Canadian Journal of Earth Sciences

A revised Coastal Sensitivity Index for Canada's marine coasts calculated using nonparametric statistics

Journal:	<i>Canadian Journal of Earth Sciences</i>
Manuscript ID	cjes-2021-0010.R1
Manuscript Type:	Article
Date Submitted by the Author:	14-May-2021
Complete List of Authors:	Hatcher, Scott; Geological Survey of Canada Atlantic Manson, Gavin; Geological Survey of Canada Atlantic
Keyword:	coastal sensitivity index, coastal management, nonparametric statistics, coastal sensitivity, coastal vulnerability
Is the invited manuscript for consideration in a Special Issue? :	Landscape & Seascape Responses to Canada's Changing Climate

SCHOLARONE™
Manuscripts

1 **A revised Coastal Sensitivity Index for Canada’s marine coasts**
2 **calculated using nonparametric statistics**

3 **Hatcher, S.V.*^{1,2} and Manson, G.K.¹**

4 ¹Geological Survey of Canada-Atlantic

5 ²B.W. Geospatial

6
7 * Corresponding author: bw.geospatial@gmail.com

8 5 Glenmore Ave., Halifax, NS, Canada, B3N1W3

9
10 **Abstract**

11 A Coastal Sensitivity Index (CSI) is a measure of the sensitivity of a coastline to its
12 physical environment which provides information useful for coastal management.
13 Traditionally a CSI is calculated as the mathematical aggregation of coastal sensitivity
14 indicators, which may include factors such as coastal material, relief, and wave energy.
15 The indicators differ depending on study area, but generally are assigned a score ranging
16 from one to five in order of increasing sensitivity. These scores are then aggregated
17 using either the square root of the product mean (the “classic” method), or the geometric
18 mean. Both of these methods are limited by mathematical assumptions, lack of
19 comparability, and the need for empirical validation. In this study, we applied an
20 alternative nonparametric method of calculation, known as μ -statistics, to Canada's
21 marine coasts to provide an improved measure of coastal sensitivity. μ -statistics, which
22 offer a mathematically sound method of aggregating ordinal indicators, have a number of
23 theoretical advantages over the classic and geometric mean methods. In practice, when
24 applied to Canada’s marine coast, we find that the μ -statistics method (1) compresses the
25 mid-range variability in the resulting sensitivity index, (2) accentuates positive and
26 negative distribution tails, and (3) minimizes propagated errors by 190% and 50%
27 respectively, compared to the classic and geometric mean methods. Additionally, the μ -
28 statistics method has a theoretical foundation that relieves the necessity to empirically
29 validate the aggregating assumptions and relies only on the assumptions inherent in the
30 scoring method. μ -statistics thus provide a new, rigorous method for the calculation of
31 coastal sensitivity indices when the underlying variables have ordinal scores.

32 *Keywords: coastal sensitivity index, coastal management, nonparametric statistics,*
33 *coastal sensitivity, coastal vulnerability*

1.0 Introduction

Coastal sensitivity is the degree to which the coast is physically affected, either adversely or beneficially, by factors such as climate change (Lemmen et al., 2016). The sensitivity of a coast depends on forcing effects such as sea-level change, sea ice cover change, currents and wave heights. It also depends on physical characteristics such as coastal materials, backshore slope, tidal range, and ground ice. Flooding during storm surges may increase due to sea-level rise, whereas erosion may increase as wave heights increase due to greater storminess or a reduction in sea ice (Shaw et al., 1998). Coastal planning necessitates the delivery of information products that communicate this complexity succinctly.

A Coastal Sensitivity Index (CSI) condenses complex information into a format that is interpretable by non-specialists (Cooper and McLaughlin, 1998). In the absence of precise modelling of coastal sensitivity, research must rely on a qualitative understanding of the multiple contributors to sensitivity. CSIs are the most common amongst a variety of tools that exist to aid coastal management decisions through quantifying complex data into an easily interpretable form (e.g. McLaughlin 2001; Abuodha and Woodroffe, 2006; Abuodha, 2009).

In the literature, CSIs have been calculated in a three step process by (1) assembling relevant physiographic datasets, (2) scoring them on a scale from least sensitive to most sensitive, (3) aggregating the scores into a single value for each segment of coast (Gornitz, 1991; Shaw et al., 1998; Thieler and Hammer-Close, 1999). The result is a number that represents the aggregation of all scores for that section of coast. Numerous CSIs have been developed at various study areas all over the world, and many variations in variables and scores for calculating a sensitivity index have been proposed (Abuodha and Woodroffe, 2006). By far the most commonly used method is that proposed by Gornitz (1991), which we here define as the “classic” method. It is calculated as the square root of the product of the indicators, divided by the number of indicators (See section 3.1.1).

