Water Quality Performance Commitment Review

Water Services Regulation Authority (Ofwat)

9 June 2023

Revision 3.0

Jacobs

Challenging today.
Reinventing tomorrow.

Client name:	Water Services Regulation Authority		
Project name:	Water Quality Performance Commitment Review		
Client reference:	PROC.01.0784	Project no:	B2436705
Revision no:	3.0	Prepared by:	Beth Fox, Andy McConkey, Steve Bosher, Zac Alexander
Date:	9 June 2023	Project manager:	Stathis Giannoustas

Document history and status

Revision	Date	Description	Author	Checked	Reviewed	Approved
1.0	02/03/2023	Preliminary draft	ZA, JM, AH	BGF	CJN	SG
2.0	10/03/2023	Updated draft	ZA, JM, AH			
2.1	17/03/2023	Updated draft	BF, AM			
2.2	24/03/2023	Full draft report	ZA	BGF	CJN/AM	SG
2.3	05/04/2023	Draft Final report	BF, AM	CJN	CJN	
3.0	09/06/2023	Final Report	AM	BF	ZA	SG

Executive summary

The Water Quality Performance Commitment aims to incentivise English and Welsh water companies to remove more phosphorus from rivers through partnership working and by enhancing performance at their wastewater treatment works. This review aimed to gather and analyse the data available for the wastewater treatment works aspect of the Performance Commitment. It also aimed to consider the Performance Commitment definition in light of the data and make recommendations to improve its effectiveness. The partnership working aspect of the Performance Commitment was not in scope.

We obtained the data required for baselining performance and assessed its quality. While we were able to form a sufficiently robust dataset, there were problems linking all of the data and many records were excluded for this reason.

Once the data were obtained, we analysed water company performance against phosphorus and dry weather flow permits for the period 2010 – 2021. We found that water companies typically operate a comfortable margin between their phosphorus permits and actual concentration. We found that the long-term average annual flow is approximately 1.2 times permitted dry weather flow, but this varies for different companies and for different years.

We identified that the default assumption for phosphorus concentration of 5.0 mg/l could cause a performance bonus for companies whenever a new site was added to the measure, and recommended options to prevent this bonus. We found that flow data were available for all sites which are likely to receive a new permit in future, so the dry weather flow assumption will probably only be used in limited circumstances.

We considered how the Performance Commitment calculation would work. We found that normalising performance by population could result in potential performance gains tailing off for high-performing companies and recommended consideration of three alternatives. We also compared three ways to calculate load but were not able to draw a conclusion from the data on which is the best.

Finally, we considered how the Performance Commitment could incentivise water companies. We consider it possible that companies will be incentivised to deliver early compared to the agreed delivery date, and this might increase the number of chemical dosing sites, and also the rate of chemical dosing sites where the permit is not set at Best Available Technology not entailing excessive cost. This is not necessarily the best value option for customers and the environment. However, it is noted that chemical dosing will be disincentivised by other performance commitments.

Ofwat has updated the definition from v0.1 based on our initial findings and has shared a revised definition with stakeholders. The revised definition is based on the Environment Act target. The revised definition will evaluate the reduction at treatment works with phosphorus limits plus phosphorus reduced through partnership working as a percentage of the overall phosphorus discharged at treatment works in 2020.

For the English companies the baseline load will be the same as for the Environment Act Targets. For Welsh companies this will need to be calculated assuming a 5.0 mg/l discharge where this is not known. This is a change from the V0.1 definition, from a 3-year average (2020-22) to simply 2020 for all baseline calculations, including in calculating reduction in phosphorus at treatment works.

Our summary dataset has been provided to Ofwat alongside this project report.

ⓘ Important note about your report

This Report is for the sole and exclusive use and benefit of the instructing party ("the Client") under the Agreement between the Client and Jacobs U.K. Limited ("the Consultant") and the liability of the Consultant is expressly limited as provided in the Agreement. No other party may use, make use of or rely on this Report or its contents unless otherwise expressly agreed in writing by the Consultant. No part of this Report may be copied or reproduced by any means without the prior written consent of the Consultant.

No liability is accepted by the Consultant for any use of this Report for purposes other than those for which it was originally prepared and provided under the Agreement. The data, information and assumptions used to develop and prepare this Report were obtained or derived from documents or information furnished by others. The Consultant has not independently verified or confirmed such documentation or information and does not assume responsibility for their accuracy or completeness. The Consultant has no obligation to update or revise this Report after its date of issue to reflect subsequent events, circumstances or transactions.

Use of this Report or any part of its contents, by any party other than the Client, shall be at the sole risk of such party and shall constitute a release and agreement by such party to defend and indemnify the Consultant and its affiliates, officers and employees from and against any liability whatsoever arising from its use of or reliance upon the Report or its contents. To the maximum extent permitted by law, such release from and indemnification against liability shall apply howsoever arising and regardless of cause including the fault, breach of contract, tort (including concurrent or sole and exclusive negligence), breach of duty (statutory or otherwise) strict liability or otherwise of the Consultant.

Contents

Executive summary i
Acronyms and abbreviations iv
1. Introduction 1
1.1 Background to the Performance Commitment 1
1.2 Project aim 1
2. Methodology 2
2.1 Data sources 2
2.2 Data management and assessment 2
2.3 Creating the unified datasource 2
2.3.1 Storage 2
2.3.2 English data 3
2.3.3 Welsh data 3
2.3.4 Data quality assessment 3
3. Analysis 6
3.1 Performance against phosphorus permit levels 6
3.2 Comparison of permitted dry weather flow with mean flow 10
4. Performance Commitment definition review 12
4.1 Recalculating the baseline using assumptions 12
4.1.1 Concentration 12
4.1.2 Flow 13
4.2 Performance normalisation 13
4.3 Calculating load 14
4.4 Other Performance Commitment design considerations 15
5. Recommendations 16
5.1 Dataset 16
5.2 Assumed values 16
5.3 Normalisation 16
5.4 Load calculation method 16
5.5 Performance Commitment effects 16
6. Revised performance commitment definition 17

Appendices

Appendix A. Excluded sites 18
Appendix B. Average annual mean actual DWF per year, from 2011-2021, for all permitted works per water company within England and Wales. 21

Acronyms and abbreviations

APR	Annual Performance Report
AMP	Asset Management Period
DWF	Dry Weather Flow
l/s	Litres/second
mg/l	Milligrams per litre
MCERTS	Monitoring certification scheme
NEP	National Environmental Programme
ODI	Outcome Delivery Incentive
P	Phosphorus
PC	Performance Commitment
PCL	Performance Commitment level
Q80	The value above which 80 % of the flow values occur
Q90	The value above which 90 % of the flow values occur
WIMS	Water Information Management System
WINEP	Water Industry National Environment Programme

1. Introduction

1.1 Background to the Performance Commitment

Wastewater contains phosphorus, a nutrient which can stimulate algal growth to the detriment of water bodies. To help control this issue, in many cases the relevant agency (the Environment Agency for England or Natural Resources Wales for Wales) specifies limits on the concentration of phosphorus which may be discharged from wastewater treatment works. These works must be built and operated to remove phosphorus, in addition to other pollutants and contaminants. Permit concentration limits typically depend on the sensitivity of the receiving waterbody to phosphorus and the population equivalent of the agglomeration the works serves. Water companies are funded to address these permit requirements.

The Westminster government has set a target to reduce phosphorus loadings from wastewater treatment by 80% from 2020 levels by 2038. Some of this might occur due to reduction in permitted concentrations or introduction of permits for sites which did not previously have them. Meeting this target could also require further reductions in phosphorus concentrations below permit levels.

The purpose of Ofwat's new River Water Quality Performance Commitment is to incentivise water companies to reduce phosphorus loadings. It will therefore also support the Westminster Government's target and lead to a healthier inland aquatic environment in England and Wales.

The Performance Commitment will be measured with two components:

- Reduction in phosphorus load from wastewater treatment.
- Phosphorus load reduction achieved by water companies working in partnership with others to achieve reductions.

It is important to note that the Performance Commitment Level (PCL) will be set to ensure that companies are not rewarded for load reductions that are funded through business plans in Period Review 2024 (PR24). This performance commitment is intended to reward companies for additional load that is removed beyond that funded through PR24

This review of the Performance Commitment covers only the reduction in load from wastewater treatment.

This part of the Performance Commitment will operate by calculating the additional phosphorus load removed from final effluent compared to a 2020 baseline, with the performance commitment levels (PCLs) accounting for the effect of reductions funded as a result of new or tightened permit requirements. The intention is for water companies to be ambitious in efficiently reducing phosphorus loadings.

This is a new Performance Commitment and, prior to it going live, Ofwat commissioned Jacobs to review its definition and the data available for it. We were provided with version 0.1 of the Performance Commitment definition dated 22 December 2022[1].

1.2 Project aim

This project's aim was to support implementation of the new Performance Commitment by completing these tasks:

- Prepare a dataset of permitted phosphorus concentration, sample data, permitted flow and actual flow for wastewater treatment works in England and Wales.
- Establish baseline effluent phosphorus load for each wastewater treatment works.
- Critically assess the available dataset for its ability to support the objective of the Performance Commitment.
- Review whether and how the Performance Commitment definition could be improved to better meet its purpose.
- Review the default assumptions for the baseline for sites where historic data are not available.
- Assess how phosphorus discharges relate to discharge permits and how this relationship should be considered in setting Performance Commitment levels.

[1] Ofwat Performance Commitment definition v0.1, 22/12/2022 River_water_quality_PC_definition.pdf (ofwat.gov.uk)

2. Methodology

2.1 Data sources

The Performance Commitment is intended to work with existing data and data collection processes. We therefore drew on existing data sources for permits and performance as shown in Table 1. For phosphorus performance we used regulatory sample records. For flow performance we used certified regulatory effluent flow.

Table 1 Data sources requested

Name	Source
Current and revoked phosphorus discharge permits	England: Environment Agency WIMS database Wales: Datamap Wales - Consented Discharges to Controlled Waters with Conditions
Phosphorus discharge performance 2010 – 2022	England: Environment Agency WIMS database. Wales (Welsh Water region): Freedom of Information Request from Welsh Water Wales (Hafren Dyfrdwy region): Freedom of Information Request from Hafren Dyfrdwy (1)
Permitted dry weather flow	England: Environment Agency WIMS database Wales (Welsh Water region): Freedom of Information Request from Welsh Water Wales (Hafren Dyfrdwy region): Freedom of Information Request from Hafren Dyfrdwy (1)
Weekly average flow	England: Freedom of Information Request from Environment Agency (2) (Environment Agency MCERTS data) Wales (Welsh Water region): Freedom of Information Request from Welsh Water Wales (Hafren Dyfrdwy region): Freedom of Information Request from Hafren Dyfrdwy (1)

Table notes

There is an overlap between the England dataset and the Welsh dataset where 42 Welsh Water sites are located in England. For all these sites we used the Welsh Water data.

(1) We did not receive the data from Hafren Dyfrdwy in time for this project.

(2) We received annual average Total Daily Flow, Q80 and Q90 data from the Environment Agency for 2011-2021.

2.2 Data management and assessment

We created a data management approach to support the project objectives through our analysis, and to enable Ofwat to undertake its own analysis on the data produced.

