Utility Regulator
ELECTRICITY GAS WATER

Water & Sewerage Services
Price Control 2015-21

Draft Determination - Annex J

Capital Maintenance Modelling

July 2014



Contents page

Water and Sewerage Services Price Control 2015-21

Annex J – Capital Maintenance Modelling

Appendices

Appendix 1	Updated PC10 Cost Models
Appendix 2	CSV Sensitivity Modelling

1.0 Introduction

1.1. Problem

1.1.1 Setting the level of capital maintenance (CM) spend is a key element of the final determination. The problem is knowing what the correct level should be.

1.1.2 The difficulty facing NI Water is a lack of robust data. It can take a number of years to see the impact of CM decisions on serviceability levels. In England and Wales (E&W), this would typically involve analysis of 10 years of cost data against 5 years of serviceability trends.

1.1.3 As yet, the company does not possess this detail. It takes time to build a store of knowledge and the correct process in place to capture and use this data.

1.1.4 Whilst the company has collected better financial data since 2007, concerns over capital expenditure still abound. These concerns result from the fact that:

a. NI Water had a backlog base maintenance allowance in the SBP (Strategic Business Plan) era;

b. At present, the company is incentivised to spend to budget. This is a problem as NI Water has less reason to reduce spend and reveal a lower level of efficient costs;

c. Year-end constraints imposed by the governance model often result in accelerated CM spend. This can mean that the minimum level of investment is not revealed; and

d. Due to the variations year on year of the capital maintenance budget, the historic levels cannot be used as a robust indication of requirements going forward.

1.1.5 The company is developing models that will improve their understanding of CM needs going forward.

1.1.6 In the absence of such an analysis, the Utility Regulator (UR) must form an opinion on what it considers to be an appropriate level of spend.

1.2. Proposed solution

1.2.1 The absence of good data has prompted the consideration of alternate options. For the most part, these options consist of using comparative company figures. The basis of the UR's approach is to look at levels of maintenance spend in England and Wales based on assets and their usage.

1.2.2 The purpose of this analysis is to determine if a relationship exists between costs and assets. This cost relationship is then applied to NI Water assets to determine predicted levels of maintenance spend.

1.2.3 The benefit of this high-level approach is that it is based on robust cost and explanatory data from well-established utilities. This should enable the UR to construct an estimate of the magnitude of required maintenance for NI Water.

1.2.4 The weakness of this approach is that it provides little detail on certain key aspects of asset maintenance e.g. state of the assets, where money should be spent etc.

1.2.5 The analysis should however provide a reasonable order of magnitude for future capital maintenance plans.

1.3. Models

1.3.1 Varieties of different model techniques are used. These range from regressions to unit rates and cost analysis. This report focuses on five different approaches:

a. **Update of the PC10 models** – This involves updating the econometric and unit cost models used at PC10;

b. **Total capex (tapex) regression using composite variables –** This method combines explanatory variables into a composite measure in order to estimate total capital maintenance;

c. **Tapex regression using density variables –** These models use a proxy density variable to establish predicted costs;

d. **Unit cost models** – Unit rates from England and Wales are established and applied to NI Water data; and

e. **Historical cost analysis** – Predicting future expenditure based on past spending decisions.

1.3.2 The regressions use a six-year average capex spend figure from 2005 to 2010. Figures have then been uplifted to 2012-13 prices[1] using COPI (Construction Output Price Index). Average expenditure is preferred in order to allow for the 'lumpy' profile of capital costs.

1.3.3 Independent variable data is taken from the 2010-11 June Return for the most part.

1.3.4 Each model provides an estimation of both average and upper quartile CM costs. The average refers to the industry mean. The upper quartile is reflective of the performance of the third (out of ten) ranked Water and Sewerage Company (WaSC).

[1] All figures in this annex are given in 2012-13 prices unless otherwise stated.

1.3.5 An alternate upper quartile is also calculated from an average of the second, third and fourth ranked WaSCs.

1.3.6 The exception is the historic cost analysis. This does not use England and Wales data to predict NI Water spend. Rather, the focus here is on predicting future costs based on historic decisions.

1.3.7 The various merits and drawbacks of each method will be discussed in the relevant section.

1.4. Other adjustments

1.4.1 The basic premise of each of the methods is to assess what England and Wales comparators would spend given NI Water assets. The problem with this approach is that there are certain factors that make Northern Ireland unique.

1.4.2 These factors need to be taken account of. To ignore them would result in either an over or under estimate of required spend.

1.4.3 For NI Water, two of the most significant factors include:

a. Regional Price Adjustment (RPA); and

b. Public Private Partnerships (PPP's).

Regional Price Adjustment

1.4.4 Northern Ireland is known to be a low cost region in terms of wages and some construction materials. The RPA is required to reflect the difference in capital costs in different regions of the UK.

1.4.5 The UR has contracted CEPA[2] to undertake a project to establish what the differential should be. Their findings suggest that general construction projects on average (including nationally procured and regional inputs) are roughly 6% cheaper in Northern Ireland.

1.4.6 Therefore, to take proper account of NI Water operating conditions, all average CM estimates derived from the models will be multiplied by 0.94.

1.4.7 The frontier companies tend to be in low cost regions as well (e.g. Yorkshire and Anglia). As a result, a 6% RPA change would not be appropriate for upper quartile estimates. The UR has chosen a 2% shift for these predictions.

1.4.8 The historic cost analysis is exempt from this adjustment. These costs are already based on Northern Ireland spend, so require no regional alteration.

[2] CEPA = Cambridge Economic Policy Associates Ltd – Annex M – Regional Price Adjustment

Public Private Partnerships

1.4.9 Another key difference for NI Water is PPP's. Unlike in England and Wales, a number of NI Water assets are run and maintained by private operators. These operators are paid via the opex unitary charge.

1.4.10 The models predict average costs based on entire company data, including PPP's. Since the private operators are responsible for this plant, it would be double counting to provide NI Water with any funds for their upkeep.

1.4.11 Using information on the construction value of the PPP non-infrastructure assets and an assumed weighted average asset life of 39 years, we have estimated an annual current cost depreciation value for the assets of £7.6m.

1.4.12 NI Water has argued that it would be unreasonable to apply this long-term average to assets that were constructed recently and include a substantive proportion of long life elements that will not be replaced in the near future. However, the Alpha and Omega assets were commissioned between 2008 and 2010 and Kinnegar WWTW was commissioned in 2004.

1.4.13 All plant will be over 10 years old by the end of PC15. NI Water's expert assessment of water and wastewater non-infrastructure asset maintenance concluded that the life of mechanical and electrical (M&E) plant is between 10 and 15 years and short life assets have a life less than 10 years.

1.4.14 Taking a cautious approach at this stage, we have assumed that 10% of M&E asset value will be replaced within PC15 and 35% of short life asset value will be replaced within PC15. This arrives at a PPP adjustment of £3.5m for PPP at the average and £2.5m at the upper quartile.

1.4.15 Findings presented for each of the alternate models tend to have the RPA and PPP adjustment included. This gives an appropriate view of what each method predicts for NI Water.

1.4.16 No specific PPP adjustment has been made for the historic cost analysis. Some of the early historic figures will include all NI Water maintenance spend. The last few years will exclude the PPP money now accounted as opex.

1.4.17 The UR has considered various methods to allow for this anomaly.

2.0 PC10 Models Updated

2.1. Rationale

2.1.1 The starting point is a rework of the PC10 capital maintenance analysis. This consisted of using the original Ofwat CM models that were discontinued in 2006-07. The models are a mix of regressions and unit costs, which provide predicted spend for different areas of the business.

2.1.2 The UR has used the same format as the original equations. The analysis has however been updated with the most recent available cost and explanatory variable data.

2.1.3 A problem with this approach concerns availability of data. Some of the variables are not collected as part of the June Return. Instead, these inputs are collected separately as part of the Capital Maintenance Econometric Return (CMER).

2.1.4 Since the models were discontinued, some data has not been collected. Without any better data, the UR has had to use variable information from the last CMER. As a result, the reliability of the findings for some of the models is brought into question.

2.1.5 In spite of this issue, the analysis has been updated. There is still some merit in the new models as most are unaffected by data problems. Unlike the other methods, they also have the advantage of looking at specific areas of spend in more detail, rather than just a total cost analysis.

2.2. Models

2.2.1 The models consist of a mixture of regressions and unit costs as shown below:

Table 2.1 – Water service models

Functional Area	Model Type	Explanatory Variables
Water Distribution Infrastructure	Log regression	Connected properties per length of main
Water Distribution Non-infrastructure	Log regression	Service reservoir and water tower capacity per pumping station capacity
Water Management & General	Log regression	Proportion of billed non-household properties
Water Resource & Treatment	Unit cost	Total connected properties

2.2.2 The sewage models are constructed as follows:

Table 2.2 – Sewerage service models

Functional Area	Model Type	Explanatory Variables
Sewerage Infrastructure	Log regression	Number of CSO's[3] per length of sewer
Sewerage Treatment	Log regression	Total number of works divided by the total load received at works
Sewerage Non-infrastructure	Unit cost	Total number of pumping stations
Sludge Treatment and Disposal	Unit cost	Weight of dry solids disposed
Sewerage Management and General	Unit cost	Number of billed properties

2.2.3 A more in-depth discussion of these models is included in Appendix 1 of this paper.

2.3. Findings

2.3.1 Using the model findings it is possible to estimate the average CM cost that would be expected of a company with NI Water assets. The results include a smearing[4] adjustment and upper quartile performance. The table also reflects the findings adjusted for PPP and RPA.

Table 2.3 – Predicted capital maintenance costs for NI Water using PC10 models (2012-13 prices)

Functional Area	NI Water Actual Spend (£m)	Average Predicted CM Spend (£m)	Upper Quartile Predicted CM Spend (£m)	Alternate Upper Quartile Predicted CM (£m)
Water	43.53	36.21	36.73	36.18
Sewage	50.04	44.57	37.30	38.19
Total	**93.57**[5]	**80.78**	**74.02**	**74.36**

Figures include an adjustment for PPP and regional prices

[3] Combined sewer overflows
[4] The smearing adjustment is explained in: http://www.uregni.gov.uk/uploads/publications/PC10_NIAUR_FD_Feb_10_-_Doc04_-_Annex_B_Cap_Maintenance_Analysis.pdf
[5] NI Water average spend figures are likely to be overestimated as they include backlog base maintenance in the early years of the SBP.

2.3.2 The results predict average maintenance spend of £81m per annum. The upper quartile expenditure is considerably lower. This is mostly due to the sewage models where the third ranked company had considerably lower spend (circa 20%) than the average.

2.3.3 The table also includes an alternate upper quartile figure. This is based on the second, third and fourth ranked WaSC's. Ultimately, there is not much of a difference in the results.

2.3.4 These model results are somewhat unreliable. A combination of poor regressions and old data make any conclusions difficult.

3.0 Composite Variables

3.1. Rationale

3.1.1 The PC10 approach split costs into different areas and has specific explanatory variables for each model. An alternate option is to model total capital maintenance by service area. The total capex (or tapex) regressions provide high-level views of CM spend using key independent variables.

3.1.2 The problem with such an approach is that there may be any number of variables that effect capital spend. There is however only a limited amount of water companies from which observations can be derived. Too many explanatory variables and too few observations can lead to ‘over-fitting’ of models and unrealistic cost predictions.

3.1.3 Problems can also arise if the independent variables are correlated with each other. This can result in errors in the models, incorrect coefficients and can magnify any other bias that might exist in the model.

3.1.4 In order to try to overcome these issues, a composite measure is constructed. This composite scale variable (CSV) is so called as it creates a new scale, which comprises the impact of its component parts into a combined value.

3.1.5 In order to generate a CSV, it must first be decided what variables will be included. These variables are then weighted in order of the importance placed on them. The sum of the weighting will add to one. Finally the weighted variables are multiplied together to create the new CSV.

3.2. Models

3.2.1 The tapex models are split by water and sewage service area. The choice of potential variables is wide.

Table 3.1 – Potential explanatory variable for CSV models

Water Explanatory Variables	Sewage Explanatory Variables
Length of mains	Length of sewers
Distribution input	Load received at treatment works
Connected properties	Sludge disposed
Reservoir capacity	Billed properties
Billed properties	Population equivalent served
Bursts	Number of WWTW’s
MEAV	MEAV

3.2.2 Whilst all the variables may have some impact on CM, two have been chosen here:

a. Main/sewer length; and

b. Billed properties.

3.2.3 Network length is obviously a key factor in determining the level of CM expenditure. Billed properties is also chosen as the variable represents connections, usage and company size.

3.2.4 Other relevant factors are not included for a variety of reasons. These include:

a. Out of date information – e.g. reservoir capacity;

b. No data – e.g. Mean equivalent asset values; and

c. Highly correlated variables measuring the same impact – e.g. billed properties and connected properties or population equivalent served.

3.2.5 With the variables chosen, the other issue that remains is to weight them according to impact. In the absence of better information, one method is to weight the variable based on England and Wales proportion of spend. For the water model mains length (water infrastructure) has a 45% weighting.

3.2.6 For the sewer model, the proportion of direct expense on infrastructure is smaller. Sewer length therefore has a 25% weighting. The form of the model variables is:

Water CSV = (Mains length ^ 0.45)*(Billed properties ^ 0.55)

Sewage CSV = (Sewer length ^ 0.25)*(Billed properties ^ 0.75)

3.2.7 Both the selection of the variables and their weighting is open to argument. By way of a sensitivity check, a number of alternate models using different weightings have been completed.

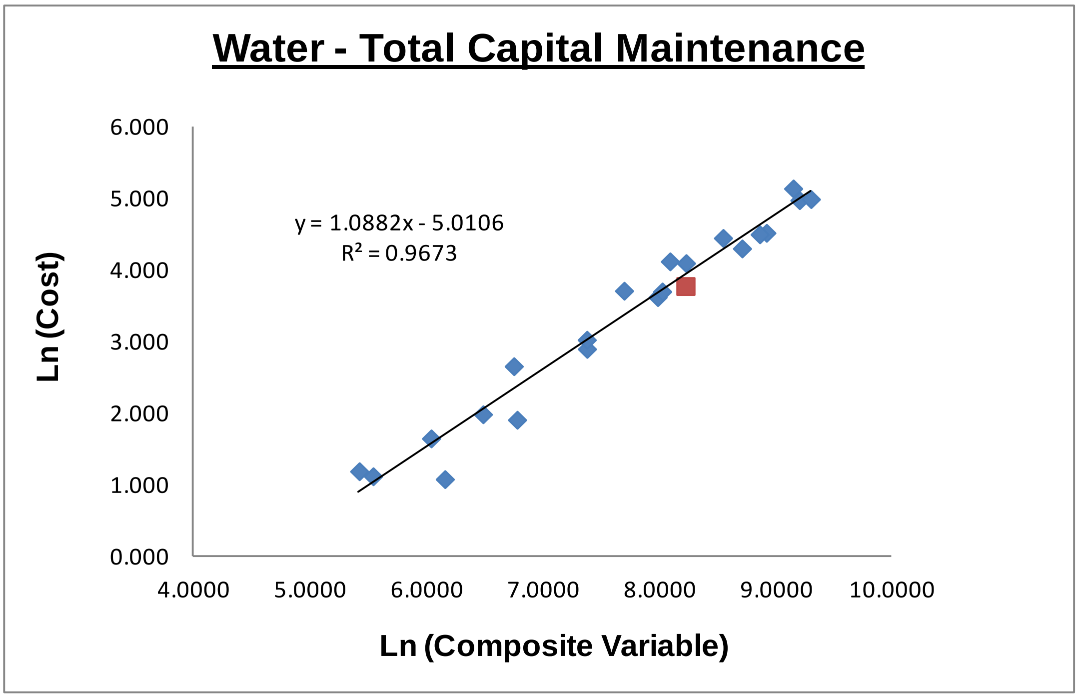
3.2.8 The form of this baseline water model is:

Table 3.2 – Tapex regression for water

Water Service:	**Water Tapex Model**	
Modelled cost:	Ln (average total water capital maintenance spend [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-5.011	0.356
Ln (csv)	1.088	0.046
Form of Model:	Ln (modelled cost) = -5.011 + 1.088 * ln {csv}	
Statistical Indicators:	Number of observations = 21	R^2 = 0.967
	Model standard error = 0.254	F test = 0.000

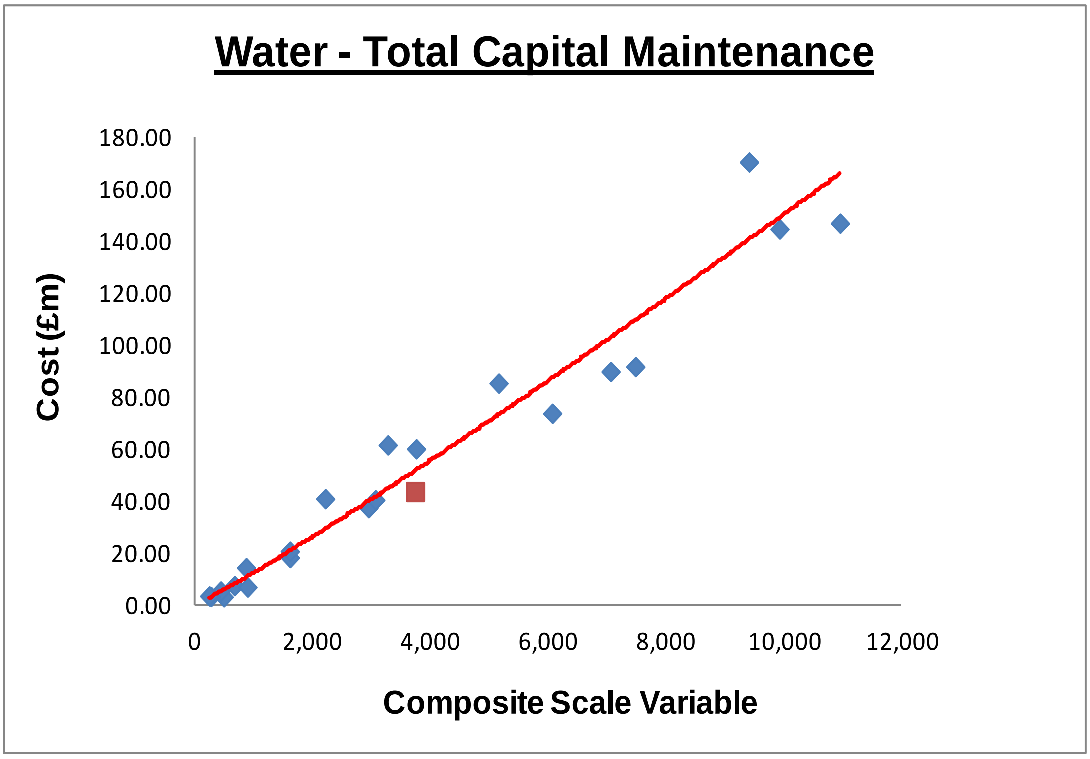
3.2.9 In graphic form, the model looks as follows.

