ASSESSMENT OF THE QUALITY OF MARINE BATHING WATERS USING DIFFERENT METHODS OF PERCENTILE CALCULATION

Diana Mance, Zrinka Vrček, Arijana Cenov, Marin Glad, Davor Mance, Darija Vukić Lušić

Abstract: *Escherichia coli* (*E. coli*) and intestinal enterococci (ENT) are the indicator bacteria used to test the microbiological quality of bathing water. In Croatia, the quality of bathing water is assessed by parametric method i.e. on the basis of the 90th and 95th percentile values which are calculated using the corresponding arithmetic mean and standard deviation of the logarithmic bacterial concentrations. There are also non-parametric methods for determining the value of a particular percentile, and one of them is the so-called Hazen method. In this work, we study how the different methods of percentile calculation affect the assessment of coastal bathing water quality. We have chosen a "problematic" location with frequently elevated bacterial concentrations as a test site.

The results show that in the case of *E. coli*, the 95th percentile values calculated using the Hazen method are significantly lower than those calculated using the parametric method (Wilcoxon rank test, $p < 0.05$), while for ENT there was no statistically significant difference. For case of annual assessment, the difference in coastal bathing water quality was statistically significant for both ENT and *E. coli* (McNemar test, $p=0.004$ and $p=0.02$ respectively). In the final classification, the difference in coastal bathing waters quality was statistically significant only in the case of *E. coli* (McNemar test, $p=0.025$).

Keywords: microbiology, bathing water quality, statistical methods, Hazen percentile method

Diana Mance, University of Rijeka, Croatia, diana.mance@uniri.hr, 0000-0001-6746-8693
Zrinka Vrček, University of Rijeka, Croatia, zrinka.vrcek@student.uniri.hr
Arijana Cenov, University of Rijeka, Croatia, arijana.cenov@zzjzpgz.hr, 0000-0002-6731-7868
Marin Glad, University of Rijeka, Croatia, marin.glad@zzjzpgz.hr, 0000-0001-9567-3897
Davor Mance, University of Rijeka, Croatia, davor.mance@uniri.hr, 0000-0002-6206-2464
Darija Vukic Lušic, University of Rijeka, Croatia, darija.vukic-lusic@zzjzpgz.hr, 0000-0003-0534-2547

Referee List (DOI 10.36253/fup_referee_list)
FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)

Diana Mance, Zrinka Vrček, Arijana Cenov, Marin Glad, Davor Mance, Darija Vukić Lušić, *Assessment of the quality of marine bathing waters using different methods of percentile calculation*, pp. 946-955, © 2024 Author(s), CC BY-NC-SA 4.0, DOI: 10.36253/979-12-215-0556-6.82

Introduction

The quality of bathing water is not only crucial for public health but also for the growth of sustainable tourism. It is influenced by both natural and anthropogenic elements [1-3]. The main objectives of the European Union's Bathing Water Directive 2006/7/EC, which regulates the management and quality of bathing water, are to preserve and protect the environment, as well as to improve quality, and protect human health [4]. *Escherichia coli* (*E. coli*) and intestinal enterococci (ENT) are the two indicators of faecal pollution used to routinely test the microbiological quality of bathing waters.

In Croatia, the quality of coastal bathing waters has been systematically monitored for more than 30 years. During this time, numerous parameters of the monitoring program have changed: regulations, microbiological criteria and test methods, limit values for individual categories of water quality, but also statistical procedures and methods of data processing. The quality of bathing waters is currently assessed on the basis of the values of the 90^{th} and 95^{th} percentiles[1] [5]. The procedure involves calculating logarithms of the bacteria concentrations, and the determination of the corresponding arithmetic mean and the standard deviation. Based on the determined values of the 95^{th} and 90^{th} percentiles and predefined intervals, the coastal bathing water quality is classified as excellent, good, sufficient or poor [3, 5].