Our workflow was:

1. Specify the data required. Identify source and acquire it, either through download from public source or Freedom of Information Request.
2. In parallel, create the database schema.
3. Copy the data to the datasource and link it.
4. Complete quality checks.
5. Produce annual summary data.
6. Proceed with analysis.

2.3 Creating the unified datasource

2.3.1 Storage

We considered several approaches for data storage. Given the quantity of data across the sources and the need for high-efficiency processing, we selected SQL Server as the storage solution. This provided several benefits, including the ability to control user access for security, the ability to work on the data directly in Python, and the ability to export data to the required final product format in Excel. It also enabled separation of raw source data from processed data and version control.

2.3.2 English data

We followed this process for the English data:

1. We standardised the water company names.
2. We linked the sample data to the permits using the permit reference, version, outlet number and effluent number.
3. We extracted the phosphorus permit and actual performance data.
4. We extracted the dry weather flow permits. To fill missing values in the Environment Agency Monitoring data we combined it with Environment Agency MCERTS data.
5. We sorted the permit and compliance data by calendar year. For each year, we found the most recent effective permit. Where a permit became effective part way through a year we ignored the previous permit during that year. Many sites have multiple phosphorus permits over the period of the study. We stored the date range for each of these to keep them as separate entities within the datasource and linked each compliance sample with the permit that was in force at the time.
6. We joined the metadata such as catchment name, waterbody name and discharge site name to the permit data.
7. For phosphorus only, we calculated the number of samples and mean, minimum and maximum performance for each permit. All samples were included in calculations, irrespective of their compliance purposes. For example, if a site had samples for both operator self-monitoring, and Urban Wastewater Treatment Directive composite samples, values from both sample purposes were included in the calculations. Where a permit had multiple outlets or effluents with samples, we recorded their performance separately. In all other cases we reported performance against permit or version.
8. We assigned the permit version to the samples based on the date of sample.
9. For the MCERTS data, we joined all the years into one table and associated them with a station number using the 2021 WIMS permit number.
10. For flow performance, we extracted mean, Q80 and Q90 for each year.
11. We identified any sites with missing compliance or permit data and flagged them for exclusion from the analysis. Appendix A lists those sites that were excluded from the analysis.
    a. Where sites had a gap in phosphorus permit data (e.g. phosphorus permit of 1.0 mg/l in 2011 and 2012, but then no reported phosphorus permit until 2020 and 2021), we flagged these sites for exclusion from any permitted load based analysis (but kept them for concentration based analysis)
    b. Where sites had a gap in Dry Weather Flow (DWF) permit data we flagged these sites for exclusion from any load-based analysis (but kept them for concentration-based analysis)
    c. Where sites were missing phosphorus compliance or flow compliance data, we flagged these for exclusion from any load-based analysis
    d. Any duplicate sites were flagged for further review and potential removal from the analysis. See section 2.3.4 for further details.

2.3.3 Welsh data

We followed a similar process to the English data with the following exceptions:

1. We linked the permit data and DWF permit data through the Permit ID. Although the flow and phosphorus results data do not have an ID field for linking, we were able to link them using the site name.
2. After manually verifying and updating any mismatches, we identified all sites with a phosphorus permit in all datasets.
3. We created a reference table to link the data using the site name from each dataset with sample point number and Permit IDs.
4. We ignored outlet numbers and effluent numbers because we didn't have performance data for them.
5. The flow data for Wales was in m^3/day averaged over the week. We assigned a start date for each week and associated each week with a permit version.
6. We averaged the flow for each permit over each year and converted it to l/s.

2.3.4 Data quality assessment

Once the data were stored, we undertook a quality assessment covering all of the imported data. We first established in how many instances we were unable to link permits with compliance and determinand data. Following this, we undertook an anomaly assessment of the data to further gauge its quality and ensure that anomalous data were excluded. We searched for:

- Duplicate records.

- Missing DWF and P permit values.
- Overlaps between England and Wales data.

The quality of data has improved between 2011 and 2021, with 2020 being the most complete dataset. Quality of data then deteriorated in 2021 again, but this is believed to be due to issues in data matching rather than evidence that the data do not exist. Our opinion is that all of the data from sites excluded from the analyses are likely to be due to issues with how completely we have been able to link the permit database to water quality compliance data and flow compliance data, rather than actual missing data. Therefore, we believe that Water Companies will be able to complete the missing data during data validation. Appendix A details the list of sites excluded from the analysis and gives the reasons for exclusions.

Table 2 below shows what proportion of sites for 2020 and 2021 are missing essential data (P compliance and/or flow data), potentially essential data (permit data), or useful contextual data.

A full breakdown of reasons why we consider the data is essential, potentially essential or contextual is given in Table A1 in Appendix A. Whilst a significant proportion of sites are excluded from the analyses for some companies (as shown in Figure 1 and Table 2 below) we consider that there is a sufficient dataset remaining after exclusions to draw robust conclusions for the Performance Commitment definition. We believe that all of the missing data can be completed by companies' during validation, making it a complete dataset for use in the PR24 process.

Figure 1 Data completeness by company, 2020 and 2021

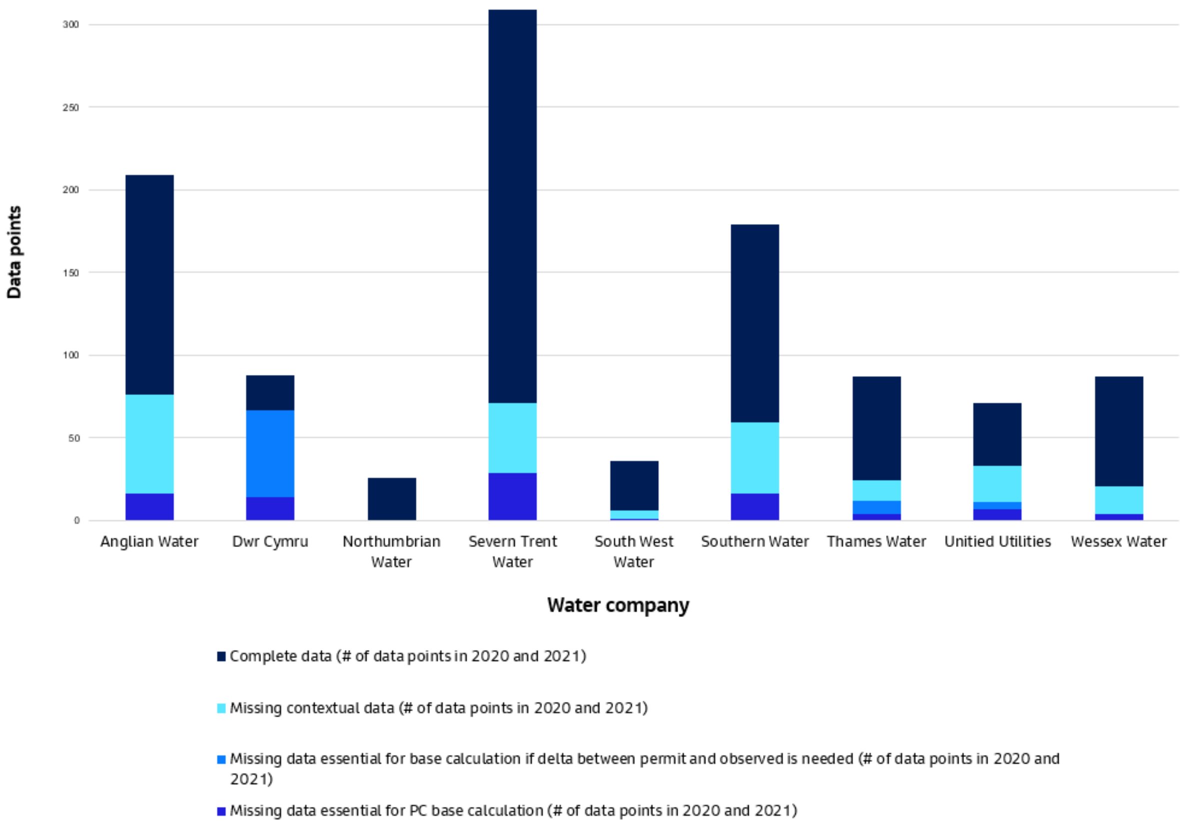


Table 2 Number of and percentage of missing data points in 2020 and 2021

Data point	Yorkshire Water	Wessex Water	Unitied Utilities	Thames Water	Southern Water	South West Water	Severn Trent Water	Northumbrian Water	Dwr Cymru	Anglian Water
Missing data essential for PC base calculation (# of data points in 2020 and 2021)	2	4	7	4	16	1	29	0	14	16
Missing data essential for base calculation if delta between permit and observed is needed (# of data points in 2020 and 2021)	0	0	4	8	0	0	0	0	53	0
Missing contextual data (# of data points in 2020 and 2021)	8	17	22	12	43	5	42	0	0	60
Complete data (# of data points in 2020 and 2021)	40	66	38	63	120	30	238	26	21	133
Total missing data (# of data points in 2020 and 2021)	10	21	33	24	59	6	71	0	67	76
Grand Total (# of data points in 2020 and 2021)	50	87	71	87	179	36	309	26	88	209
Missing data essential for PC base calculation (% of data points in 2020 and 2021)	4.0%	4.6%	9.9%	4.6%	8.9%	2.8%	9.4%	0.0%	15.9%	7.7%
delta between permit and observed is needed (% of data points in 2020 and 2021)	0.0%	0.0%	5.6%	9.2%	0.0%	0.0%	0.0%	0.0%	60.2%	0.0%
Missing contextual data (% of data points in 2020 and 2021)	16.0%	19.5%	31.0%	13.8%	24.0%	13.9%	13.6%	0.0%	0.0%	28.7%
Total missing data (% of data points in 2020 and 2021)	20.0%	24.1%	46.5%	27.6%	33.0%	16.7%	23.0%	0.0%	76.1%	36.4%

3. Analysis

3.1 Performance against phosphorus permit levels

We established how companies tend to perform against their permits.

Figure 2 shows which companies treated phosphorus in 2020 and the range of their permit limits. It shows that some companies will have greater exposure to the wastewater treatment element of the Performance Commitment than others, since performance is driven by outperformance against permits.