65

66 With the exception of Cogswell et al. (2018) and McLaughlin (2001), little work has
67 focused on the validity of methods for aggregating underlying indicators of coastal
68 sensitivity (step 3 above). McLaughlin (2001) noted that the adoption of the classic
69 method is a pragmatic approach, therefore accepting its mathematical assumptions in
70 order to retain simplicity. Specifically, in adopting the classic method there is an
71 assumption of equal intervals for each ordinal score, and in aggregating them there are
72 assumptions regarding their relative importance, cross-correlation, and functional
73 relationship to the independent variable (Wittkowski et al., 2004). We explore the likely
74 risks inherent in adopting methods that sum ordinal factors in calculating CSIs. Assuming
75 equal magnitude intervals between ordinal ranks can introduce error where that
76 assumption is unable to be validated. In other fields where ordinal ranks are aggregated
77 using a mean, the assumption of equal intervals can be validated using empirical data.
78 However, due to the subjective nature of coastal sensitivity studies, equivalent validation
79 is absent from the literature (McLaughlin 2001).

80

81 Subtle changes in the aggregation method can have dramatic effects on the resulting
82 index (Gibb, 1992; Cogswell et al., 2018). Nevertheless, the originally published
83 aggregation method proposed by Gornitz (1991) continues to be the dominant method in
84 use (Abuodha and Woodroffe, 2006; Cogswell et al., 2018). This is likely for two
85 reasons: (1) the 1991 paper was the first method published and was subsequently cited as
86 the source of the method, and (2) it concentrates on higher sensitivity coastlines, a goal
87 for coastal managers. Cogswell et al. (2018) use the true geometric mean, which is an
88 alternative to the Gornitz (1991) method, but validation of the mathematical assumptions
89 introduced by aggregating ordinal scores that are assumed to be linear remains relatively
90 unexplored. Furthermore, the lack of empirical data with which to validate CSIs means
91 that replication and standardization across geographies remains out of reach, which
92 compounds the original problem. A method that is conservative in its introduction of
93 mathematical assumptions, that retains the explanatory power of the traditional CSI
94 method, and that provides opportunities to enhance and globalize the analysis, would
95 advance our understanding of coastal sensitivity.

In this paper we borrow from a branch of unbiased estimation statistics that addresses the problem of aggregating ordinal rank indicators. The method, named μ -statistics, is used in medical science to look at exposure risk to certain diseases using clinical indicators (Wittkowski, 1992; Wittkowski et al., 2004). Their main advantage is that aggregation of ordinal indicators can be done with only one assumption: sensitivity increases with rank (Wittkowski et al., 2004). This single assumption is in contrast to the assumptions inherent in the classic method of calculating a CSI: that score magnitude is linear, and that all indicators are equally weighted. Our approach is to apply the μ -statistics method to the entire Canadian marine coastline (Fig. 1) using CanCoast datasets (Manson et al., 2019). We will then compare the results to two previously published methods and discuss the implications of adopting the μ -statistics method.

2.0 Methods and data

We employ a two-step methodology to compare results of CSIs calculated by three methods described below. First we use simulated data, following the methodology of Cogswell et al. (2018), to test the sensitivity to assumed uncertainty for each method. Second, we apply all three methods to the CanCoast national database to test the interpretability of each method using “real world” examples.

2.1 Methods of aggregating indicators

We investigated three methods to aggregate CSI indicators which we refer to as: 1) the classic method (CSI_c), 2) the geometric mean method (CSI_g), and 3) the μ -statistics method (CSI_μ).

2.1.1 The classic method (CSI_c)

Gornitz and Kanciruk (1989) and Gornitz (1991) originally proposed the classic method (CSI_c , Eq. 1), which was later implemented by Shaw et al. (1998) in the Canadian

context. As stated by Cogswell et al. (2018), it is likely that Gornitz (1991) mistakenly called the CSI_c a geometric mean. If so, the CSI_c formula is best described as the square root of the product of indicators (x) divided by the number of indicators (n). Despite this confusion surrounding the calculation, the CSI_c method was chosen for comparison in this study because of its widespread acceptance in the literature (see McLaughlin, 2001, and Abuodha, 2009, for thorough discussions of the ubiquity of this method in the literature).

$$CSI_c = \sqrt{x_1 x_2 \dots x_n / n}$$

Equation 1

2.1.2 The geometric mean method (CSI_g)

Cogswell et al. (2018) proposed the geometric mean (CSI_g), defined as the n^{th} root of the product of variable scores (Eq. 2). This method was chosen for comparison because of its simplicity, and because it has previously been favourably compared to the CSI_c (Cogswell et al., 2018).

$$CSI_g = \sqrt[n]{x_1 x_2 \dots x_n}$$

Equation 2

2.1.3 The μ -statistics method (CSI_μ)

The μ -statistics method (CSI_μ) is a nonparametric method derived from a branch of unbiased estimation statistics (Hoeffding, 1948; Wittkowski et al., 2004; Morales et al., 2008). The method was designed to provide a way to aggregate ordinal rank variables in a statistically robust way that reduces the risks inherent in using the arithmetic or geometric mean (Wittkowski et al., 2004).

