Pembroke Abstraction Renewal Supporting Information. Arithmetic and Geometric Mean

Document no: GBR01/117389562_1
Revision: 1

Jacobs Document Number: JUKL/B2386292/LIC/R07
Pembroke Marine Monitoring

Executive summary

This report has been prepared in support of the Abstraction Licence Renewal for Pembroke Power Station. It is part of a suite of documents to provide the relevant information and analysis requested by NRW.

This report is provided to respond to the following request by NRW to RWE:

Provision of annual entrapment estimates using arithmetic means rather than geometric means: This has not yet been provided. It is acknowledged that work has been done by Jacobs to present comparison between use of geometric and arithmetic means for impingement numbers. This shows the consistent underestimate in entrapment estimates from using the geometric mean. The entrainment data also apparently uses the arithmetic mean – it is not clear why an inconsistent approach is used between the two datasets.

In this report the term entrapment is used to mean to whole effect of cooling water abstraction (encompassing both impingement and entrainment), impingement refers to the portion of the catch that is retained on the screens and entrainment refers to the portion of the catch that passes through the screen. In the annual reporting analysis is undertaken on each portion individually to look at trends, before entrapment as a whole is assessed.

This report responds to the NRW request by presenting the annual impingement estimates based on an arithmetic mean (Section 1), and a justification for the methods used in the calculation of annual entrainment (Section 2). Section 3 also provides for NRW a comparison of the use of geometric mean and arithmetic mean in impingement calculations in the context of annual trends. To date, annual assessments of impingement have been made on scaled data using a geometric mean owing to the inherent variability within the monthly data and the geometric mean providing more realistic estimates of entrapment.

Whilst the individual numbers presented using the geometric and arithmetic means are different, the analysis undertaken in the annual reporting reviews community trends and comparisons within entrapment with the wider monitoring data set and also in wider populations. The trends are still the same and the conclusions made in the reporting of the Pembroke Environmental Monitoring Programme do not change. In addition, the conclusions made in the Stage 2 Report to Inform Habitats Regulations Assessment also remain valid in that *the continued abstraction as already authorised for Pembroke Power Station (alone or in combination with other plans or projects) will not have an adverse effect on the integrity of any European designated sites in view of those sites conservation objectives.*

Contents

Tables

Figures

1. Introduction

This report has been prepared as part of the series of documents to provide the relevant information and analysis requested by NRW in relation to the Abstraction Licence renewal for Pembroke. The suite of documents provides the relevant information requested by NRW in their letter dated June 2024, reference PAN-025790.

1.1 Abstraction

Water is abstracted throughout the year from Pennar Gut for non-evaporative cooling of RWE Generation UK plc's (RWE) Pembroke Power Station. Cooling water is drawn from Pennar Gut, at the mouth of the Pembroke River, and discharged back into Milford Haven. The current licence (see below) is due to expire on the 31st March 2025.

The existing licence (22/61/06/0156) was originally granted by Environment Agency Wales (EAW) on the 3rd February 2009, and reissued by National Resource Wales (NRW) on the 21st November 2014 to reflect the change in name resulting from their reorganisation. The licence allows for the following maximum quantities of water to be abstracted from Pennar Gut, Pembroke Dock (NGR SM9365402652):

- 144,000 cubic metres per hour
- 3,456,000 cubic metres per day
- 1,200,000,000 cubic metres per year
- at an instantaneous rate not exceeding 40 cubic metres per second.

1.2 Document Aim

The aim of this report is to provide the additional information requested by NRW in their PAN-025790 letter, specifically:

"Provision of annual entrapment estimates using arithmetic means rather than geometric means: This has not yet been provided. It is acknowledged that work has been done by Jacobs to present comparison between use of geometric and arithmetic means for impingement numbers. This shows the consistent underestimate in entrapment estimates from using the geometric mean. The entrainment data also apparently uses the arithmetic mean – it is not clear why an inconsistent approach is used between the two datasets."

This document provides the following elements to respond to NRWs request:

- Annual impingement estimates using arithmetic mean.
- Explanation for the use of arithmetic mean for entrainment estimates.

