04 December 2024

# EPA review for water and sewerage companies - consultation response

## Background and objectives

This document sets out Water UK's response to the Environment Agency's targeted stakeholder engagement supporting its review of the Environmental Performance Assessment (EPA).

Water UK is the trade association for the UK's water industry and includes all regulated water and wastewater companies in England, Northern Ireland, Scotland, and Wales.

Water company activities have a significant impact on the environment. This impact must be measured properly and made transparent to the public – not just to build accountability and trust, but also because this helps industry, regulators, and others to identify trends and fix problems.

The EPA is the single most important mechanism for understanding and articulating companies' overall environmental impact. It also shapes the decisions of companies themselves, with incentives and penalties set on the basis of its scores.

For this system to work well, it must use metrics that are:

- clear – there should be no doubt about whether relative or absolute performance is improving (or deteriorating), and the precise impact it is attempting to measure;
- consistent – comparisons between companies and over time must be fair, and;
- meaningful – the metric should be set in a way that as closely as possible measures 'real world' impacts.

Where the EPA's measurements (and therefore conclusions) vary from these principles, it becomes less useful as a tool for understanding the interaction between the water industry and environmental goals. This may lead to the wrong conclusions and the wrong prioritisation of action.


Company number: 03539600
Registered in England and Wales
Registered office address: as above

These three principles inform the rest of our response, which is split into two parts: firstly, an overview of cross-cutting issues, where we provide feedback on overall features of the proposed EPA methodology, and a second more specific section on individual measurements.

We have real concerns about whether many of the proposals put forward by the Environment Agency meet the three objectives we have set out; however, for the two metrics where the EA is proposing some outcome-based indicators, we wish to acknowledge the potential for real benefits these might provide.

## Cross-Cutting issues

### Removal of Category 4 - no impact claims for minor discharges

The consultation document states in section 4.1.1 that 'the provision in the guidance that allowed no impact claims for minor discharges to be set at a Category 4 incident classification will be withdrawn from 1 January 2026 onwards. These incidents will now be set at a Category 3 classification as a minimum'.

This marks a departure from the previous approach described in the Common Incident Classification Scheme (CICS)[1] guidance. This guidance notes that substantiated incidents that result in no environmental impact or where the impacts cannot be confirmed should be recorded as Category 4. This guidance also states that 'any spillage or discharge of noxious, poisonous or polluting matter to surface waters or groundwater will be presumed to have an impact,' with the responsibility lying on the operator to prove otherwise.

The proposed change means that incident involving a discharge will be, at minimum, a Category 3 incident, regardless of whether it has any impact at all. Water companies suggest removal of Category 4 will increase the number of pollution incidents by 100%.

Past work conducted between the EA and water companies has found degrees of inconsistency in the EA's approach to assessing and classifying incidents. We are expecting the EA's review of its guidance on classifying incidents ('pollution incident self-reporting and recording guidance' also known as 16.02 guidance) to go some way in addressing this issue, however we remain concerned that companies will be affected differently by the removal of Category 4 which will further exacerbate regional variations.

We welcome the subsequent clarification received from the EA that a consultation will be carried out on this change as part of a more general consultation on the update of the 16.02 guidance. However it is concerning nonetheless as it will lead to two misperceptions: first, that harm is being done to waterbodies

[1] https://www.ofwat.gov.uk/wp-content/uploads/2017/12/20171129-Incidents-and-their-classification-the-Common-Incident-Classification-Scheme-CICS-23.09.16.pdf#:~:text=This%20document%20describes%20what%20an%20incident%20is%20and%20the%20two-tier

when no such harm exists; and second, that performance has deteriorated when it has not (because the data between years will be non-comparable).

By doubling the number of total pollution incidents without adjusting the thresholds, the number which fall into either Amber or Red categories will automatically and significantly increase despite no change in overall harm or performance. This will have four negative knock-on impacts:

- First, it will undermine public confidence in the EPA as a tool that meaningfully records and articulates harm. It will no longer accurately demonstrate the extent of water companies' 'real world' environmental impact and so reduce accountability and the information available to the public.
- Second, public trust in the sector's performance will be undermined without any real cause or deterioration to underlying performance.
- Third, as the gap grows between EPA scores and 'real world' harm, companies and their individual managers will be increasingly incentivised to focus their efforts and budgets on achieving a synthetic benchmark rather than reducing pollution. This is likely to result in worse real-world pollution.
- Finally, we expect that an unfair approach on this issue could affect negatively investor's sentiment. This is important because it could make it more difficult to secure all of the fundings needed for the delivery of the environment programme.