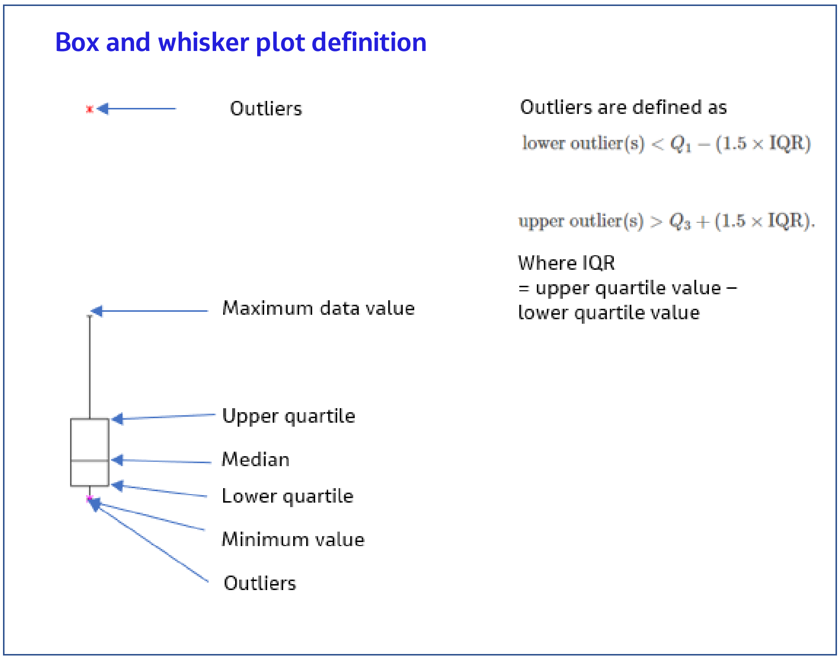
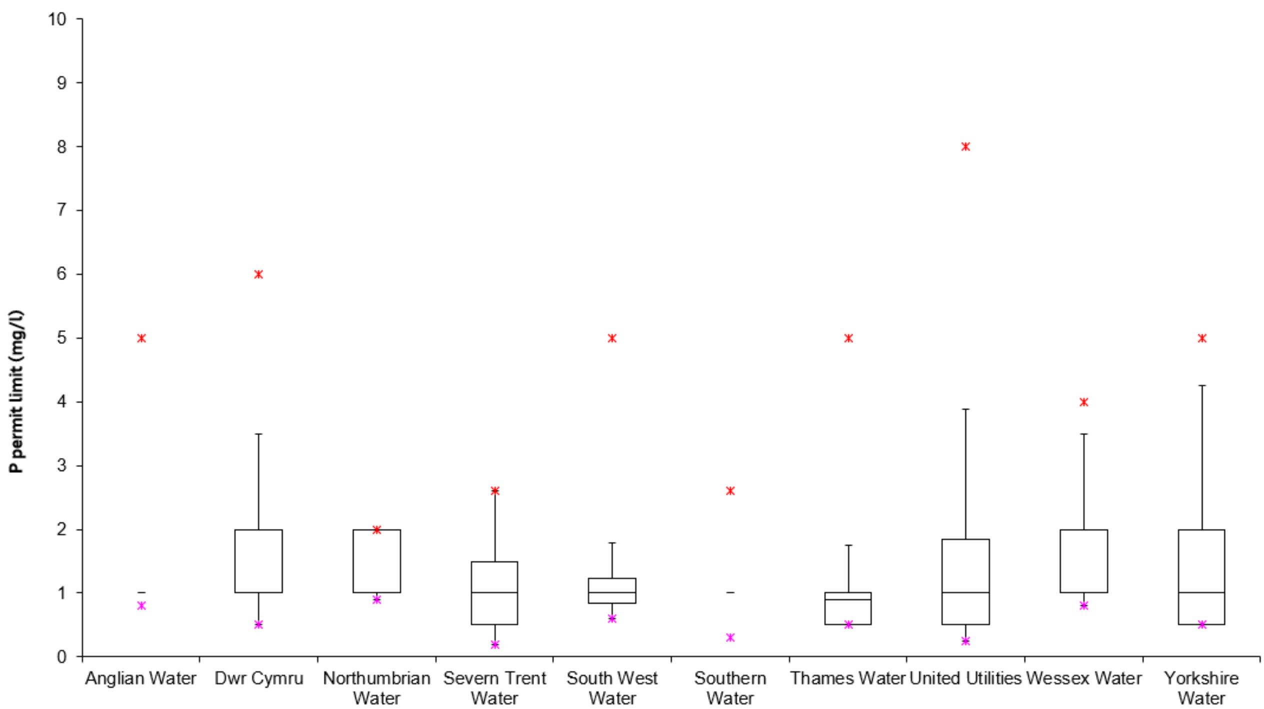


Figure 2 Permit limit variability (in 2020) by company



This chart does not show the amount of work needed by companies to outperform, because this also requires consideration of the scale and type of permitted treatment works.

For companies in England and Wales, mean performance against permit levels is generally stable (Figure 3). A breakdown of this by company (shown in Figure 4 and Table 3) shows that there is some variation by company, but usual practice is to operate a comfortable margin between the actual and permitted concentrations. This not intended to allow a detailed assessment of each company; it is to illustrate the degree of variance behind the average performance.

Figure 3 Mean P permit outperformance across all water companies in England and Wales from 2011 to 2021

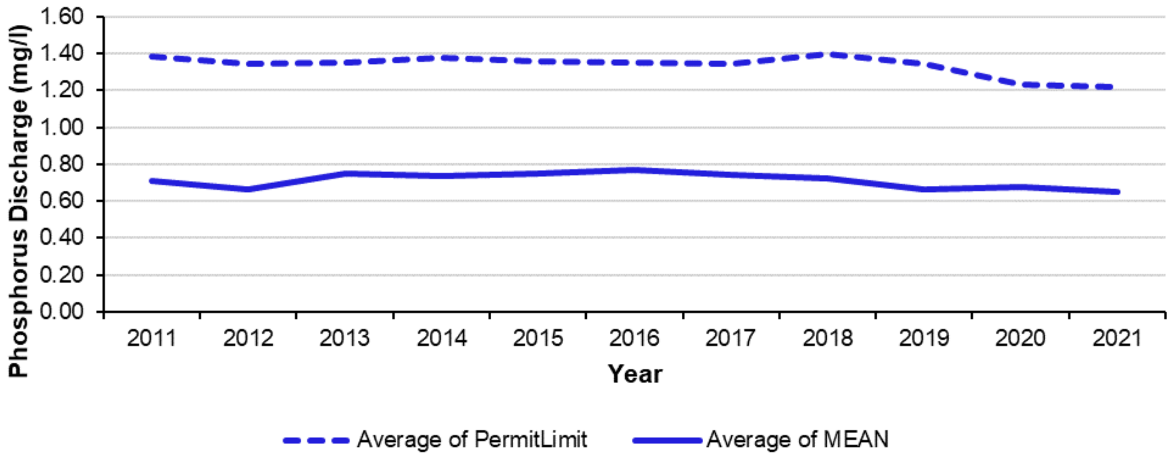


Figure 4 Mean P permit outperformance from 2011 to 2021 for each water company in England and Wales

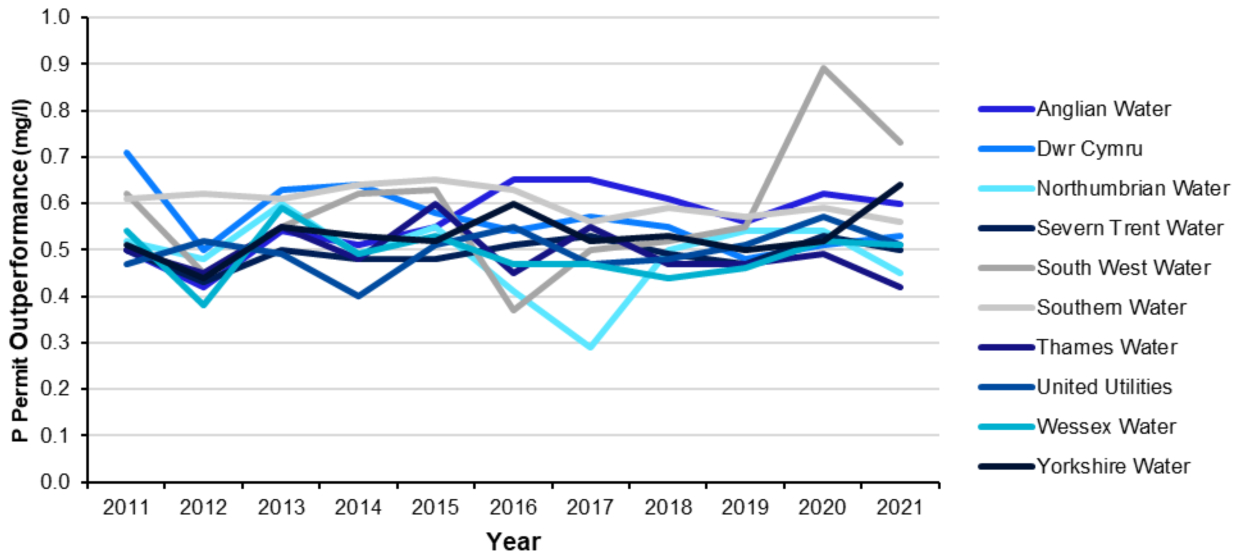
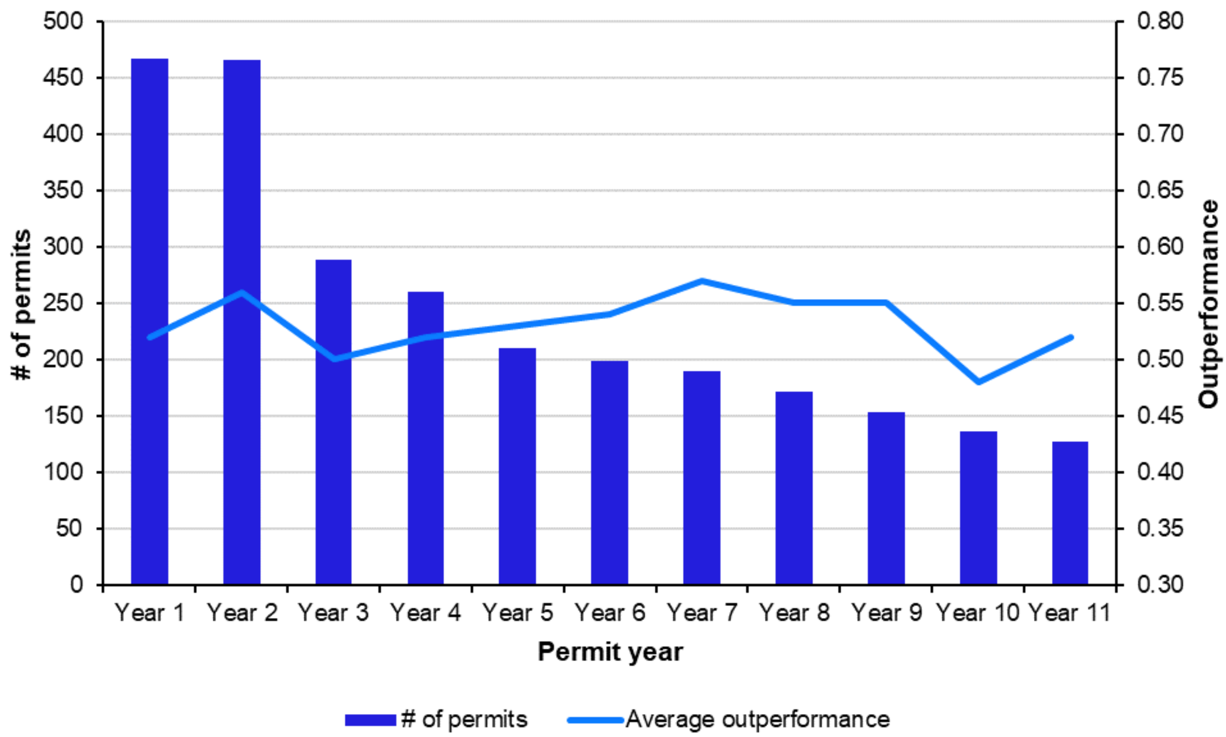