By analysing different potential methods, this report provides a robust justification for the assessments adopted as part of the renewal application; these methods are also aligned with those used in the Pembroke Environmental Monitoring Programme.

1.3 Terminology

In the context of cooling water system effects, the following definitions apply:

- Entrapment: the entry into the cooling water system of aquatic organisms caused by the ingress of water. The term implies that the organism is unable to resist capture owing either to poor or no swimming ability, or to its failure to interpret the water intake as a hazard. The term can encompass both impingement and entrainment or may alternatively be used interchangeably to describe either aspect.

- Impingement: the retention of entrapped organisms on cooling water intake screens that are employed to prevent debris entering the cooling water heat exchangers. To become impinged, organisms must be large enough to be retained by the screen meshes (usually includes e.g. juvenile/adult fish, macroinvertebrates such as shrimps, crabs and large molluscs, and marine plants/algae).
- Entrainment: the passage of entrapped organisms that penetrate the cooling water screens and are returned back to the estuary via the pumps, heat exchangers and other components of the cooling water circuit (typically zooplankton, including ichthyoplankton, and phytoplankton). Note: the size break-point between impingement and entrainment depends on the size of mesh openings selected, the orientation of organisms to the cooling water screens, and the extent to which the screen performance is affected by debris present.

2. Annual Impingement Estimates

Annual estimation of entrapment is already carried out as part of the Pembroke Environmental Monitoring Programme in accordance with methods agreed with NRW. In addition, for the purpose of the renewal application, NRW has requested annual estimation of impingement to be based on arithmetic rather than geometric means.. This data is now provided in Table 1 below.

Table 1 Annual Impingement estimates of abundance and biomass calculated using the arithmetic mean.

Year	Abundance	Biomass		
		(g)	(kg)	(t)
2012	1,204,713	2,212,073.67	2,212.07	2.21
2013	698,436	2,693,042.50	2,693.04	2.69
2014	1,740,022	3,037,620.06	3,037.62	3.04
2015	823,254	2,926,773.76	2,926.77	2.93
2016	1,706,712	4,381,414.47	4,381.41	4.38
2017	107,766	994,244.16	994.24	0.99
2018	309,242	1,724,327.02	1,724.33	1.72
2019	126,833	721,005.91	721.01	0.72
2020	133,187	924,528.90	924.53	0.92
2021	181,155	1,156,798.86	1,156.80	1.16
2022	170,274	807,810.90	807.81	0.81
2023	166,494	829,184.70	829.18	0.83

3. Entrainment Calculations

As noted in the information request from NRW the entrainment data already supplied in support of the renewal application is calculated using an arithmetic mean rather than a geometric mean. The entrainment forms one portion of entrapment and represents the fraction of the abstracted water that passes through the cooling water screens. The entrainment survey is undertaken using a pump sampler where water is sampled for a period of 24 hours at a measured flow rate. Surveys are completed between April and August inclusive, with between one and two samples collected each month (nine samples per annum). The method of sampling means that the catches are standardised to the quantity of water sampled and then scaled for abstraction. Therefore, whilst the sampling flow and cooling water abstraction varies, the entrainment is calculated based on sampled water, influences such as fish behaviour do not influence entrainment, and therefore the data is on a more consistent basis than impingement. Owing to the more standardised sampling pattern and only one to two samples being collected each month it is most appropriate to use an arithmetic mean in this instance. In addition, the entrainment sampling is focused during the peak season for entrainment of fish larvae meaning that the data generally shows a more normal distribution (as opposed to impingement that tends to be more skewed with the peaks and troughs in abundance seen throughout the year) and as such it is more appropriate to use an arithmetic mean.

4. Geometric and Arithmetic Mean Comparison

In order to provide a comparison of the methods that have been used in the assessment of entrapment, both geometric and arithmetic means of the impingement survey data were calculated within R (R Core Team, 2024) utilising the tidyverse package (Wickham *et al.*, 2019) and then subsequently visualised in Sigmaplot. An ANOVA test was performed using JASP (JASP Team, 2024) to check for significant differences between the years with post-hoc testing where appropriate to determine the source of the differences. This was repeated for arithmetic and geometric means of both abundance and biomass.