Finally, given the consultation being undertaken on EAs' 16.02 guidance and the potential for large changes in how pollution incidents are recorded and tested for (which goes beyond Category 4 incidents and also speaks to changes such as the inclusion of dry day spills), it is premature for the consultation to set thresholds for red, amber and green categorisations. Setting of these thresholds should rather be considered once the guidance has been updated, so it can proceed with more facts in hand concerning the overall regime on pollution incidents.

## Gradual strengthening of thresholds

The Environment Agency's approach to the EPA is generally to tighten thresholds of company performance over time, to ensure that standards are being raised on a continual basis.

Whilst it is positive that tightening thresholds should increase performance, the binary nature of passing or failing a threshold (especially for core metrics) will lead to headline consequences that do not necessarily reflect the environmental impact of shortcomings.

Although we are sympathetic to increasing performance expectations over time on important issues such as prevention of serious pollution incidents, tightening thresholds across a high number of metrics has the disadvantage of preventing the tracking of absolute progress over time, as current results are not comparable with historical, and it is thus very challenging for water companies to demonstrate the progress

and improvements made. Further, the thresholds for several metrics are at or close to 100% or 0, meaning very small and sometimes trivial failures can majorly impact the overall summary star rating.

By changing thresholds and constantly moving the goalposts, the EA creates challenges for companies who constantly need to recalibrate their expectations and adapt to new challenges, unfortunately this can lead to company expertise not being deployable for continuous improvement (for example, where measures shift considerably) and creates disincentives to improve performance year-to-year, as true comparison becomes much more difficult.

## Comparing performance on pollution incidents

We note that the EA is looking at improving the consistency and fairness in how it measures pollution incidents by exploring different methods for normalisation (a method by which performance of large and small companies can be equated).

We recommend the EA goes beyond examining normalisation and instead considers including a risk-based measure that would enable a finer distinction of pollution incidents based on their impacts, moving away from the simplicity of a three-category system. Moving to a more risk-based measure would enable consideration of factors beyond impact, for example the sensitivity of the receiving water course and how the company reacted. This brings in a wider set of considerations than are currently made in the categorisation of pollution incidents (which are restricted to persistence, extent and seriousness of effects) and would allow more consideration and focus to be placed on the operational environment (e.g. whether an incident happened in a well-used water course).

## Overlap and clarity of metrics

As detailed in the sections below, some proposed metrics overlap and use the same data to produce different assessments. As such, one water company could fail under two individual metrics for the same incident (e.g. the metrics on 'discharge permit compliance' and 'descriptive permit compliance at numeric sites').

We note that the EA intends to select a subset of all metrics it is consulting on. We ask that care be taken to ensure the final metrics do not overlap and that every metric genuinely adds value and avoids 'double jeopardy' that could disproportionately weight results in a way that reduces the EPA's reflectiveness of real-world impacts.

We also note in our feedback below that the wording or articulation of several metrics is unclear. We propose that the EA do further work to articulate these in more detail (e.g. RNAGs, phosphorus, waste management, abstraction, and impounding) and then use these metrics for shadow reporting initially. The outcome of 1-2 years of shadow reporting can be used to select which metrics to keep in the final EPA and which metrics drop or continue as shadow reporting only.

## Core metrics

In light of the extensive changes being proposed to calculation of the total pollution incidents metric, we have doubts about its inclusion as a core metric as any deficiencies which mean that many companies trend to a 'red' categorisation will take on an even bigger significance, blocking achievement of 4-star performance.