Table 3 Reported average phosphorous concentration / permitted phosphorus concentration by company

Year	Anglian Water	Dŵr Cymru	Northumbrian Water	Severn Trent Water	South West Water	Southern Water	Thames Water	United Utilities	Wessex Water	Yorkshire Water
2011	0.50	0.71	0.52	0.50	0.62	0.61	0.50	0.47	0.54	0.51
2012	0.42	0.50	0.48	0.43	0.45	0.62	0.45	0.52	0.38	0.44
2013	0.54	0.63	0.60	0.50	0.55	0.61	0.55	0.49	0.59	0.55
2014	0.51	0.64	0.49	0.48	0.62	0.64	0.48	0.40	0.49	0.53
2015	0.55	0.58	0.55	0.48	0.63	0.65	0.60	0.51	0.53	0.52
2016	0.65	0.54	0.41	0.51	0.37	0.63	0.45	0.55	0.47	0.60
2017	0.65	0.57	0.29	0.53	0.50	0.56	0.55	0.47	0.47	0.52
2018	0.61	0.55	0.50	0.49	0.52	0.59	0.47	0.48	0.44	0.53
2019	0.56	0.48	0.54	0.47	0.55	0.57	0.47	0.51	0.46	0.50
2020	0.62	0.51	0.54	0.53	0.89	0.59	0.49	0.57	0.52	0.52
2021	0.60	0.53	0.45	0.50	0.73	0.56	0.42	0.51	0.51	0.64

To find out how performance against permit changes with permit age we rebased performance to have the first year of a permit or a permit tightening as Year 1. We found that there is no significant correlation between the number of years a permit has been established and the company's outperformance of it (Figure 5).

Figure 5 Permit outperformance trend across the duration of the permit by permit year



A further exploration of performance against permit-by-permit concentration showed that tighter permits generally lead to a narrower spread of performance (Figure 6). We chose bands of 0.5 mg/l for this analysis because this gave a sufficient number of permits within each band, and because permits are often set in 0.5 mg/l intervals.

Figure 6 Treatment works performance across England and Wales for sites with P permits and sampling by P permit range.

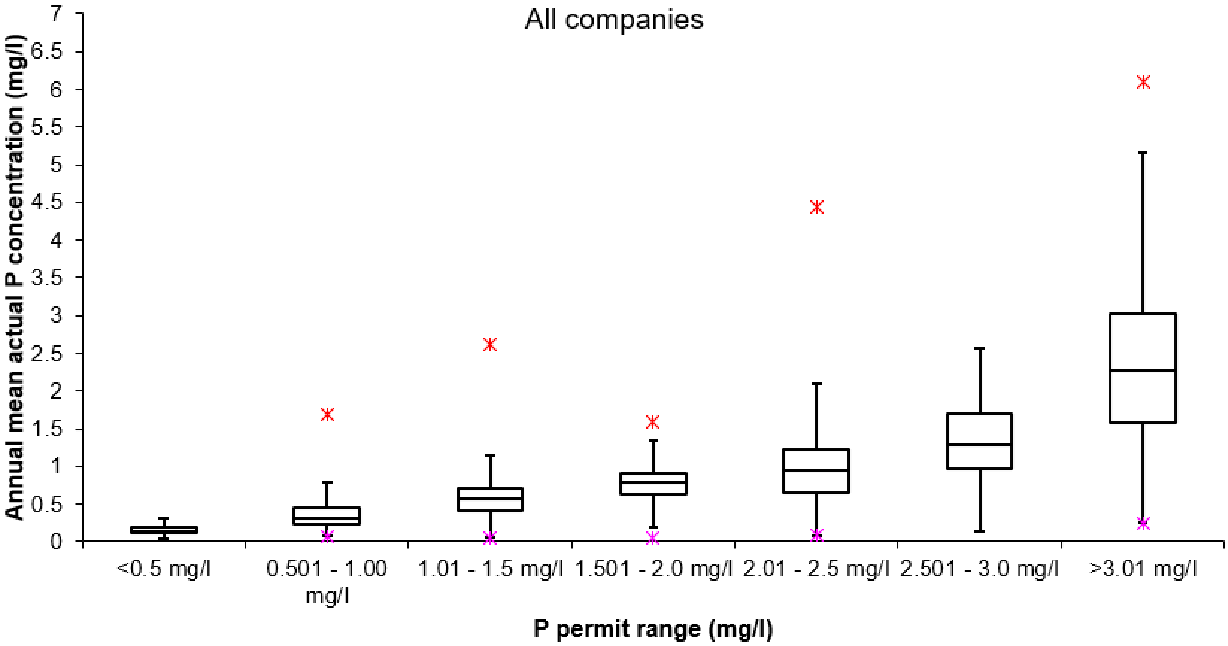


Figure notes

1 – Only sites with missing P permit and/or P compliance data was excluded from this analysis

2 – Where no max outlier is shown, this is because the max outlier = the top of the whisker

We also looked at how relative outperformance varied by permit band for all companies and found that the relative outperformance was relatively stable by permit band. This is shown in Figure 7.

Figure 7 Treatment works relative performance across England and Wales for sites with P permits and sampling by P permit range

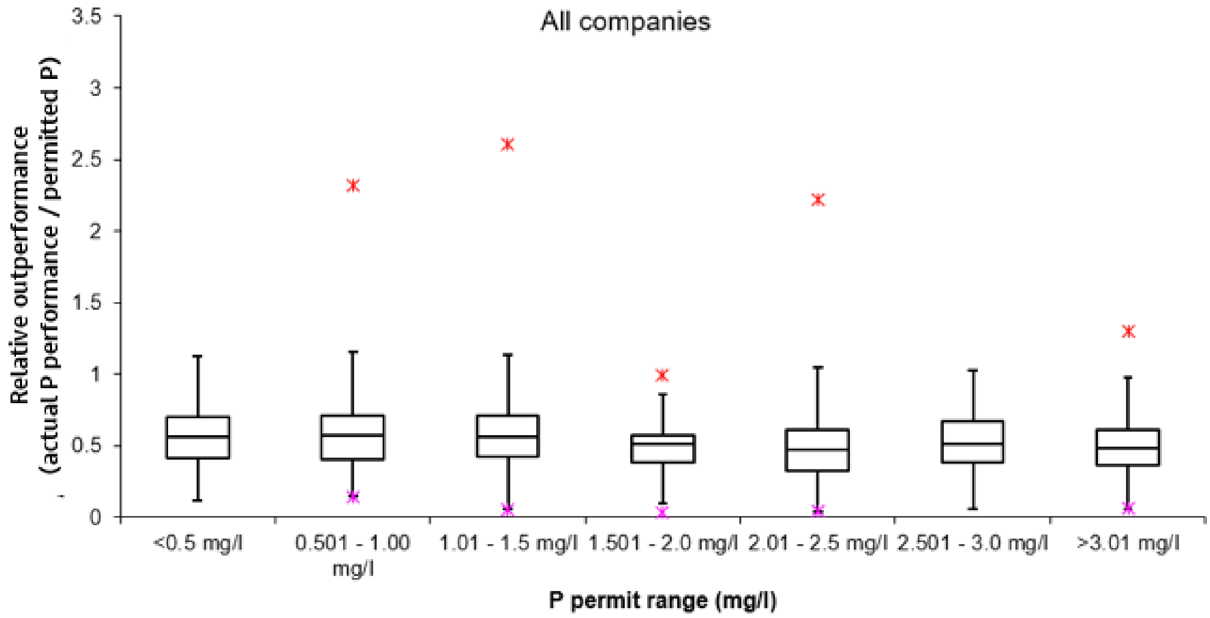


Figure notes

1 – Only sites with missing P permit and/or P compliance data was excluded from this analysis

2 - Where no max outlier is shown, this is because the max outlier = the top of the whisker

3 - Relative outperformance = average of (Mean P in mg/l from compliance data / P permit) calculated for all sites for all years.

3.2 Comparison of permitted dry weather flow with mean flow

The assumption in the Performance Commitment definition is that the mean flow average is 1.2 times the permitted dry weather flow. This will be used to backfill site performance where flow was unknown when the baseline was established. While the long-term average is close to 1.2, there is significant variation between companies (Figure 8). We compared mean flow with dry weather flow across all sites to determine whether this assumption is representative. The data are shown in Table 4. See Appendix B for average annual mean flow by water company for all permitted works in England and Wales from 2011 to 2021.

Figure 8 Average annual flow against permitted dry weather flow for all companies 2011 – 2021.

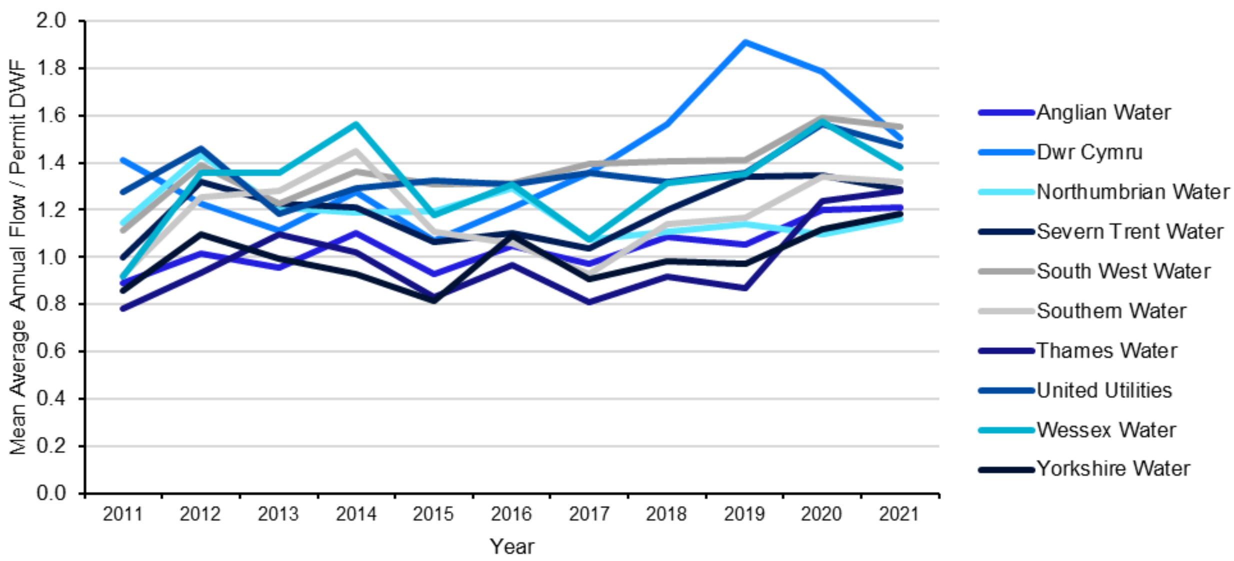


Table 4 Reported average flow / permitted dry weather flow by company

Year	Anglian Water	Dŵr Cymru	Northumbrian Water	Severn Trent Water	South West Water	Southern Water	Thames Water	United Utilities	Wessex Water	Yorkshire Water
2011	0.9	1.4	1.1	1.0	1.1	0.9	0.8	1.3	0.9	0.9
2012	1.0	1.2	1.4	1.3	1.4	1.3	0.9	1.5	1.4	1.1
2013	1.0	1.1	1.2	1.2	1.2	1.3	1.1	1.2	1.4	1.0
2014	1.1	1.3	1.2	1.2	1.4	1.4	1.0	1.3	1.6	0.9
2015	0.9	1.1	1.2	1.1	1.3	1.1	0.8	1.3	1.2	0.8
2016	1.1	1.2	1.3	1.1	1.3	1.1	1.0	1.3	1.3	1.1
2017	1.0	1.4	1.1	1.0	1.4	0.9	0.8	1.4	1.1	0.9
2018	1.1	1.6	1.1	1.2	1.4	1.1	0.9	1.3	1.3	1.0
2019	1.1	1.9	1.1	1.3	1.4	1.2	0.9	1.4	1.4	1.0
2020	1.2	1.8	1.1	1.3	1.6	1.3	1.2	1.6	1.6	1.1
2021	1.2	1.5	1.2	1.3	1.6	1.3	1.3	1.5	1.4	1.2

We approximated size bands from the flow data. This showed that smaller sites tend to operate further below their dry weather flow permit than larger sites (Table 5).