4.1 Geometric and Arithmetic Mean

Figure 1 shows the geometric as well as arithmetic mean of total abundance (per 10^6 m^3) of fish, with the equivalent for biomass (per 10^6 m^3) shown in Figure 2. Each figure is displayed by month across the entire monitoring period of 2012 to 2023. While it is evident that both versions of the mean accurately visualise the variation in catch size throughout each year (e.g. peak abundance observed during winter months) there is a difference between the estimates based on the method of extrapolation. The seasonal and temporal (within and between year) variations in overall impingement is not unexpected as the susceptibility of fish is strongly influenced by the abundance and distribution of species and size classes within the vicinity of the cooling water intake.

The geometric mean exhibits a far more stable behaviour, while the arithmetic mean is more erratic as point events influence the arithmetic mean in a greater capacity. Due to the intended random nature of the impingement sampling (i.e. 40 samples distributed throughout the year rather than targeting of specific months and tides), it is likely that outlier data will occur during some surveys throughout the year – this inherent variability introduces skew into the data set. As a result, the overarching trend is more accurately represented by a geometric mean. It should be noted that the geometric mean does not always report a lower figure when compared with the arithmetic mean (as can be seen in Figure 2). In addition, the biomass and abundance don't naturally correlate as the biomass can be skewed by the presence of a small number of large fish.

4.1.1 Statistical analysis

Statistical tests were undertaken to look at the trends in the data sets using the geometric and arithmetic means. A direct comparison of the two would not be appropriate as it will show statistical differences that are not relevant; the differences are reflective of the different methods, rather than anything of ecological significance. Therefore, between year differences were looked at for the two methods and the resulting trends were compared to see if they were similar.

Differences in the impingement abundance and biomass values were tested using a Kruskal-Wallis one way Analysis of Variance (on ranks) on all replicate data to identify significant differences between years. This nonparametric test was used as the data was not normally distributed and is consistent with the agreed approach used in the Pembroke annual environmental monitoring programme. Analysis of fish abundance data using Kruskal Wallis analysis of variance did not show a significant difference between years for either geometric or arithmetic mean as the p values were above the threshold value for significance ($p<0.05$) (Table 2). It should be noted that the Kruskal Wallis H value for the two sets of data was similar (arithmetic mean $H = 16.974$ $p = 0.109$, geometric mean $H = 17.843$, $p = 0.085$). This indicates a similar degree of variance within both datasets which would be expected as both methods for calculating the mean will show the variance in a similar way.

For biomass the test values were again similar (arithmetic mean $H = 19.862$ $p = 0.047$, geometric mean $H = 18.828$, $p = 0.064$). As the results for arithmetic mean fall just below the threshold for statistical significance ($p<0.05$) this shows a significant difference between years for that dataset. Further testing showed that this variance was accounted for by biomass values for earlier years of the study (2013 to 2016) being higher than later years (2019 to 2022).

While geometric and arithmetic means are different mathematical functions and as such will produce different results, both will show the variance within the data similarly. The similarity of the outputs of the analysis of variance indicates that this is the case which can be seen by similar H statistics (Table 2). Any change, variation or trend within the data would be detectable whichever method was used.

Table 2. Results of Kruskal-Wallis (H) tests which identified whether geometric and arithmetic mean abundance and biomass varied significantly between years. Significant where p ≤ 0.05 (in bold). The statistically significant different pairwise comparisons are also provided.

	Degrees of Freedom[1]	Kruskal-Wallis H Statistic[2]	p-value[3]	Statistically significant different pairwise comparisons
Abundance geometric mean	11	17.843	0.085	
Abundance arithmetic mean	11	16.974	0.109	
Biomass geometric mean	11	18.828	0.064	
Biomass arithmetic mean	11	19.862	**0.047**	2013 – 2016 significantly higher than 2019 2014 – 2016 significantly higher than 2020 2015 -2016 significantly higher than 2021 2016 significantly higher than 2022

The method adopted in the Pembroke entrapment reports is to represent the 'typical' values entrapped by using the geometric mean rather than the arithmetic mean in calculations as it better represents the central tendency particularly in datasets that have a degree of skew. It has been shown that screen-catch data has a characteristic positive skew, therefore the use of the arithmetic mean in scaling catches is biased (Turnpenny *et. al*., 1983). Furthermore, many fish species swim in groups (schools) which are randomly distributed within the environment and fish themselves are randomly distributed within the groups. As a result, seine net, trawl and impingement catches are often not normally distributed, rather they exhibit random contagious distributions (probability distribution that exhibits a clustering effect).