The method described for determining the value of a particular percentile is also known as the parametric method, as the arithmetic mean and the standard deviation are used for the calculation. These are the measures of central tendency and variability for data sets that are distributed according to a normal distribution. Since bacterial data are almost never normally distributed, the question arose as to whether this method is suitable for analysing this type of data. There are several non-parametric statistical techniques (statistical methods that do not assume a normal distribution of the data) to determine the value of a particular percentile, and each has advantages and disadvantages of its own. The New Zealand Ministry of the Environment, for example, uses the Hazen method to determine the value of a particular percentile [6]. In this paper, we have compared the 90^{th} and 95^{th} percentile values, as well as the classification of coastal bathing water quality, based on calculations using the parametric method and the Hazen method. Since the Hazen method is ultimately intended to replace the parametric method, this type of comparison is necessary to determine whether possible changes in the water classification are due to actual changes in water quality or to a change in the method used to calculate the percentiles.

Materials and Methods

In Primorje-Gorski Kotar County (Croatia), the quality of the sea is regularly tested at approximately 300 locations during the bathing season. In most cases, the

[1] Percentiles are generally described as the proportion of values in a set of data that fall below a given number. The 90^{th} and 95^{th} percentile values therefore indicate the lower limits below which 90 % and 95 %, respectively, of the total data in a given set can be found.

sea proved to be of excellent quality [2]. Sites with the highest water quality are not interesting for our analysis, so we have chosen a "problematic" site where microbiological contamination is not uncommon. The data were obtained for samples collected during the bathing season from 2009 to 2022. The seawater for recreation is sampled on average every 14 days at a specific location during the bathing season. Therefore, in most cases (11 times, 78.6 %) there were 10 analysed samples per year at a given location. The maximal number of samples per year was 13. A total of 146 samples were analysed for each of the bacteria of interest. Both *E. coli* and ENT were cultured according to the standardized methods and analysed using the membrane filtration technique [7, 8]. The results (bacteria concentrations) are expressed in the number of colony-forming units in 100 mL of water (CFU/100 mL).

In the parametric method for determining the values of the percentiles, we used the procedure described in the regulation [5]. Calculation of the logarithm of the concentration value[2], calculation of the corresponding arithmetic mean (μ) and standard deviation (σ), and finally determination of the value of the 90th percentile (1), and the 95th percentile (2):

$$90^{th}\ \text{percentile}=\text{antilog}(\mu+1.282\cdot\sigma) \quad (1)$$

$$95^{th}\ \text{percentile}=\text{antilog}(\mu+1.65\cdot\sigma) \quad (2)$$

We used an interactive calculator to calculate percentile values based on the Hazen method [9]. The percentile calculations were carried out at the end of each bathing season (annual assessment) and on the basis of data from the current and the three previous bathing seasons (final assessment). The classification of the coastal bathing water is according to Table 1.

Table 1 – Standards for the assessment of coastal bathing water quality according to national regulation [5].

	Bathing water quality			
Indicator (CFU/100 mL)	Excellent	Good	Sufficient	Poor
intestinal enterococci	≤ 100*	≤ 200*	≤ 185**	> 185**
Escherichia coli	≤ 150*	≤ 300*	≤ 300**	> 300**

*based on the value of 95th percentile
**based on the value of 90th percentile

The statistical software jamovi [10] and RStudio [11] were used for the statistical analysis, and the results were interpreted at a significance level of 5%. The Shapiro-Wilk test was used to test the normality of the distribution of the bacterial

[2] If the bacterial concentration is 0 CFU/100 mL, the concentration is regarded as the detection limit, in our case, 3 CFU/100 mL.

concentration data. As it was found that the data were not normally distributed, the median was used as a measure of central tendency and the interquartile range (IQR) as a measure of variability. The Mann-Whitney U test was used to compare the mean values of ENT and *E. coli* concentrations. The Wilcoxon rank sum test for paired samples was used to compare the mean percentile values calculated by the parametric and Hazen methods. Coastal bathing water classifications were compared with the McNemar test.

