\*\*Comprehensive Methodology for Industrial Water Demand Analysis (Industrial Report Version)\*\*

The following methodology details the procedures systematically applied in the industrial water demand analysis, specifically tailored for an industrial report format:

1. \*\*Data Collection:\*\*

- Industrial water withdrawal, circular water use, and cooling water data were gathered from official Thuringian statistical sources for the years 2004, 2007, 2010, 2013, 2016, and 2019.

- Corresponding Gross Value Added (GVA) economic data were also collected.

2. \*\*Weather Data Integration:\*\*

- Obtained regional weather data (mean temperatures, summer temperature, and the frequency of hot days) from the German Weather Service (DWD).

- Given the limited availability of certain detailed weather variables (hot days, summer temperatures) from 2011 onward, analyses utilizing these specific variables were limited to the 2010-2019 period, thus ensuring the accuracy of results.

3. \*\*Data Preprocessing and Merging:\*\*

- Standardized datasets by replacing missing, non-numeric, and irrelevant administrative region data.

- Merged industrial water withdrawal data with weather and economic data using common keys ('Kreis' and year).

4. \*\*Variable Transformation and Calculation:\*\*

- Calculated the intensity of industrial water use by relating water withdrawals to the number of companies and economic output (GVA).

- Conducted logarithmic transformations of water withdrawal and GVA to normalize data distributions.

4. \*\*District Classification by Circular Water Use:\*\*

- Districts were categorized based on circular water usage intensity using the single-use to total water use ratio:

- Districts exceeding the 75th percentile threshold of this ratio were classified as "Low Circular Use" (indicating lower efficiency in water reuse).

- Districts below this threshold were classified as "High Circular Use."

- The purpose of this classification was to assess if districts characterized by higher single-use water intensity significantly impact total water withdrawals.

5. \*\*Econometric Modeling:\*\*

- Performed panel regression modeling employing only Fixed Effects (within) and Pooling methods, deliberately omitting Random Effects models to avoid cross-district interactions and instead focus exclusively on within-district variations.

- Separate regressions for "High Circular" and "Low Circular" district groups evaluated differential water withdrawal impacts.

6. \*\*Cooling Water Demand Modeling:\*\*

- Developed regression models analyzing the impact of various weather parameters (summer temperatures and hot days) individually to select the most suitable weather-related predictors for cooling water demand.

6. \*\*Visualization and Model Interpretation:\*\*

- Visualized data through time-series and stacked-area plots, including state-level aggregated trends and district-level detailed visualizations.

- Generated plots to visually assess the performance of regression models, comparing observed versus predicted water use and visualizing regression coefficients with confidence intervals for intuitive interpretation.

7. \*\*Technological Change Indicators:\*\*

- Initial attempts to calculate technological change (TC) using recycled wastewater data were unsuccessful due to negative and inconsistent values.

- Subsequently explored water intensity-based TC indicators but identified potential multicollinearity issues since water intensity is inherently linked to both dependent and independent variables, thus potentially biasing regression outcomes.

8. \*\*Cooling Water Usage Models:\*\*

- Analyzed multiple weather variables individually (mean summer temperatures, hot days above specific temperature thresholds) through separate panel regressions to identify the most relevant climatic predictor of cooling water usage.

9. \*\*Base Year and Data Standardization:\*\*

- Applied a base-year standardization approach, setting 2004 as the baseline to assess relative economic and water usage growth clearly over time.

10. \*\*Documentation and Reproducibility:\*\*

- Saved the final datasets and results systematically in structured formats (RDS and CSV files) to enable transparency and reproducibility.

- Structured regression outcomes into summary tables for easy reference and dissemination through Word documentation.