SRN-DDR-012: Risk Appendix

Draft Determination Response

August 2024
Version 1.0

from Southern Water

Contents

WATER for LIFE from Southern Water

1. Key risk driver evidence, data input, methodology

1.1. Total pollution incidents

1.1.1. Evidence and data inputs

Pollution risk drivers

Analysis determining the drivers of pollution incidents for all WaSCs using the Environment Agency (EA) data on pollution incidents is shown below:

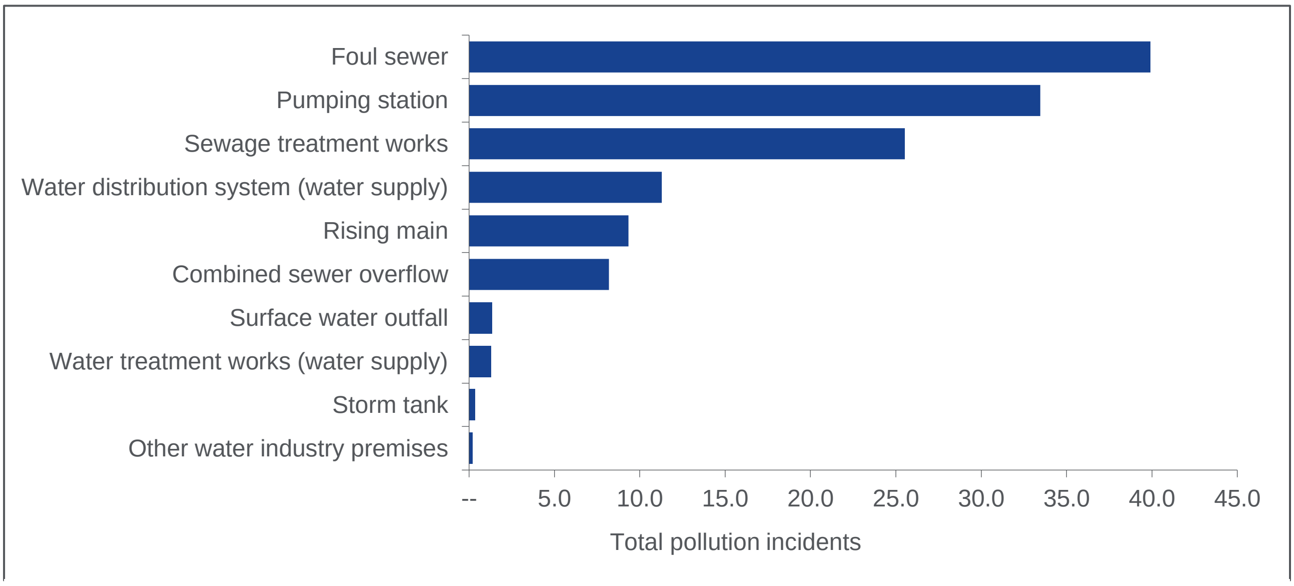

Figure 1:Sector total incidents by cause (2020-2022)

Naturally, pollutions incidents are asset driven. When an asset – such as a pumping station, rising main, or treatment work – fails, pollution incidents may occur. As a result, any notional company consideration must not be biased by the asset health of any one company. However, there is no robust dataset of asset health available for the sector which could inform a more granular notional company analysis that incorporates asset health. Resultingly, our notional company analysis focuses on the relationship between precipitation as a key driver of pollution incidents.

Using our data, the key drivers of pollution incidents are as follows. Data bars are coloured individually to show the four key drivers – electrical, mechanical, sewer blockage, and rising main – as well as all other incident types which we grouped into “Other” in further analysis.

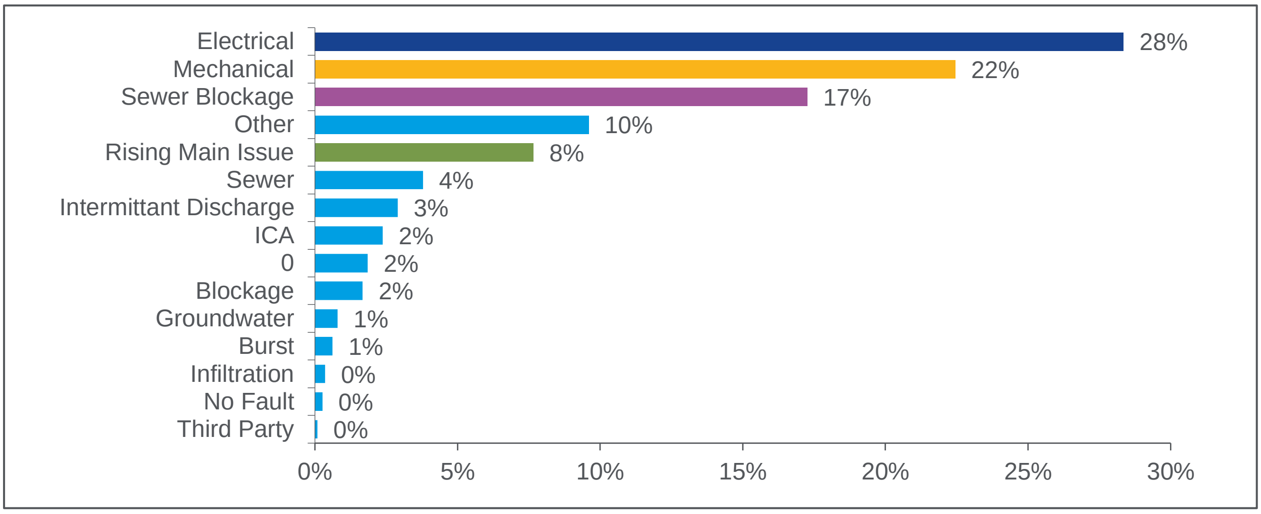

Figure 2: Proportion of pollution incidents (2020-2022) by cause

We analysed the relationship between the four key incident types and precipitation. Precipitation data was sourced from our equipment to measure rainfall across the region covered by Southern Water.

Our analysis separately considered incidents occurring in 2020-2022 and those occurring in 2023 due to the pollution incident reduction plan (PIRP) we undertook in 2023 which materially reduced the number of incidents. The improvement plan focused on wastewater treatment works, pumping stations and rising mains. This plan would have a distorting impact on the statistical model unless the improvement plan was a model input.[1]

Resultingly, the robustness of any statistical analysis using data combined across 2020-2023 would be impacted by the improvement plan. Instead, we (1) used 2020-2022 data to inform correlations and regression modelling due to the longer time period compared with only 2023, and (2) used 2023 data as a cross-check and to understand the impact asset improvement has on the pollution incident-precipitation relationship.