Figure 3.1 – Tapex regression for water



3.2.10 The chart shows the data in log format. The red dot represents NI Water. In real terms, the analysis illustrates a clear relationship with predicted costs represented by the red line.

Figure 3.2 – CSV model - predicted versus actual costs



3.2.11 The statistics show a high degree of correlation between costs and the explanatory variable. The R^2 stat suggests that the model does not suffer from omitted variable bias.

3.2.12 The CSV coefficient is positive and greater than one. This means that maintenance costs increase as mains length and billed properties rise. In a log model, the coefficient is interpreted as the percentage change in cost given a percentage increase in the predictor variable.

3.2.13 In this case, a 5% rise in the CSV would lead to a 5.45% increase in predicted capital maintenance. This is calculated as follows:

1.05 ^ 1.0882 = 1.0545 = 5.45%

3.2.14 If billed properties increased by 5% but there was no change in the network length, this would result in a 2.96% cost rise.

1.05 ^ 0.55 = 1.0272 ^ 1.0882 = 1.0296 = 2.96%

Sewage CSV model

3.2.15 The sewage model results are shown below:

Table 3.3 – Tapex regression for sewage

Sewage Service:	**Sewage Tapex Model**	
Modelled cost:	Ln (average total sewage CM spend [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-2.088	0.965
Ln (csv)	0.818	0.117
Form of Model:	Ln (modelled cost) = -2.088 + 0.818 * ln {csv}	
Statistical Indicators:	Number of observations = 10	R^2 = 0.859
	Model standard error = 0.220	F test = 0.000

3.2.16 In graphic form, the model appears a good fit for the data. In this case, the chart shows the line of best fit with 95% confidence intervals. NI Water is represented by the red marker.

Figure 3.3 – Tapex regression for sewage

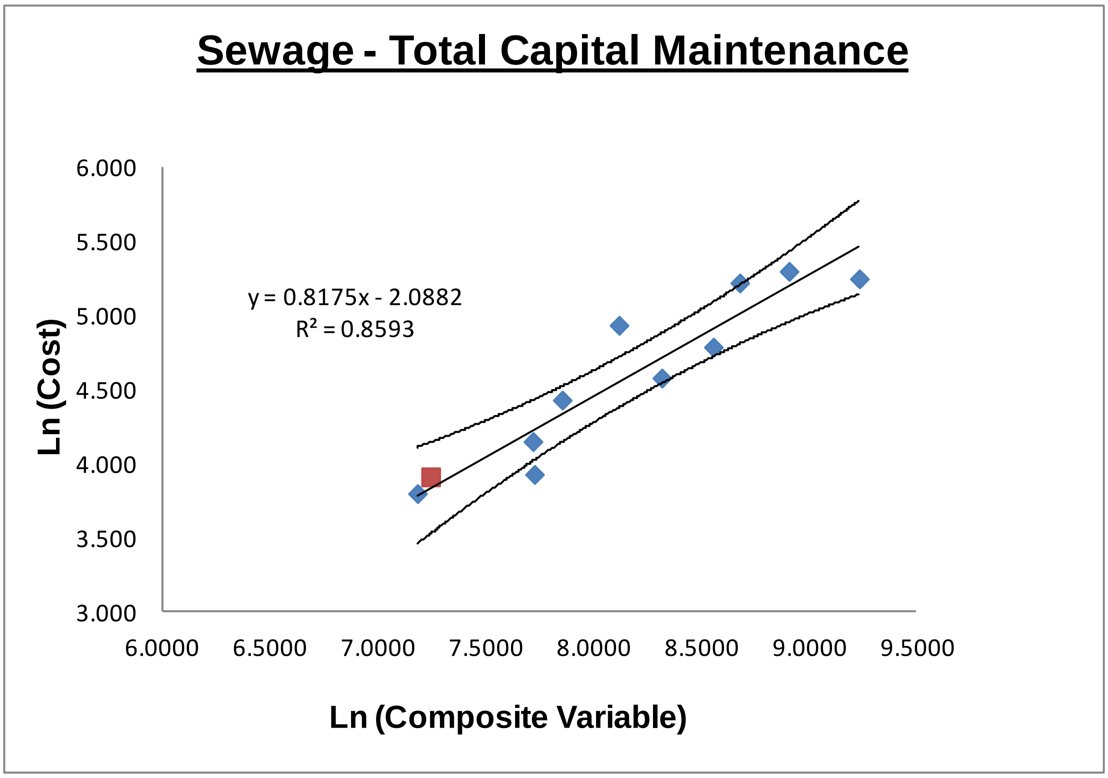
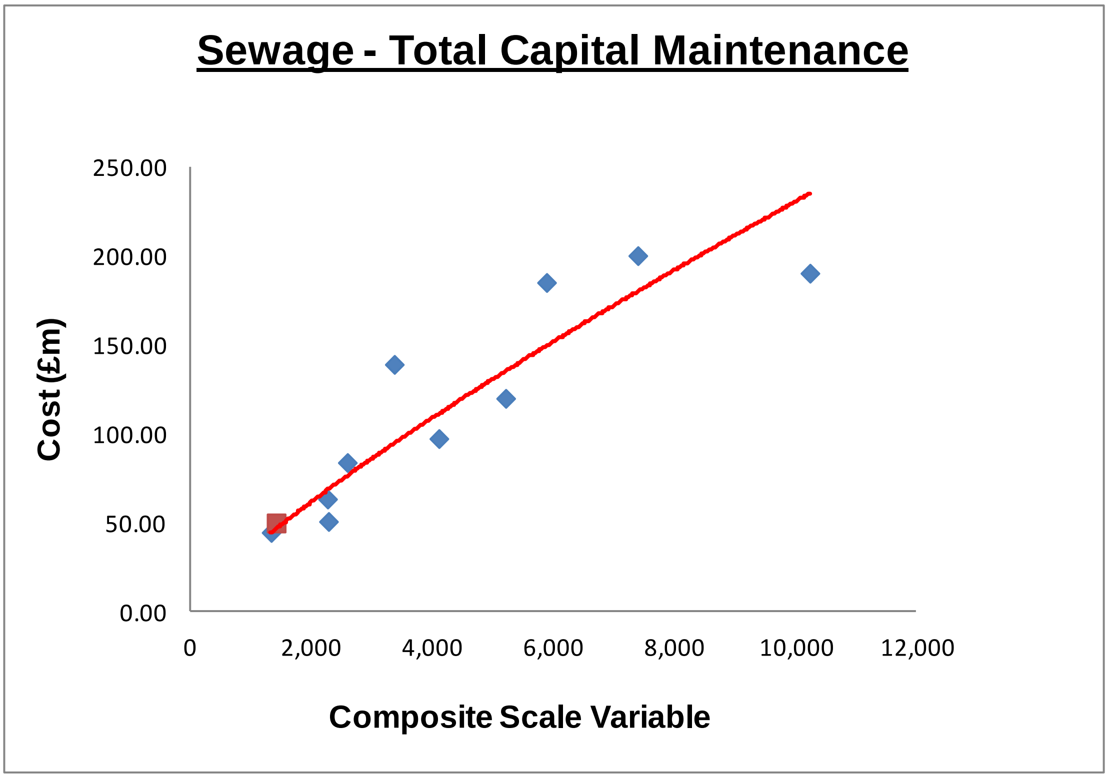


Figure 3.4 – CSV model - predicted versus actual costs



3.2.17 Figure 3.4 shows each companies position with predicted costs from the regression represented by the red line. The statistical results for this model are encouraging. The variable is strongly significant and the predictor appears to have a strong relationship with CM costs.

3.2.18 The coefficient is positive so has a direct relationship with costs. The coefficient means that capital maintenance spend will rise less than the size of the increase to the business.

3.2.19 For instance, in this case a 5% increase in sewer length and billed properties will result in a predicted capital maintenance increase of 4.07%.

$$1.05 \wedge 0.818 = 1.0407 = 4.07\%$$

3.2.20 There are only 10 observations for the sewerage model. Normally this would be on the low side. However, there is only one independent variable and the model appears to fit the data quite well.

3.3. Findings

3.3.1 Both the choice of variables and their weightings have assumptions impacting on results. The results of the two regressions are however positive.

3.3.2 The data appears to fit the models well. The R^2 statistics appears to show that the combined variable explains most of the cost movement. This indicates that the regressions do not suffer from omitted variable bias.

3.3.3 Whilst ideally there would be more observations on the sewage side, the results are good here as well.

3.3.4 In terms of predicted CM cost, the CSV models give the following results:

Table 3.4 – CSV regression results (2012-13 prices)

Functional Area	NI Water Actual Spend (£m)	Average Predicted CM Spend (£m)	Upper Quartile Predicted CM Spend (£m)	Alternate Upper Quartile Predicted CM (£m)
Water CSV Model	43.53	47.92	42.74	42.68
Sewage CSV Model	50.04	42.69	38.54	37.22
Total	**93.57**	**90.62**	**81.28**	**79.90**

Figures include an adjustment for PPP and regional prices

3.3.5 The findings suggest average predicted spend in the region of £91m or closer to £80m per annum at the upper quartile.

3.3.6 Whilst this represents the base case, it is recognised that there is no definitive selection method for either variables or their weights.

3.3.7 In order to test the variability of the findings to changes in the modelling, a number of sensitivity tests were run. These tests included the following:

a. CSV Model A – Base case;

b. CSV Model B – Network length / Billed Properties (50%:50% weight);

c. CSV Model C – Network length / Billed Properties (75%:25% weight);

d. CSV Model D – Network length / Billed Properties (25%:75% weight);

e. CSV Model E – Network length / Billed Properties / Usage (33%:33%:33% weighting); and

f. CSV Model F – Network length / Billed Properties / Usage (30%:50%:20% weighting).

3.3.8 The sensitivity tests vary both the weights and the variables used. Results for the models show quite large deviations from the base case.

Table 3.5 – CSV sensitivity testing results (2012-13 prices)

Functional Area	NI Water Actual Spend (£m)	Average Predicted CM Spend (£m)	Upper Quartile Predicted CM Spend (£m)	Alternate Upper Quartile Predicted CM (£m)
CSV Model A	93.57	90.62	81.28	79.90
CSV Model B	93.57	100.19	88.58	88.33
CSV Model C	93.57	122.82	108.42	108.70
CSV Model D	93.57	81.78	73.72	72.49
CSV Model E	93.57	96.71	86.45	86.71
CSV Model F	93.57	90.47	81.27	80.73

Figures include an adjustment for PPP and regional prices

3.3.9 All the models give good results in terms of statistical significance. There is however, a reasonable level of divergence in predicted costs.

3.3.10 In order to estimate which is the most robust regression, a couple of indicators are used. The first is the coefficient of determination (R^2), which is a goodness-of-fit measure. The higher the R^2 the better the data fit, although increasing the numbers of independent variables tends to increase the R^2 value.

3.3.11 The second is the standard error (root mean squared error) which is a measure of the dispersion of the forecast values from actual observations. The lower the error the closer the forecasts are to actuals.

3.3.12 Using both methods, Models A, D and F prove to be the most robust. Model C is the least robust of the options.

3.3.13 Each of the models and their findings are set out in further detail in Appendix 2.

4.0 Density Variables

4.1. Rationale

4.1.1 The next option is to model costs based on population density. It is anticipated that the type of operating environment will have a major impact on companies CM costs.

4.1.2 Population density will affect the number of assets a company has, their usage, size and ultimately cost.

4.1.3 Northern Ireland has a dispersed population. While this may prove a difficulty in terms of building a cost efficient network, the case is not so clear when it comes to maintaining it. The Independent Water Review Panel recognized this fact when it stated,

> *"Northern Ireland is characterised by a dispersed and large rural population.......This clearly impacts on costs/capital investment levels as it is much more cost effective to construct a network which will serve the needs of a dense population than of a dispersed population.*
>
> *However, whilst construction costs will be higher maintenance costs may be lower where the network serves a dispersed population as it is often more costly to undertake maintenance in an urban environment and higher levels of maintenance may be required."*[6]

4.1.4 The UR has attempted to construct models with density as a predictor variable to test this conclusion. Whilst no specific population density figures are available by company, proxy variables can be used.

4.1.5 Such an analysis is further justified as population density has been used previously to determine CM in historic Ofwat models.

4.2. Models

4.2.1 Like the CSV modelling, these regressions focus on each service area separately. For water, the proxy variable for density is properties per main. This mirrors the variable previously used in the water infrastructure model.

4.2.2 Regressing this variable against total water capital maintenance does not give a statistically reliable model. This means an absence of robust conclusions. However, the coefficient is positive, suggesting that costs will rise as density increases.

---

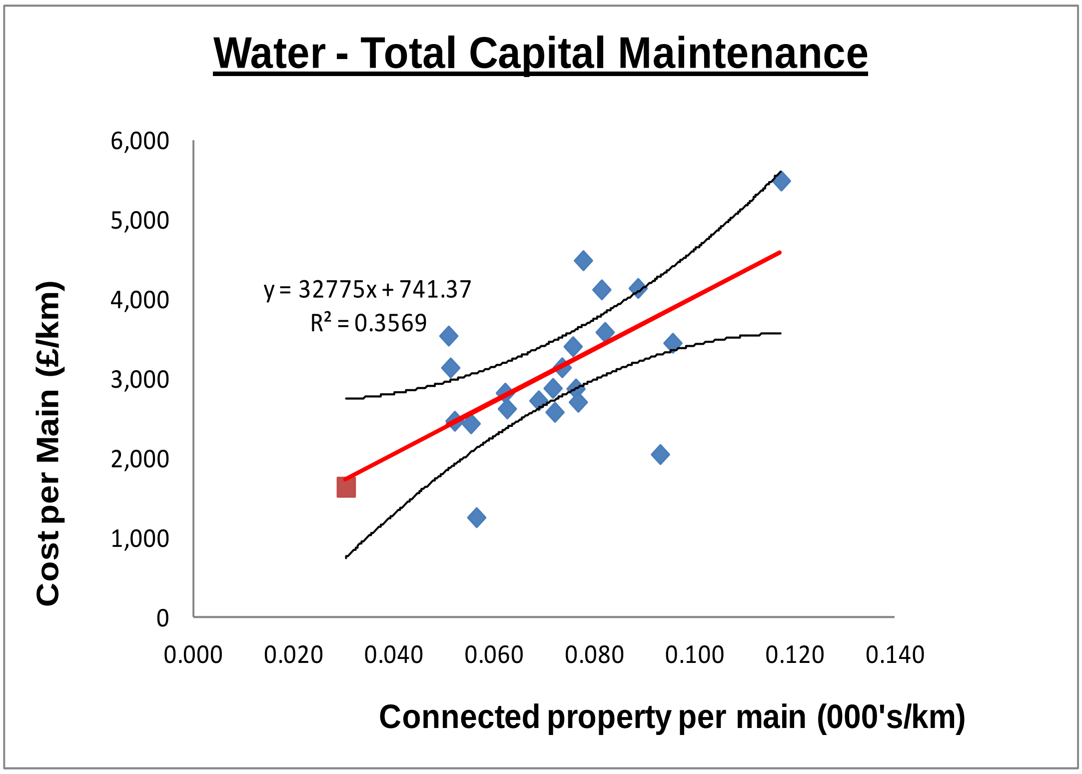
[6] IWRP, Strand One Report, Technical Annex, Costs and Funding

4.2.3 An alternative is to model unit costs as the dependent variable. Using cost per main as opposed to total costs, we do find an observable relationship.