There is significant variation in the ratio between permitted dry weather flow and annual average flow. This variation exists between companies, between years and between size of works.

Table 5 Reported average flow / permitted dry weather flow by approximate works size band

Year	Band 1	Band 2	Band 3	Band 4	Band 5	Band 6
2011	1.0	1.1	1.0	0.9	1.1	1.0
2012	1.5	1.4	1.3	1.1	1.3	1.1
2013	1.4	1.5	1.3	1.0	1.1	1.1
2014	1.6	1.3	1.4	1.1	1.2	1.2
2015	1.2	1.0	1.2	1.0	1.2	1.0
2016	1.4	1.1	1.2	1.0	1.2	1.1
2017	1.4	1.1	1.1	1.0	1.3	0.9
2018	1.3	1.3	1.3	1.1	1.3	0.9
2019	1.3	1.4	1.3	1.2	1.4	1.0
2020	1.6	1.4	1.4	1.3	1.5	1.1
2021	1.5	1.3	1.4	1.3	1.5	1.1

4. Performance Commitment definition review

4.1 Recalculating the baseline using assumptions

In version 0.1 of the Performance Commitment:

- The baseline is the annual average phosphorus load for permitted sites in the base period (2020 to 2022).
- Performance is the difference between actual load in the base period and actual load in the performance year.
- When a site receives its first permit, it will be added to the baseline as if it had been permitted in 2020.
- If actual load in 2020 cannot be calculated because phosphorus concentrations or flow were not measured to regulatory standards, assumed values for flow (1.2 times permitted dry weather flow) and concentration (5.0 mg/l) will be used.

We considered whether these assumptions were supported by the data and what effects they might have. These considerations apply both to newly-permitted sites and to updated permits.

4.1.1 Concentration

We found no instances of sites having phosphorus compliance data while there was no permit in effect, indicating that the phosphorus concentration assumption is likely to be used every time a site receives its first phosphorus permit.

Because there were no samples without permits, we were not able to test the assumption of 5.0 mg/l against regulatory reported sample data. However, we have found no evidence to disprove the 5.0 mg/l assumption.

The effect of this assumption will be to create a reduction in apparent load once the permit comes into effect, from 5.0 mg/l to the new performance level. Companies are separately funded to meet the new permit (usually through WINEP or NEP), so this represents potential double funding of the improved performance.

There are two potential mechanisms to prevent this double counting:

Set the Performance Commitment based on permit level rather than 5.0 mg/l

This would use the permit in 2020 (or subsequently amended before 2025) as the baseline from which load reduction would be assessed. This would require the baseline to re-assessed every year as new permits come into force, and a changing baseline could make performance results look erratic.

Set the Performance Commitment based on 5.0 mg/l, and use the Performance Commitment Level to prevent double counting

Setting the Performance Commitment Level to account for already-funded performance would show the total improvement in performance including reductions from permit compliance, but only reward performance where companies had gone further than this.

For example, a hypothetical company with one treatment works is funded to introduce a consent ready for the start of 2029 at 4.0 mg/l. To reflect stretching performance, the Performance Commitment Level could be set based on a concentration of 2.5 mg/l. This would create a performance expectation of outperformance of 40% against the permit and a total reduction in load of 50%.

This could bring the following possible outcomes for the company:

- If it did not deliver a reduction, it would have revenue reduced by underperformance payments.
- If it delivered the reduction on time, it would be performance neutral.
- If it outperformed and delivered on time but with a total reduction of 60% it would receive outperformance for the extra 10%.
- If it delivered a 60% reduction in 2026 and 2027 it would receive additional outperformance for those two years, before falling back to 10% from 2029.

This would require the companies and the Environment Agency to agree a realistic regulatory delivery date, and not just set the regulatory date in the last year of the AMP.

4.1.2 Flow

Average total daily flow is available for all sites that are likely to have a P permit imposed in the future, and the data are reported to the Environment Agency annually. This means that the flow assumption will probably not be used to calculate actual performance. We recommend that actual average total daily flow is used, except in the rare circumstance where MCERTS data is not available. If this is the case, the assumption of 1.2 times the permitted dry weather flow would need to be used.

4.2 Performance normalisation

In version 0.1 of the definition, performance is expressed as a load per head of population. This normalisation is to allow comparison between the companies and ensure that companies with a large population are not unfairly advantaged compared to those with small population, because total phosphorus load is generally correlated with population. This normalisation also has the effect of reducing the incentive to outperform at lower loads. For example, halving load from 1.0 kg to 0.5 kg per head is a 0.5 kg/h performance improvement. Halving it again is only a 0.25 kg/h improvement.

This effect will tend to flatten performance at low consents rather than pushing further outperformance, focusing companies' efforts on reductions on sites which have the highest loads. An alternative Performance Commitment definition which will continue to reward outperformance at lower loads and still allows comparison between companies is to measure the percent reduction since 2020. This would be simpler to calculate because it requires fewer inputs but has the drawback of ignoring the scale of the reduction.

Under any definition, it will be more difficult for companies which had low permitted concentrations in 2020 to outperform.

Also under either definition, it will be more difficult for companies with many permitted works to outperform. For example, a company with ten permitted sites would need to achieve an average reduction across all of them, while a company with just one site would only need to achieve the same proportional reduction on that one. If the outcome is focused on actual load removed, and the ODIs are calibrated on load removed (i.e. £ per kg), this risk will be managed.

These effects are demonstrated in Table 6. In this scenario, all three companies achieve the same percent reduction in load, but to do that Company B has had to reduce load by ten times more units than Company A. Company C started from a lower baseline so it has only achieved half the performance of the other two, when normalised by population. This would need to be corrected by company specific Performance Commitment Limits and Outcome Delivery Incentive rates.

Table 6 Fictional comparison of company performance

Company	Number of permitted works	Population	Baseline load	Reported load	Performance normalised by population	Percent improvement	Load reduction
Company A	1	10	2	1	0.10	50%	1
Company B	5	100	20	10	0.10	50%	10
Company C	1	10	1	0.5	0.05	50%	0.5

In terms of achieving the environment outcomes, total reduction in load is more important than the comparison between companies, although that competition between companies could lead to more reduction in load.

To go further towards rewarding outcomes rather than outputs, the measure could be adapted to incentivise reductions where they are most likely to cause environmental benefits. For example, discharges into sensitive waterbodies or waterbodies which fail to achieve good environmental status due to phosphorus could be given a greater weighting. We note that environmental impacts are already part of the Environment Agency's assessment of what works should have P permits.

4.3 Calculating load

Load is the product of flow and concentration. Calculating an average load from a limited number of samples presents challenges.

The perfect calculation would have continuous flow rates multiplied by continuous phosphorus measurements. This is not practicable. What we have is:

- All phosphate samples taken at a site, irrespective of their compliance purpose (for example concentration taken as a spot sample either 12 or 24 times per year and UWWTD composite samples).
- Mean daily flow averaged over each week for Wales.
- Average annual flow (l/s) for England.

There will be many different flow and concentration conditions represented in the data. Daily flow varies significantly over timescales of less than a week, depending mainly on rainfall and catchment response times. Also, higher flow rates lead to dilution of the load and thus a lower phosphorus concentration. For individual sites, this variation and the relatively small number of samples contribute to large apparent variance in the calculated load year-to-year. High volume/low concentration flows require less work to achieve low effluent P concentrations and could therefore lead to higher loads, but we were not able to find any correlation between annual rainfall and load.

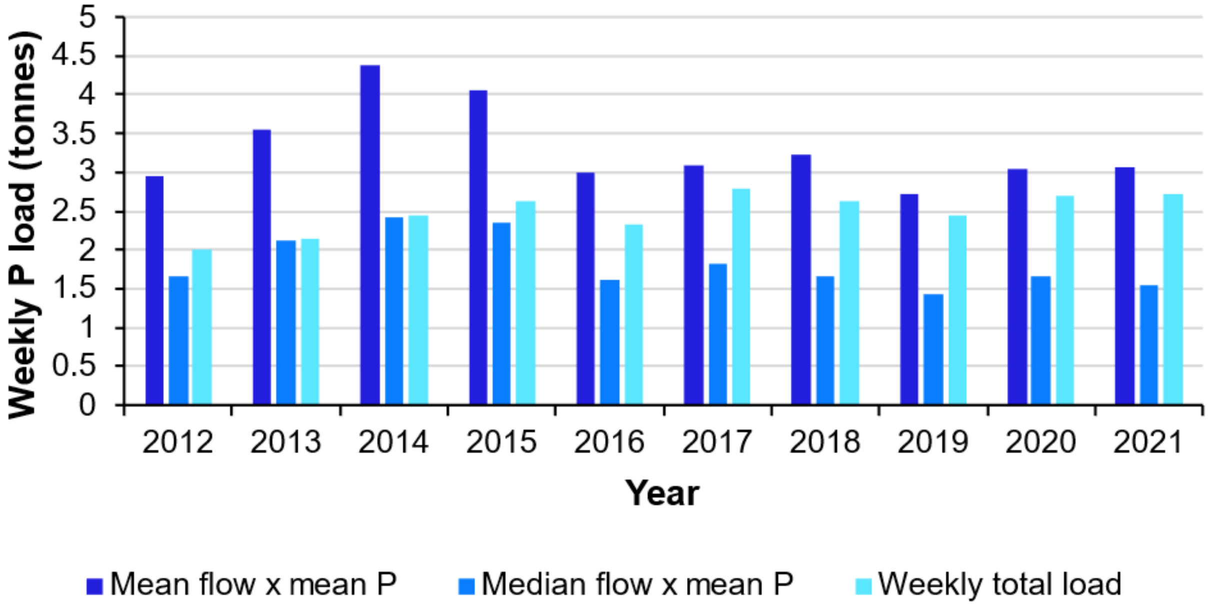
Our view is that, at the company level, the number of samples is probably acceptable because sampling is randomised, and fluctuations in individual site reading will most likely be averaged out.

