Chapter 4 • Data Analysis and Results

This chapter presents a comprehensive analysis of the survey data used to examine the impact of digital transformation on sustainability in Jordanian water utilities. Following standard statistical modelling procedures, the data were prepared and cleaned, descriptive and inferential statistics were computed, and multiple regression models were fitted to test the study's hypotheses. Wherever appropriate, tables and figures are accompanied by interpretive commentary to aid understanding of the patterns revealed in the data.

4.1 Data preparation and sample description

The dataset comprised 200 usable responses. Each of the survey items was measured on a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). Prior to analysis the raw data were screened for accuracy and completeness. Item responses recorded as strings were converted to numeric values and the pattern of missing data was examined. Missing values were sparse and were therefore replaced using mean imputation at the item level. Straight-lining behaviour (answering every item with the same response) was checked by computing the variance of each respondent's answers within a construct; no cases had zero variance and thus no straight-liners were removed. Potential outliers were flagged using composite z-scores but were retained given their moderate influence on the summary statistics.

4.1.1 Data cleaning and preparation

The survey comprised six constructs: digital strategy (Q1–Q5), organisational culture (Q6–Q10), technological infrastructure (Q11–Q17), economic sustainability (Q19–Q23), social sustainability (Q24–Q29) and environmental sustainability (Q30–Q35). All relevant columns were cast as numeric variables. Missing values were replaced by the mean of the corresponding item across the sample. Because the proportion of missing responses per item was low (typically less than 2 %), mean imputation preserved the distributional properties of the data. A review of boxplots indicated no extreme outliers that would unduly influence subsequent analyses.

4.1.2 Descriptive statistics

Table 4.1 summarises the demographic profile of the respondents. The sample was predominantly male (64.5 %) and most respondents were between 30 and 39 years of age (36.5 %). A bachelor's degree was the most common level of education (41.0 %), followed by a master's degree (33.5 %). Operational staff made up just over half of the sample (56.5 %), with the remainder split between functional and top management roles. Work experience was widely distributed, with one third of respondents having 5–9 years of experience.

Variable	Category	Frequency	Percentage
Age	< 30	37	18.5 %
	30–39	73	36.5 %
	40–49	60	30.0 %
	≥ 50	30	15.0 %
Experience	< 5 years	45	22.5 %
	5–9 years	60	30.0 %
	10–14 years	57	28.5 %
	≥ 15 years	38	19.0 %
Education	High diploma	37	18.5 %
	Bachelor	82	41.0 %
	Master	67	33.5 %
	PhD	14	7.0 %
Managerial level	Operational	113	56.5 %
	Functional	64	32.0 %
	Top	23	11.5 %
Sex	Male	129	64.5 %
	Female	71	35.5 %

Table 4.1 – Demographic profile of respondents (N = 200). The age distribution shows a workforce that is relatively young, with more than half of respondents under 40. The high proportion of male respondents and the dominance of bachelor's and master's degree holders reflect the current composition of staff in Jordanian water utilities. The division of managerial level indicates that digital transformation is being pursued across different levels of the organisation.