Table 4.1 – Density regression for water

Water Service:	**Water Density Model**	
Modelled cost:	Average total water CM spend [£] divided by mains length [Km]	
Explanatory Variables:	Coefficient	Standard Error
Constant	741.37	756.94
Connected properties [000] divided by mains length [Km]	32,775	10,092
Form of Model:	Cost per main [£/km] = 741.37 + 32,775 * (connected properties / mains length)	
Statistical Indicators:	Number of observations = 21	R^2 = 0.357
	Model standard error = 753.13	F test = 0.004

Figure 4.1 – Density regression for water



4.2.4 Whilst the analysis is by no means perfect, a relationship does exist. The R^2 stat is quite low but the density variable is significant. The graph shows a positive correlation. This suggests that maintenance per km of main is more expensive the more urban the population served.

4.2.5 The regression predicts the cost per main based on density. On a unit cost basis, NI Water has one of the lowest levels of actual spend per km of main (£1,637/km versus an average of £3,311/km).

4.2.6 The model anticipates this result as Northern Ireland has a very low connection density. As a result, the regression predicts CM of £1,730/km for NI Water.

Sewage density model

4.2.7 The same analysis was completed for sewage costs. Cost per sewer was predicted based on density (i.e. properties per sewer).

4.2.8 Whilst a relationship was observed, the results were not considered realistic. On a cost per sewer basis, NI Water is quite low but not really an outlier. In terms of density, they are very much an outlier. The result of this outlier status was predicted unit costs much lower than could reasonably be expected.

4.2.9 As this model could not be relied upon, an alternate was sought. The result was the following model.

Table 4.2 – Density regression for sewage

Sewage Service:	**Sewage Density Model**	
Modelled cost:	Ln (Average total water CM spend [£] divided by number of wastewater treatment works [nr])	
Explanatory Variables:	Coefficient	Standard Error
Constant	11.067	0.135
Ln (Billed properties [000] divided by WWTW [nr]	0.850	0.098
Form of Model:	Ln (Cost per works [£/nr]) = 11.067 + 0.850 * ln (billed properties / WWTW)	
Statistical Indicators:	Number of observations = 10	R^2 = 0.904
	Model standard error = 0.213	F test = 0.000

4.2.10 In this model, cost per works is the dependent variable. Properties per treatment works is used as the proxy for density. The red marker represents NI Water, while the red line indicates predicted costs for different levels of density.

4.2.11 The model looks as follows in graphical form.

Figure 4.2 – Density regression for sewage

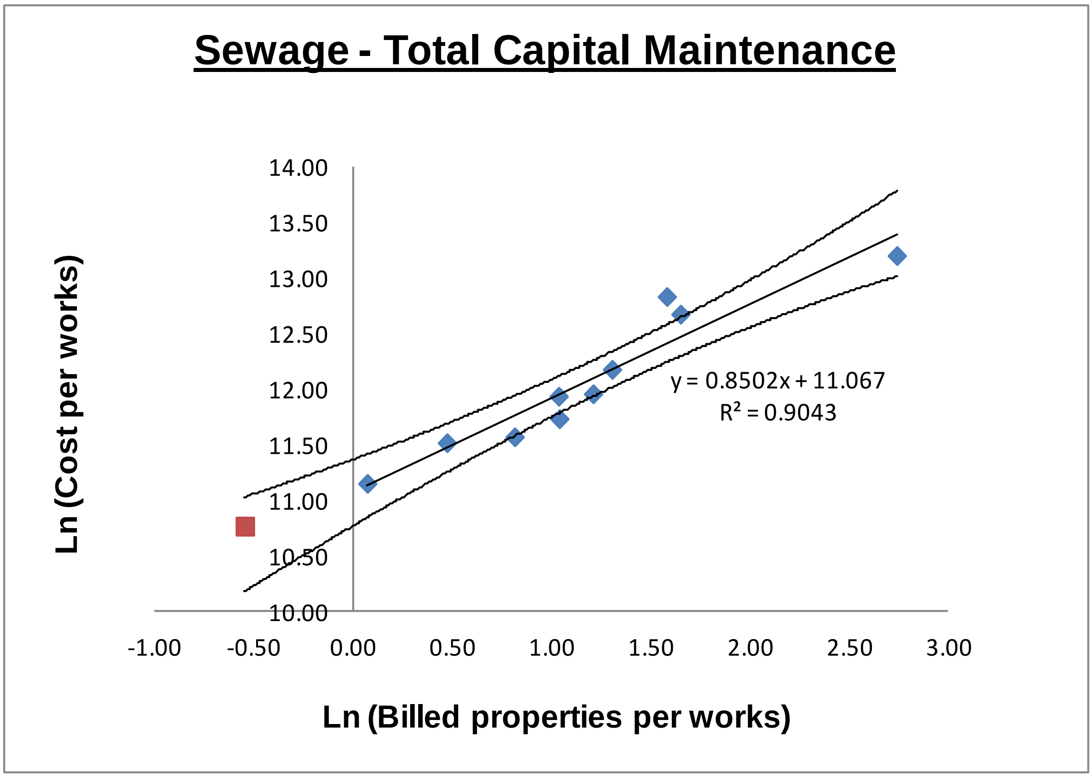
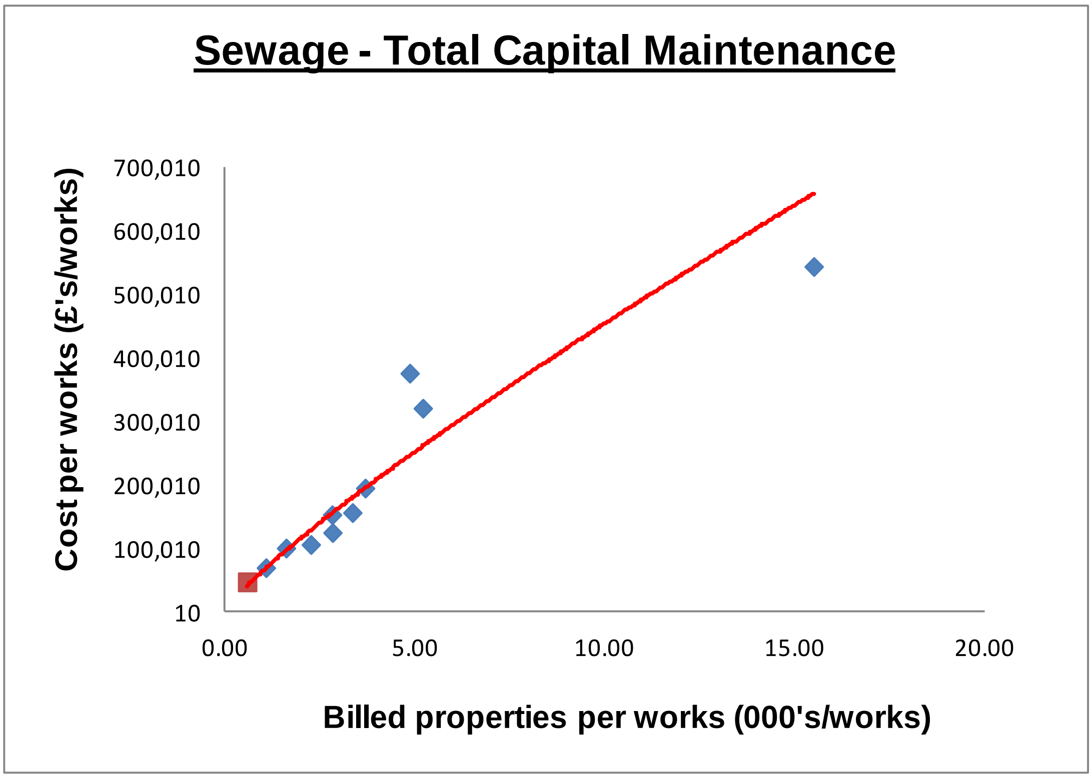


Figure 4.3 – Density model - predicted versus actual costs



4.2.12 A very clear pattern emerges on this occasion. The cost per works rises with density. This is unsurprising as more people results in larger works, more asset usage and hence higher maintenance costs.

4.2.13 Whilst there are only ten observations, the model appears a good fit for the data. The density variable is strongly significant, suggesting that it is a good predictor of unit costs.

4.2.14 A coefficient of less than one in a log model also indicates economies of scale. This means that maintenance cost will rise by a lower percentage than the increase in density.

4.2.15 NI Water has a particularly low density. This reflects the rural network and a high proportion of small treatment works. On a unit cost basis the average CM cost per works (£185,000/works) is almost four times that of NI Water's spend (£47,600/works).

4.2.16 The model does however predict this to be the case as population density is so much lower in Northern Ireland. From the model, the estimated CM costs are £40,200/works for NI Water.

4.3. Results

4.3.1 The result of using the density variables is shown in the table below.

Table 4.3 – Density model regression results (2012-13 prices)

Functional Area	NI Water Actual Spend (£m)	Average Predicted CM Spend (£m)	Upper Quartile Predicted CM Spend (£m)	Alternate Upper Quartile Predicted CM (£m)
Water Density Model	43.53	41.61	40.92	41.72
Sewage Density Model	50.04	38.86	33.26	33.57
Total	**93.57**	**80.47**	**74.18**	**75.28**

Figures include an adjustment for PPP and regional prices

4.3.2 The analysis suggests average CM expenditure in the region of £81m compared with NI Water actual spend of £94m per annum.

5.0 Unit Costs

5.1. Rationale

5.1.1 Besides regressions, unit cost analysis is a standard form of benchmarking company performance. The method is a simple calculation of total cost divided by the variable impacting maintenance spend.

5.1.2 Individual company values are compared against a weighted average taken from the total industry. Whilst simple to perform, the process is less robust than OLS regression as it fails to account for economies of scale.

5.1.3 This paper has considered a range of unit costs. These include:

a. Billed properties;

b. Network length; and

c. Usage i.e. distribution input and load received at works.

5.1.4 The three variables are considered influential in affecting CM spend. Properties will reflect the population served. Network length indicates asset size while usage will obviously influence maintenance needs.

5.1.5 The analysis has again focused on a high-level total cost by service area. Other variables could have been chosen, or costs broken down into further segments. This may be the basis of further work if required.

5.2. Models

5.2.1 The first of the unit cost models is billed properties. The comparison of both water and sewage costs per billed property is shown below.

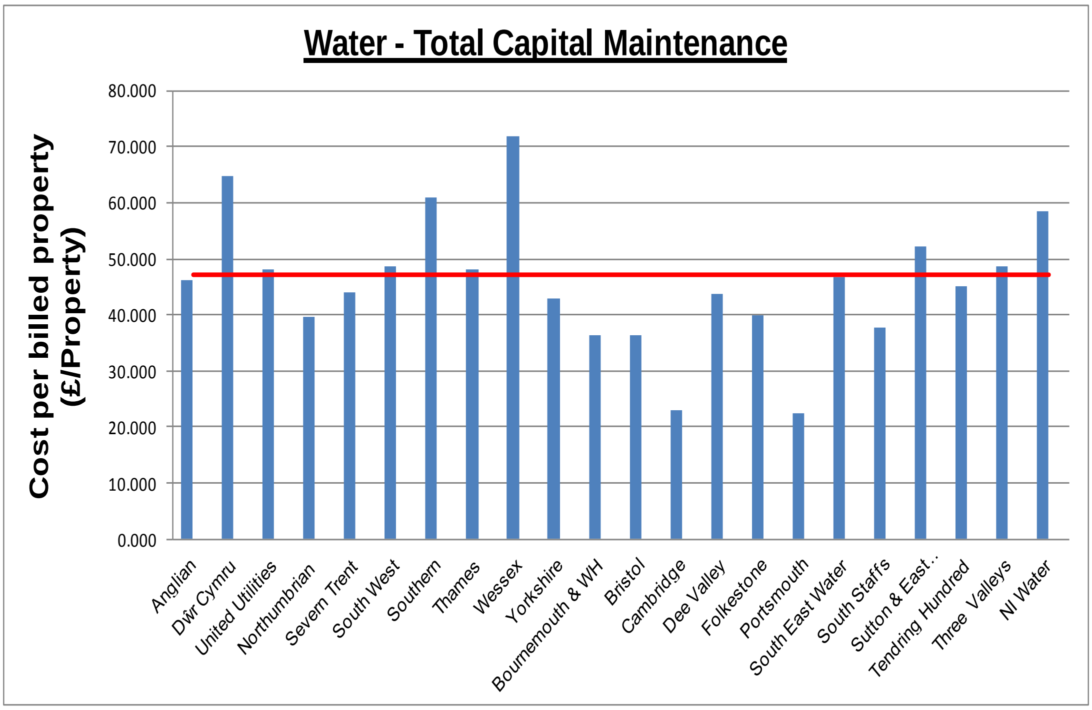
Table 5.1 – Billed properties unit costs – water and sewage (2012-13 prices)

Unit Cost	Water	Sewage
NI Water cost per billed property (£/property)	58.35	82.27
Weighted average cost per billed property (£/property)	47.23	51.21

5.2.2 The figures indicate over expenditure compared against the average. This is particularly true on the sewage side. NI Water spends over £30 per property more on CM than its counterparts do.

5.2.3 The water costs per property are shown in graph form below.

Figure 5.1 – Billed properties unit costs – water



5.2.4 It is interesting to note that some of the lowest unit costs are found among the small water only companies.

5.2.5 The results may not tell the full story for NI Water. It can be seen that other more rural companies such as Welsh Water or Wessex also have high unit costs. This may suggest that using billed properties is a disadvantageous comparison for rural companies.

5.2.6 The alternative is to look at cost per network length i.e. main or sewer. The result of this analysis paints a different picture.

Network length unit costs

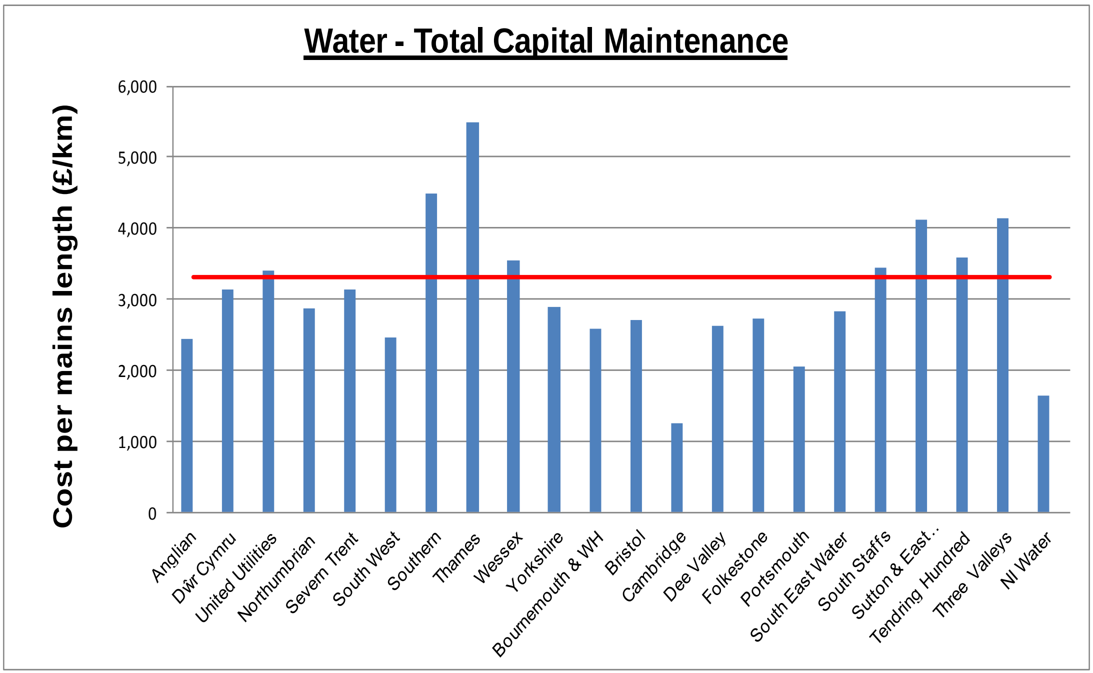
5.2.7 On a cost per main or cost per sewer basis, the results are very different. Comparison against average unit costs is provided in the table below.