There are at least three different ways to calculate the annual load from the available data, and we tested whether at the company level there is a significant difference in the method used.

- Mean: We multiplied the mean concentration by the mean weekly flow and converted the result to kg/year.
- Median: We multiplied the mean concentration by the median weekly flow for each site and year and converted the result to kg/year.
- Weekly total: We multiplied the sampled phosphorus discharge by the daily flow value for the week the sample was taken, divided the result by the number of samples taken that year, and converted the result to kg/year for each site.

The results of our analysis are shown in Figure 9. We did not calculate results for England because the weekly total option is not possible with the data available (total daily flow yearly average). This chart only contains data for the 14 remaining sites for which we had all years' data.

Figure 9 Results for three types of load calculation



The analysis shows that the different calculation methods yield a different result, and the relation between those results undergoes some change year to year. However, from the data we have, it is not possible to draw a conclusion about the most suitable method to calculate load, because we do not have continuous measurements to compare against. Average total daily flow is reported to the Environment Agency (for English sites) annually, and it is recommended therefore that the definition uses average rather than median flow to calculate load.

4.4 Other Performance Commitment design considerations

We considered how water companies operate their sites and how the Performance Commitment could incentivise them. Our analysis has confirmed that water companies already tend to operate with a comfortable margin between permit level and performance, although this margin differs between companies. Our engineering experience and practice confirms that sites are designed for this.

Based on this, it is our opinion that it would be challenging for companies to further improve their performance against permit in a short time-period except by increasing chemical dosing. Biological phosphorus treatment is relatively small-scale in the UK and does not readily achieve better performance than designed, leaving chemical dosing as the main option for improving performance in a short time-period.

It is important that the performance commitment can incentivise innovation rather than just increased chemical dosing. However, there is the potential that rewarding early delivery will tend to early delivery of existing technology (chemical dosing), rather than innovation. This risk is managed to some extent by the fact that treatment works that are converted to chemical dosing for P-removal would require revised Environmental Permits for the dosing chemicals that are added, therefore the Environment Agency and Natural Resources Wales have some control over the programme of early delivery of chemical dosing for P removal.

There is also a risk that the performance commitment might skew investment in additional P removal where it provides the greatest financial reward rather than the greatest benefit to customers and the environment. But this is still better than the alternative of no performance commitment, where the incentive is to deliver at least cost to meet the permit level.

Version 0.1 of the definition partially addresses the risk of increased chemical dosing through the requirement for the treatment method used to be the best long-term approach for customers and the environment, taking into account a wide range of factors including the long-term resilience of the supply chain. Depending on how this is interpreted and the type and quality of evidence accepted, this condition could rule out additional chemical dosing in some or all cases. We also note that chemical dosing will be disincentivised through the Greenhouse Gas Emissions PC for which the embedded emissions of chemicals should be reported as well as any process emissions.

We also note that it could be theoretically possible for water companies to voluntarily adopt new phosphorus permits at relatively high concentrations and use these as an easy way to outperform the performance level through chemical dosing, although this would require the complicity of the Environment Agency so would be unlikely.

5. Recommendations

5.1 Dataset

The baseline dataset is sufficiently large and representative to support the conclusions and recommendations made here. The quality of the baseline dataset has improved between 2011 and 2021, and we expect the quality of the future baseline dataset to continue to improve. Data quality could be further improved with collaborative work with water companies, the Environment Agency and Natural Resources Wales. We believe that a significant proportion of sites excluded from the analyses are likely to be due to issues with how completely we have been able to link the permit database to water quality compliance data and flow compliance data, rather than actual missing data.

5.2 Assumed values

The Performance Commitment (version 0.1) contains assumed values for flow (1.2 times permitted dry weather flow) and phosphorus concentration (5.0 mg/l) which can be used when actual data are not available for works. We found no evidence to support or contradict these values.

5.3 Normalisation

Based on version 0.1 of the Performance Commitment definition, we recommended that the Performance Commitment measure should not be normalised to avoid compounding the disincentive for companies with tight permits to stretch their performance. However, in the current draft of the Performance Commitment that has been consulted with stakeholders, we are satisfied that this issue is resolved. This is subject to the appropriate Performance Commitment Level and Operational Delivery Incentive rates being set to avoid rewarding already-expected performance.

To go further towards rewarding outcomes rather than outputs in the future, the measure could be adapted, for PR29 and beyond, to incentivise reductions where they are most likely to cause environmental benefits. For example, discharges into sensitive waterbodies or waterbodies which fail to achieve good environmental status due to phosphorus could be given a greater weighting.

5.4 Load calculation method

We tested whether at the company level there is a significant difference in the method used to calculate the annual load, mean, median or weekly total. We only received data to do this from Welsh Water therefore, from the data we have, it is not possible to draw a firmer conclusion about the most suitable method to calculate load.

5.5 Performance Commitment effects

We recommend that careful consideration is given to the behaviours that the Performance Commitment could incentivise, in particular increased use of chemical dosing and adoption of ‘easy to meet’ voluntary permits.

Consideration should be given in determining the date at which the Performance Commitment assumes a permit is applied. WINEP and NEP often have regulatory delivery dates of the last year of the AMP, despite the expectation that delivery would be profiled over the AMP. This could lead to companies being rewarded for delivering on their expected profile because of the mismatch between expected delivery date and regulatory delivery date.

We also note that there are currently multi-year delays on new and revised regulatory permits being issued by the Environment Agency, and companies should not be penalised for late issue of permits.

6. Revised performance commitment definition

Based on our initial findings in our draft reporting, Ofwat has proposed a revised definition which has been consulted with stakeholders. The revised definition is:

Step 1 Calculate phosphorus load in the year at each treatment works with a phosphorus limit:

Phosphorus discharged (kg/year) = mean concentration x mean daily flow x 365

Step 2 Reduction in phosphorus emissions =

Phosphorus discharged from treatment works in 2020
less
Phosphorus discharged from treatments works in the year.

What happens if we don't know what happened in the base period?

- If average concentration is not known in the base period use 5.0mg/l.
- If flow is not known in the base period use dry weather flow multiplied by 1.2.

Step 3 Adjust for any additional phosphorus prevented from entering rivers from partnership working.

Step 4 Divide by the total load discharged from treatment works in 2020.

Step 5 Compare the reduction to the performance commitment level set at PR24 and calculate any underperformance or outperformance.

Appendix A. Excluded sites

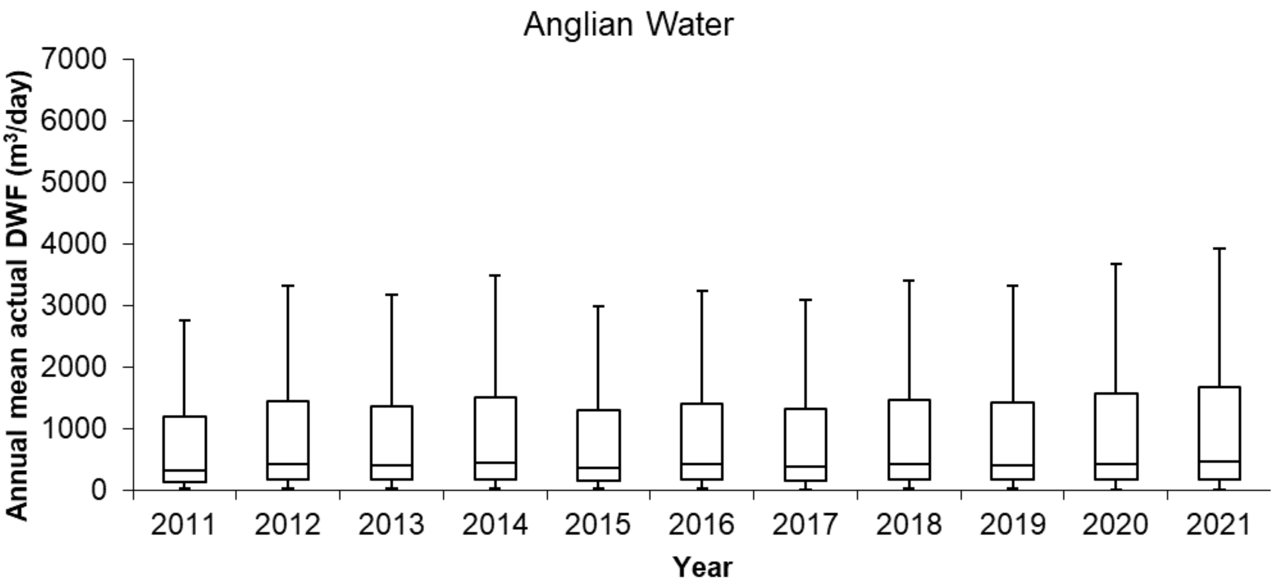
Table A1: Summary of excluded sites with reasons

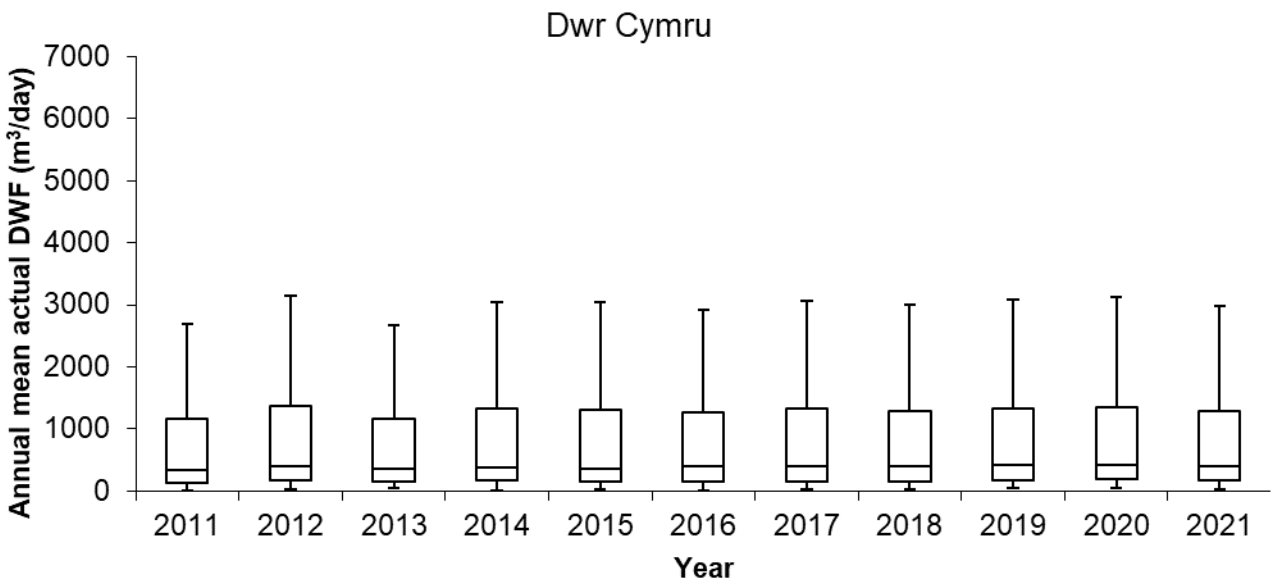
Missing parameter	Yorkshire Water	Wessex Water	United Utilities	Thames Water	Southern Water	South West Water	Severn Trent	Northumbrian Water	Welsh Water	Anglian Water	Comments	Do the data exist	Are the data essential
Flow - year			3		2		7				This is essential to calculate actual load	Unknown	Yes
P compliance year	2	4	4	4	13	1	20			16	This is essential to calculate actual load	We believe that many, if not all, of these data points do exist.	Yes
P permit site			2								Self explanatory	Probably	Yes
DWF permit and actual DWF/Q80									2		Self explanatory	Probably	Yes
P Compliance year and actual DWF/Q80									14		Self explanatory	Probably	Yes
Flow - site	8	2	11	4	4	2	7			2	This means that some data is missing for some years at this site. This site cannot be used in annual trend plots	Unknown	No
P compliance - site		15	9	8	39	3	35			58	This means that some data is missing for some years at this site. This site cannot be used in annual trend plots	We believe that many, if not all, of these data points do exist.	No

