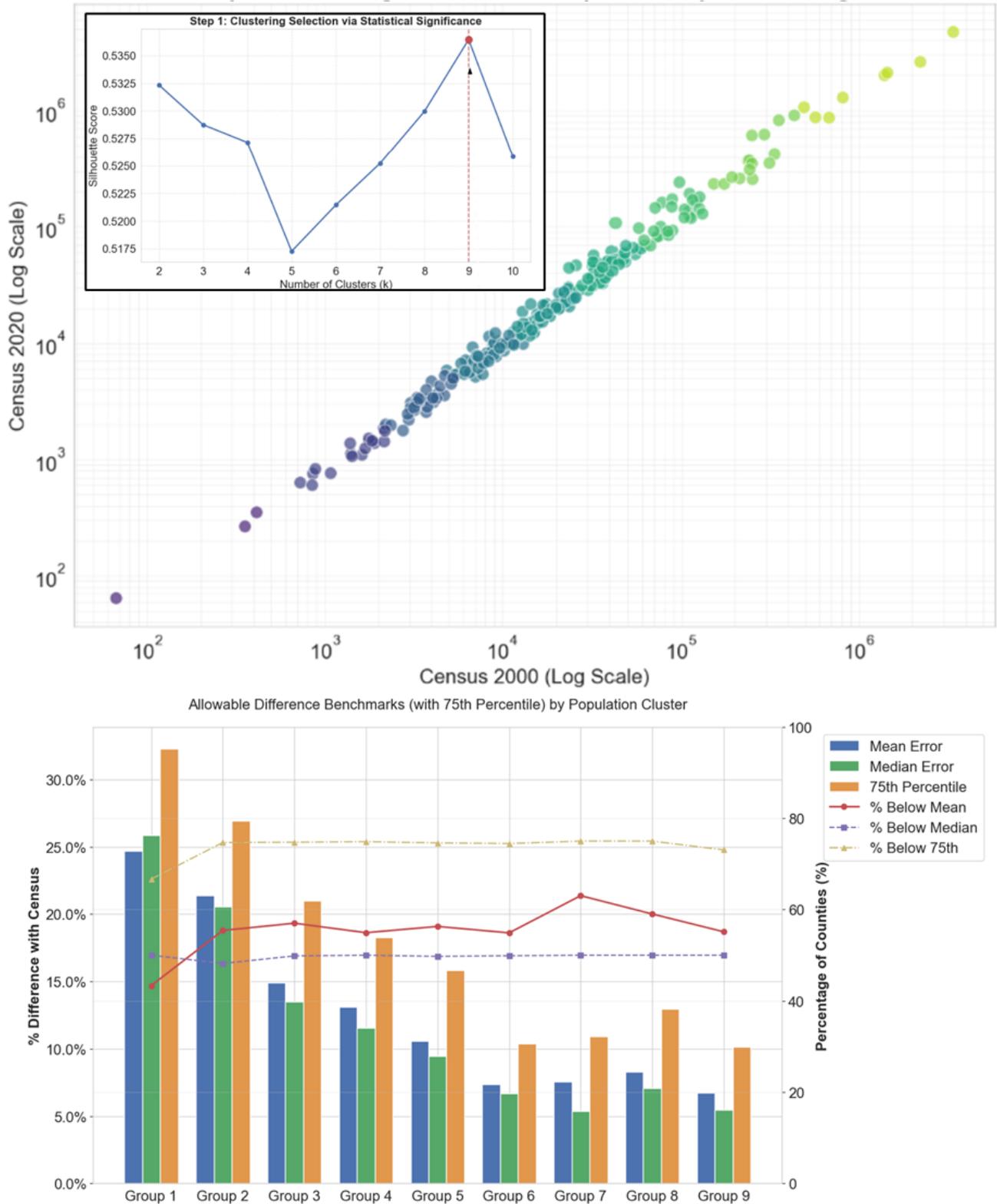
Figure 15: Allowable threshold calculation. We first use Silhouette scores to find statistically sound k-means clusters. Then the percentage difference from census was calculated for each cluster. We calculated the mean, median, and the 75th percentile based on the percent difference from census of each county in each cluster.

Table 1: Population ranges and percentage of difference with census along with the corresponding percent of counties that are considered with the defined cluster threshold. 67-75 % of counties in each cluster fall with in the 75th percentile of % difference with census. Outliers have been removed when calculating the 75th percentile in clusters 1, 6 and 9.

Cluster	Population range	75th Percentile Diff with Census	Mean Diff with Census	Median Diff with Census	% Counties Below Mean Diff.	% Counties Below Median Diff.	% Counties Below 75th Percentile Diff.
1	64 to 350	*32%*	25%	26%	43%	50%	**67%**
2	600 to 1,848	27%	21%	21%	55%	48%	**75%**
3	1,770 to 5,237	21%	15%	13%	57%	50%	**75%**
4	5,133 to 12,164	18%	13%	12%	55%	50%	**75%**
5	11,587 to 29,268	16%	11%	9%	56%	50%	**75%**
6	28,409 to 68,632	10%	7%	7%	55%	50%	74%
7	68,755 to 241,067	11%	8%	5%	63%	50%	**75%**
8	233,479 to 906,422	13%	8%	7%	59%	50%	**75%**
9	865,657 to 4,731,145	*10%*	7%	5%	55%	50%	73%