We further note that three of the four core metrics are proposed to be adopted by Ofwat as Common Performance Commitments (CPCs) for 2025-2030, meaning their achievement or failure will already have (sometimes large) financial implications. Taking pollution incidents as an example, in PR19 water companies received £110.6 million of net penalties for 2020-21 to 2023-24. Ofwat is still finalising the precise penalty rates for PR24 but its draft determinations in July 2024 indicated penalties of approximately £366,000 for every Category 3 pollution incident and £2.1 million for every Category 1 or 2 pollution incident. The EA should consider whether compounding the existing monetary and reputational penalties for poor performance on these measures is proportionate and consider adopting an approach where scores are determined as part of a rounded performance assessment.

## EPA should have a 5-star rating

While the EA traditionally uses a four-star rating, this arrangement continues to show some limitations and is the basis for misunderstandings.

Our experience shows that stakeholders can misunderstand the 4-star rating system without a 'neutral' middle point. As they expect a 5-star rating system (the most commonly used scale for assessments of this kind, reflecting humans' universal use of base ten numeric systems), they assume that 4 stars are not the highest rating.

The use of a 5-star rating system is standard in many sectors, for example, the Citizens Advice assessment of energy suppliers (5-star rating)[2] and Energy Performance Certificates (A-G is a 5-point scale) and customer ratings of products in many retailers (e.g. Amazon, John Lewis).

A 5-star rating system would also provide more nuance in the overall assessment of company performance. It would reflect the complexity of assessments when the number of metrics is set to grow and gives more chances for companies to improve gradually. We also ask that the labels used be unequivocal (see proposed labels in the table below).

[2] https://www.citizensadvice.org.uk/consumer/your-energy/get-a-better-energy-deal/compare-domestic-energy-suppliers-customer-service/

| Star Rating | Description |
|---|---|
| 5 | Excellent performance |
| 4 | Good performance |
| 3 | Average performance |
| 2 | Insufficient performance |
| 1 | Poor performance |

## Metric-specific feedback

### Serious pollution incidents metric

We note that this is one of the three metrics that Ofwat proposes adopting as a Common Performance Commitment for 2025-2030.

The EA proposes to remove the previous glide path (for years 2021-25, see figure below) that was provided for performance improvements and require that all companies have zero serious pollution incidents from 2027 (2026 data) to be rated green, one serious pollution incident for amber, with red triggered by two or more serious pollution incidents.

Although most companies project that they will meet the target of zero serious pollution incidents by the end of AMP8, this will not be possible from the beginning of the period as improvements will take some time to be delivered. Companies therefore need to see a glide path that accounts for the delivery profile of ambitious improvement schemes in AMP8.

**Fig 1: Thresholds for Serious Pollution Incident Performance (Red/Amber/Green) in the current EPA**

The RAG threshold 5-year glide path, 2021-2025:

| | Anglian Water<br>Severn Trent Water<br>Thames Water<br>United Utilities | Dŵr Cymru Welsh Water<br>Northumbrian Water<br>Southern Water<br>South West Water<br>Wessex Water<br>Yorkshire Water |
|---|---|---|
| Year 1 & 2<br>(2021 & 2022 data) | <=3 | <=1 |
| | 4 or 5 | 2 or 3 |
| | >=6 | >=4 |
| Year 3 & 4<br>(2023 & 2024 data) | <=2 | <=1 |
| | 3 or 4 | 2 |
| | >=5 | >=3 |
| Year 5<br>(2025 data) | <=1 | 0 |
| | 2 or 3 | 1 |
| | >=4 | >=2 |

## Total pollution incidents metric

We note that this is one of the three metrics that Ofwat proposes adopting as a CPC for 2025-2030.

The thresholds are left unchanged compared to 2025. We are concerned that the proposed thresholds do not reflect the upward pressure on total pollution incidents caused by the redefinition of pollution incidents and that the EA's inclusion of dry-day spills and the removal of Category 4 incidents will make the proposed 'green' threshold for total pollution incidents impossible to achieve.

With these changes set to be phased over two years - inclusion of dry days spills from 2025 reporting and the proposed removal of Category 4 incidents in 2026 - we will see two consecutive years of increasing total pollution incidents, irrespective of company actions. This will ultimately undermine the work that both water companies and the EA are doing to reduce the risks and impacts of pollution incidents by splitting focus at a time when companies should address the most significant issues.