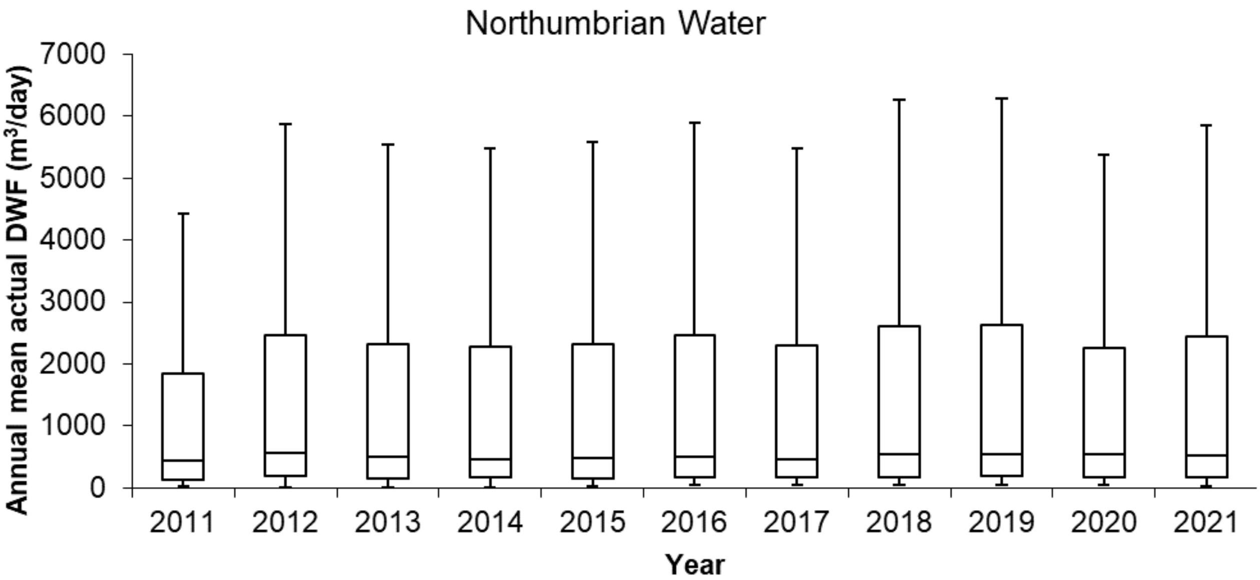
Missing parameter	Yorkshire Water	Wessex Water	United Utilities	Thames Water	Southern Water	South West Water	Severn Trent	Northumbrian Water	Welsh Water	Anglian Water	Comments	Do the data exist	Are the data essential
Duplicate England / Wales data (Dwr Cymru only)									11		Duplicate DCWW sites in both English and Welsh data	NA	No
Actual DWF/Q80									51		We cannot calculate permitted load without the actual Q80 because we use the ratio of Q80 to permitted DWF to calculate the 'permitted' average flow. Although we don't have this data, it is available.	Yes	It is essential if the performance commitment needs to calculate the delta between actual and permitted load
DWF permit value			3	6							Self explanatory	Unknown	It is essential if the performance commitment needs to calculate the delta between actual and permitted load
DWF permit value			1	2							Same as above	Unknown	It is essential if the performance commitment needs to calculate the delta between actual and permitted load

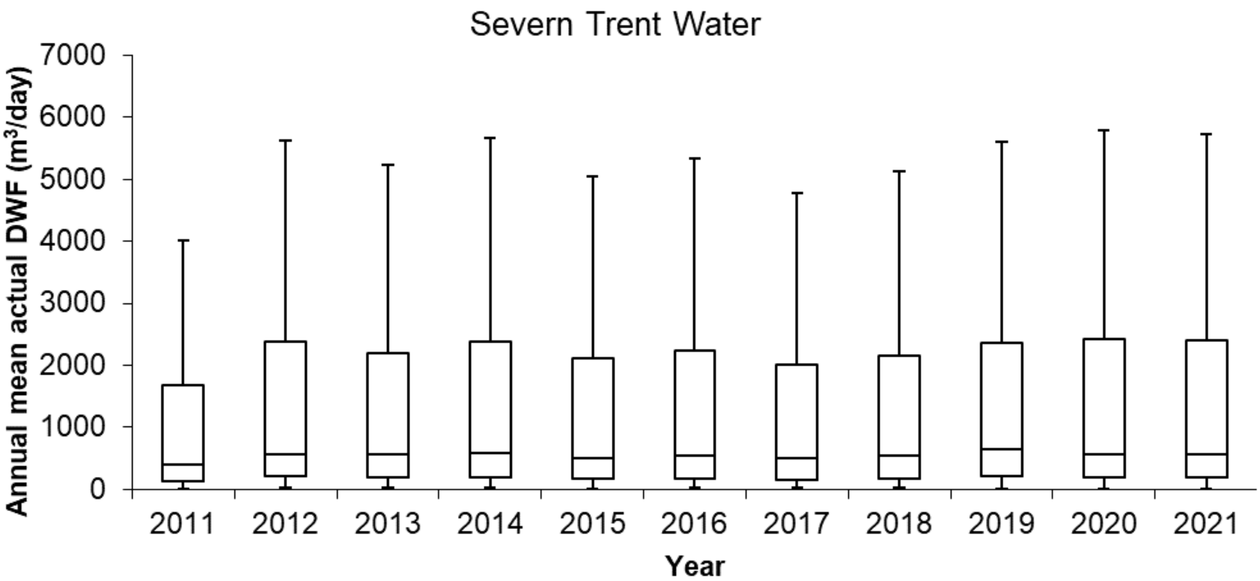
Missing parameter	Yorkshire Water	Wessex Water	United Utilities	Thames Water	Southern Water	South West Water	Severn Trent	Northumbrian Water	Welsh Water	Anglian Water	Comments	Do the data exist	Are the data essential
P compliance year and DWF permit value							2				Self explanatory	Probably	It is essential if the performance commitment needs to calculate the delta between actual and permitted load
P compliance year and flow year					1								

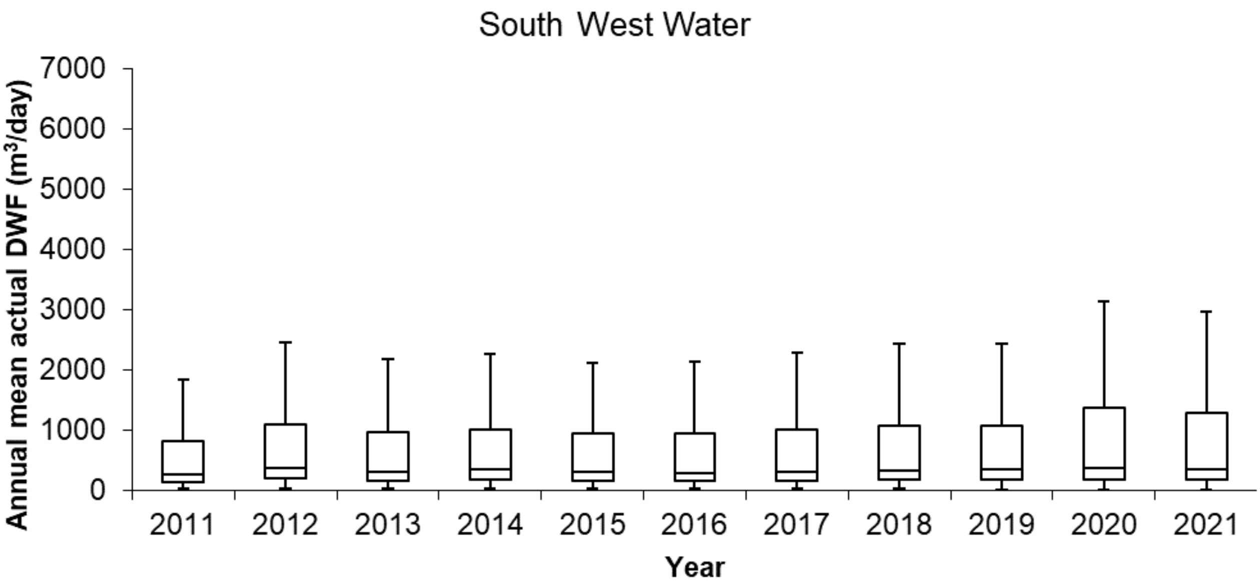
Appendix B. Average annual mean actual DWF per year, from 2011-2021, for all permitted works per water company within England and Wales.