Table 5.2 – Network length unit costs – water and sewage (2012-13 prices)

Unit Cost	Water	Sewage
NI Water cost per main/sewer (£/km)	1,637	2,933
Weighted average cost per main/sewer (£/km)	3,311	3,613

5.2.8 Using main/sewer length as the denominator, NI Water now becomes one of the most efficient operators. This is demonstrated most starkly in the water graphic.

Figure 5.2 – Network length unit costs – water



5.2.9 NI Water spends over 50% less per kilometre of main on water CM as the industry average. For sewage, the figure is almost 19% less. This suggests either under spend or very efficient expenditure.

5.2.10 Again, the figures may be somewhat confusing. This method treats all mains the same. In reality, there are large differences in both their size and cost. Comparison is difficult as NI Water has a large proportion of small mains that would not be expected to incur much cost.

5.2.11 This is difficult to compare with a company like Thames. They have many large mains, which incur much higher maintenance costs. Using the industry average on this occasion could result in an over estimation of CM requirements.

Usage unit costs

5.2.12 A final alternative is to assess costs based on usage. Usage is reflected by distribution input for water and the load received at works for sewage.

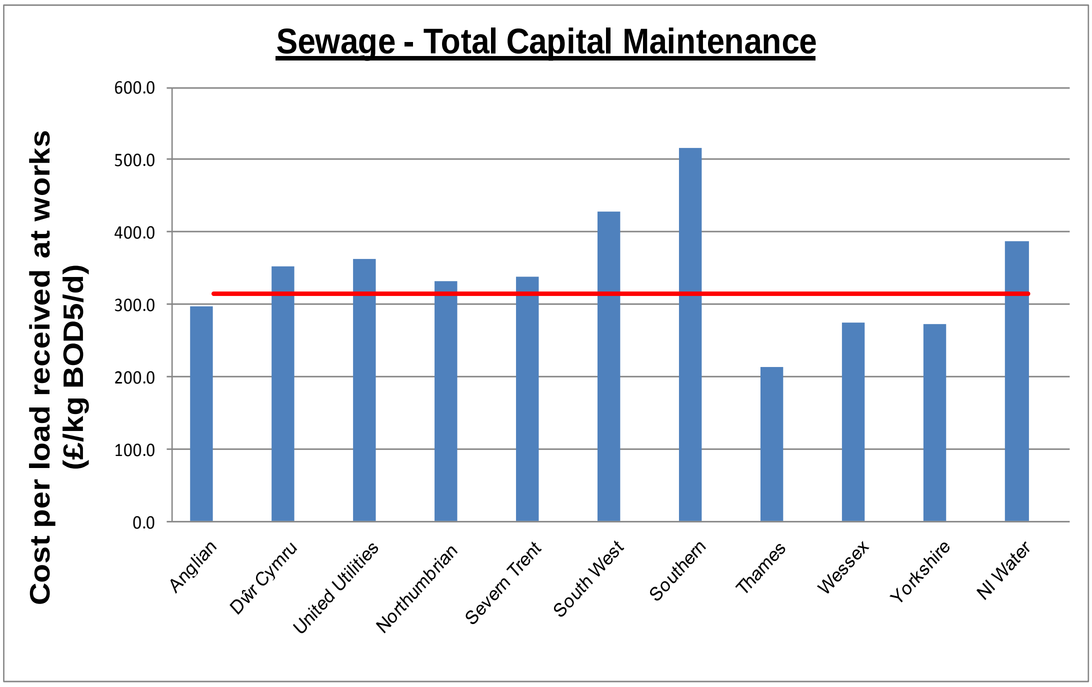
5.2.13 Usage is considered important. The more an asset is used, the more it will deteriorate and hence the more capital will be required to maintain it. Results for this unit cost are provided below.

Table 5.3 – Usage unit costs – water and sewage (2012-13 prices)

Unit Cost	Water	Sewage
NI Water cost per DI/load (£/Ml) or (£/kgBOD5/d)	190.77	388.0
Weighted average cost per DI/load (£/Ml) or (£/kgBOD5/d)	208.46	314.2

5.2.14 The results are mixed on this occasion. NI Water under spends on water CM but over spends on a cost per load basis for sewage.

Figure 5.3 – Usage unit costs – sewage



5.2.15 On both occasions, the results are quite close to the weighted average. This contrasts with the other unit cost approaches where NI Water tends to be somewhat of an outlier.

5.3. Results

5.3.1 There is a wide range of predicted costs generated from the unit cost methods described above. Using the weighted average industry unit costs to predict CM for NI Water gives the following results.

Table 5.4 – Unit cost results – billed properties (2012-13 prices)

Unit Cost Model	NI Water Actual Spend (£m)	Average Predicted CM Spend (£m)	Upper Quartile Predicted CM (£m)	Alternate Upper Quartile Predicted CM (£m)
Water model – Billed prop	43.53	31.48	30.98	31.16
Sewage model – Billed prop	50.04	27.63	26.67	26.48
Total Base Maintenance	**93.57**	**59.11**	**57.65**	**57.64**

Figures include an adjustment for PPP and regional prices

Table 5.5 – Unit cost results – network length (2012-13 prices)

Unit Cost Model	NI Water Actual Spend (£m)	Average Predicted CM Spend (£m)	Upper Quartile Predicted CM (£m)	Alternate Upper Quartile Predicted CM (£m)
Water model – Mains length	43.53	81.13	73.71	76.92
Sewage model – Length	50.04	56.30	47.53	47.24
Total Base Maintenance	**93.57**	**137.43**	**121.24**	**124.16**

Figures include an adjustment for PPP and regional prices

Table 5.6 – Unit cost results – usage (2012-13 prices)

Unit Cost Model	NI Water Actual Spend (£m)	Average Predicted CM Spend (£m)	Upper Quartile Predicted CM (£m)	Alternate Upper Quartile Predicted CM (£m)
Water model – DI	43.53	43.07	40.74	41.21
Sewage model – Load	50.04	36.45	33.47	34.60
Total Base Maintenance	**93.57**	**79.52**	**74.20**	**75.82**

Figures include an adjustment for PPP and regional prices

5.3.2 The size of the estimated range demonstrates the difficulty with using unit costs for predicting total maintenance levels. Perhaps greater clarity could be brought by analysing costs that are broken down further by functional area.

5.3.3 Whilst recognising problems with each method, it is difficult to choose one unit cost over another. Rather than giving one preference, an alternative is to take an average of the findings.

5.3.4 A simple average of the three methods would give average and upper quartile predicted costs of roughly £92m and £85m respectively.

6.0 Historic Cost Analysis

6.1. Rationale

6.1.1 Unlike the other methods, the historic cost analysis does not rely on England and Wales data. Rather, future CM forecasts are derived from historic spends.

6.1.2 The basic rationale is the assumption that there is an underlying stability in the historic data. While spend may fluctuate from year to year, a stable trend can be estimated by smoothing the data.

6.1.3 As long as historic spend has been sufficient to maintain assets in the past, the underlying trend should be a reasonable predictor of future needs.

6.1.4 This section considers a number of different techniques used to smooth the data. The trend is then used to predict future CM needs. The different models include:

a. **Simple average** – Mean of the CM spend for the entire dataset. Each observation has equal weight;

b. **Moving average –** A moving average is an approach that uses most recent data values to predict the next period. Older data will not have any influence;

c. **Weighted moving average** – This method is the same as the moving average only gives greater weighting to more recent data; and

d. **Exponential smoothing** – A form of weighted average with a declining weight placed on each observation the older it is.

6.1.5 The problem with each of these methods is the reliability of the historic data. As mentioned in the introductory section, there are a number of causes for concern with the data.

6.1.6 The issue of backlog base, spend to budget and year-end constraints fail to provide assurance that historic spend is at correct levels.

6.1.7 These problems are further complicated by the PPP issue since these costs are in some years of the data but not others.

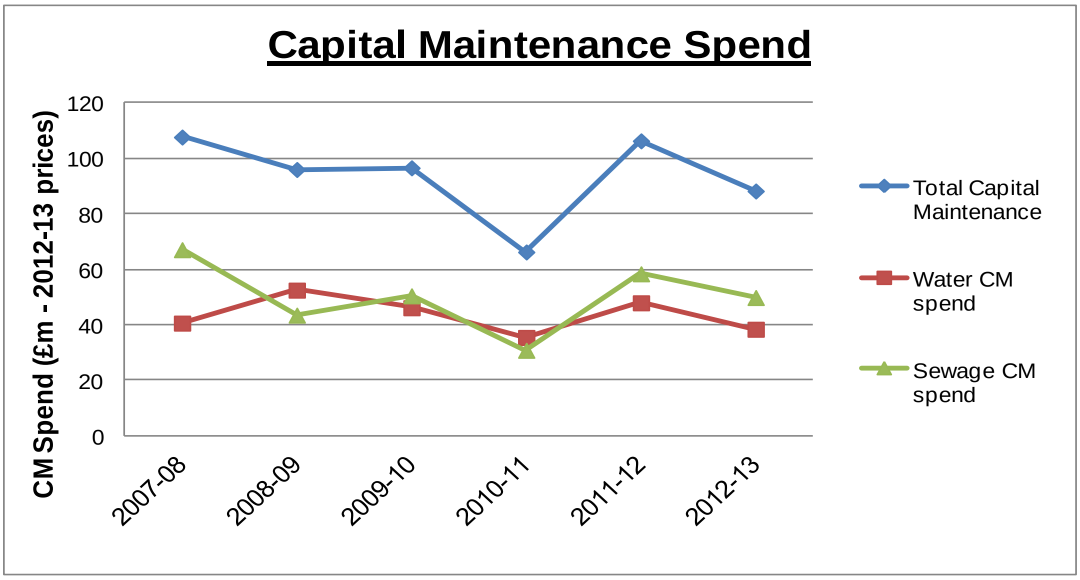
6.1.8 A potential problem may also be the size of the dataset. It is questionable whether six years of financial records are enough to determine an underlying trend. This then makes future predictions problematic.

6.1.9 Concerns over the information place a question mark on any analysis that makes future predictions using it. The UR has however run these models in order to see the outcome of adopting such an approach.

6.2. Models

6.2.1 The levels of CM spend split by water, sewage and total cost is shown below.

Figure 6.1 – Historic capital maintenance for NI Water

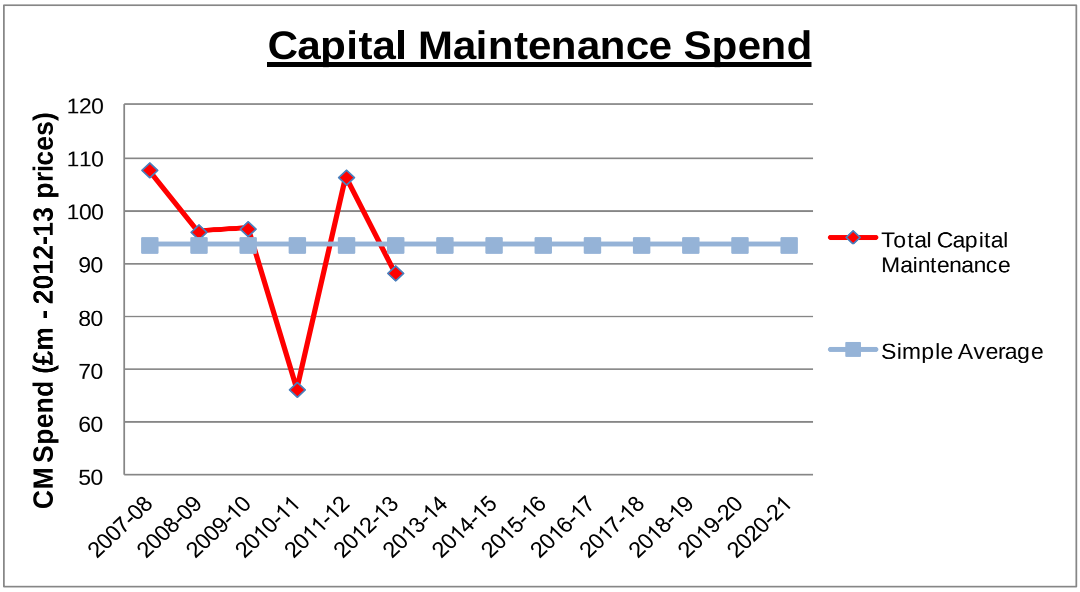


6.2.2 In real terms, spending has been quite stable. There may be a slight downward trend, but it is not immediately obvious. The decrease may be accounted for by the removal of PPP asset maintenance in later years. However, there does not appear to be a step change as might have been expected.

Simple average

6.2.3 The most obvious method of smoothing the data to establish a trend is to take a simple average. The results show predicted spend of £93.6m per annum.

Figure 6.2 – Simple average predictions



6.2.4 While simple and easy to understand, the raw average is perhaps not ideal in this scenario. The mean gives equal weighting to each observation. The issue with this is that spend on all assets are accounted for in the SBP years. In the last three years, the PPP assets are excluded.

6.2.5 To give equal weighting to each observation may overstate the long-term trend. By way of an alternate, a moving average is used.

Moving averages

6.2.6 The moving average, as it is known, consists of a series of averages based on sub-sets of the data. The smoothed data depends on the size of these sub-sets. A two-point moving average means that the next year (t_1) is calculated based on the previous two observations i.e. ($\frac{t + t_{-1}}{2}$).

6.2.7 A worked example illustrates this more clearly.

Table 6.1 – Two-point moving average (2012-13 prices)

Year	NI Water Actual CM Spend (£m)	Moving Total (n=2) (£m)	Moving Average Forecast (£m)
2007-08	107.75	NA	NA
2008-09	96.05	NA	NA
2009-10	96.65	107.75 + 96.05	101.90
2010-11	66.30	96.05 + 96.65	96.35
2011-12	106.39	96.65 + 66.30	81.47
2012-13	88.29	66.30 + 106.39	86.34
2013-14	97.34	106.39 + 88.29	97.34
2014-15	92.81	88.29 + 97.34	92.81

Predicted spend is given in red text

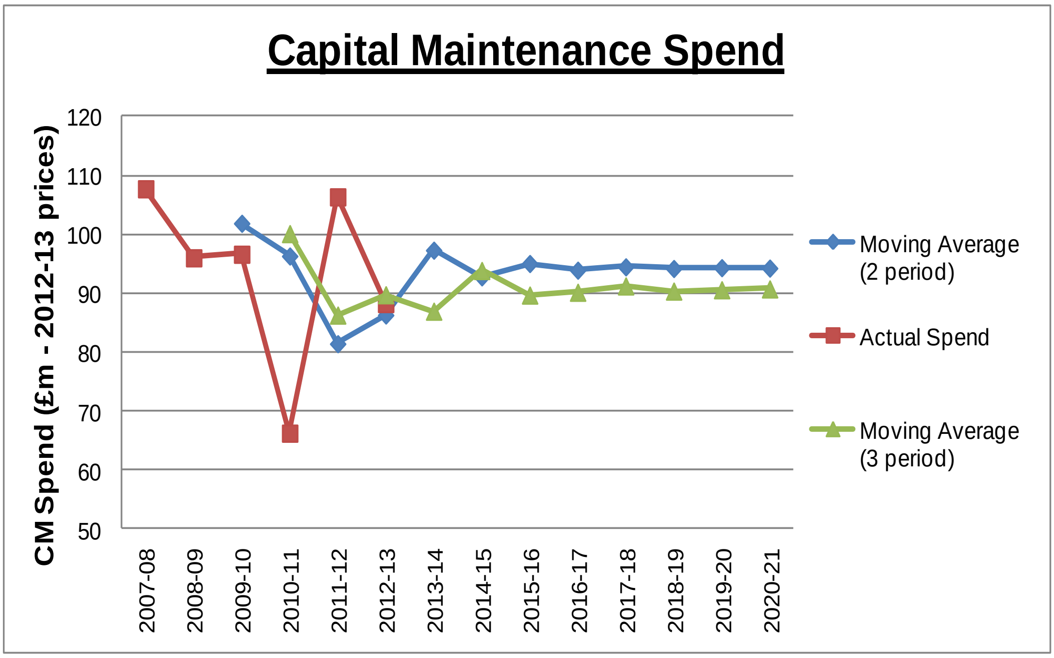
6.2.8 The table shows how the moving average is calculated. The benefit of such an approach is that older data is ignored when predicting future spend. In essence, the smaller the time sub-set, the more weight is given to recent observations.

6.2.9 This is helpful for this particular dataset as the later figures excluding PPP are more relevant to future forecasts.

6.2.10 The problem with this method is that small time-periods may not reflect the underlying trend. The limited size of the data sample restricts the moving averages to a two-point and three-point analysis. It is questionable whether this is an appropriate sample size to forecast forward.

6.2.11 The chart below illustrates the two moving average predictions.

Figure 6.3 – Moving average predictions



6.2.12 The graph shows that the three-point average is less responsive to the fluctuations in actual spend than the two period approach.

6.2.13 The two-point average generates predicted spend of £94.4m per annum across PC15. The three-point average gives a value of £90.5m over the same period.