[1] Pollution Incident Reduction Plan (southernwater.co.uk)

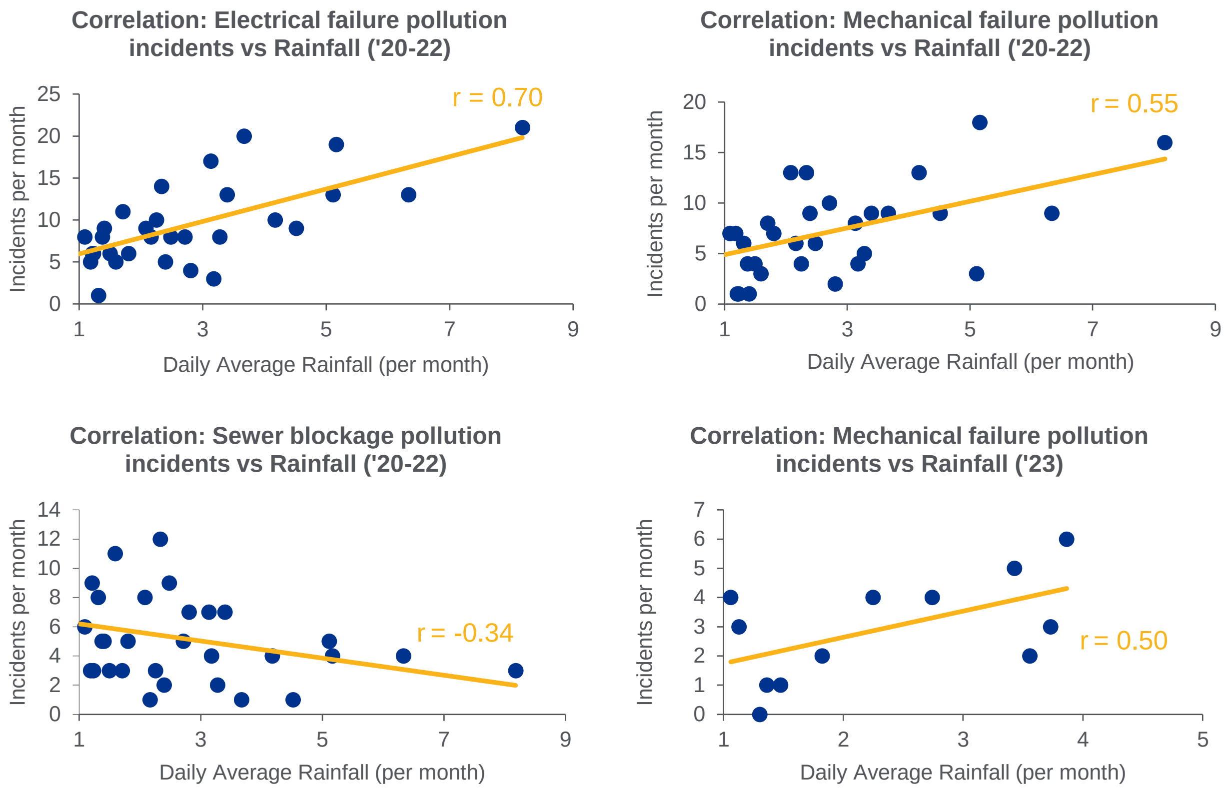

Figure 3: a,b,c,d: Common pollution incident causes vs rainfall

There are clear positive correlations between rainfall and pollution incidents caused by electrical failures, mechanical failures, and rising mains issues. This is due to high levels of rainfall increasing the load on machinery which leads to failure. In contrast, incidents caused by sewer blockage display a negative correlation with precipitation as these occur most frequently when there is insufficient water volume to wash away potential blockages.

We combined all other incident types into a single “Other” category and found a 0.55 correlation with precipitation. Resultingly, we found 83% of incident causes to have a positive relationship with precipitation, with the remaining 17% being sewer collapses with a negative correlation.

We performed similar analysis using 2023 data to cross check the relationship between precipitation and incidents under a scenario of improved asset health. The relationships derived using 2023 data are comparable to those using 2020-2022 data, with the exception of sewer blockage incidents. These incidents display a positive correlation with precipitation in 2023, in contrast to the negative relationship previously determined. This is due to the efficacy of the improvement plan throughout the year, with total incidents decreasing and reducing the explanatory power of rainfall.