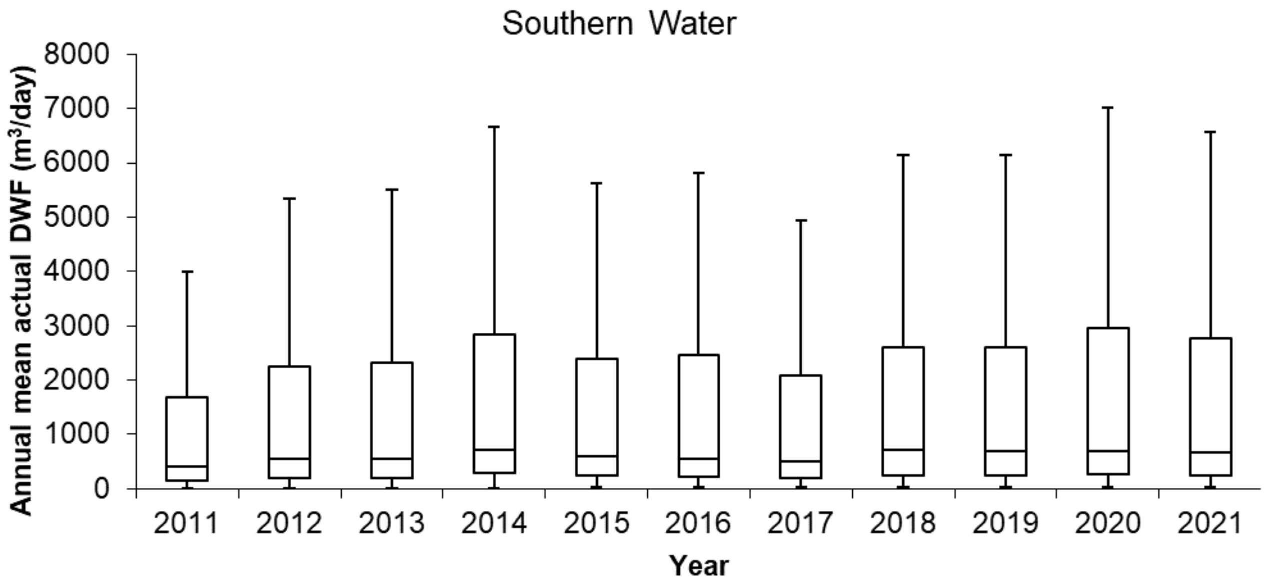


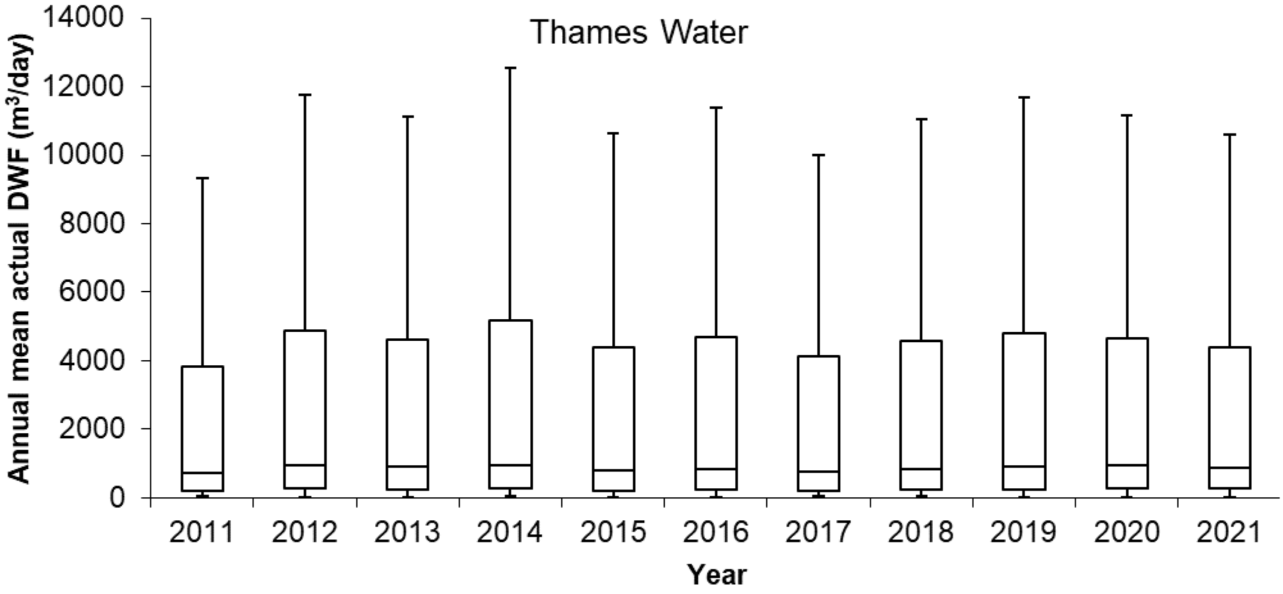


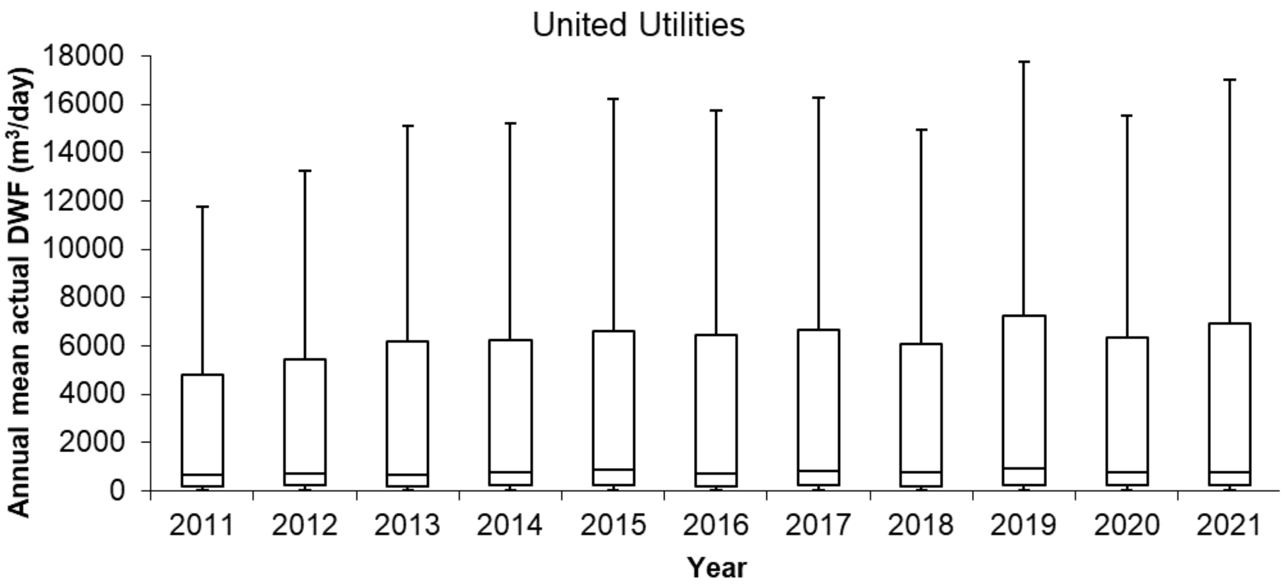

Northumbrian Water
Annual mean actual DWF (m3/day)
7000
6000
5000
4000
3000
2000
1000
0
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
Year

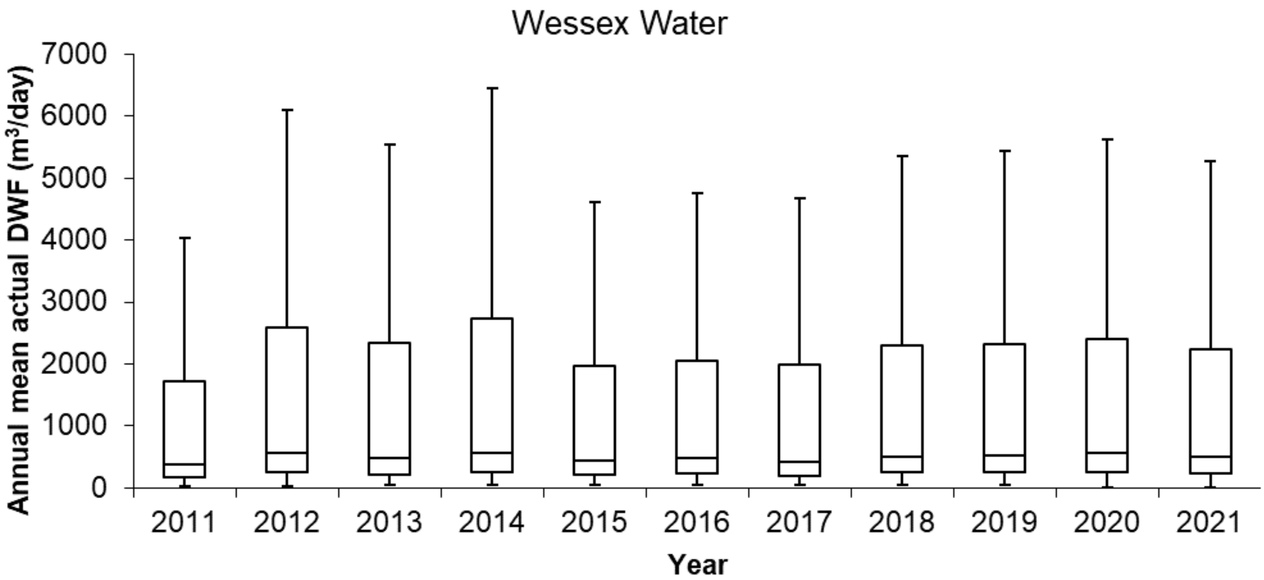

Severn Trent Water
Annual mean actual DWF (m3/day)
7000
6000
5000
4000
3000
2000
1000
0
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
Year

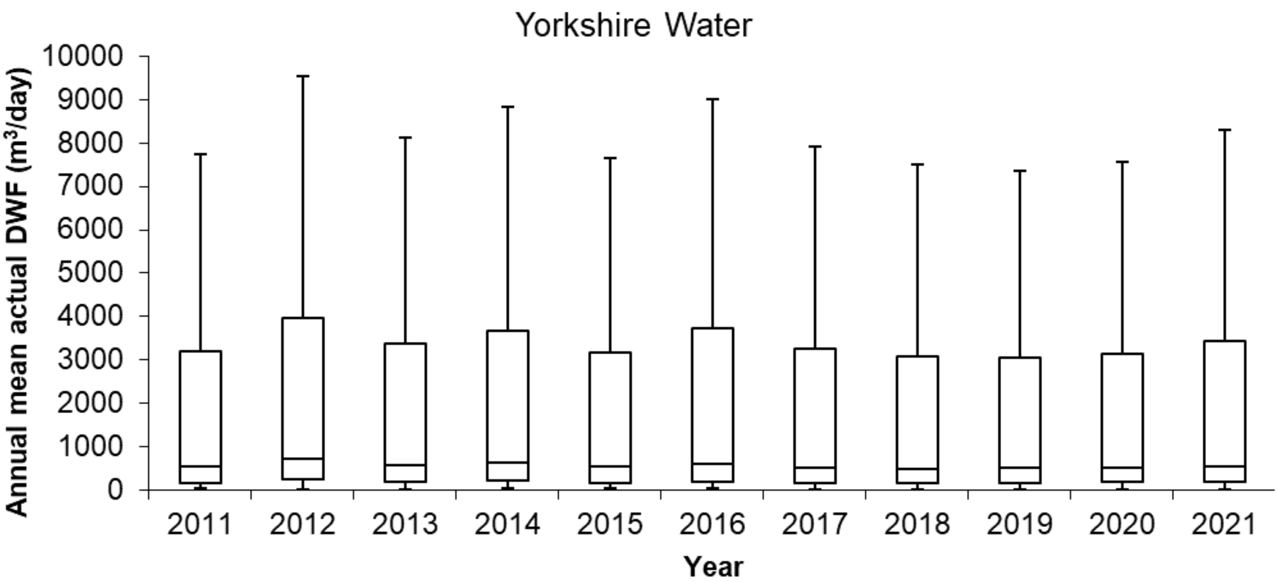

South West Water
Annual mean actual DWF (m3/day)
7000
6000
5000
4000
3000
2000
1000
0
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
Year


Southern Water
Annual mean actual DWF (m3/day)
8000
7000
6000
5000
4000
3000
2000
1000
0
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
Year


Thames Water
Annual mean actual DWF (m3/day)
14000
12000
10000
8000
6000
4000
2000
0
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
Year


United Utilities
Annual mean actual DWF (m3/day)
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
Year


Wessex Water
Annual mean actual DWF (m3/day)
7000
6000
5000
4000
3000
2000
1000
0
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
Year


Yorkshire Water
Annual mean actual DWF (m3/day)
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
Year