Results

The median value of the ENT concentration is 18 CFU/100 mL, with a minimum value of 3 CFU/100 mL (lower detection limit) and a maximum value of 290 CFU/100 mL. For *E. coli*, the median is 22 CFU/100 mL, and the corresponding minimum and maximum values are 3 CFU/100 mL and 520 CFU/100 mL. The results of the Mann-Whitney U-test (p=0.02; Table 2) show that the values for *E. coli* are significantly higher than the concentration values for ENT.

Table 2 – Descriptive statistics for the concentrations (in CFU/100 mL) of intestinal enterococci and *Escherichia coli*. The significance of the difference between the mean bacterial concentrations was tested with the Mann-Whitney U test. IQR – interquartile range.

Microbiological parameter	median	IQR	Minimum	Maximum	p-value (Mann-Whitney U test)
enterococci	18	39.8	3	290	0.02*
E. coli	22	93.5	3	520	

*statistically significant

Table 3 – Intestinal enterococci: Median values of the 90th and 95th percentiles for annual and final (current season and three previous seasons) assessment, calculated using parametric and Hazen methods. The significance of the difference between the medians was tested using the Wilcoxon rank test for paired samples. IQR – interquartile range.

Percentile	Assessment	Method	Median (CFU/ 100 mL)	IQR (CFU/ 100 mL)	p-value (Wilcoxon rank test)
90th	annual	parametric	85	47.75	0.263
		Hazen	82.5	41.25	
	final	parametric	78	24	0.197
		Hazen	80	4	
95th	annual	parametric	125.5	84.5	0.055
		Hazen	86	65	
	final	parametric	126	33.5	0.197
		Hazen	122	44	

The comparison of the values of the 90th and 95th percentiles for the annual and final assessment shows that in the case of ENT there is no statistically significant difference between the values determined using the parametric and Hazen methods (Table 3). In the case of *E. coli*, the difference proved to be significant, but only in the case of the 95th percentile value (Table 4).

Table 4 – *Escherichia coli*: Median values of the 90th and 95th percentiles for annual and final (current season and three previous seasons) assessment, calculated using parametric and Hazen methods. The significance of the difference between the medians was tested using the Wilcoxon rank test for paired samples. IQR – interquartile range.

Percentile	Assessment	Method	Median (CFU/ 100 mL)	IQR (CFU/ 100 mL)	p-value (Wilcoxon rank test)
90th	annual	parametric	195	71	0.572
		Hazen	156	118.3	
	final	parametric	198	44	0.365
		Hazen	185	53.5	
95th	annual	parametric	345	149.3	0.03*
		Hazen	186	132	
	final	parametric	345	63	0.005*
		Hazen	260	83	

*statistically significant

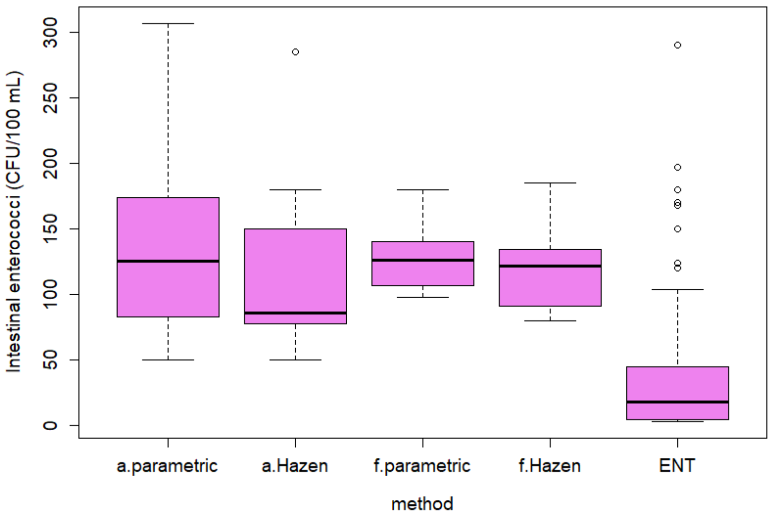

Figure 1 – Intestinal enterococci 95th percentile for: a.parametric – annual assessment calculated using the parametric method, a.Hazen – annual assessment calculated using the Hazen method, f.parametic – final assessment calculated using the parametric method and s.Hazen – final assessment calculated using the Hazen method. ENT – the intestinal enterococci concentration is given for comparison.