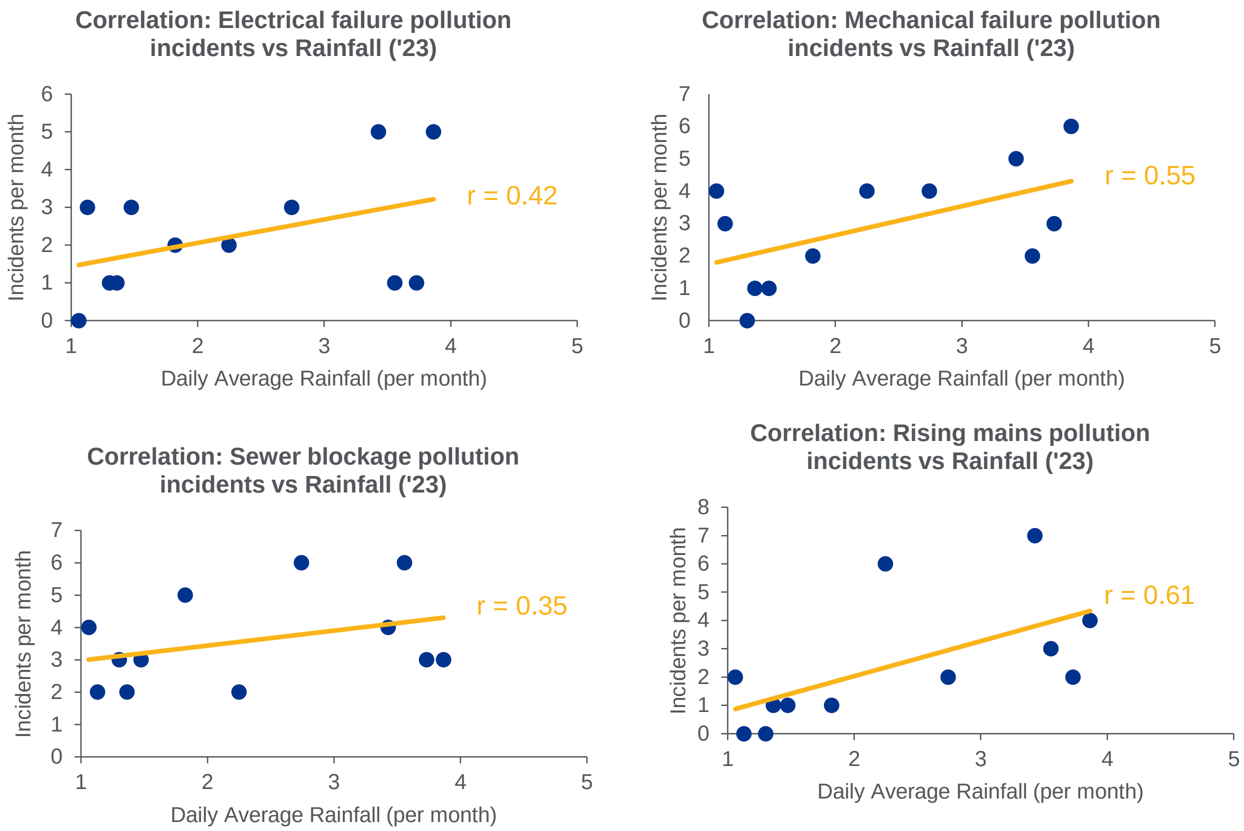

Figure 4: Common pollution incidents causes vs rainfall (2023 data)

Whilst the relationships between rainfall and incidents were derived with our data, they are highly probable to be applicable to all companies in the sector given they hold true both before and after the targeted asset improvement plan. Resultingly, the relationships derived are appropriate for application to the notional company and indicate its exposure to penalties relating to the pollution incidents ODI is highly dependent on precipitation, which is outside of the control of the company.

The link between rainfall and pollution incidents may not only relate to general level of rainfall. We have considered the following rainfall factors which may influence incidents:

1) **General level of rainfall** - captured as the daily average rainfall.

2) **Surges in rainfall** – periods of drought drying up the ground followed by large amounts of rainfall may cause incidents as the ground cannot readily absorb rainwater.

3) **Persistence of rainfall** – long periods of rainfall will increase ground water and cause flooding and further strain on wastewater assets.

The effect of general level of rainfall and persistent rainfall is captured in the above correlations as we have taken a daily average for rainfall. Thus factors (1) and (3) and measured together in our analysis.

To consider factor (2), the impact of surges, we defined a surge where there is a 3-day period where rainfall is at least 2.0x higher than the average from the previous 14-day period. An example is shown in the below figure.

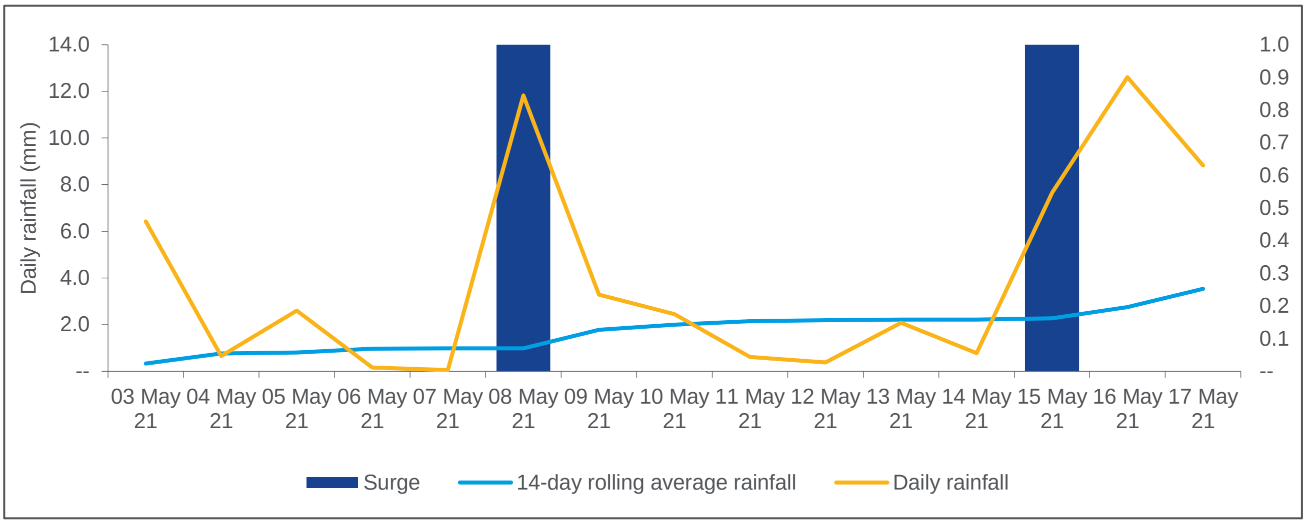

Figure 5: Rainfall surge definition

The below graph and table demonstate how the two precipitation variables - (1) daily average rainfall per month which captures general and persistent rainfall, and (2) surges in rainfall – evolve over 2020-2022.

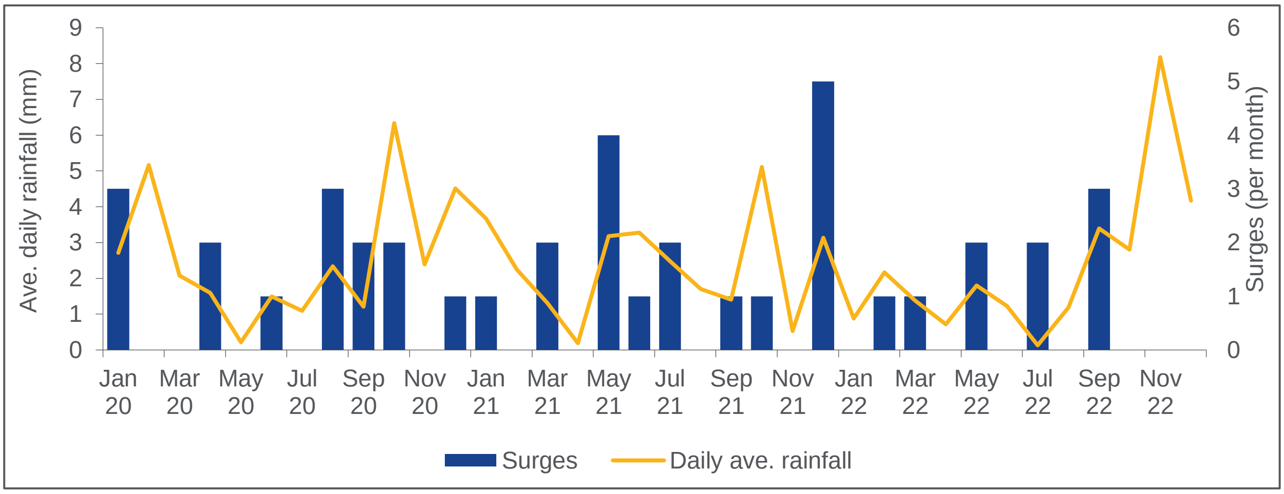

Figure 6: Precipitation data

	2020 - 2022			
Precipitation variable	**Max**	**Min**	**Mean**	**Total**
Daily ave. rainfall	8.2	0.1	2.4	n/a
Surges	5.0	0.0	1.1	40.0