## Self-reporting metric

Water companies are strongly committed to tackling incidents quickly and minimising possible harm. Doing so relies on their receiving reports directly and quickly, which can then be passed to regulators. However, the EA encourages members of the public to report incidents to them through their new online tool[3] which could cause delays in crucial information reaching water companies and potentially increase the impact of pollution incidents. Secondly, the metric is defined very narrowly, which may lead to misunderstanding. For

[3] Reporting water pollution to get easier with new online service – Creating a better place

example, a company is not credited with self-reporting an incident, even if it has, whenever the same incident raised concurrently by a customer. So long as it does not weaken the incentive on the company to self-report, it is important that this measure is improved to allow for concurrent identification of pollutions by multiple parties, to ensure fairness. This will improve the metric's accuracy.

We would urge EA to adopt an approach that allows for companies to gain credit for reporting an incident even when it is reported concurrently by a customer. We would also like EA to update its messaging to let customers know they can make reports directly to their company, rather than going via its online tool.

## Discharge permit compliance metric

We note that this is one of the three metrics that Ofwat proposes adopting as a CPC for 2025-2030.

Industry would like to highlight the overlaps between this metric and the new proposed descriptive permit compliance metric. As per our explanation above, we are not in favour of overlapping metrics and urge EA to keep only one in the revised methodology.

Water UK notes the Labour Party's election manifesto commitment to "…ensure independent monitoring of every outlet,"[4] and supports having a strong system of safeguards in place to ensure that the sampling programme shows an accurate and holistic picture of the quality of final effluent. Concerning the EA's proposal, further details are needed to understand how sample frequency will be articulated in the context of no-flow events and to ensure the metric does not unintentionally penalise companies. For example, we are keen to see that failures are not recorded but rather rescheduled in instances where samples legitimately cannot be taken (e.g. where unforeseen or emergency maintenance work is being carried out at the WwTWs).

## Water Industry National Environment Programme (WINEP) and National Environment Programme in Wales scheme delivery metric

Water UK supports the proposed continuation of the 'cumulative reporting approach' to tracking delivery of environmental improvement projects listed in the EAs 'WINEP' programme. With the sharp increase in ambition from companies for investment, it is vital that customers are able to understand the improvements being paid for through their bills, and we strongly encourage EA to consider ways to show how the improvements delivered by water companies remove environmental harm.

## Supply Demand Balance Index metric

The supply-demand balance index (SDBI) is a metric that measures the extent of a company's water supply risk based on a numerical assessment of its water resource systems' resilience to drought. While this is

[4] Change Labour Party Manifesto 2024

important information, Water UK recommends that the EA remove SDBI from the EPA because it is a measure of how companies are managing risks to customer supply, rather than an environmental measure and thus not consistent with the other metrics in EPA. We note that the shadow metric on abstraction and impounding licence compliance is being retained, which is a better measure of how company activity to underpin water resources (i.e. abstraction) affects the environment (e.g. river flow).

If the EA wants to retain the tracking of this measure, it should be part of the reporting under the annual review of the Water Resources Management Plans (WRMPs). Please find a position paper on SDBI in Appendix A.

## Descriptive permit compliance at numeric sites metric

The Environment Agency sets numeric discharge quality limits in permits to ensure discharges comply with water quality objectives. Currently, a WwTW fails compliance under its discharge permit metric only if its treated sewage effluent discharge fails to comply with the discharge quality emission limit. This new metric proposes to expand compliance to also consider compliance with flow limits and compliance with descriptive conditions for all discharges that are associated with the operation of the discharges covered by quality limits. One example of a descriptive condition failure would be the company failing to follow an operating technique as required by the permit, which had the potential to result in a significant pollution.

Water UK considers that this metric should remain focused on quantitative numeric limits for the following reasons:

- This metric duplicates the discharge permit compliance metric and could lead to an unbalanced situation where one compliance breach leads to a water company failing two metrics.
- The metric relies on the individual assessment of compliance performed during inspections and is thus open to subjectivity and has a high potential for inconsistency. This is not a sound nor desirable basis for a metric.
- Descriptive condition breaches from discharges are covered by pre-existing enforcement mechanisms via the compliance classification scheme process.
- The thresholds are set for the whole 2026-2030 period, whereas a glide path would have been more appropriate.