In ecology this distribution frequently occurs due to the uneven distribution of resources such as food or because of mutual attraction (e.g. fish school for protection from predators). It is therefore common practice, where there is sufficient data, to use the geometric mean when investigating abundance and fluctuations in fish populations (Hutchings, 1996). The distribution of fish species within the impingement and fisheries surveys are varied with both small groups of individuals and large schools of fish observed (referred to as contagious distribution). This observation means that, where appropriate to do so the geometric mean has been used to determine indices of abundance (i.e., where data are highly varied). This provides a more representative estimate of population abundance which can then be compared temporally.

When assessing the potential effects of power station operation, there are a number of factors that are reviewed. An entrapment pressure report is prepared to look at the effect of cooling water abstraction, but

[1] Degrees of Freedom denotes the number of sampling years in the study minus one. In this instance there are 12 sampling years, so the degrees of Freedom are 11.

[2] The H statistic is the output of the Kruskal Wallis test and denotes the level of similarity.

[3] Significant differences are denoted by a p value which is less than or equal to 0.05.

this is also supported by a wider fish report where the whole fish community is assessed (larval, juvenile and adult). These reports work in unison with data from both supporting each other, and patterns observed across the whole programme used to assess the data. Therefore, whilst the individual impingement numbers presented using the geometric and arithmetic means are different, the analysis undertaken in the monitoring reports, looking at community trends and comparisons use the whole picture of what is happening within the local community and also wider populations remains the same. Amending the data to show scaled impingement data based on an arithmetic mean would not amend the conclusions of the Pembroke environmental monitoring reports as any observed population trends and change in communities observed from the Haven fish data are not ecologically significant. There is no evidence of significant community change that is attributable to power station operations and the conclusions made in the reporting of the Pembroke Environmental Monitoring Programme do not change.

This is also relevant for the assessments supporting the abstraction licence renewal. There is no indication of community change resulting from operation of the power station and therefore the method of calculation of the entrapment figure does not change the conclusion of the Habitats Regulations Assessment with respect to the conservation objectives for the SAC; where is it considered that the continued abstraction will not have an adverse effect on the integrity of any European designated sites in view of those site's conservation objectives.