Table 1: Rainfall data summary

We performed correlation analysis between each incident root cause and surges and found rising electrical issues to have a correlation of 0.13.

Forward looking precipitation data

We accessed Met Office publicly available UKCP18 climate projections to gather potential daily rainfall between 2025-2030 under RCP8.5 for Southern Water regions.

1.1.2. Methodology

Determining geographical areas for precipitation data

For a notional company in the southeast of England, the geographical areas served have been assumed to align with ours. There are a total of 119,166 postcodes served by 381 wastewater sites across 212 districts[2]. We therefore grouped each postcode district using the K-means clustering algorithm based on the coordinates[3] of the district centroids. For each group, we determined the district which best represented the centre point of that clustered. Note, district coordinates were rounded to the nearest 5km to coincide with the granularity at which MET office weather data is available. This is demonstrated below where each district is represented in blue, with the centroid clusters in orange.

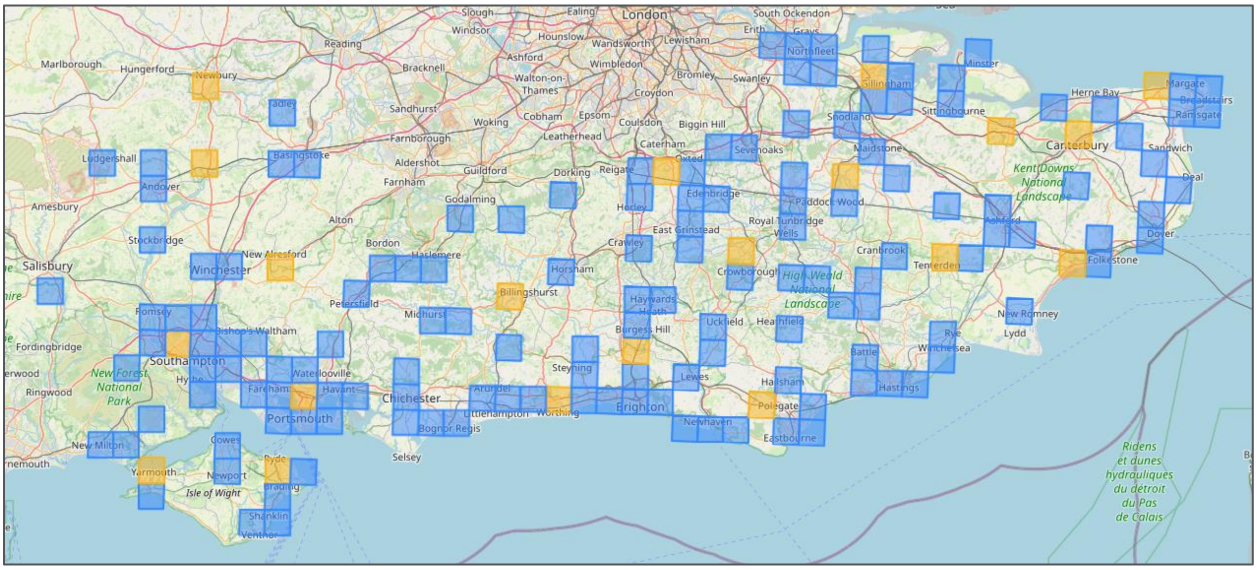

Figure 7: Map of MET office regions covered by Southern Water

This approach facilitates using the central district as a representative area for which weather data can be obtained. To obtain the weather data for each grid individually is beyond the scope of this analysis and given long-term weather data is likely to have minimal variance between two adjacent areas, using a grouping's central district is an appropriate proxy.

1.1.3. Pollution and precipitation data transformation

Analysis on the relationship between pollution incidents and precipitation could be performed at a cluster level. However, at this granularity incidents will have other strong explanatory factors, namely asset health, which are more applicable to the actual company. Therefore, our notional company analysis takes the average weather data across the MET office grids described above and compares these to pollution incidents. The preceding methodology for determining grids for which to obtain precipitation data ensures an appropriate proxy for the whole area served by Southern Water.

Pollution incidents were divided into three distinct groups based on the relationship between root cause and precipitation, as demonstrated in Pollution Risk Drivers section. These groups were as follows:

---

[2] A district is defined as the leading part of a postcode in either Letter-Letter-Number or Letter-Letter-Number-Number format. E.g., the postcodes GU6 8JE and BN26 5TA are in districts GU6 and BN26 respectively.

[3] Our analysis used the Easting Northing coordinate system to align to the precipitation scenario data available from the MET office. Note, this is a cartesian coordinate system – i.e. a 2D project – in contrast to the 3D geodesic projection of latitude-longitude.

	Group	Root causes	% Total Incidents	Correlation with Average Rainfall	Correlation with Surges
1	Explained by level of rainfall	Mechanical, Rising mains issues, Other	54%	0.75	-0.03
2	Explained by level of rainfall and surges	Electrical	29%	0.71	0.13
3	Not directly explained by rainfall factors	Sewer blockage	17%	-0.36	0.36

Table 2: Pollution risk driver groups

As group 3 does not have a direct link to precipitation it was excluded from the scope of this analysis. The factors in scope explain 83% of total incidents.

We performed analysis collating data on a monthly basis, i.e., pollution incidents per month, average daily rainfall in month, and total surges per month. This was due to the number of incidents. Smaller time periods, e.g. a week, may yield more precise analysis however sacrifice statistical accuracy where there are insufficient incidents occurring during each period from which to derive robust relationships.