Weighted averages

6.2.14 A weighted average approach is essentially the same as the moving average. The difference here is that a weighting applies to the data points depending on how old they are.

6.2.15 For instance, in the two-point moving average illustrated above, the forecast period is a simple average of the previous two figures. In a weighted average, more emphasis is placed on the current time-period and less on the year before.

6.2.16 Use of such an approach is appropriate if it were felt that recent data is more relevant. For the CM spending dataset there is no obvious reason why 2012-13 figures are better for prediction than the year before.

6.2.17 However, it is expected that the company is becoming more efficient over time. Under these circumstances, it may be better to place more emphasis on recent figures. Failure to do so might incorporate more inefficient expenditure into future forecasts.

6.2.18 There are no set rules for determining the weights each year should carry. In the absence of any overriding logic, we have chosen the following weights.

a. Two-point weighted average = 70%-30% weighting; and

b. Three-point weighted average = 50%-30%-20% weighting.

6.2.19 The calculations for the two-point weighted average are below.

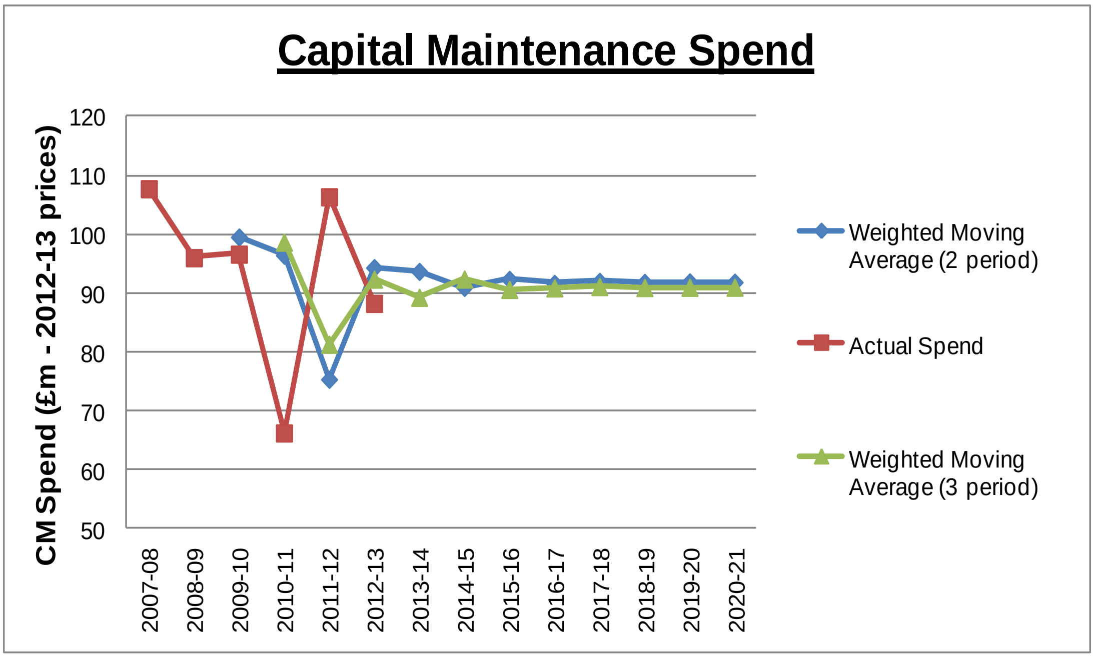
Table 6.2 – Two-point weighted average (2012-13 prices)

Year	NI Water Actual CM Spend (£m)	Moving Average Calc (n=2) (£m)	Moving Average Forecast (£m)
2007-08	107.75	NA	NA
2008-09	96.05	NA	NA
2009-10	96.65	0.3*(107.75) + 0.7*(96.05)	99.56
2010-11	66.30	0.3*(96.05) + 0.7*(96.65)	96.47
2011-12	106.39	0.3*(96.65) + 0.7*(66.30)	75.40
2012-13	88.29	0.3*(66.30) + 0.7*(106.39)	94.36
2013-14	93.72	0.3*(106.39) + 0.7*(88.29)	93.72
2014-15	91.00	0.3*(88.29) + 0.7*(93.72)	91.00

Predicted spend is given in red text

6.2.20 The forecasts of the weighted average approaches are in the chart below.

Figure 6.4 – Weighted average predictions



6.2.21 The two-point weighted average forecasts spend of £92m per annum. The alternate method gives a value of £91m p.a. in 2012-13 prices across the PC15 period.

Exponential smoothing

6.2.22 Unlike the weighted averages, this technique does not completely ignore early data. Rather, the exponential smoothing takes account of all data but the weighting reduces the older the data is.

6.2.23 Forecast figures for the next period are based on the current year actual and the current year forecast weighted by a smoothing factor (α). The smoothing factor is a value between zero and one.

6.2.24 The smaller the value of the smoothing factor, the less responsive it is to change. The modeller sets the smoothing factor depending on what is hoped to be achieved.

6.2.25 The formula for exponential smoothing is as follows:

$$F_{t+1} = \propto y_t + (1{-}\propto)F_t$$

Where:

F_{t+1} = forecast for the next period

$\propto$ = smoothing constant

y_t = observed value in period t (current period)

F_t = old forecast for period t

6.2.26 The table below illustrates how the formula works in practice. Since the first forecast is not know, either the first observation is chosen or an average of the data.

Table 6.3 – Exponential smoothing (2012-13 prices)

Year	NI Water Actual CM Spend (£m)	Smoothing Calc (α = 0.1) (£m)	Exponential Smoothing Forecast (£m)
2007-08	107.75	93.57	93.57
2008-09	96.05	0.1*(107.75) + 0.9*(93.57)	94.99
2009-10	96.65	0.1*(96.05) + 0.9*(94.99)	95.09
2010-11	66.30	0.1*(96.65) + 0.9*(95.09)	95.25
2011-12	106.39	0.1*(66.30) + 0.9*(95.25)	92.35
2012-13	88.29	0.1*(106.39) + 0.9*(92.35)	93.76
2013-14	93.21	0.1*(88.29) + 0.9*(93.76)	93.21
2014-15	93.21	0.1*(93.21) + 0.9*(93.21)	93.21

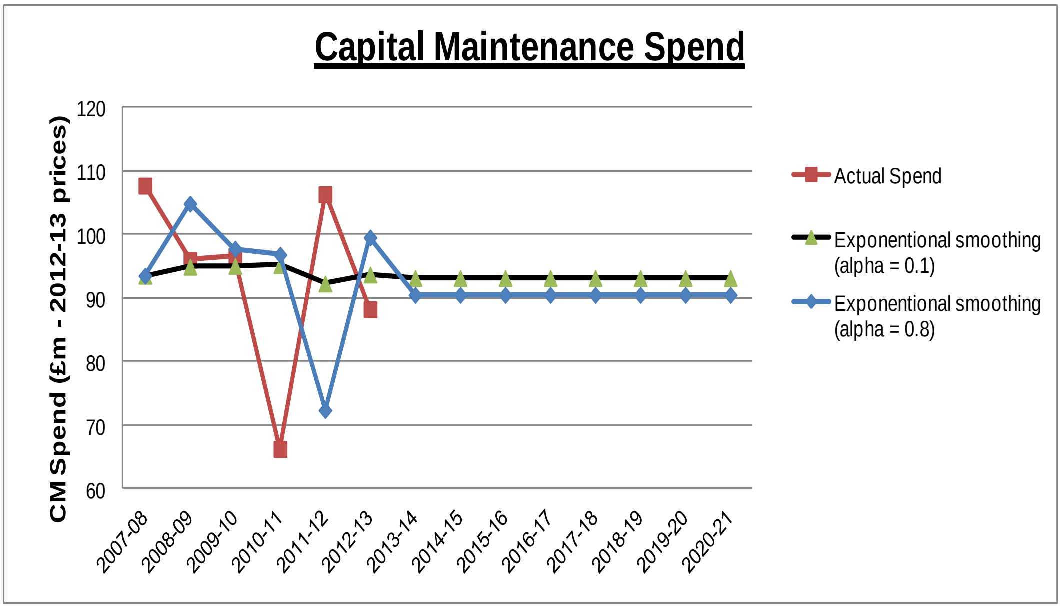
Predicted spend is given in red text

6.2.27 The formula shows that the future forecast is the function of the current value and the current forecast. The current forecast is however a function of previous actual values.

6.2.28 This is known as exponential smoothing as the weighting of each historic actual value increases by the power of one the older the data.

6.2.29 The UR assessed the forecasts using both a high and low smoothing factor i.e. α = 0.8 and α = 0.1 respectively. The results are shown in the graph below.

Figure 6.5 – Weighted average predictions



6.2.30 The graph shows the difference the smoothing factor makes. A low alpha value is a lot less responsive to changes in the data. The low alpha predicts future expenditure at £93.2m per annum. The alternate gives an estimate of £90.6m.

6.3. Results

6.3.1 The table below provides forecast expenditure for each of the methods.

Table 6.4 – Historical cost analysis results (2012-13 prices)

Method	Forecast Spend (£m)	Mean Absolute Deviation
Simple Average	93.6	10.9
Two-Point Moving Average	94.4	15.5
Three-Point Moving Average	90.5	14.0
Two-Point Weighted Average	92.0	17.5
Three-Point Weighted Average	91.0	15.7
Exponential Average (α = 0.1)	93.2	10.9
Exponential Average (α = 0.8)	90.6	16.7
Overall Average of Methods	**92.2**	**N/A**

6.3.2 Predictions for future spend range from £90.5m to £94.4m per annum in real terms. The average of each method gives an overall prediction of £92.2m.

6.3.3 The table also includes a column for the mean absolute deviation. This figure gives an indication of how good each method is at forecasting observed values.

6.3.4 None of the methods seems overly robust. In this case, where there is no discernible upward or downward trend, it appears that the simple and exponential averages are the best predictors of future costs. Both give a value of just over £93m per annum.

7.0 Conclusions

7.1. Results and conclusions

7.1.1 The various methods each provide an estimate of capital maintenance. Some of the models are more convincing than others.

7.1.2 The drawbacks of the PC10 models and the unit cost analysis are important. The tendency therefore would be to place more emphasis on the CSV models and the density regressions.

7.1.3 Results for each of the methods are provided below.

Table 7.1 – Model results – total for water and sewage (2012-13 prices)

Models	NI Water Actual Spend (£m)	Average Predicted CM Spend (£m)	Upper Quartile Predicted CM Spend (£m)	Alternate Upper Quartile Predicted CM (£m)
PC10 models	93.57	80.78	74.02	74.36
CSV Base case	93.57	90.62	81.28	79.90
Density variable regression	93.57	80.47	74.18	75.28
Unit costs – average	93.57	92.02	84.37	85.87
Historic cost analysis	93.57	92.20	N/A	N/A

Figures include an adjustment for PPP and regional prices

7.1.4 Average predicted costs for each method tend not to be that different (up to 14% less) from what NI Water is actually spending to date. This difference is even less if backlog base maintenance is removed from historic NI Water spending.

8.0 Appendix 1 – Update PC10 Models

8.1. Models

8.1.1 The rationale for the structure of the PC10 models is provided in Annex B of the final determination. The link below includes this detail. http://www.uregni.gov.uk/uploads/publications/PC10_NIAUR_FD_Feb_10_-_Doc04_-_Annex_B_Cap_Maintenance_Analysis.pdf

8.1.2 This document does not intend providing full explanation of the models. Rather, the focus is on showing the results of updating the regressions with more recent data.

Water infrastructure

8.1.3 The water infrastructure regression shows the relationship between unit costs and connection density.

Table 8.1 – Water infrastructure updated PC10 model

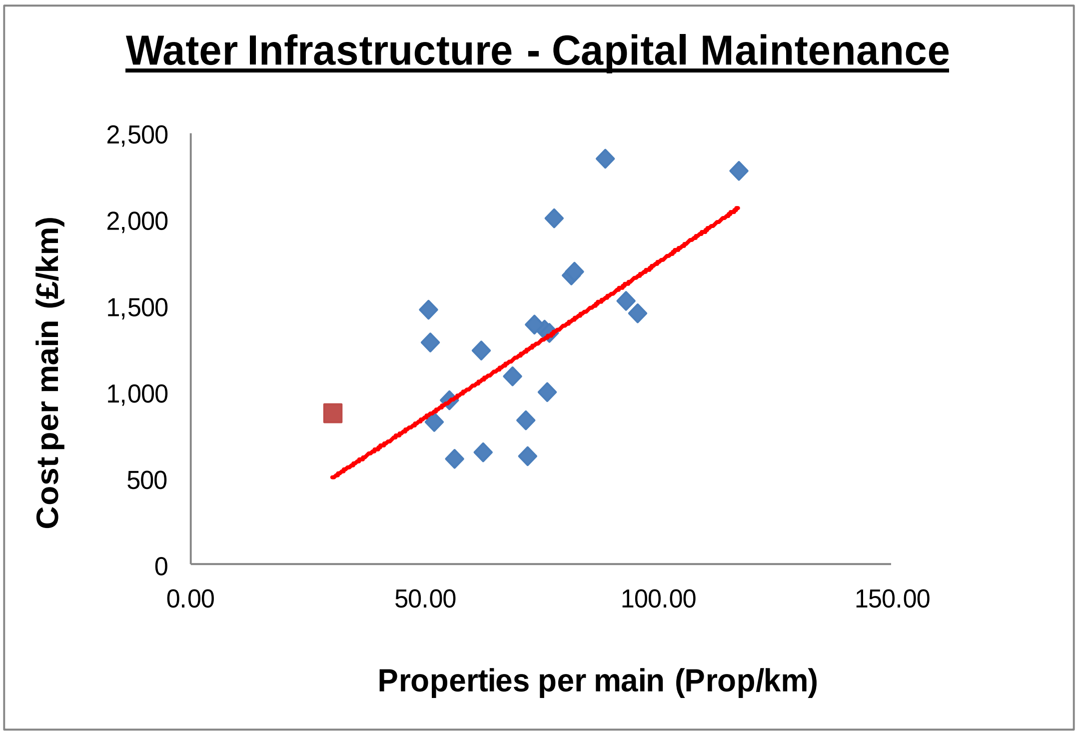
Water Service:	Water infrastructure updated PC10 model	
Modelled cost:	Ln (annual average water distribution infrastructure functional cost [£m], divided by length of main [km])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-3.938	0.873
Ln (number of connected properties [000's], divided by length of main)	1.047	0.330
Form of Model:	Ln modelled cost = -3.938 + 1.047 * ln {connected props / length of main}	
Statistical Indicators:	Number of observations = 21	R^2 = 0.347
	Model standard error = 0.329	F test = 0.005

8.1.4 The regression shows a positive correlation between costs and density. The independent variable is statistically significant so does in part help to explain infrastructure costs.

8.1.5 An R^2 value of 0.35 does however suggest that other factors need to be accounted for in order to fully explain this spend.

8.1.6 Whilst the regression is in log format, the graph below shows the relationship in real terms. As previously, the red mark represents NI Water. The red line is the predicted costs derived from the regression for each level of connection density.

Figure 8.1 – Water infrastructure – predicted versus actual costs



Water non-infrastructure

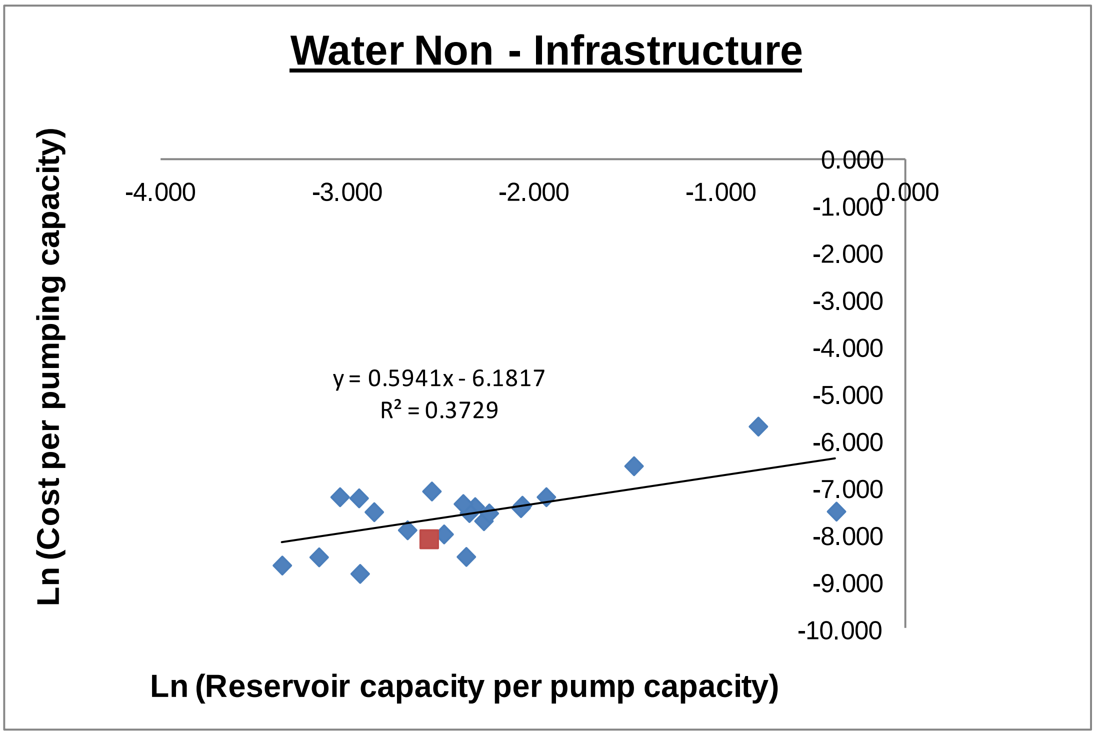
8.1.7 The non-infrastructure model attempts to explain unit costs by virtue of storage capacity.