Water UK considers that, as proposed, the metric is not mature enough to be included in the EPA.

## Storm overflow metric

This is a newly proposed metric that will record and score company performance on the operational availability of storm overflows and the percentage that have associated spill data.

While this metric is sound in theory, more work is required to establish whether the threshold of more than 95% of EDMs being available over 90% of the time (the minimum level for ‘amber’) is achievable. For

example, it is estimated that following the Public Switched Telephone Network (PSTN) being switched off, 60% of the UK's water and wastewater assets will rely solely on 4G communications which are more prone to outages. As EDM monitors are reliant on communications networks, a company could fail this metric because network coverage is not constant – the risk of this occurring is magnified in areas where signal strength is poorer (e.g. remote locations) so companies that operate largely in the countryside will be disadvantaged through no fault of their own. In addition, increases in storms and severe weather over recent years has led to power and communication outages across the 4G network meaning that at various times water and sewerage companies have lost visibility of hundreds of assets at a time. For example, the "Beast from the East" event in March 2018 led to a freeze-thaw event that not only caused physical damage to water infrastructure but also resulted in 4G network disruptions. The thresholds should be revised to acknowledge and account for this variability in communications signals and the risks of extreme weather.

Finally, given the proposed target for the metric is that 'by 2050, no storm overflows will be permitted to operate outside of unusually heavy rainfall or to cause any adverse ecological harm', we would welcome introduction of a definition for 'unusually heavy rainfall' to ensure consistency between this metric and others and prevent differing interpretations.

## Wastewater treatment works annual flow permit limit compliance metric

Water UK welcomes this new metric; however, we would like to raise some issues with its definition. The fact that the metric is trying to combine several types of permitted flow limit within one metric needs to be clarified, and we recommend that the EA focuses on one metric only.

We note that the EA has removed maximum flow volume limits from some of its more recent permits. Consequently, and in light of the focus on Flow to Full Treatment (FFT) triggered by the recent Ofwat/EA investigation, we recommend that this metric focus on FFT as a metric against which compliance with annual flow requirements is measured.

## Waste management metric

Water UK has serious concerns about both aspects of this metric, which is intended to rate company performance as regards waste permits and unpermitted waste activities.

Aspect A assesses the compliance of waste operations with environmental permit conditions, most of which are Industrial Emissions Directive (IED) permits. While the EA has requested that all IED sites are compliant by March 2025, because of the timing of IED permits being issued, the ongoing discussions around specific improvement conditions required and the lead time needed to get the work done at site level, this metric will be impossible to achieve. We also note that this metric will end up penalising water companies on something they will have already been penalised for via the Compliance Classification Scheme assessment framework.

Aspect B, related to WaSCs' self-reporting of the total sites they operate that perform a regulated facility activity without an environmental permit, is an enforcement measure and, as such, is not related to performance.

Water companies operate a wide range of facilities, from complex anaerobic digestion sites covered by IED permits to more basic storage sites covered by basic standard-rule storage permits. This variation in complexity needs to be considered in the performance assessment.

We are also uncertain how this metric would be used in practice. The wording in the consultation document is unclear, for example:

- it is stated that 'it does not require formal enforcement action to be taken for an unpermitted activity to amount to a breach of the Environmental Permitting Regulations and therefore illegal.' Does this mean that a decision from an inspection officer (which is not a formal enforcement measure) will be sufficient to make an activity illegal?
- The consultation states that a facility subject to a duly made application would not be treated as operating illegally; however, there are many instances where water companies submit a permit application, and significant time elapses before the regulator looks at the application and determines it is duly made.
- There is no definition provided for what a 'redundant' lagoon is, and ambiguity such as this should be avoided.
- The metric's scope is unclear, but as it refers to "facilities which undertake permitted biological treatment processes," we assume that operations from exempt facilities under the Environmental Permitting Regulations are outside the metric's scope.

In addition to the issues with the measures included in Aspects A&B, the thresholds described in the consultation document penalise water companies with fewer permits as those with 32 or fewer permits would either score green or red as when turned into percentages, each failure is more significant (as 31 out of 32 permits within band A would mean a score of 96.8%, which is less than 97%).