1.1.4. Regression

We performed regression analysis on groups 1 and 2 using 2021 and 2022 data. We excluded 2023 due to the asset improvement programme having a significant influence on the number of incidents as discussed in Pollution risk drivers section. We used only those precipitation variables found to be materially explanatory as independent variables, i.e. average rainfall for group 1 and both average rainfall and surges for group 2.

	Group	Constant	Average Rainfall	Rainfall Surges	R^2
1	Explained by level of rainfall	8.9	3.2	N/A	0.57
2	Explained by level of rainfall and surges	3.9	1.9	0.2	0.49

Table 3: Pollution regression coefficients

The R^2 values resulting from the regression imply there are other explanatory factors for pollution incidents, however in combination with the correlations between incidents and precipitation demonstrate there is a clear relationship. The positive regression coefficients demonstrate this relationship is positive, i.e. that increased rainfall is likely to increase the number of pollution incidents.

We have used this regression model with the precipitation data per the MET office's forward-looking scenario for 2025-2030 to determine the expected impact on pollution incidents under that scenario. As the MET office data represents a base case, we also considered a number of precipitation scenarios and determined the implied pollutions impact for the notional company. Scenarios were informed by analysis of historic data and are as below. The percentages derived were applied to the MET office base scenario.

- **More rainfall.** We used the P10 scenario from the CMIP6 ensemble climate model projections, available from the world bank, to determine a plausible scenario of a 15% increase in rainfall in AMP8[4]. We also considered half this impact, a 7.5% increase in total rainfall.
- **More seasonal rainfall.** Rainfall is expected to increase in the winter and decrease in the summer[5]. To estimate a plausible increase in winter rainfall, we the CMIP6 historic data set from 1950-2014 and determined the P10 rainfall to be 25% than the long-term average. We therefore considered (1) a scenario of 25% increase in winter rainfall in AMP8 and, (2) a scenario with half the rainfall increase, 12.5%, during AMP8 winters.
- **More surges.** To capture increased volatility of rainfall in the scenario analysis, we used days with >50mm of rainfall from the CMIP6 SSP585 P10 scenario as a proxy for surge patterns. We derived a

[4] Climate Change Knowledge Portal Data Catalogue, World Bank (climateknowledgeportal.worldbank.org)
[5] Changing climate risk in the UK: A multi-sectoral analysis using policy-relevant indicators - ScienceDirect, section 2.6

57% increase in surges from AMP7 to AMP8. We therefore used 50% and 25% in our scenario analysis.

Under each scenario, we calculated a % change in incidents and applied this to the standardised ODI target for AMP8 to derive a £m impact using both the incentive rate of £1.45m per the Draft Determination. Financial impacts were then scaled to be presented as a proportion of notional company AMP8 regulated equity.

Note that this methodology for calculating financial impact assumes baseline performance as the ODI target. In reality, given these targets represent substantial reductions from AMP7 levels a baseline underperformance is likely and therefore the financial impact of these scenarios would be worse than those presented in this analysis.

1.2. Named storms

1.1.5. Evidence and data inputs

Until FY23, pollution incidents as a result of named storms were excluded from reportable pollution incidents. However, this exclusion has now been revoked such that any future incidents as a result of named storms will contribute to a water companies ODI performance. This creates additional risk for the company.

We submitted a Freedom of Information request to the Environment Agency regarding incident exclusions due to named storms across the water sector[6], with the following results. Note, all incidents were category 3.

Storm / Company	Arwen (2021)	Eunice (2022)	Franklin (2023)
Anglian Water (ANH)	0	1	0
Welsh Water (WSH)	0	0	0
Northumbrian Water (NES)	40	0	0
Severn Trent (SVE)	22	0	0
South West Water (SWB)	0	0	0
Southern Water (SRN)	0	85	0
Thames Water (TMS)	0	7	0
United Utilities (UUW)	12	3	1
Wessex Water (WSX)	0	7	0
Yorkshire Water (YKY)	0	1	0
Total	**74**	**104**	**1**

Table 4: AMP7 Pollution incidents excluded due to being from named storms

To understand the relationship between incidents due to these storms and the geographical areas affected, we used the MET office reports for Arwen[7] and Eunice[8]. Storm Franklin was excluded due to only resulting in "Yellow" warnings in England and Wales, with Northern Ireland seeing "Amber" warnings[9] and therefore no material pollution incidents. See the below charts, compared with the company area map[10], which show clearly:

- **Storm Arwen** most severely impacted the northeast coast of England, an area predominantly served by NES. "Yellow" warnings were issued for the northwest and parts of the midlands, hence some incidents incurred by SVE and UUW. As this storm did not impact the south of England, Yorkshire or Wales no resulting pollution incidents were reported by the other WaSCs.

[6] This data is available to be used per the Open Government Licence.
[7] "Storm Arwen, 26 to 27 November 2021", Met Office (2021)
[8] "Storms Dudley, Eunice and Franklin, February 2022", Met Office (2022)
[9] "Storm Franklin named", Met Office (2022)
[10] Contact details for your water company - Ofwat

- **Storm Eunice** most severely impacted the southeast and part of the southwest of England. Most of the areas served by Southern Water were subject to "Red" warnings which resulted in 85 pollution incidents. Other companies experienced incidents; however these were to a lesser extent due to their geographical location. However, TMS and WSX were the next most impacted companies as they operate in the area impacted by Storm Eunice. Notably less than Southern Water's.