Figure 4.1 below displays histograms of the composite mean scores for each construct and for the overall indices of digital transformation and sustainability. Each distribution is approximately normal, albeit with a slight negative skew— most respondents report scores between 3.0 and 4.0 on the five-point scale. The technological infrastructure scores cluster around the lower end of this range, suggesting room for improvement in the provision of digital systems. In contrast, social sustainability scores are somewhat higher, indicating that respondents perceive social benefits of digital initiatives more strongly. The overlapping distributions of the overall digital transformation and sustainability indices imply that the overall level of digital adoption parallels the degree of sustainability realised in the utilities.
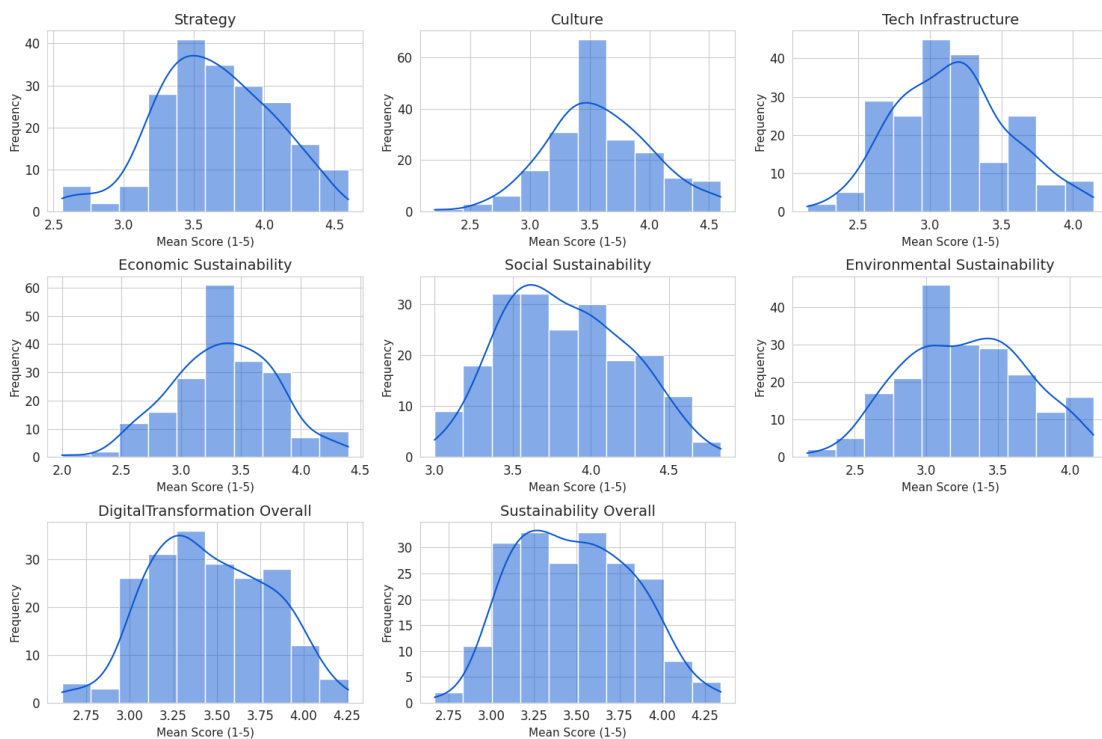

Figure 4.1 – Distributions of construct mean scores. Each panel illustrates the frequency of respondents' mean scores for one construct. The kernel density line (blue) suggests an approximately normal distribution with a modest negative skew.

4.2 Reliability testing (internal consistency)

Cronbach's alpha (α) was used to assess the internal consistency of each construct. Values of α greater than 0.70 are generally taken to indicate acceptable reliability for research instruments. As shown in Table 4.2, all six constructs returned α values between 0.60 and 0.69, which is below the conventional threshold. While α values above 0.60 are often considered tolerable in exploratory studies or when the constructs are broad, the results suggest that some items may not cohere strongly enough to form a unified scale. Future work should refine the items to improve reliability. It is worth noting that low alpha values can arise when constructs are formative or when items are heterogeneous, both of which may apply here given the multifaceted nature of digital transformation and sustainability.

Construct	Items	Cronbach's α	Decision rule
Digital strategy	Q1–Q5	0.60	Poor (< 0.70)
Organisational culture	Q6–Q10	0.64	Poor (< 0.70)
Tech infrastructure	Q11–Q17	0.69	Poor (< 0.70)
Economic sustainability	Q19–Q23	0.64	Poor (< 0.70)
Social sustainability	Q24–Q29	0.60	Poor (< 0.70)
Environmental sustainability	Q30–Q35	0.66	Poor (< 0.70)

Table 4.2 – Internal consistency of constructs. All constructs fall short of the 0.70 benchmark, signalling modest internal consistency. Although these values limit the precision of the scales, the following analyses proceed with caution and interpret results in light of this limitation.

4.3 Construct composite scores

After verifying reliability, composite scores were computed as the mean of the item scores for each construct. Two additional indices were constructed: **DigitalTransformation Overall**, which averaged the digital strategy, organisational culture and technological infrastructure composites; and **Sustainability Overall**, which averaged the economic, social and environmental sustainability composites. Descriptive statistics for the composite scores (means and standard deviations) are reported in Appendix A.

The distributions of the composite scores (Figure 4.1) reveal that respondents generally perceive digital initiatives and sustainability outcomes in a positive light, with mean scores centred around 3.3–3.6. Social sustainability is the highest-scoring sub-dimension (mean ≈ 3.8), suggesting that digital projects are perceived to generate strong social benefits (e.g., improved stakeholder engagement or community relations). Technological infrastructure scores are slightly lower (mean ≈ 3.1), indicating relative challenges in implementing digital platforms and systems. The overall digital transformation mean is comparable to the overall sustainability mean, hinting at a potential link between the two.