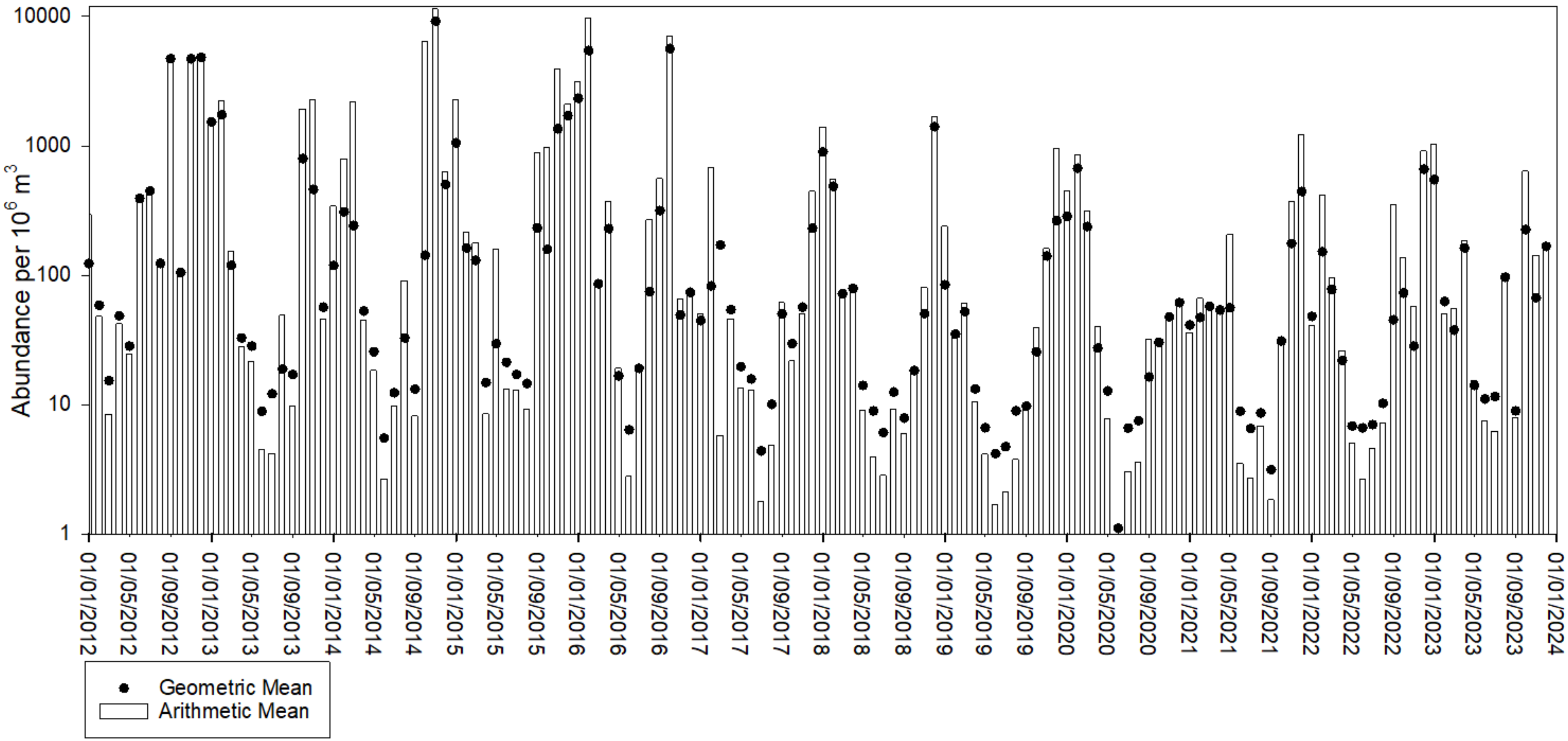

Figure 1. The geometric and arithmetic mean of total monthly abundance recorded per 1×10^{6}m^3 of water abstracted during impingement surveys between 2012-2023.

