Technical note

Project:	UKWIR Resilience Metrics	**To:**	The WRLTPF Steering Group
Subject:	Drought Resilience Metric: Development of a *'Certainty Grade'*	**From:**	Helen Gavin, Doug Hunt (Atkins) *Helen.gavin@atkinsglobal.com*
Date:	22/01/2018	**Version**	4

Table of Contents

1 Introduction 1
2 Spatial focus of the metric 1
3 Semi-Qualitative 'Certainty Grade' 2
3.1 Existing Ofwat confidence grade approach 2
3.2 Proposed amendment for the Drought Resilience Metric 2
3.2.1 Methodology Grade: the rigour or sophistication of the drought definition process 2
3.2.2 Risk Score: how close each company may come to implementing restrictions 3
3.3 The 'Acceptability' of the Resulting Certainty Grade 4
4 Reporting the Drought Resilience Metric 5
5 Conclusion 5

Tables

Table 1 The different spatial ways in which the metric can be calculated 2
Table 2 The methodological rigour or sophistication of the drought definition process 2
Table 3 Calculating the surplus or deficit for the 1:200 drought event 3
Table 4 Risk Label Definitions 4
Table 5 'Acceptability' colour band definitions 5
Table 6 Compatibility of the 'Acceptability' colour and the Certainty Grade 5

1 Introduction

Ofwat's proposed metric for drought resilience was discussed in a meeting of the WRLTPF Steering on 25 October 2017. The proposed metric is "*the percentage of the population the company serves that would experience severe supply restrictions (e.g. standpipes or rota cuts) in a 1 in 200 year drought".*

A key feature of the discussion centred on how companies could ascribe a degree of confidence in their reported value for the metric, given sources of uncertainty, including:

- data uncertainties, for example whether the company has a robust understanding of deployable output (DO) for all its sources, which would affect the impact of drought upon water supply; and
- the different methodologies that could be used to define a 1 in 200 drought in the future, given that droughts vary in intensity, duration, frequency, and geographical location, plus climate change has, or will cause, droughts to become more frequent and severe, invalidating traditional return period calculations.

These uncertainties mean that it is not possible to assign a quantitative level of certainty to a value ascribed to this metric. However, a semi-qualitative '*Certainty Grade'* could be used instead, adapted from the established Ofwat 'confidence grade' process.

Such an approach was subsequently developed, and further discussed in the WRLTPF Steering Group meeting on 24 November 2017. The approach was largely accepted, but some amendments were suggested.

Having considered the Steering Group's comments, this technical note presents the amended approach.

2 Spatial focus of the metric

As stated above, the proposed metric is "*the percentage of the population the company serves that would experience severe supply restrictions (e.g. standpipes or rota cuts) in a 1 in 200 year drought".* There are two ways to calculate this metric from a spatial context: for each water resource zone (WRZ) or as a summary for the company. Table 1 sets out the attributes of each approach.

Overall it is recommended that the calculations are done at the WRZ level, as there is no reasonable mathematical way in which the WRZ Certainty Grades can be collated into a single one.

Table 1 The different spatial ways in which the metric can be calculated

Approach	Advantage	Disadvantage
Zonal level	• Different zones may have a different level of resilience to the design drought. Applying the calculation on a zonal basis will help companies to identify more vulnerable zones. • The results from the most vulnerable zone could be reported for the metric allowing any improvements over time to be easily observed	• Calculating the metric for each zone is more burdensome • The metric asks for the percentage of the population affected; reporting on a single (worse case) zone may underestimate the true affected population
Company level	• Only one calculation is needed • Gives an overview value that is less sensitive to changes that arise from (small scale) methodological changes which may mask real improvements in resilience. • Better represents what may happen operationally during such a drought episode in which a water company would try to supply water to a zone in deficit from another zone in surplus before implementing restrictions	• May hide the detailed picture i.e. obscure the vulnerability of any particular individual zones to restrictions • Difficult to combine into a single value as each letter in the Certainty Grade represents a different methodology.