As both Aspects A&B relate to enforcement rather than performance and we recommend this metric is not included in the final EPA assessment framework. However, if the EA is determined to keep this metric, it should be used only as a shadow metric and not until its definition and assessment are more mature. We also recommend making the wording more straightforward and less open-ended (e.g., 'regulated facilities include but are not limited to' is likely to lead to a lack of consistency).

## Phosphorus and Reasons for Not Achieving Good (RNAG) metrics

Water UK strongly welcomes the inclusion of outcome-based metrics, which we have previously called upon the EA to introduce. However, we note that both metrics have yet to be fully defined and expect that

further exchanges will be necessary to ensure the obligation, calculation, frequency, and thresholds are set adequately for them.

On the **Phosphorus metric**, we note a partial overlap with the WINEP delivery metric, as reducing the phosphorus load is one of WINEP's objectives. Ofwat has also proposed a Common Performance Commitment that tracks the delivery of phosphorus improvements, with financial penalties associated with late delivery. As such, the EPA metric is a third measure linked to the same activity, and we question whether it adds significant value.

To ensure the metric captures performance accurately and does not simply reflect load reduction requirements made on companies, load reduction should be calculated based on a fixed average flow from preceding years or a multiplier of recorded dry weather flow (e.g., 1.2x) and EPA targets based on the percentage of expected load reduction delivered.

On the **RNAG metric**, we need clarification on how the specific RNAGs would be matched with water companies. Furthermore, while the EA can add RNAGs to the datasets between formal EPA updates, the RNAGs dataset is only published once per year, so tracking progress against RNAGs will be challenging. The inclusion of the RNAG metric is positive, but work is needed to agree on its baseline and the process that the EA and water company would follow to ensure an RNAG is addressed. It is also very important that the metric is linked to the water companies' investment plans.

Overall, further work is needed on both of these metrics as they are not mature or precise enough. However, if they are correctly framed, we believe they can provide more depth to the environmental performance assessment. We are keen to work closely with the EA to support the development of these metrics.

## Conclusions

While supporting the concept of environmental reporting, we are concerned that the EPA in its current format has some limitations. We welcome further engagement beyond this consultation with the EA to continue to develop some of the newer metrics that need more consideration (including waste management metric, phosphorus and RNAGs, storm overflow metrics) and in particular address potential issues with the more subjective metrics which are being proposed for inclusion (e.g., descriptive measures metric) which could, without careful thought, lead to less fair and more biased assessments.

Outside of the new metrics, the EA has proposed several changes to EPA (e.g. the removal of Category 4 pollution incidents and glide paths for pollution incident reduction) that will change the baseline industry is working to and serve to make water company performance look as though it is stagnating or deteriorating when it is not. The outcome is likely to be further public concern and focus on water company performance that could, in turn, manifest in focus being taken away from dealing with the most serious incidents towards making improvements in areas of less consequence. This would be an unfortunate unintended outcome of

the proposed changes. We urge EA to consider how it creates a scorecard that reflects improvements and highlights only genuine and serious failures.

Considering the above observation, we are disappointed that the EA has not incorporated the requests from water companies and Water UK for a balanced scorecard approach which includes new indicators on issues such as carbon emissions and biodiversity net gain and thus captures a more fulsome picture of environmental performance. While the EA has moved toward our request for more outcomes-based measures, only two, RNAGs and phosphorus, are proposed, and they require significant development to be made fit for purpose. We note it is some time before the new EPA will be rolled out, so would be happy to facilitate further dialogue around the broadening of the scope of measures that it considers, should the EA wish to rethink.

Finally, we are concerned that the creation of a reporting system for one sector in isolation ignores the impacts from other sectors on water pollution and contrasts with the relatively low scrutiny of other sectors that cause diffuse pollution. If environmental harm is to be tackled holistically, all parties that contribute to it need a similarly bright light shone on their performance. In the continued absence of this, pressure will not grow on other stakeholders to improve their environmental practices, and the efforts of the water industry alone will not be sufficient to reverse the tide of water quality in England and Wales.

## Annex A – SDBI paper

Annex A - Water UK SDBI paper.pdf