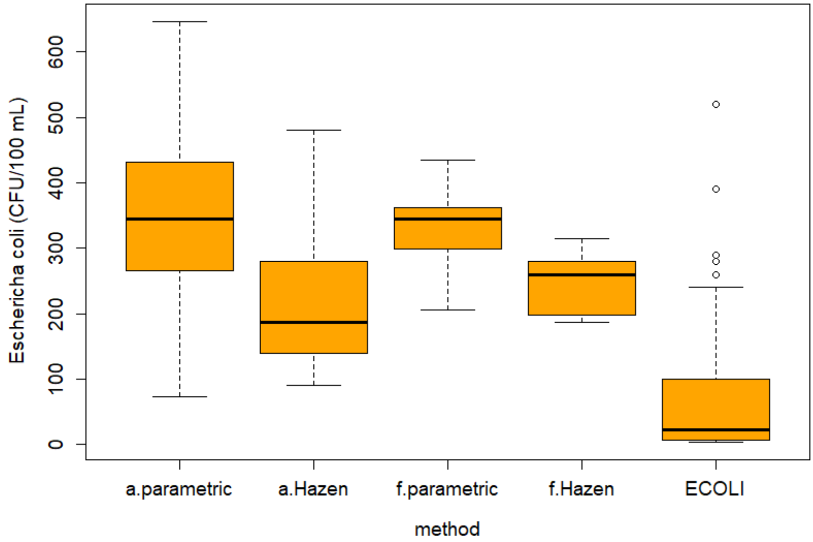

Figure 2 – *Escherichia coli* 95th percentile for: a.parametric – annual assessment calculated using the parametric method, a.Hazen – annual assessment calculated using the Hazen method, f.parametic – final assessment calculated using the parametric method and s.Hazen – final assessment calculated using the Hazen method. ECOLI – the *Escherichia coli* concentration is given for comparison.

Figures 1 and 2 show comparisons of the 95th percentile values calculated using the parametric and Hazen methods for ENT and E. coli, respectively

According to Table 1, we determined the coastal bathing water quality for a annual and final assessment based on the concentration of ENT (Figure 3) and *E. coli* (Figure 4). For further statistical analysis and due to the small amount of data, we have combined the categories "excellent" and "good" in the "higher" category and the categories "sufficient" and "poor" in the "lower" category. In the annual classification, the difference in classification proved to be statistically significant both in the case of ENT (McNemar test, p=0.004, Figure 5) and in the case of *E. coli* (McNemar test, p=0.02, Figure 6). In the final classification, the difference was statistically significant only for *E. coli* (McNemar test, p=0.025), where the quality of the seawater was rated better than the quality determined by the parametric method in 45.5 % of cases.

Intestinal enterococci / year	annual assessment		final assessment	
	parametric	Hazen	parametric	Hazen
2009	S	G		
2010	E	E		
2011	E	G		
2012	G	G	G	G
2013	G	E	G	G
2014	E	E	E	G
2015	G	E	G	G
2016	E	E	G	E
2017	G	G	G	E
2018	G	E	G	E
2019	S	G	G	G
2020	E	E	G	G
2021	G	E	G	G
2022	S	P	G	G

Figure 3 – Coastal bathing water quality (annual and final assessment) based on the values of the 90th and 95th percentiles of intestinal enterococci concentrations calculated using the parametric and Hazen methods. E- Excellent; G – Good; S – Sufficient; P – Poor according to the Table 1.

Escherichia coli /year	annual assessment		final assessment	
	parametric	Hazen	parametric	Hazen
2009	E	E		
2010	S	G		
2011	E	E		
2012	S	G	G	G
2013	S	G	G	G
2014	S	E	G	G
2015	S	G	S	G
2016	S	G	S	G
2017	G	S	S	G
2018	G	G	S	G
2019	P	P	S	S
2020	E	G	S	S
2021	S	G	S	S
2022	S	G	S	G

Figure 4 – Coastal bathing water quality (annual and final assessment) based on the values of the 90th and 95th percentiles of *Escherichia coli* concentrations calculated using the parametric and Hazen methods. E- Excellent; G – Good; S – Sufficient; P – Poor according to the Table 1.