3 Semi-Qualitative 'Certainty Grade'

It should be noted that the term '*Certainty Grade'* has been used here to make it clear that, although it is similar to the 'confidence grade' used by Ofwat reporting, it has some fundamental differences that need to be borne in mind by companies, regulators and stakeholders.

3.1 Existing Ofwat confidence grade approach

Box 1 outlines the Ofwat approach, in which a letter and number are selected to expresses the current level of reliability and accuracy in the reporting number respectively. These sections are combined to form a confidence grade. Ofwat set a minimum confidence grade that should be attained by a company for a given metric, with a target confidence grade of A2, A3, B2 (or better) expected in most cases. Companies would be expected to improve reporting performance over time, leading to the target confidence grade or higher. Where the reported confidence grade is below the minimum or expected target, or a deterioration from previous reported levels, the company would need to set out the reason(s) why, together with an action plan for improvement. The reported confidence grade should express the current status of the data, not the intended future status following any improvements or actions.

3.2 Proposed amendment for the Drought Resilience Metric

As set out above, there are two dimensions that contribute to the Ofwat confidence grade: Reliability and Accuracy. For the Drought Resilience Metric '*Certainty Grade'*, the two dimensions can be repurposed as follows.

3.2.1 Methodology Grade: the rigour or sophistication of the drought definition process

The first dimension would reflect the sophistication of the method used to derive a 1:200 year drought event. For example, more certainty would be placed in a sophisticated stochastic modelling approach incorporating latest findings from research on climate change effects on rainfall and hydrological processes, using many years of simulations, than in a method that relies upon extrapolation from historic, short, time-series data which may not include severe drought episodes. Table 2 shows how the Ofwat reliability band has been repurposed to express the sophistication of the drought definition process.

Table 2 The methodological rigour or sophistication of the drought definition process

Method Grade	Description
A	Use of drought event data from latest research outputs that have employed global climate model ensemble runs to simulate droughts, such as those arising from the MaRIUS Project within the NERC Drought and Water Scarcity Programme[1],
B	Use of perturbed data from existing observed weather datasets, using stochastic processes, in a "weather generator" approach used with UKCP09 Regional Climate and Future Flows (2012[2].)
C	Extrapolation from limited sample of historic company data
D	Use of arbitrary or 'yardstick' deviations based on historic observations

[1] http://www.nerc.ac.uk/research/funded/programmes/droughts/

[2] Future Flows Project (2012). www.ceh.ac.uk/our-science/projects/future-flows-and-groundwater-levels

Box 1 Ofwat Reliability and Accuracy Bands, and resulting Confidence Grade

The confidence grade is a combination of reliability and accuracy bands, following the steps and tables below.

- **Step 1**: Assign a Letter to indicate the **reliability** of the dataset used to report the metric, e.g. "A" (highly reliable)
- **Step 2**: Select a Number to reflect the **accuracy** of the reported number to the true value, e.g. "2" (estimated to be within +/- 5%)
- **Step3**: Combine the results of Step 1 and 2 to produce the **confidence grade**, e.g. "A2" (a highly reliable data estimated to be within 5% of the true value). Note that some reliability and accuracy band combinations are considered to be incompatible (marked by n/a in the table below).

Reliability Bands

Label	Description
A	Sound textual records, procedures, investigations or analysis properly documented and recognized as the best method of assessment
B	As A, but with minor shortcomings. Examples include old assessment, some missing documentation, some reliance on unconfirmed reports, some use of extrapolation
C	Extrapolation from limited sample for which Grade A or B data is available
D	Unconfirmed verbal reports, cursory inspections or analysis

Accuracy Bands

Label	Accuracy to or within +/-	but outside +/-
1	1%	
2	5%	1%
3	10%	5%
4	25%	10%
5	50%	25%
6	100%	50%
X	Accuracy outside +/- 100 %, zero or small numbers or otherwise incompatible	

Compatible Confidence Grades

Accuracy Band	Compatible Confidence Grades			
	Reliability Band			
	A	B	C	D
1	A1	n/a	n/a	n/a
2	A2	B2	C2	n/a
3	A3	B3	C3	D3
4	A4	B4	C4	D4
5	n/a	n/a	C5	D5
6	n/a	n/a	n/a	D6
X	AX	BX	CX	DX

3.2.2 Risk Score: how close each company may come to implementing restrictions

The second dimension would cover how close a company would come to implementing Emergency Drought Order ('Level 4') restrictions (rota cuts and standpipes) on its customers doing a 1:200 drought event.