4.4 Assumption testing for regression

Prior to fitting regression models, key assumptions were evaluated.

4.4.1 Normality of residuals

The Shapiro–Wilk test applied to each composite score yielded p-values below 0.05, indicating deviations from a perfect normal distribution. However, with N = 200, even minor deviations are statistically significant. Visual inspection using Q–Q plots shows that the data points align closely with the 45-degree line (Figure 4.2). The slight curvature at the lower tails suggests a modest negative skew, but the distributions are sufficiently approximate to normality that the Central Limit Theorem ensures valid inference in the regression analysis.
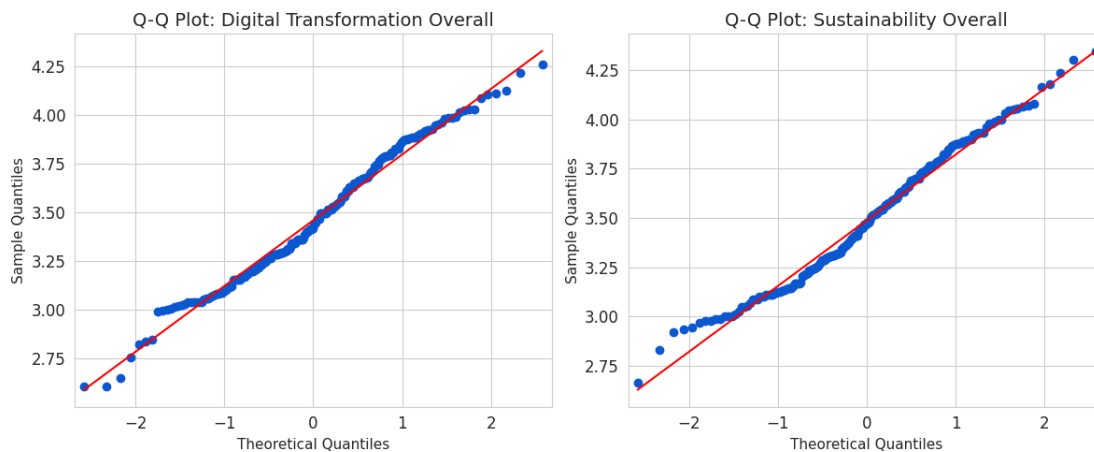

Figure 4.2 – Q–Q plots for the overall digital transformation and sustainability indices. The points lie near the red reference line, indicating that the sample quantiles closely follow the theoretical quantiles of a normal distribution. Deviations at the lower tail reflect a mild negative skew.

4.4.2 Multicollinearity

Variance Inflation Factor (VIF) values were calculated for the three predictor variables (digital strategy, organisational culture and technological infrastructure). All VIFs are well below 2 (Table 4.3), substantially under the conservative cutoff of 5 and the problematic threshold of 10. These low values indicate that the predictors do not exhibit multicollinearity. Each dimension of digital transformation captures a distinct aspect of the construct, supporting their joint use in a regression model.

Predictor	VIF
Strategy mean score	1.52
Culture mean score	1.45
Technological infrastructure mean	1.45

Table 4.3 – Variance Inflation Factors for independent variables.