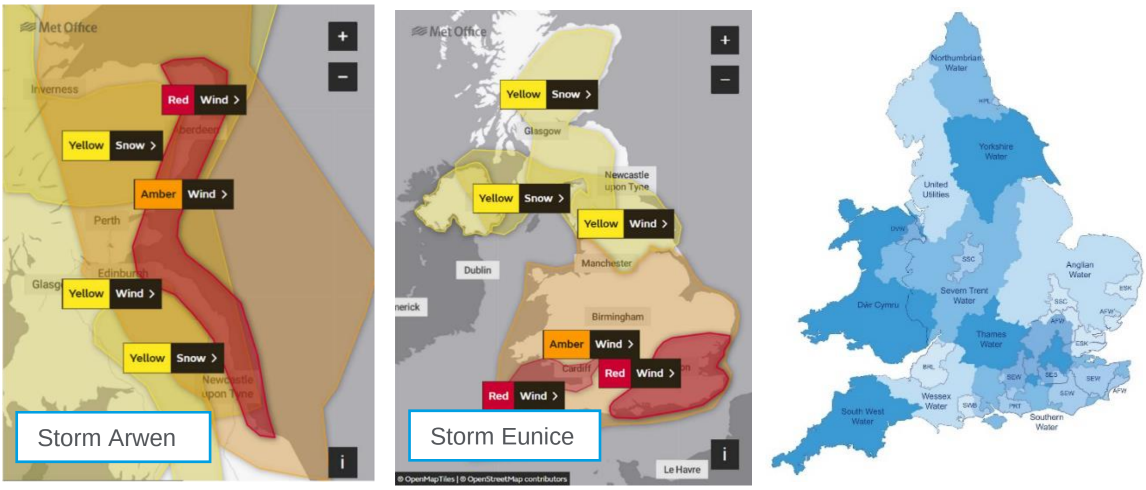

Figure 8: Impact of named storms Arwen and Eunice

These data demonstrate the exposure to named storms is a notional company issue, with 8 out of the 10 WaSCs experiencing pollution incidents as a result. As incidents due to named storms are no longer excluded from ODI reporting in 2023, this increases the risk to which the notional company is exposed.

1.1.6. Methodology

Per the freedom of information request submitted, pollution incidents in 2021 and 2022 were most notably caused by storms Arwen (74 incidents) and Eunice (104 incidents) respectively. These storms impacted 8 of the 10 WaSCs. Our analysis considers only those companies with operational area materially impacted by Arwen and Eunice, and we have therefore omitted the incidents incurred by companies ANH, WSH, SWB, and YKY. We calculated the average incidents incurred due to named storms:

Company	Storm	Pollution Incidents
Northumbrian Water (NWL)	Arwen	40
Severn Trent (SVE)	Arwen	22
Southern Water (SRN)	Eunice	85
Thames Water (TMS)	Eunice	7
United Utilities (UUW)	Arwen	12
United Utilities (UUW)	Eunice	3
Wessex Water (WSX)	Eunice	7
Average		**25**

Table 5: Total pollution incidents caused by Storm Arwen and Storm Eunice

This indicates, on average, named storms result in 25 pollution incidents. However, this likely understates the true risk exposure as TMS, UUW and WSX only had small operational areas impacted by Arwen and Eunice. These storms resulted in "Red" warnings for material amounts of NES and SRN's operational areas – the average incidents incurred for these companies was 63. These averages represent a proxy for notional company performance in the event of a named storm as these incident data are across six companies with diversified asset health and operational strategies.

From the sector data we derived two scenarios, (1) a moderate named storm impact as represented by the total average of 25 incidents, and (2) a severe named storm impact where the operating region is substantially covered by a "Red" warning, represented by the average incident count of 63 from NES and SRN during Arwen and Eunice respectively.

As part of the PR24 business plan submission, we proposed targeted total incidents of 581 across AMP8. For the notional company, a named storm could increase this by 25 incidents under scenario 1, a 4.3% increase, or alternatively by 63, a 10.8% increase, if “Red” warnings are in place for most the southeast region as in scenario 2.

The incentive rate on pollution incidents proposed by Ofwat is £1.45m per pollution incident per 10,000km of sewers. Average sewer length in AMP8 is expected to be 40,300 km. Therefore, 27 incidents relating to named storms translated into 6.2 standardised incidents with a £m impact of £9.7m. In the severe scenario where most of the southeast region is covered by a “Red” warning, such a named storm could induce 15.6 standardised incidents with a financial impact of £22.6m. As a proportion of notional company AMP8 regulated equity these financial impacts are 0.06% and 0.13% of RoRE respectively.

1.3. Serious pollution incidents

1.1.7. Evidence and data inputs

We obtained data from the EA on pollution incidents across the WaSCs in England and found a total of 145 serious incidents between 2020 and 2022 with an annual average per WaSC of 5 serious incidents. Whilst serious pollution incidents may have relationships with actual company characteristics such as asset health, this indicates that serious pollution incidents are a notional company issue.

Whilst serious pollution incidents are relatively infrequent, they are difficult to predict and therefore difficult to mitigate. This is due to high variability in (1) geographic location, as demonstrated by the below map showing serious pollution incidents from 2020-2023 across the region served by Southern Water; and (2) asset type, as demonstrated by EA data on serious pollution incidents in the sector from 2020-2022. Note, we considered geographical location of Southern Water incidents only due to unavailability of sector data.

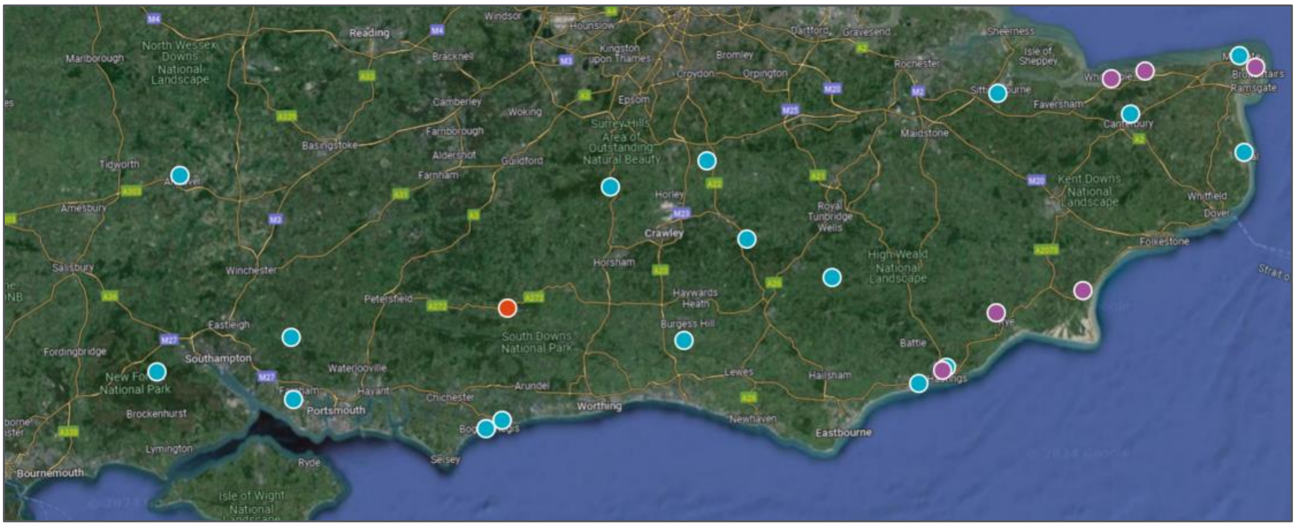

Figure 9: Map of Southern Water serious pollution incidents 2020-2022

Company	Foul sewer	Sewage treatment works	Rising main	Pumping station	Other	Total incidents
	Serious pollution incidents (2020-2022)					
Anglian	10	8	10	4	3	35
Northumbrian	0	2	0	0	0	2
Severn Trent	1	2	0	2	1	6
South Western	6	2	4	1	0	13
Southern	6	3	4	8	0	21
Thames	18	15	7	0	2	42
United Utilities	0	1	0	0	0	1
Wessex	10	1	1	0	2	14
Yorkshire	2	1	4	3	1	11
Total	**53**	**35**	**30**	**18**	**9**	**145**
Median	**6**	**2**	**4**	**1**	**1**	**13**