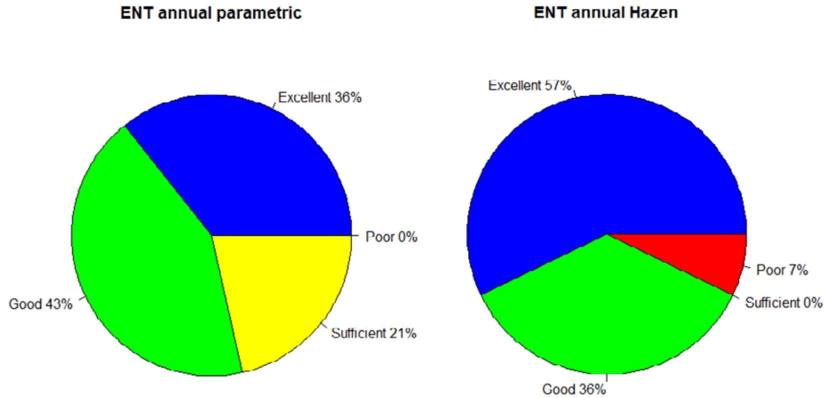

Figure 5 – Changes in coastal bathing water quality (annual assessment based on the concentration of intestinal enterococci - ENT) as a result of different methods for calculating the 90th and 95th percentiles: parametric method (left) and Hazen method (right).

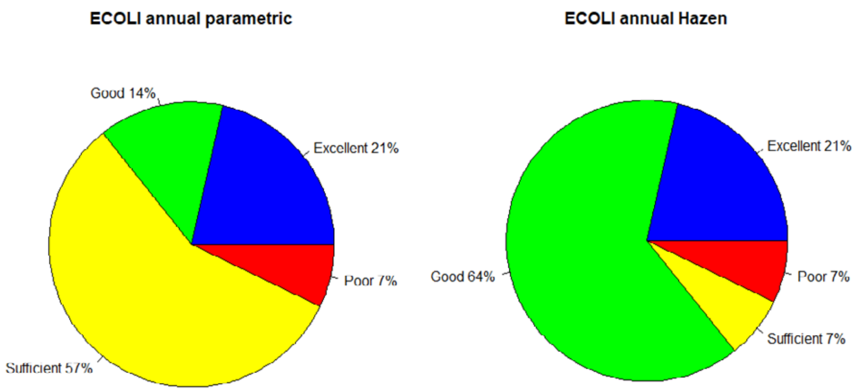

Figure 6 – Changes in coastal bathing water quality (annual assessment based on the concentration of *Escherichia coli* - ECOLI) as a result of different methods for calculating the 90th and 95th percentiles: parametric method (left) and Hazen method (right).

Discussion

In Croatia a programme for the coastal bathing water quality monitoring has been in place for more than three decades, providing us with a long continuity of data. In addition to real changes in the environment, changes in the methodology for analysing microbiological indicators, changes in assessment criteria and statistical data processing also have an impact on the classification of beaches. These factors must be taken into account when assessing the state of the environment and monitoring long-term trends.

The data on ENT and *E. coli* concentrations were analysed over 14 consecutive bathing seasons at a site with frequently elevated levels of these indicator bacteria. It was found that the *E. coli* concentrations were significantly higher than the ENT concentrations.

As the 90th and 95th percentile values are used to determine the quality of coastal bathing waters, the values calculated using the parametric (standard) method and the Hazen method were compared. It turned out that these values are significantly lower in the case of the 95th percentile values of *E. coli* calculated with the Hazen method.

In the case of coastal bathing water quality, we see that the Hazen method has a significantly higher number of results with better water quality compared to the parametric method.

From a public health perspective, the underestimation of bathing water quality is a minor problem, considering that the main objective of water quality monitoring is to protect public health. On the other hand, it is important not to overestimate the risk of transmission of infectious diseases when using recreational waters, as the implementation of remedial measures can lead to significant economic damage.