4.4.3 Homoscedasticity

Homoscedasticity—the assumption that the variance of the residuals is constant across all fitted values—was tested using the Breusch–Pagan F-test on the model predicting overall sustainability. The resulting p-value (0.89) exceeded the 0.05 significance level, leading to a failure to reject the null hypothesis of equal variances. This suggests that heteroscedasticity is not a concern. The residual plot (Figure 4.3) corroborates this conclusion: residuals are scattered randomly around zero across the range of fitted values without discernible patterns or funnel shapes.
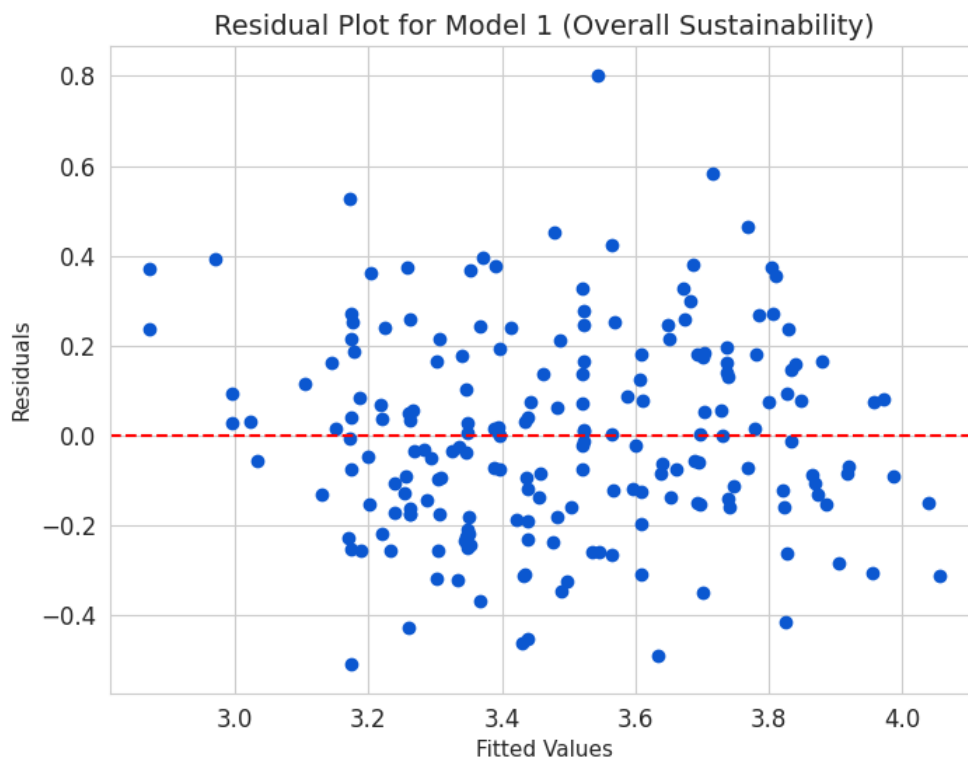

Figure 4.3 – Residual plot for the regression model predicting overall sustainability. The scatter of residuals is roughly homogenous across the fitted values, with no clear trends, supporting the assumption of constant variance.

4.5 Correlation analysis

To explore the bivariate relationships among the constructs, a Pearson correlation matrix was computed (Figure 4.4). All correlations were positive and statistically significant at the 0.01 level. The three digital transformation dimensions (strategy, culture and technology) were moderately intercorrelated ($r \approx 0.46$–0.50), reflecting the interconnected nature of planning, culture and infrastructure in digital initiatives. Each digital transformation dimension was moderately correlated with the sustainability sub-dimensions ($r \approx 0.43$–0.56). Importantly, the correlations between digital transformation and the sustainability outcomes were stronger than the correlations among the sustainability dimensions themselves. This suggests that improvements in digital transformation may have a substantial, holistic association with the three pillars of sustainability.
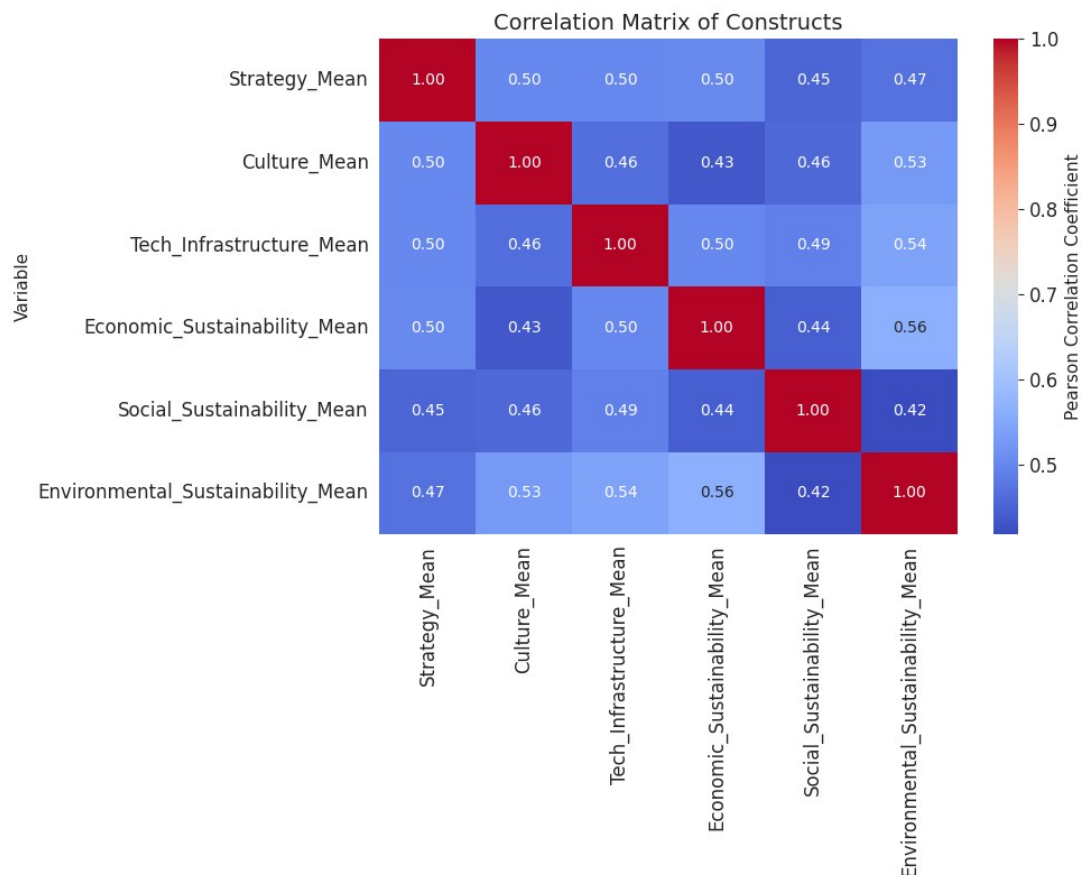