Table 6: Serious pollution incidents across the water sector, 2020-2022

We used our data to determine the months in which serious incidents most frequently occur, as presented in the below graph. Our data has been used as a proxy for the notional company in absence of granular monthly data for the sector. This suggests serious incidents occur most frequently during summer and autumn, and least frequently in winter. There is a positive relationship with temperature and a negative relationship with rainfall, i.e. serious pollution incidents are more likely to occur when temperatures are higher and rainfall is lower.

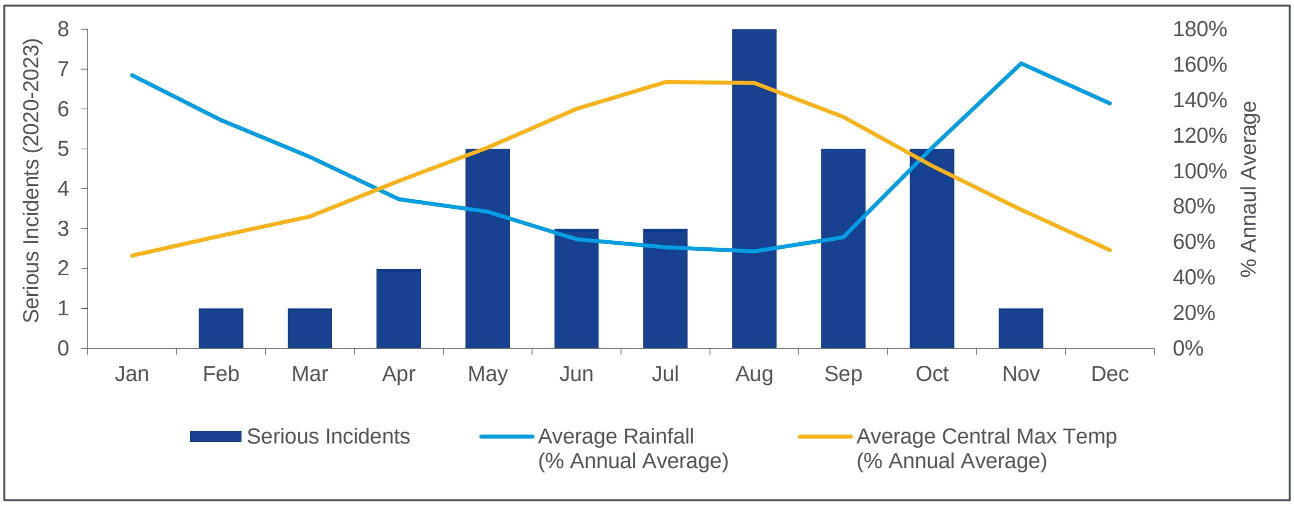

Figure 10: Serious Pollution Incidents by Month (Southern Water, 2020 - 2023)

1.1.8. Methodology

Whilst statistical analysis is limited by the small population of serious incidents, we performed correlations analysis using our pollutions data. Company specific data was used as sector data is not available with the required granularity. We would expect the relationship between weather and serious incidents to hold true for the notional company.

Correlations between serious incidents and rainfall were calculated on a quarterly basis between 2020-2022 and were determined to be -0.33 with rainfall and 0.64 with maximum temperature. Hotter weather and less rainfall results in less dilute pollutants.

To understand how operational improvement may improve performance on the serious pollutions ODI, we compared our performance on total pollution incidents and serious pollution incidents between 2020-2022 with 2023 where we implemented an asset improvement plan mainly focusing on pumping stations and rising main calming. We used our data in lieu of the required information on improvement plans undertaken by other companies in the sector.

Company	Total Pollution Incidents	Serious Pollution Incidents
2020-2022 annual average performance	377	5
2023 performance	225 - 234 (subject to EA decision)	12

Table 7: Impact of PIRP on serious pollution incidents

This demonstrates that whilst our improvement initiative materially improved total pollution incident, there was no such improvement to serious pollutions and in fact we experienced more in 2023. This could be because serious incidents occurred from assets or at sites which weren't subject to improvement, however 7 serious incidents in 2023 were at water pumping stations or rising mains, asset types in scope of the improvement plan. Overall, this suggests that serious pollutions are very difficult to mitigate with operational improvements and therefore regulatory mitigations may be required.

1.4. Correlations

1.1.9. Evidence and data inputs

The correlation analysis on PCs was undertaken based on our own data across nearly all PCs. Data was consistently sourced from AMP7 performance data at the lowest level of frequency feasible: weekly, monthly, or quarterly. Some data was available geographically and we grouped this by county: Hampshire, Isle of Wight, Kent, Sussex and to a limited degree Surrey and Wiltshire.

Bathing water quality, river water quality, and discharge permit compliance were excluded from this analysis as the reporting is only available on an annual basis. Further analysis is needed to determine how discharge permit compliance could be analysed on an incident basis given the different ways a site level failure can occur that are not consistent across sites.

We would reasonably expect a relationship between discharge permit compliance and river water quality however these metrics were excluded from this analysis. Ofwat defines river water quality as the per cent reduction in P-level found in site effluent. The relationship is based on the permit levels for phosphorous in place at the baseline year 2019/20 which determined the available scope for reducing the effluent P-level and an estimate for reduced phosphorous entering rivers from other efforts with third parties[11]. This analysis was included in the Risk Technical Annex citing that sector average P-level permits per wastewater treatment site was 1.22mg/L[12] and the AMP8 permit levels are expected to tighten further for the sector. For reference, our permit levels for phosphorous are tightening by 25% in 2025 ahead of AMP8. This degree of tightening will drive the performance on river water quality given the baseline year is 2019/20 and complying with the permits would likely result in achieving the river water quality PC target as well. To achieve lower P-levels, companies could increase ferric dosing to remove greater amounts of phosphorous from the effluent and improve performance on river water quality while increasing the risk they breach the permit level of iron in the effluent. The makes the relationship between discharge permit compliance and river water quality more dynamic and dependent on both ferric and P-level permits, the techniques available to remove phosphorous and level of phosphorous and iron present in the influent. This dynamic requires further analysis and may impact notional company performance. However, our analysis has left this correlation a nil in the absence of supporting data.