Conclusion

Parametric statistical methods are considered to be more rigorous and intuitive statistical methods compared to non-parametric methods. A prerequisite for the application of parametric methods is the normal distribution of the analysed data. In the case of bacteriological water pollution, this requirement is usually not met. Therefore, it is expected that the non-parametric (Hazen) method may become the new standard for the assessment of coastal bathing waters quality.

By comparing the results obtained with the parametric and Hazen methods, we aim to determine whether a change of method affects the final results of the coastal bathing water quality classification. Our results show that the change of method significantly affects the results of the quality assessment. In fact, we can expect the Hazen method to provide data indicating higher water quality.

The results reported in this article should be interpreted with caution as only one site was studied. Including a larger number of sites in this type of study would lead to more reliable results.

Acknowledgements

This work was funded by the UNIRI projects uniri-iskusni-prirod-23-177; uniri-iskusni-drustv-23-163 and uniri-iskusni-biomed-23-245 (https://www.croris.hr/projekti/projekt/10322)

References

[1] Mance D., Mance D., Vukić Lušić D. (2018) - *Marine pollution differentiation with stable isotopes of groundwater*, Sci. J. Mar. Res. 32 (1), 80 – 87.

[2] Mance D., Mance D., Vukić Lušić D. (2021) - *Environmental and anthropogenic factors affecting coastal bathing water quality: preliminary study for Primorje-Gorski Kotar County (Croatia)*, Sci. J. Mar. Res. 35 (1), 57 – 68.

[3] Mance D., Mance D., Vukić Lušić D. (2022) - *Managing water commons using mediator variables to bridge the gap between environmental factors and anthropogenic pollution indicators*, Ninth International Symposium "Monitoring of Mediterranean Coastal Areas: Problems and Measurement Techniques", Livorno, Italy, 14–16 June 2022, pp. 515–524.

[4] *Directive 2006/7/EC of the European Parliament and of the Council of 15 February 2006 concerning the management of bathing water quality and repealing Directive 76/160/EEC*. (2006) Off. J. Eur. Union 2006, L 64, 37–51.

[5] Official Gazette 73/08 (2008) - *Regulation on sea bathing water quality.* Croatian Ministry of Environment and Energy. https://narodne-novine.nn.hr/clanci/sluzbeni/2008_06_73_2426.html (in Croatian)

[6] Ministry of the Environment, New Zealand (2003) - *Microbiological water quality guidelines for marine and freshwater recreational areas*, https://environment.govt.nz/publications/microbiological-water-quality-guidelines-for-marine-and-freshwater-recreational-areas/

[7] Jozić S., Vukić Lušić D., Ordulj M., Frlan E., Cenov A., Diković S., Kauzlarić V., Fiorido Đurković L., Stilinović Totić J., Ivšinović D., Eleršek N., Vucić A., Peroš-Pucar D., Unić Klarin B., Bujas L., Puljak T., Mamić M., Grilec D., Jadrušić M., Šolić M. (2018) - *Performance characteristics of the temperature-modified ISO 9308-1 method for the enumeration of Escherichia coli in marine and inland bathing waters*, Marine Pollution Bulletin, 135, 150 – 158.

[8] ISO 7899-2 (2000) - *Water quality – Detection and enumeration of intestinal enterococci – Part 2: Membrane filtration method (ISO 7899-2:2000; EN ISO 7899-2:2000)*. International Organization for Standardization, Geneva.

[9] Ministry of the Environment, New Zealand (2013) - *Hazen percentile calculator*, https://environment.govt.nz/publications/bathewatch-user-guide/hazen-percentile-calculator/

[10] The jamovi project (2024). jamovi (Version 2.5) [Computer Software]. Retrieved from https://www.jamovi.org

[11] RStudio Team (2020). RStudio: Integrated Development Environment for R (Version 2023.12.1.402) [Computer Software]. RStudio, PBC, Boston, MA. Retrieved from http://www.rstudio.com/.