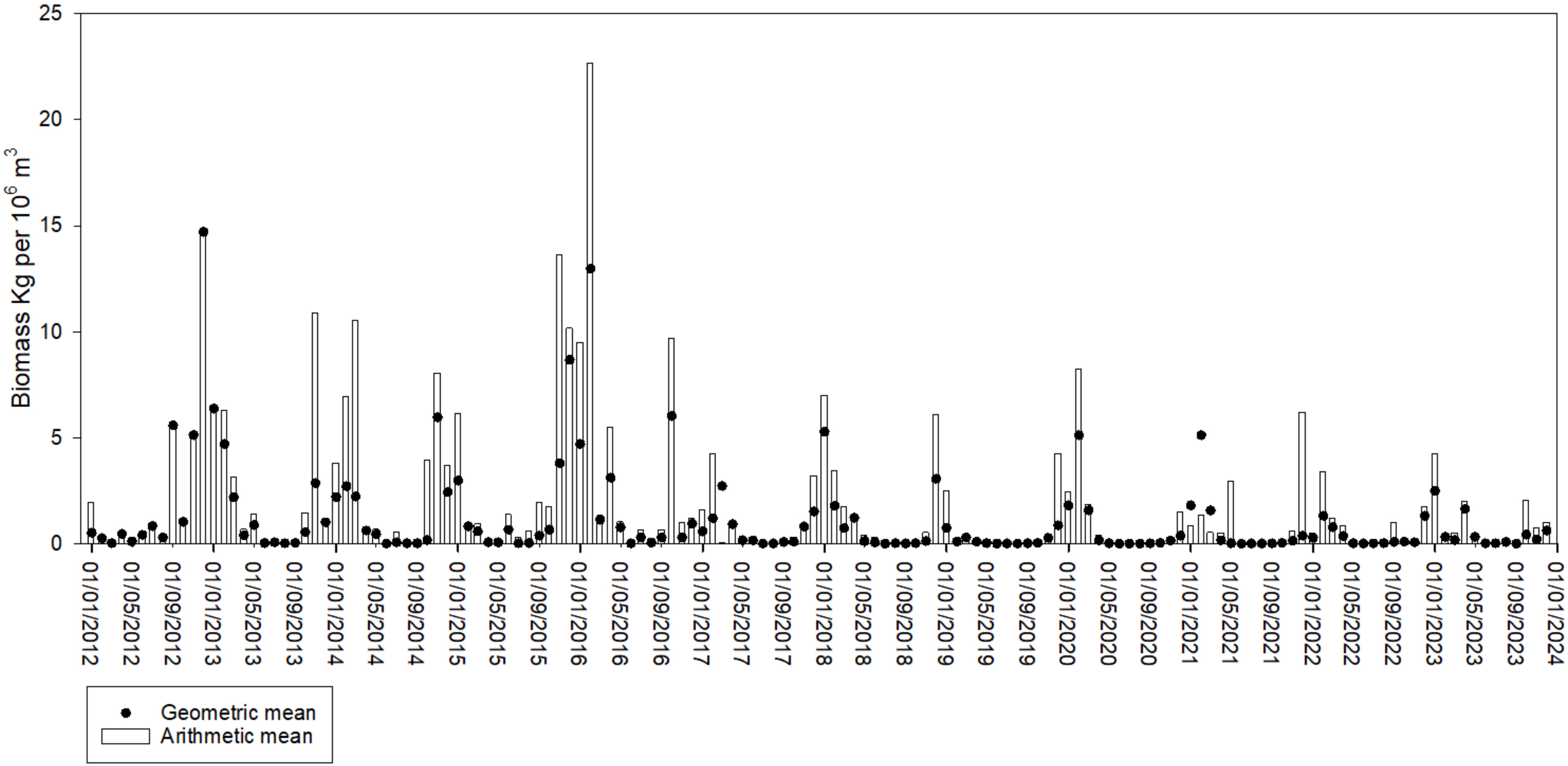

Figure 2. The geometric and arithmetic mean of total monthly biomass recorded per $1\times10^6 m^3$ of water abstracted during impingement surveys between 2012-2023.

5. Conclusions

This report has provided the response to NRW's request for use of arithmetic mean in the calculation of annual impingement. The impingement calculations using an arithmetic mean are presented in Section 2 and a justification for the use of arithmetic mean used for entrainment calculations provided in Section 33. Section 4 shows a graphical comparison of impingement numbers calculated using a geometric and arithmetic mean. Whilst entrapment estimates using the arithmetic mean are generally higher than those calculated using the geometric mean, the trends shown by the two methods are consistent (Section 3). Regardless of the method used, the wider Pembroke environmental monitoring programme has not shown an ecologically significant shift in communities outside of natural variation. Therefore, conclusions drawn from the use of an arithmetic mean would be consistent with those presented in the monitoring reports and the documents prepared to support the abstraction licence renewal. As such is it considered most appropriate to continue to analyse the data sets as has been done to date (geometric means in the analysis and scaling of impingement data, and arithmetic mean in the analysis and scaling of entrainment data).

6. References

Hutchings, J.A. (1996). Spatial and temporal variation in the density of northern cod and a review of hypotheses for the stock's collapse. *Canadian Journal of Fisheries and Aquatic Science*. 53, 943 – 962.

JASP Team. (2024). JASP (Version 0.19.0) [Computer software].

R Core Team. (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

Turnpenny, A.W.H., Utting, N.J., Milner, R.S. and Riley, J.D. (1983). The effect of fish impingement at Sizewell A Power Station, Suffolk, on North Sea fish stocks. Central Electricity Generating Board, Sizewell Public Inquiry Support Document No. TPRD/L/3270/R88, 28 pp.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686.