[11] Performance commitment definition - River water quality (ofwat.gov.uk) p3
[12] srn57-risk_redacted.pdf (southernwater.co.uk) p66

The below table indicates the lowest level frequency of reporting incorporated into the correlation analysis for each PC and whether the information could be split geographically into county or remained at the regional level (i.e. the Southern Water region):

Performance commitment	Frequency of report	Geographic level data
Pollution Incidents	Weekly	County
Serious Pollution Incidents	Weekly	County
Internal sewer flooding	Weekly	Region
Sewer Collapse	Weekly	Region
External Sewer Flooding	Weekly	Region
Water Supply Interruptions	Weekly	County
Leakage	Weekly	County
PCC	Monthly	Region
CRI	Weekly	County
Storm Overflows	Weekly	County
CMEX	Quarterly	Region
DMEX	Quarterly	Region
Business demand	Monthly	Region
Unplanned outage	Weekly	County
Customer contacts on water quality	Weekly	County
Mains repairs	Weekly	County

Table 8: PC data availability

We then considered the units for the correlation analysis. Because not all the ODI definitions necessarily lend themselves well to correlation analysis, we identified the underlying driver of the PC performance and modelled these. The below table summarises where we made these simplifying adjustments to the data to allow for correlation analysis:

Performance commitment	Correlation analysis input	PC definition
Leakage	Ml/day	Per cent reduction from baseline of the three-year average
PCC	Ml/day	Per cent reduction from baseline of the three-year average
Storm Overflows	Number of spills	Number of spills / CSO + unmonitored spills adjustment
CMEX	CES score	CES + CSS score
Business demand	Ml/day	Per cent reduction from baseline of the three-year average

Table 9: Adjustments to PCs to facilitate correlations analysis

Leakage, PCC and business demand were not feasible to analyse as a three-year average given the limited data points and so was translated to actual leakage or consumption per day. The storm overflows unmonitored spill adjustment is not incident driven and added only as an annual figure. The adjustment is not related to the same underlying risk drivers as spills. Therefore, we excluded to avoid distorting the results. The CMEX score was decomposed into the CES and CSS, and we focused on the CES. The CSS is likely driven by customer service levels provided by our customer facing teams and not driven by the same risk factors as other PCs, while the CES score is driven by customers' general perception of our business and is inherently related to our PC performance, especially environmental performance.

Any other PCs with standardising metrics were unstandardised to facilitate this analysis. The standardised values did not create a material difference as most standardising adjustments (like km of sewer) remained relatively constant year to year.

Finally, we also considered average daily rainfall and average daily temperature in our analysis to better understand where PCs had common risk drivers. This data was sourced from weather stations we have around the southern water region.

1.1.10. Methodology

Once the data was formatted and arranged, we calculated the Pearson correlation coefficient between each PC for each time interval and geographic split. The results of the correlation analysis were considered for each combination of PC at each level of granularity available to understand the dynamics of the relationship.

Then, we considered each relationship against a materiality threshold to eliminate low strength relationships that we felt were not sufficiently different from zero. The final results were then considered against the possible explanations from an operational perspective to justify the relationship was not due to coincidence in the data and that the actual company risk was not driving the relationship. The below provides further detail to each step of the analysis:

1) The correlations were calculated at the quarterly, monthly, and weekly level plus the monthly by county level where possible. This resulted in 368 distinct correlation coefficients across 120 different relationships between any two PCs.

2) All 120 combinations of PCs were then considered individually across all available frequency and geographical variations. For example the relationship between total pollutions incidents and serious pollutions incidents was considered as follows:

Frequency – Geographic split of correlation	Correlation coefficient
Quarterly – Region	-0.40
Monthly – Region	-0.11
Weekly – Region	0.08
Quarterly – County	0.36
Monthly – County	0.24

Table 10: Frequency - Geographic split of correlations

Given the relationship between total pollutions and serious pollutions is reasonably expected to be positive given that all serious pollutions are also total pollution incidents, we sought to break down the data to most granular level possible. Because one incident counts as both, the lowest possible geographical split and most frequent time interval would most accurately capture the relationship. In general, as most of the data was at the incident level and expected to coincide in both time and geography, this approach was applied to almost all relationships.

While a correlation cannot be checked for statistical significance, we sought to validate the results of our analysis a variety of ways to ensure any correlations identified could withstand scrutiny from an informed third party and would be applicable to a notional company. Statistical significance can only be tested for predicted values and correlations do predict future values, rather indicate the direction and strength of the relationship between two variables. The following steps include those validations to increase the robustness of our results.

3) Firstly, we applied a materiality threshold of +/-0.20 to eliminate weak relationships that we felt were difficult to justify as materially different from zero. This eliminated 57 correlations from our analysis.

4) Then we considered the logical underpinning of the relationship from an operational perspective and validated each relationship with our internal team. Where there was no common risk driver or logical relationship between two PCs, we excluded these relationships regardless of the mathematical results. For example the DMEX score was not found to be logically related to any other PC from either a direct relationship or a common risk driver, and these results were excluded. This removed a further 49 correlations.

5) Finally, we analysed the remaining results with a logical underpinning that were materially different non-zero and considered whether the actual company risk was influencing the relationship. The relationships excluded were as follows:

   a) **CMEX / Water supply interruptions:** given our performance on water supply interruptions it's likely the strength of the relationship is influenced by our actual company performance. While customers likely do consider supply interruptions when completing the CES survey across the country, this is unlikely to make as material an impact as our data suggests due to our position in sector performance. Therefore, we excluded to remain conservative.

   b) **PCC / Business demand:** due to the Covid-19 pandemic PCC increased drastically and business demand decreased sharply as the government-imposed lockdowns preventing people from going into the office. This drove a strong negative correlation between PCC and business demand. Since we do not expect the relationship to hold in the future given the return to work we have excluded this relationship.

The remaining relationships were considered valid for a notional company like Southern Water and were included in our analysis of notional company risk as such. More detail on each relationship can be found in the results subsection.