The Risk Score can be interpreted as follows:

- For a WRZ with a surplus supply/demand balance at the 1 in 200 year level, the score represents the amount and reliability of this surplus water; and
- For a WRZ with a negative supply/demand balance (i.e. it is not resilient to the 1 in 200 year event) then the score represents a view of the level of deficit.

The calculation involves a number of steps for each WRZ, and makes use of data in WRMP Table 10, as shown in Table 3.

Table 3 Calculating the surplus or deficit for the 1:200 drought event

Step	Detail	Output
Step 1A	For each WRZ, calculate the volume available for the 1:200 drought event: *WAFU minus outage minus demand plus Table 10 (Drought Interventions) benefits*	The surplus or deficit volume at 1:200 year drought
Step 1B	Express the output from Step 1A as a percentage of the Target Headroom volume, from WRMP Table 10	An '*Optimistic'* percentage of the surplus or deficit relative to the WRZ Target Headroom
Step 2A	Taking the outputs from Step 1A, determine how much of the surplus or deficit depends on supplies and operational interventions outside of the company's direct control, such as bulk supplies (where there is uncertainty over the volume that might be delivered during a severe regional event); Drought Permits and Drought Orders.	The surplus or deficit volume at 1:200 year drought within the company's direct control
Step 2B	Express the output from Step 2A as a percentage of the Target Headroom volume, from WRMP Table 10	A '*Pessimistic'* percentage of the surplus or deficit relative to the WRZ Target Headroom

As shown in shown in Table 3, two percentage values are calculated:

- The '*Optimistic*' or 'absolute' surplus or deficit relative to WRZ Target Headroom; and
- The '*Pessimistic'* or 'worst case' surplus relative to WRZ Target Headroom, which does not include water from bulk supplies[3] or Drought Permits/Orders

These outputs are then used to select an appropriate Risk Label which is relevant to the magnitude of the surplus or deficit, using Table 4. In contrast to the Ofwat accuracy confidence grade table (see Box 1), where the label of "1" defines the highest accuracy or smallest margin, here the descriptive order of the labelling is reversed so that "1" relates to the greatest or most secure surplus, or the largest deficit. This is done to preserve the norm that a label of '1' is the most aspirational label.

Table 4 Risk Label Definitions

Risk Label	Resilient WRZs, i.e. have a surplus	Non-resilient WRZs, i.e. have a deficit
1	'Pessimistic' surplus % is > 125% of Target Headroom	'Optimistic' Deficit >= 125% of Target Headroom but 'Pessimistic' Deficit <=100% of Target Headroom
2	'Pessimistic' surplus % <125% but still > 100% of Target Headroom	'Optimistic' Deficit >= 100% of Target Headroom but 'Pessimistic' Deficit within 50% to 100% of Target Headroom
3	'Optimistic' surplus % >= 100% of Target Headroom, but 'Pessimistic' surplus only 50% to 100% of Target Headroom	'Optimistic' Deficit within 50% to 100% of Target Headroom but 'Pessimistic' Deficit within 0% to 50% of Target Headroom
4	'Optimistic' surplus % >= 100% of Target Headroom, but 'Pessimistic' surplus only 0% to 50% of Target Headroom	'Optimistic' Deficit within 0% to 50% of Target Headroom but 'Pessimistic' Deficit within 0% to 50% of Target Headroom
5	'Optimistic' surplus % >= 100% of Target Headroom, but the 'Pessimistic surplus' shows an SDB deficit	'Pessimistic' Deficit within 0% to 50% of Target Headroom AND there are potential interventions not included in WRMP Table 10
6	'Optimistic' surplus % >= 100% of Target Headroom, but the 'Pessimistic surplus' shows an SDB deficit AND there is a significant reliance on higher risk interventions to generate the initial surplus	'Pessimistic' Deficit within 0% to 50% of Target Headroom AND there are no allowances for TUBs, NEUBs or Permits/Orders in WRMP Table 10
X	*Not used*	

3.3 The 'Acceptability' of the Resulting Certainty Grade

A final element is now introduced.