Table 8.2 – Water non-infrastructure updated PC10 model

Water Service:	Water non-infrastructure updated PC10 model	
Modelled cost:	Ln (annual average water distribution non-infrastructure functional cost [£m], divided by pumping station capacity [kW])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-6.182	0.428
Ln (service reservoir and water tower capacity [Ml], divided by pumping capacity [kW)	0.594	0.177
Form of Model:	Ln modelled cost = -6.182 + 0.594 * ln {storage capacity / pumping station capacity}	
Statistical Indicators:	Number of observations = 21	R^2 = 0.373
	Model standard error = 0.577	F test = 0.003

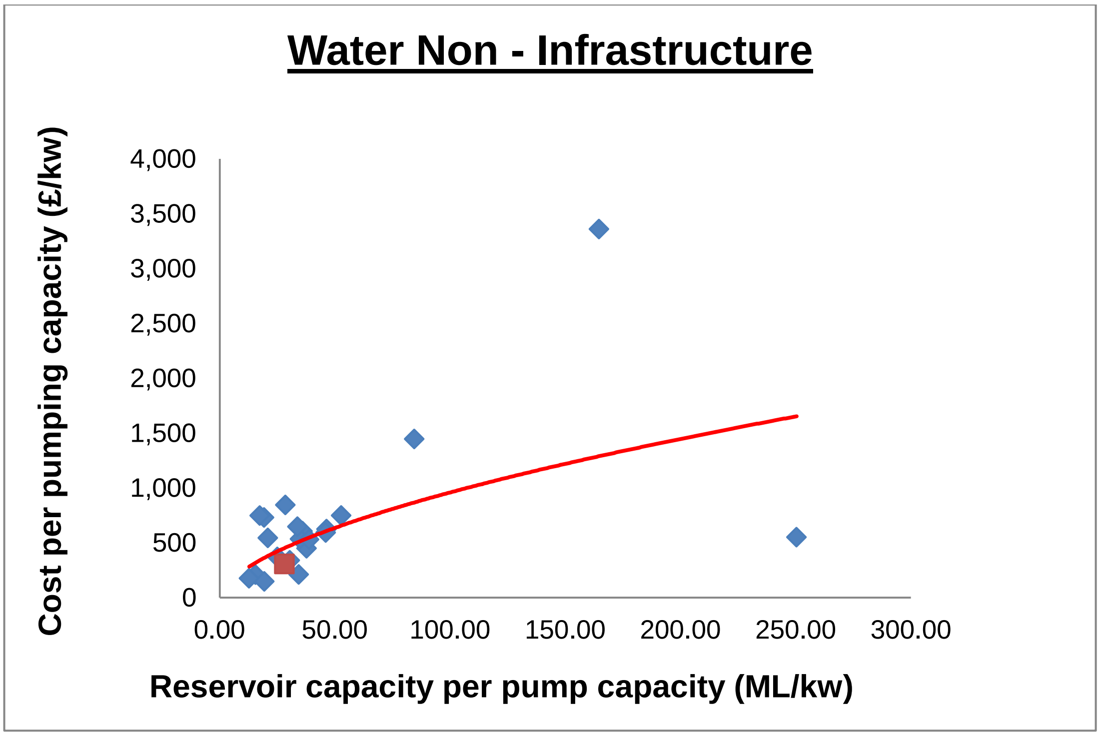
8.1.8 Results show a positive relationship between unit costs and storage capacity as expected. The variable is significant but the R^2 value is reasonably low. In graphical form, the model is as follows.

Figure 8.2 – Water non-infrastructure updated PC10 model



8.1.9 The log model shows that a relationship exists. In real terms, the analysis is as follows:

Figure 8.3 – Water non-infrastructure – predicted versus actual costs



8.1.10 Whilst there appears to be a relationship, predicted costs are being skewed by one outlier. The robustness of this model may be improved by removing this company.

Water management and general

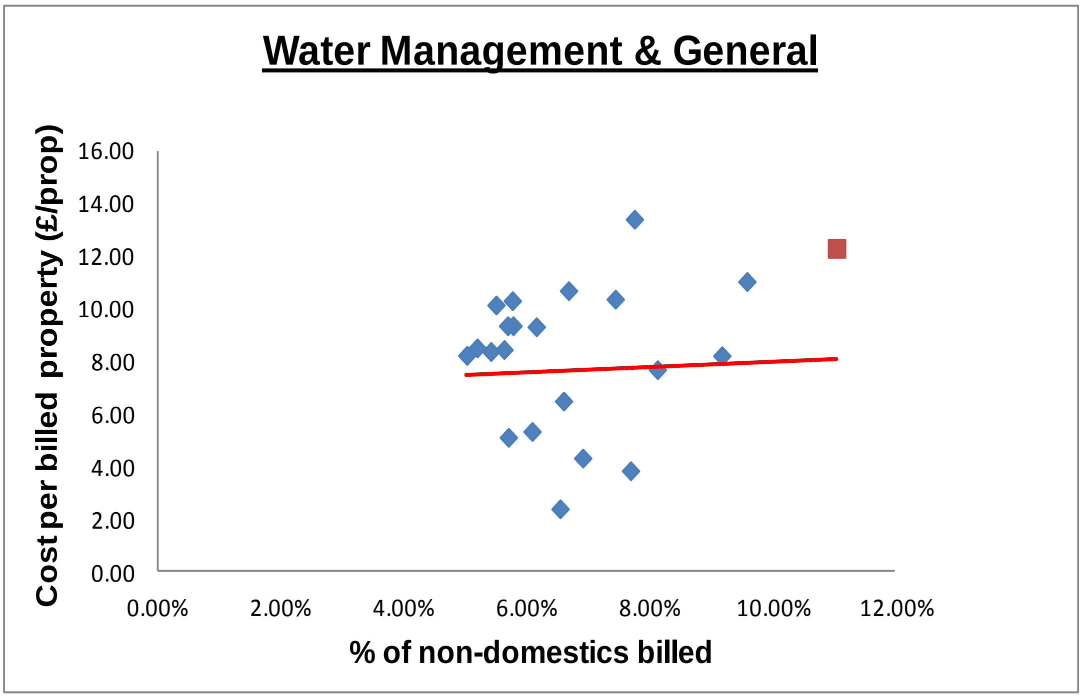
8.1.11 This model predicts management and general unit costs based on the proportional size of non-households billed. The model is as follows:

Table 8.3 – Water management and general updated PC10 model

Water Service:	**Water management and general updated PC10 model**	
Modelled cost:	Ln (annual average water M&G cost [£m], divided by billed properties [000's])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-4.967	0.498
Proportion of billed non-household properties	1.323	7.425
Form of Model:	Ln modelled cost = -4.967 + 1.323 * proportion of non-household properties	
Statistical Indicators:	Number of observations = 21	R^2 = 0.002
	Model standard error = 0.425	F test = 0.860

8.1.12 In real terms, the graphical view shows a lack of observable relationship.

Figure 8.4 – Water management and general - predicted versus actual

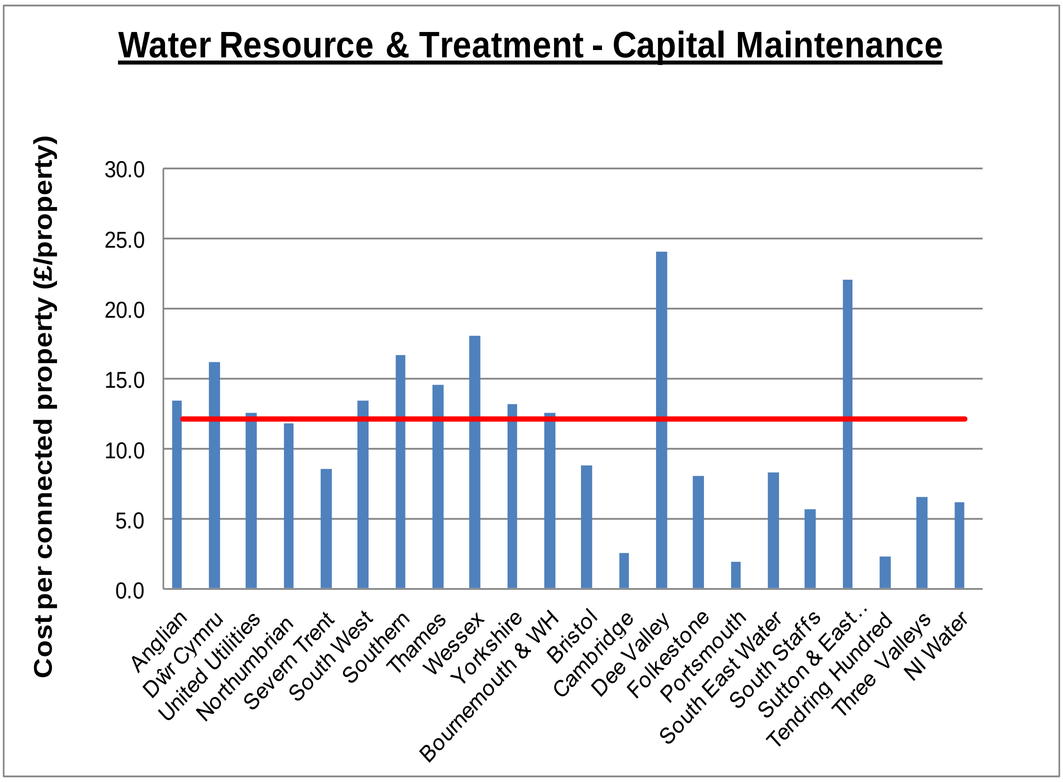


8.1.13 This is not a reliable model. The explanatory variable is not significant and the model fails to explain CM costs in this area. Findings for this particular model cannot be relied upon.

Water resources and treatment

8.1.14 Resources and treatment spend is predicted using a unit cost model. Connected properties are the denominator.

Figure 8.5 – Water resources and treatment unit costs model



8.1.15 The graphs show that NI Water's unit spend is almost half the industry average. This may indicate either efficiency or under investment.

Sewage infrastructure

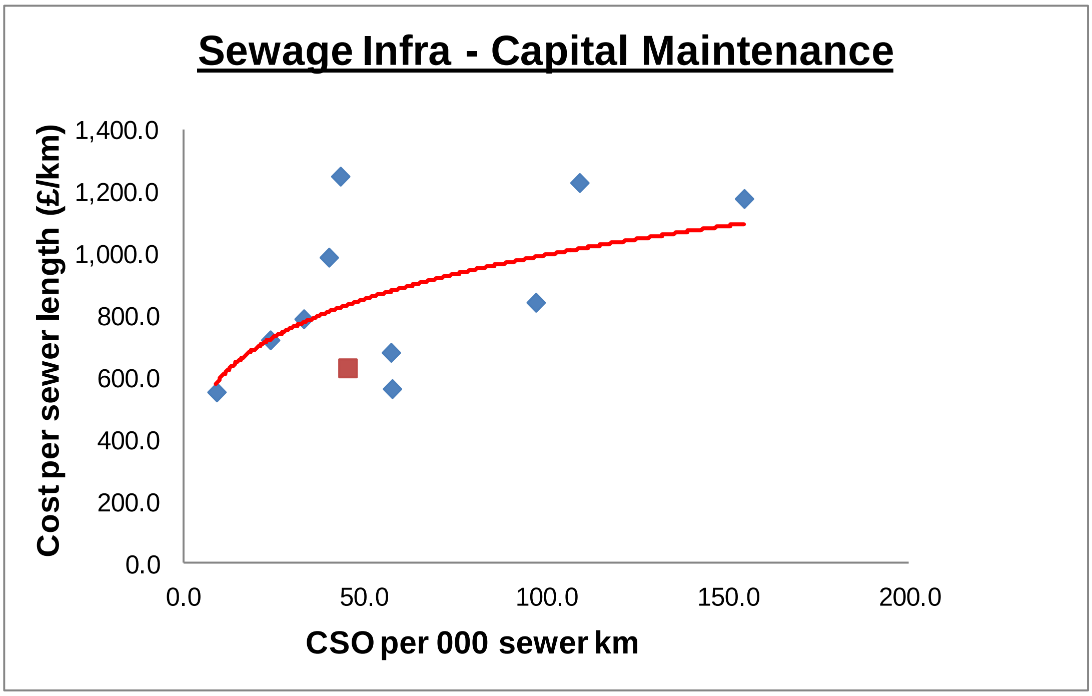
8.1.16 In this regression, unit costs are modelled against the number of combined sewer overflows (CSO's) per sewer length. The rationale is that CSO's are generally larger and more costly to maintain than foul sewers.

8.1.17 The form of the model is as follows:

Table 8.4 – Sewage infrastructure updated PC10 model

Sewage Service:	Sewage infrastructure updated PC10 model	
Modelled cost:	Ln (annual average sewerage infrastructure functional cost [£m], divided by sewer length [km])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-6.398	0.333
Ln (number of CSO's divided by sewer length [km)	0.225	0.106
Form of Model:	Ln modelled cost = -6.398 + 0.225 * ln {CSO's / length of sewer [km]}	
Statistical Indicators:	Number of observations = 10	R^2 = 0.358
	Model standard error = 0.262	F test = 0.068

Figure 8.6 – Sewage infrastructure - predicted versus actual costs



8.1.18 The model shows a definite correlation. The R^2 value of 0.36 does however imply that the explanatory power of the model could be better.

Sewage treatment

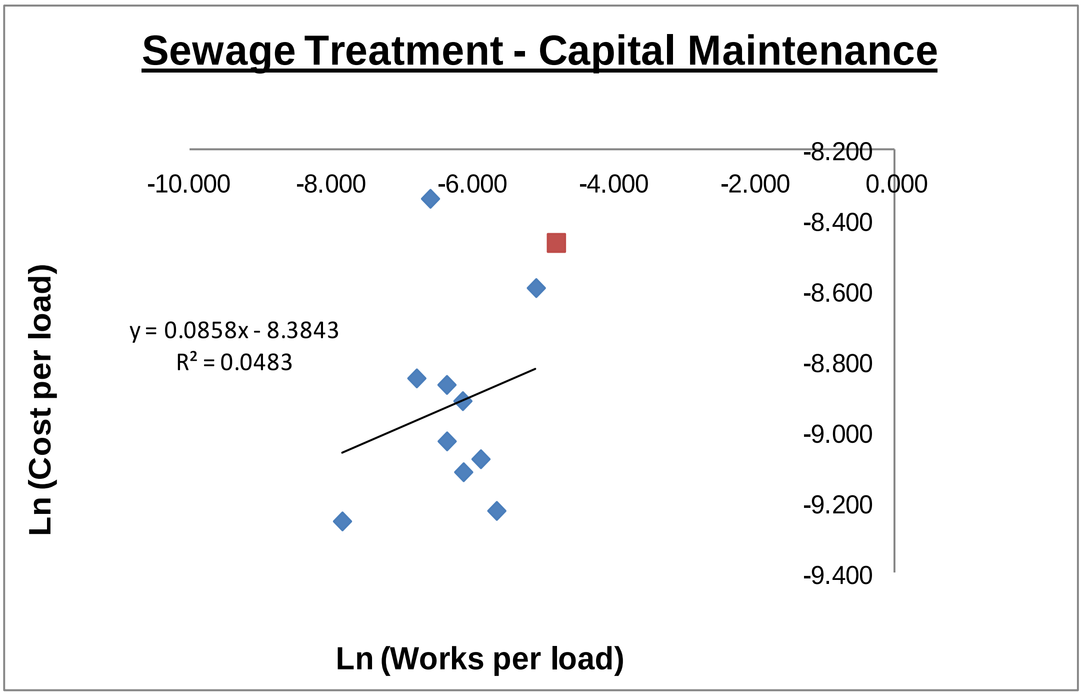
8.1.19 The treatment model predicts costs per load received. The explanatory variable is the number of works per load.

8.1.20 The reasoning is based on the assumption that smaller works require more maintenance spend per load treated than larger, more efficient works. Therefore as the number of treatment works per load increase, so too will base maintenance (positive coefficient).