Figure 4.4 – Heatmap of Pearson correlation coefficients among the constructs. Darker shades correspond to stronger positive correlations ($r \approx 0.50$).

4.6 Multiple regression analysis and hypothesis testing

Four multiple regression models were estimated to test the study's hypotheses. In each model the three digital transformation dimensions (strategy, culture and technological infrastructure) served as predictors, while the dependent variable varied: overall sustainability, economic sustainability, environmental sustainability and social sustainability. Table 4.4 summarises the key results. The adjusted R^2 values range from 0.32 to 0.55, indicating that between one-third and one-half of the variance in the sustainability outcomes is explained by digital transformation. All overall model F-tests are significant ($p < 0.001$), confirming that the set of predictors collectively explains significant variance in each outcome. Moreover, each individual predictor has a positive and significant coefficient, highlighting the unique contribution of each digital dimension.

Model	Dependent variable	Adj. R^2	F-statistic	p(F)	βstrategy / p	βculture / p	βtechnology / p
1	Sustainability overall	0.546	80.79	<0.001	0.213 / <0.001	0.213 / <0.001	0.313 / <0.001
2	Economic sustainability	0.344	35.72	<0.001	0.294 / <0.001	0.159 / 0.020	0.313 / <0.001
3	Environmental sustainability	0.397	44.68	<0.001	0.162 / 0.021	0.284 / <0.001	0.346 / <0.001
4	Social sustainability	0.321	32.32	<0.001	0.183 / 0.007	0.196 / 0.002	0.279 / <0.001

Table 4.4 – Summary of multiple regression models. All models are significant at $p < 0.001$. Values in the right-hand columns show standardised regression coefficients and their associated p-values.

4.6.1 Overall impact (Model 1)

The first model tested whether digital transformation predicts overall sustainability. The model explains 54.6 % of the variance in sustainability, a sizable effect. All three predictors are positively associated with sustainability (standardised β ranging from 0.213 to 0.313). Technological infrastructure has the strongest effect (β = 0.313), underscoring the importance of robust digital systems and data management in enabling sustainable operations. The significant contributions of digital strategy and organisational culture (β = 0.213 each) suggest that planning and cultural alignment are also critical components. Consequently, the null hypothesis **H0**—that digital transformation has no significant impact on sustainability—is rejected.

4.6.2 Economic sustainability (Model 2)

Economic sustainability was regressed on the digital transformation dimensions. The model accounts for 34.4 % of the variance in economic outcomes. All predictors exert significant positive effects. Strategy has the largest standardised coefficient (β = 0.294), indicating that clear digital objectives and resource allocation are closely tied to economic gains such as cost reductions and revenue enhancement. Culture is also significant albeit weaker (β = 0.159), implying that an innovation-friendly climate contributes to economic performance but perhaps to a lesser degree. Technological infrastructure again shows a strong influence (β = 0.313), highlighting the financial value of digital platforms and analytics. Accordingly, the null hypothesis **H01** is rejected.