While Ofwat set a minimum or aspirational confidence grade for particular metrics, this is not the approach proposed here, as companies may need or wish to adopt more or less sophisticated methods to determine the impact of drought, depending upon their particular situation. Instead, the company selects the 'acceptability' of the Certainty Grade by stating whether the methodology used is appropriate to the risk for the Company and the WRZ. This is done by a choosing a particular colour band that suitably defines the Certainty Grade, as shown in Table 5. Only certain colour bands are compatible with particular Certainty Grades, as shown in Table 6.

The final form of the Drought Metric Certainty Grade would therefore range from:

[3] This will be open to some interpretation, but needs to take into account the hydrological nature of the supply and the robustness of current sharing agreements during severe drought events.

- **A1, Blue**: which represents a sophisticated approach using latest rigorous data or modelling advances where the available 'pessimistic' surplus is greater than 125% of the WRZ Target Headroom; to
- **D6, Red**: which represents a rudimentary method or very unreliable data, and where the WRZ suffers a volume deficit equating to all or half of the WRZ's target headroom.

Table 5 'Acceptability' colour band definitions

Colour	Acceptability definition
Blue	Very certain, no need to review unless there is a large change to the SDB
Green	Certain; approach and margin are acceptable, but there may be some benefit in reviewing either the method or the role of transfers, Orders and Permits on the SDB.
Amber	Some uncertainty; the classification of the WRZ is relatively uncertain and the method and/or the assessment of transfers, Orders and Permits should be reviewed at the next WRMP.
Red	Significant uncertainty; the adopted method is not appropriate and further work on the method and assessment of transfers, Orders and Permits is required.

Table 6 Compatibility of the 'Acceptability' colour and the Certainty Grade

Risk Score	Methodology Grade			
	A	B	C	D
1	A1	B1	C1	D1
2	A2	B2	C2	D2
3	A3	B3	C3	D3
4	A4	B4	C4	D4
5	A5	B5	C5	D5
6	A6	B6	C6	D6

4 Reporting the Drought Resilience Metric

The number (percentage) reported for the Metric represents the company as a whole. It is determined by performing the calculations set out in Section 3 to identify which (if any) WRZ(s) would be in deficit and therefore the number of customers that would experience severe supply restrictions. The number of affected customers in each WRZ is then totalled and expressed as a percentage against the total population served by the company.

The Certainty Grades are computed on a WRZ level: however only one can be used to accompany the Drought Metric number. For simplicity, it is recommended that the lowest Certainty Grade of all the WRZs is selected to accompany the Drought Metric number.

It is also recommended that a brief narrative accompanies the reported Drought Metric, to summarise the spread of results at a WRZ level, and provide a commentary on relevant factors such as whether the lowest Certainty Grade is indicative of the wider resilience of the company, and (where relevant) what steps are being actioned to reduce the number of customers at risk of severe water supply restrictions during a 1:200 drought event.

5 Conclusion

This technical note sets out a way to calculate and assign a semi-qualitative 'Certainty Grade' to the Drought Resilience Metric. The recommended approach is as follows:

a) The Certainty Grade evaluation is applied at the WRZ level;
b) The Certainty Grade comprises a 'methodology' letter (A-D) plus 'risk' score (1-6) and 'acceptability' colour band (Blue, Green, Amber, Red);
c) The 'methodology' letter describes the sophistication of the method used to determine the 1: 200 drought;
d) The 'risk' score represents the surplus or deficit volume, and how much is within the control of the company;
e) The 'acceptability' colour band states the company's view on the Certainty Grade for each WRZ;
f) The final form of the Certainty Grade combines the three dimensions, e.g. "A1, Blue"; or "C4 Amber" etc.;
g) The lowest WRZ Certainty Grade accompanies the reported Drought Metric to give the company view; and
h) The Drought Metric is supported by a short commentary to describe the spread of WRZ Certainty Grades and other relevant information.