Table 8.5 – Sewage treatment updated PC10 model

Sewage Service:	**Sewage treatment updated PC10 model**	
Modelled cost:	Ln (annual average sewage treatment functional cost [£m], divided by total load received at treatment works [kg BOD_5/day])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-8.384	0.851
Ln (total number of works divided by total load received at treatment works [kg BOD_5/day])	0.086	0.135
Form of Model:	Ln modelled cost = -8.384 + 0.086 * ln {number of works / total load received at sewage treatment works [kg BOD_5/day]}	
Statistical Indicators:	Number of observations = 10	R^2 = 0.048
	Model standard error = 0.295	F test = 0.542

Figure 8.7 – Sewage treatment updated PC10 model



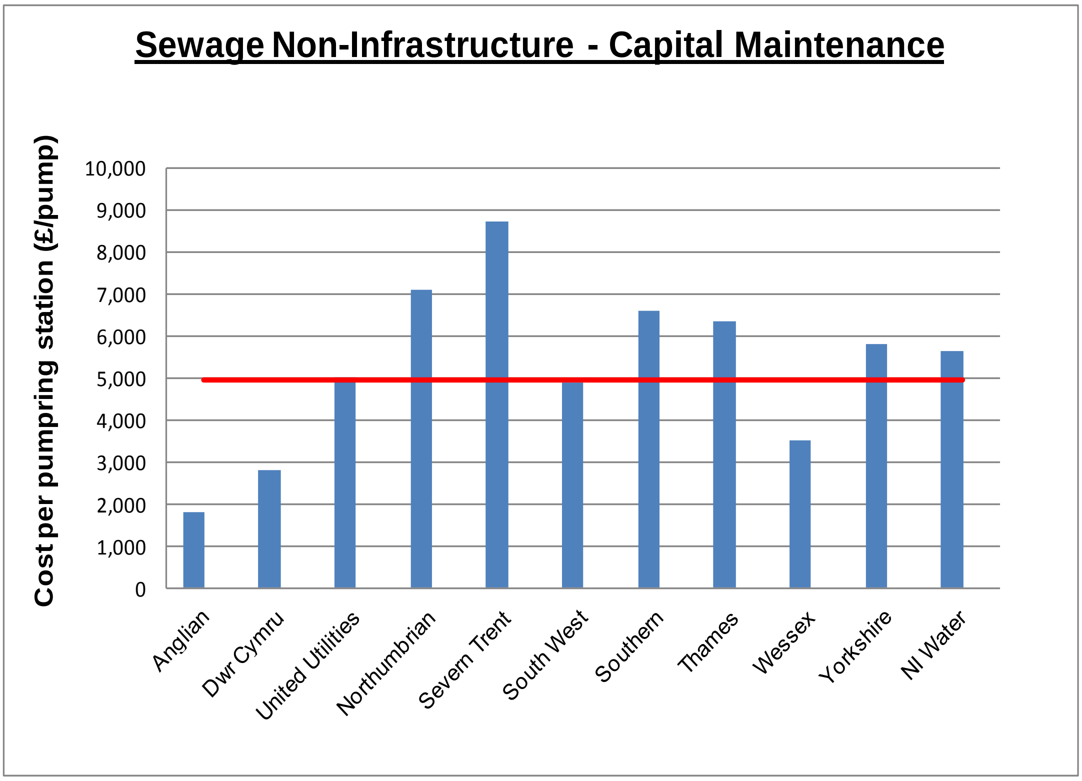
8.1.21 The update of this model no longer gives adequate insight into expected costs in this area. Whilst the sign of the coefficient is as expected, no assurance can be provided from this, as the variable is statistically insignificant.

8.1.22 The graph shows the variability in the observations and the failure of linear regression to explain unit costs. This is borne out by the low R^2 value. The findings of this regression cannot be relied upon.

Sewage non-infrastructure

8.1.23 Non-infrastructure costs are predicted using pumping stations as the explanatory variable. This takes the form of a unit cost model with a weighted industry average.

Figure 8.8 – Sewage non-infrastructure updated PC10 model



8.1.24 NI Water is slightly above the England and Wales average on a unit cost basis. This results in slightly lower predicted costs than what is currently being spent.

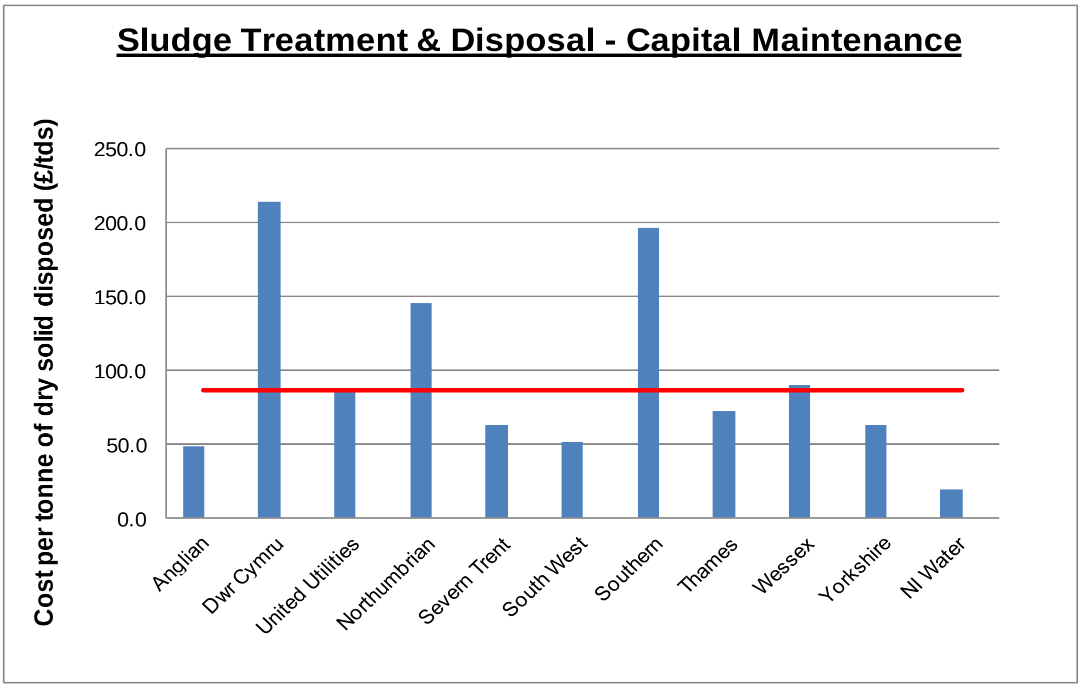
Sludge treatment and disposal

8.1.25 Capital maintenance in this area is predicted using a unit cost model. The denominator used is the amount of sludge disposed.

8.1.26 For NI Water this is a somewhat false model. The PPP contractors are responsible for all sludge disposals in Northern Ireland. They are also obliged to maintain the assets associated with disposal e.g. incinerators.

8.1.27 The service is paid for through the opex unitary charge. Consequently, the requirement for CM spend in this area is low.

Figure 8.9 – Sludge treatment and disposal updated PC10 model

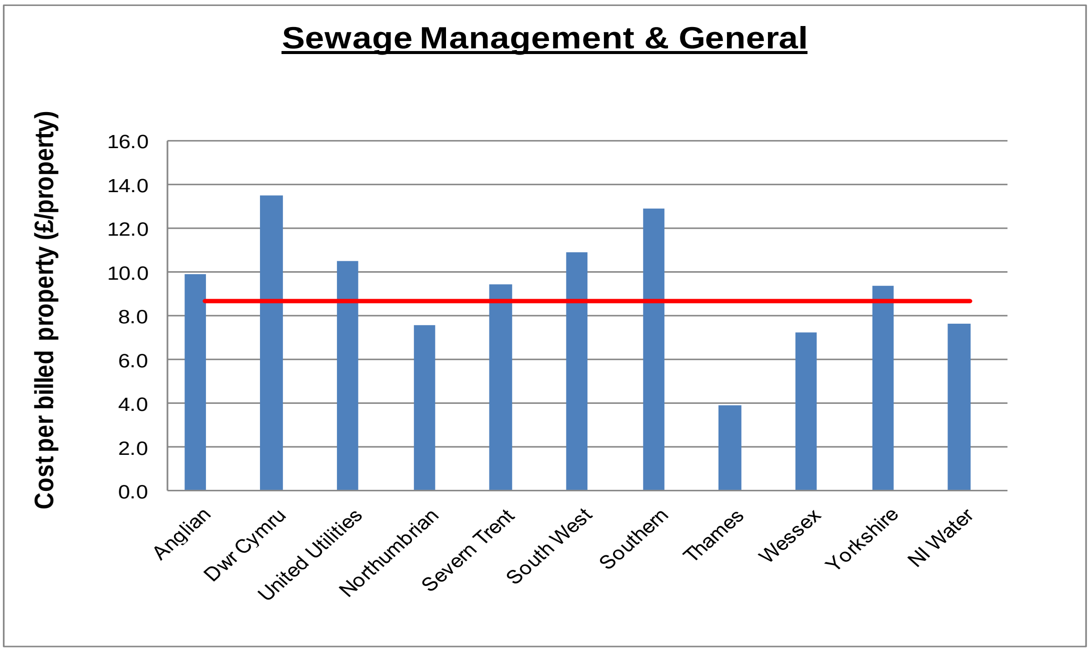


8.1.28 As anticipated, the NI Water unit cost is well below the industry average and all the other comparator companies. This indicates that a PPP adjustment is necessary.

Sewage management and general

8.1.29 Management and general expenditure is predicted using a billed property unit cost model.

Figure 8.10 – Sewage management and general updated PC10 model



8.1.30 The graphic illustrates that NI Water is close to the industry average.

8.2. Findings

8.2.1 The split of modelled and predicted costs by functional area gives average predicted costs of £89m. The table further includes estimates of upper quartile expenditure.

Table 8.6 – Predicted capital maintenance costs for NI Water using PC10 models (2012-13 prices)

Functional Area	NI Water Actual Spend (£m)	Average Predicted CM Spend (£m)	Upper Quartile Predicted CM Spend (£m)	Alternate Upper Quartile Predicted CM (£m)
Distribution Infrastructure	23.40	13.91	13.37	13.18
Distribution Non-Infra	5.97	10.19	9.80	9.65
Management & General	9.18	6.44	6.20	6.11
Resources & Treatment	4.99	9.74	9.36	9.23
Water Total	**43.53**	**40.27**	**38.73**	**38.17**
Sewerage Infrastructure	10.73	14.54	11.63	11.90
Sewerage Treatment	27.27	20.25	16.19	16.57
Sewerage Non-infrastructure	6.72	5.89	4.71	4.81
Sludge Treatment and Disposal	0.70	3.22	2.57	2.63
Management and General	4.62	5.27	4.21	4.31
Sewage Total	**50.04**	**49.16**	**39.31**	**40.22**
Total Base Maintenance	**93.57**	**89.43**	**78.04**	**78.38**

8.2.2 After adjusting for PPP and regional prices, predicted spend looks more like £81m per annum and £74m at the upper quartile. These estimates are however not totally reliable given the uncertainty around some of the regressions and the data.

Table 8.7 – Updated PC10 models predicted spend (2012-13 prices)

Updated PC10 models	NI Water Actual Spend (£m)	Average Predicted CM Spend (£m)	Upper Quartile Predicted CM (£m)	Alternate Upper Quartile Predicted CM (£m)
Water models	43.53	36.21	36.73	36.18
Sewage models	50.04	44.57	37.30	38.19
Total Base Maintenance	**93.57**	**80.78**	**74.02**	**74.36**

Figures include an adjustment for PPP and regional prices

9.0 Appendix 2 - CSV Sensitivity Modelling

9.1. Models

9.1.1 The rationale and form of the CSV modelling has already been explained. However, there is recognition that both the choice and weighting of explanatory variables is open to judgement and uncertainty.

9.1.2 As a result, the UR has run a number of sensitivity tests. These include making adjustment for the weight attached to each variable. They further involve the inclusion of more variables.

Table 9.1– CSV sensitivity testing

Model	Model Description
CSV Model A	Base case
CSV Model B	Network length / Billed Properties (50%:50% weight)
CSV Model C	Network length / Billed Properties (75%:25% weight)
CSV Model D	Network length / Billed Properties (25%:75% weight)
CSV Model E	Network length / Billed Properties / Usage (33%:33%:33% weighting)
CSV Model F	Network length / Billed Properties / Usage (30%:50%:20% weighting)

CSV Model B

9.1.3 Model B uses the same variables as the base case but allocates a simple 50%:50% weighting to each variable. Results are as follows:

Table 9.2 – Tapex regression for water – CSV Model B

Water Service:	CSV Model B	
Modelled cost:	Ln (average total water capital maintenance spend [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-5.154	0.364
Ln (csv)	1.088	0.046
Form of Model:	Ln (modelled cost) = -5.154 + 1.088 * ln {csv}	
Statistical Indicators:	Number of observations = 21	R^2 = 0.967
	Model standard error = 0.255	F test = 0.000

9.1.4 The regression does not differ greatly from the base case. The graph format shows the strong relationship between costs and the explanatory variable.

Figure 9.1 – Tapex regression for water - CSV Model B

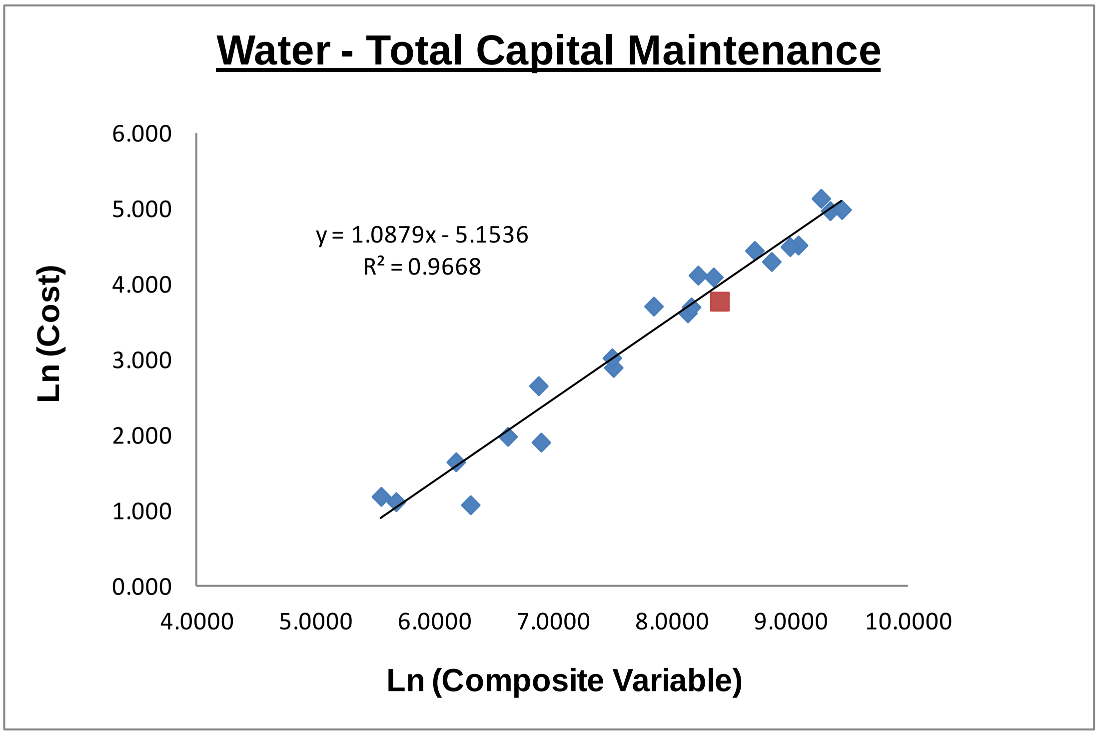
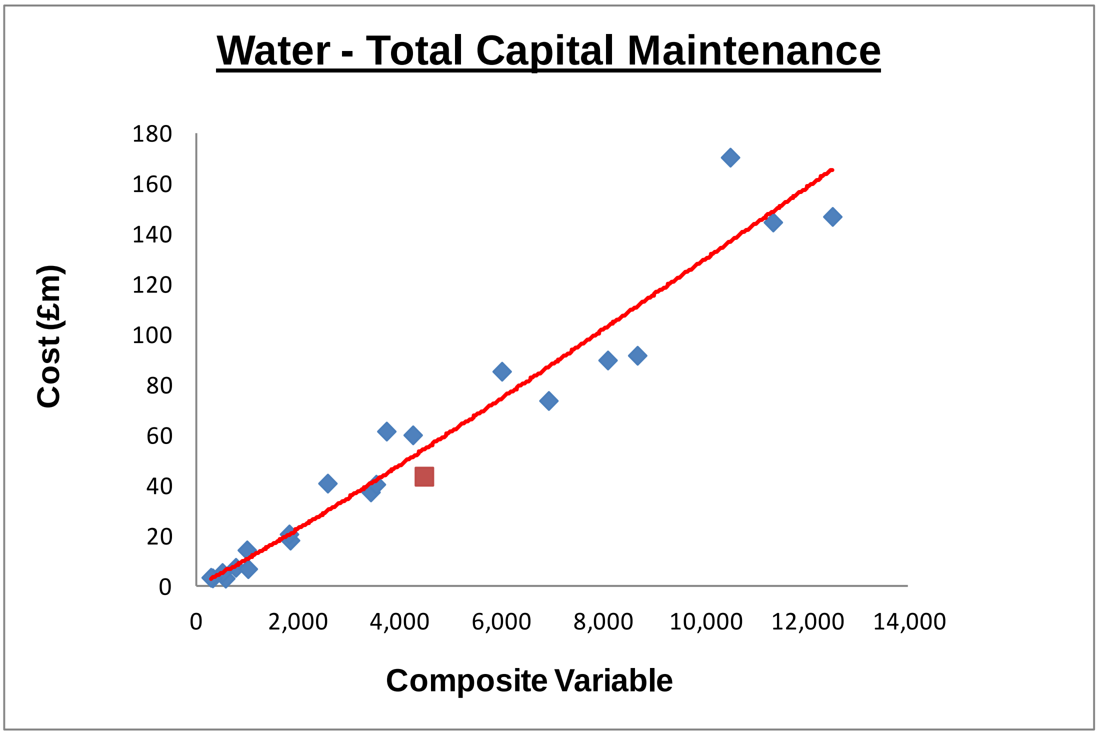