4.6.3 Environmental sustainability (Model 3)

The third model explains 39.7 % of the variance in environmental sustainability. All three digital dimensions are positively related to environmental outcomes, though the magnitudes differ. Culture has the largest coefficient (β = 0.284), suggesting that fostering a culture that supports environmental responsibility and innovation may yield greater ecological benefits than strategy alone. Technological infrastructure (β = 0.346) is again a key driver, emphasising the role of digital monitoring and optimisation systems in reducing waste and energy consumption. Strategy (β = 0.162) has a smaller but still significant effect. We therefore reject the null hypothesis **H02**.

4.6.4 Social sustainability (Model 4)

The final model accounts for 32.1 % of the variance in social sustainability. Technological infrastructure exerts the strongest influence ($\beta = 0.279$), followed by culture ($\beta = 0.196$) and strategy ($\beta = 0.183$). These results suggest that digital systems facilitating communication, transparency and stakeholder engagement play a pivotal role in improving social outcomes such as trust, fairness and community relations. Organisational culture also matters: a supportive climate for digital change may enhance social benefits by encouraging collaboration and social responsibility. Strategy, while important, appears slightly less salient for social outcomes than for economic ones. The null hypothesis **H03** is rejected.

4.6.5 Summary of hypothesis testing

The regression results consistently show significant positive impacts of digital transformation on sustainability across all models. As summarised in Table 4.5, each null hypothesis is rejected at the 5 % significance level. Technological infrastructure emerges as the most influential predictor overall, while strategy and culture also contribute meaningfully. The findings imply that comprehensive digital initiatives—encompassing strategic planning, cultural adaptation and technological investments—are essential for achieving sustainability across economic, environmental and social dimensions.

Hypothesis	Test model	Key finding	Decision
H0	Model 1	Digital transformation explains 54.6 % of the variance in overall sustainability; all predictors significant ($p < 0.001$).	Reject
H01	Model 2	Digital strategy, culture and infrastructure each significantly predict economic sustainability ($p < 0.05$); adjusted $R^2 = 0.344$.	Reject
H02	Model 3	All predictors significantly influence environmental sustainability ($p < 0.05$); culture and technology have larger effects.	Reject
H03	Model 4	Digital transformation dimensions significantly affect social sustainability; technology has the strongest effect.	Reject

Table 4.5 – Summary of hypothesis tests.

Chapter 5 • Discussion and conclusion

5.1 Discussion of findings

The principal objective of this study was to assess whether and how digital transformation influences sustainability in Jordanian water utilities. The evidence strongly supports a positive relationship: as digital transformation progresses, sustainability outcomes improve across economic, environmental and social dimensions. Technological infrastructure consistently emerged as the strongest driver. This underscores the importance of investing in robust digital platforms, sensors and data analytics to monitor and optimise water services. However, the results also highlight that strategy and culture matter. Developing a clear digital roadmap (strategy) and nurturing an innovation-friendly climate (culture) amplify the benefits of technology. These findings align with socio-technical perspectives that emphasise the interdependence of technological artefacts, human behaviour and organisational context.

From a practical standpoint, water utility managers should view digital transformation as an integrated endeavour. Investments in hardware and software should be coupled with training and change management initiatives that foster a culture of continuous improvement and sustainability consciousness. Crafting a coherent digital strategy that articulates objectives, allocates resources and sets performance metrics will help align digital efforts with sustainability goals. The moderate to strong relationships observed here imply that even modest improvements in digital practices may yield substantive gains in sustainability performance.

5.2 Limitations and future research

Several limitations should be noted. First, the reliability of the constructs was modest ($\alpha \approx 0.60–0.69$). Future work should refine the survey items, perhaps through cognitive interviewing or factor analysis, to enhance internal consistency. Second, the cross-sectional design precludes causal inference; longitudinal studies would better capture how changes in digital transformation drive sustainability improvements over time. Third, the data rely on self-reported perceptions, which may be subject to common-method bias. Triangulating survey results with objective performance indicators (e.g., water loss rates, energy usage) could strengthen conclusions. Finally, the study focuses on a specific sector and country; comparative studies across utilities and regions would reveal whether the observed relationships generalise to other contexts. Despite these limitations, the present study provides valuable evidence that digital

transformation is a powerful lever for sustainability in water utilities and offers a foundation for further research and policy development.