Figure 9.2 – CSV Model B – predicted versus actual water costs



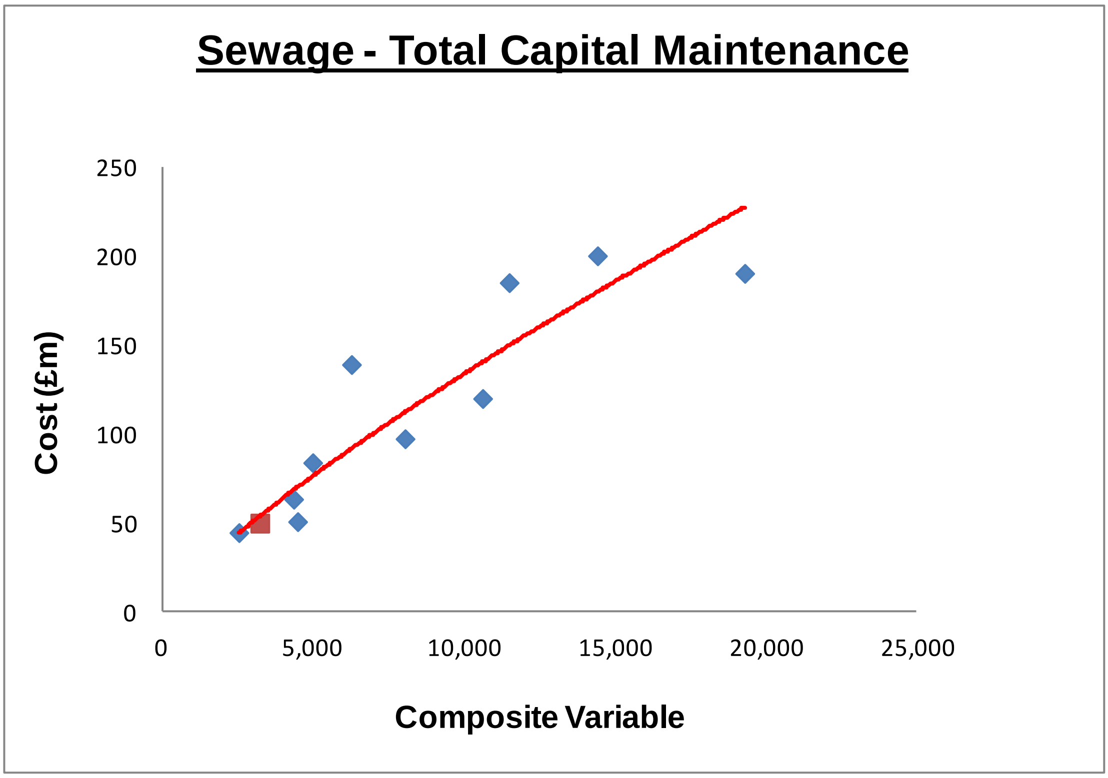
9.1.5 For sewage, the equal weighting also provides a good regression.

Table 9.3 – Tapex regression for sewage – CSV Model B

Sewage Service:	CSV Model B	
Modelled cost:	Ln (average total sewage capital maintenance spend [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-2.555	1.087
Ln (csv)	0.809	0.122
Form of Model:	Ln (modelled cost) = -2.555 + 0.809 * ln {csv}	
Statistical Indicators:	Number of observations = 10	R^2 = 0.846
	Model standard error = 0.230	F test = 0.000

9.1.6 Results indicate a high degree of correlation between costs and the explanatory variable. Whilst the regression is in log format, the graph below compares predicted against actual costs. The red point signifies NI Water and the red line is the predicted cost line.

Figure 9.3 – CSV Model B – predicted versus actual sewage costs



CSV Model C

9.1.7 Model C gives a 75% weighting to network length and a 25% weighting to properties. There does not appear to be any intuitive reason for such a weighing given the higher correlation between costs and properties. The model has however been run as a sensitivity check.

9.1.8 Results are as follows:

Table 9.4 – Tapex regression for water – CSV Model C

Water Service:	**CSV Model C**	
Modelled cost:	Ln (average total water capital maintenance spend [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-5.845	0.420
Ln (csv)	1.084	0.049
Form of Model:	Ln (modelled cost) = -5.845 + 1.084 * ln {csv}	
Statistical Indicators:	Number of observations = 21	R^2 = 0.962
	Model standard error = 0.272	F test = 0.000

Table 9.5 – Tapex regression for sewage – CSV Model C

Sewage Service:	**CSV Model C**	
Modelled cost:	Ln (average total sewage CM spend [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-2.988	1.220
Ln (csv)	0.798	0.127
Form of Model:	Ln (modelled cost) = -2.988 + 0.798 * ln {csv}	
Statistical Indicators:	Number of observations = 10	R^2 = 0.831
	Model standard error = 0.242	F test = 0.000

9.1.9 The report has not included any graphs for these models as they follow a similar pattern to the other CSV regressions.

9.1.10 Statistically the results are very similar to other models, if not quite as good. The R^2 stat on both occasions is slightly lower than the base case as is the significance of the independent variable.

CSV Model D

9.1.11 Model D again uses network length and billed properties as the composite variable. This time the weighting is 25% networks and 75% properties.

9.1.12 Again, the coefficients and their significance are similar to the other models. The form of the regressions is:

Table 9.6 – Tapex regression for water – CSV Model D

Water Service:	**CSV Model D**	
Modelled cost:	Ln (average total water CM spend [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-4.425	0.331
Ln (csv)	1.088	0.046
Form of Model:	Ln (modelled cost) = -4.425 + 1.088 * ln {csv}	
Statistical Indicators:	Number of observations = 21	R^2 = 0.967
	Model standard error = 0.254	F test = 0.000

Table 9.7 – Tapex regression for sewage – CSV Model D

Sewage Service:	**CSV Model D**	
Modelled cost:	Ln (average total sewage CM spend [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-2.088	0.965
Ln (csv)	0.818	0.117
Form of Model:	Ln (modelled cost) = -2.088 + 0.818 * ln {csv}	
Statistical Indicators:	Number of observations = 10	R^2 = 0.859
	Model standard error = 0.220	F test = 0.000

9.1.13 Both of the models provide a sound basis for predicting CM spend. Comparatively speaking, Model D appears to be on a predictive par with the base case models.

CSV Model E

9.1.14 This regression introduces usage (i.e. distribution input or load) as a part of the composite variable. Usage is considered important as the more an asset is used the faster it is likely to depreciate.

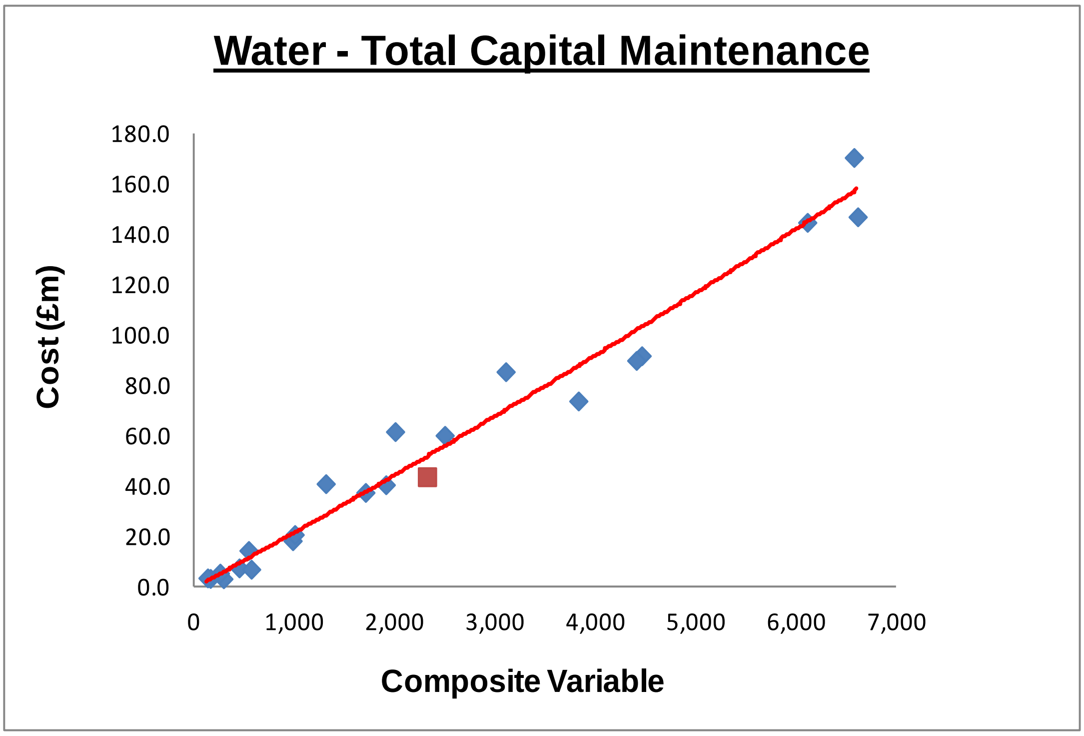
9.1.15 On this occasion, each of the three explanatory variables has been allocated an even weighting.

9.1.16 Results for the water model are:

Table 9.8 – Tapex regression for water – CSV Model E

Water Service:	CSV Model E	
Modelled cost:	Ln (average total water CM spend [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-4.352	0.341
Ln (csv)	1.070	0.047
Form of Model:	Ln (modelled cost) = -4.352 + 1.070 * ln {csv}	
Statistical Indicators:	Number of observations = 21	R^2 = 0.965
	Model standard error = 0.263	F test = 0.000

Figure 9.4 – CSV Model E – predicted versus actual water costs



9.1.17 The model appears statistically robust. The R^2 is high, while the composite variable is strongly significant. It is interesting however to note that the addition of the new variable does not seem to have increased the explanatory power compared with the base case.

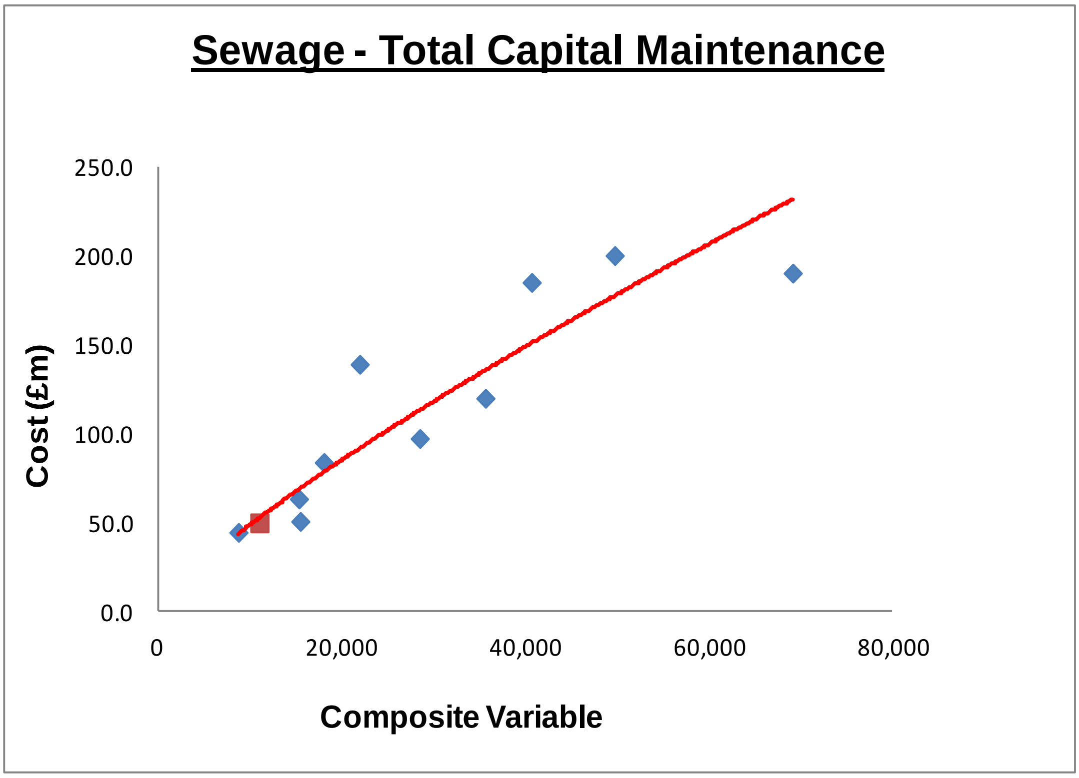
9.1.18 This may be due to the fact that properties and distribution input are closely correlated. Thus, adding usage into the CSV makes little difference to the regressions.

9.1.19 For sewage, the results are similar.

Table 9.9 – Tapex regression for sewage – CSV Model E

Sewage Service:	CSV Model E	
Modelled cost:	Ln (average total sewage CM spend [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-3.567	1.220
Ln (csv)	0.809	0.120
Form of Model:	Ln (modelled cost) = -3.567 + 0.809 * ln {csv}	
Statistical Indicators:	Number of observations = 10	R^2 = 0.850
	Model standard error = 0.227	F test = 0.000

Figure 9.5 – CSV Model E – predicted versus actual sewage costs



9.1.20 The model has statistically significant results but does not appear to improve on the base case scenario.

CSV Model F

9.1.21 Model F uses the same three components to the composite variable. This time the weightings are 30%:50%:20% for network length, properties and usage respectively.

9.1.22 There is no underlying logic behind this weighting. The analysis is merely designed to test if changing the weights will improve the robustness of model predictions.

9.1.23 The tables below illustrate the regression results.

Table 9.10 – Tapex regression for water – CSV Model F

Water Service:	**CSV Model F**	
Modelled cost:	Ln (average total water CM spend [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-4.383	0.336
Ln (csv)	1.077	0.046
Form of Model:	Ln (modelled cost) = -4.383 + 1.077 * ln {csv}	
Statistical Indicators:	Number of observations = 21	R^2 = 0.966
	Model standard error = 0.258	F test = 0.000

Table 9.11 – Tapex regression for sewage – CSV Model F

Sewage Service:	**CSV Model F**	
Modelled cost:	Ln (average total sewage CM spend [£m])	
Explanatory Variables:	Coefficient	Standard Error
Constant	-2.981	1.116
Ln (csv)	0.812	0.119
Form of Model:	Ln (modelled cost) = -2.981 + 0.812 * ln {csv}	
Statistical Indicators:	Number of observations = 10	R^2 = 0.854
	Model standard error = 0.224	F test = 0.000

9.1.24 This methodology proves slightly better than Model E, but there is no increase in predictive power compared to the base case.

9.2. Results

9.2.1 All the CSV regressions have good statistical properties. The coefficients and their significance do not vary greatly from model to model.

9.2.2 However, when looking at the predicted costs derived from each method, a reasonable level of deviation is found.

9.2.3 However, when comparing the predictive power of each model it can be seen that Models A, D and F rank slightly better than their counterparts do.

9.2.4 Results of each model are given below.

Table 9.12 – CSV sensitivity testing results (2012-13 prices)

Functional Area	NI Water Actual Spend (£m)	Average Predicted CM Spend (£m)	Upper Quartile Predicted CM Spend (£m)	Alternate Upper Quartile Predicted CM (£m)
CSV Model A	93.57	90.62	81.28	79.90
CSV Model B	93.57	100.19	88.58	88.33
CSV Model C	93.57	122.82	108.42	108.70
CSV Model D	93.57	81.78	73.72	72.49
CSV Model E	93.57	96.71	86.45	86.71
CSV Model F	93.57	90.47	81.27	80.73

Figures include an adjustment for PPP and regional prices