An example best explains the μ -statistics method as applied in the context of CSI calculation. The sample dataset and index calculations are shown in Figure 1. The initial dataset has three indicators (I_1 , I_2 , and I_3 in Fig. 2a), each of which has an ordinal rank score from 1 to 5. The method is a simple sorting algorithm. A superiority function (Eq. 3) determines whether one row is “superior” (or “more sensitive”) than another. For row

$X_{j'}$ to be superior to another row X_j , all of its indicators must be greater than or equal to the corresponding indicator in the other row, and at least one indicator must be greater (Wittkowski et al., 2004).

$$X_j < X_{j'} \Leftrightarrow \{\forall i = 1, \dots, I X_{ji} \leq X_{j'i} \wedge \exists i = 1, \dots, I X_{ji} < X_{j'i}\} \quad \text{Equation 3}$$

This superiority function is used to sort the dataset into a partial ordering. It is a partial ordering because there are large groups of rows that cannot be sorted because superiority between them cannot be determined (see group 2 in Figure 2b). Following sorting, a scoring function ($I(X_j < X_{j'})$ in Eq. 4) is used to calculate the final index. In this function, a value of 1 is given for every comparison where $X_{j'}$ is superior, a value 0 for every comparison where superiority is ambiguous, and -1 where X_j is inferior. The final index is simply the sum of the results of this scoring function (Eq. 5, see the μ Score column in Figure 2c, as well as an implementation of the algorithm in the R computer language in the supplementary data file). The final index value represents the number of inferior rows subtracted from the number of superior rows (Eq. 5). An index value of 0 means that there is an equal number of inferior rows and superior rows for that row. Likewise, a positive index value indicates that that row is superior to more rows than it is inferior.

$$I(X_j < X_{j'}) = \begin{cases} 1 & \text{if } X_j < X_{j'} \\ 0 & \text{if } X_j \text{ and } X_{j'} \text{ cannot be ordered} \\ -1 & \text{if } X_{j'} < X_j \end{cases} \quad \text{Equation 4}$$

$$CSI_\mu = \sum_{j'} I(X_j < X_{j'}) \quad \text{Equation 5}$$

2.2 Simulated coastal segments

Following Cogswell et al. (2018), we created a simulated database to compare results derived using the three methods. Six variables were randomly sampled from a normal distribution with a mean randomly selected between 1 and 100 and a standard deviation of 1. Each variable contained 1000 samples ($n=1000$). These variables were then binned

187 using equal intervals and assigned sequential scores from 1 to 5. We used the three
188 methods of aggregation to create simulated CSI scores. We ran this generation process 30
189 times, and we aggregated summary statistics (mean, median, min, max, skewness, and
190 kurtosis) for each run using the arithmetic mean.

191
192 We tested the propagation of error through each of the three methods using the simulated
193 dataset. We assume a reasonable error of ± 1 score for each indicator. This error estimate
194 is justified given the arbitrary precision inherent in binning the continuous variables. For
195 example, a tidal range value of 4.0 m may in reality be 4.1 m, leading to an erroneous
196 rating of 1 instead of 2 for that section of coast. A difference of 0.1 m is well within
197 potential error of estimates of national-scale variables such as tidal range. In the absence
198 of validation data, we feel it reasonable to test the sensitivity of each method to a shift of
199 all ratings by ± 1 , thereby calculating the upper and lower bounds of the index due to
200 expected error limits. Upper and lower limits of the data were calculated by either adding
201 or subtracting 1 for each indicator. In cases where the addition or subtraction took the
202 score below 1 and above 5, we used the limit. We then calculated the CSI using the three
203 methods for (1) the lower limit of error, (2) the original data, and (3) the upper limit of
204 error. We then calculated the mean percent error over the 30 runs for each of the three
205 methods. Next we calculated the percent error that propagated through each method as an
206 average of the upper and lower bounds.

207
208 To further illustrate the method: (1) we created a simulated dataset and calculated a CSI
209 for each segment of coast. A single segment of coast may have a CSI value of 5.0 using
210 method 1. (2) we then modified that simulated dataset by either adding 1 (for upper
211 bounds) or subtracting 1 (for lower bounds), and CSI scores are recalculated. Looking at
212 the same segment as above, the lower bounded CSI value might be 4.5 and the upper
213 bounded CSI value might be 5.5. (3) We then calculated the percent error for upper and
214 lower bounded CSI values (10% upper bounded, 10% lower bounded), and the average of
215 these two values yields the percent error for that segment (10%). (4) We then averaged
216 these percent errors for the entire simulated dataset to yield an average percent error for
217 that dataset and method. (5) We repeated steps 1 through 4 for each of the three methods.

(6) We repeated the process again 30 times to yield an average percent error for each method over 30 simulations. This measurement is meant to communicate the resulting uncertainty introduced by the method alone given identical initial conditions and a reasonable expected amount of error. The method with the lowest percent error is considered the most conservative, and therefore the most appropriate for use in situations without validation data.

2.3 Application to a national coastal database

In order to compare each method in a real-world context, we applied the three methods to a national database (CanCoast) of coastal physiographic variables for Canada (Manson, et al., 2019). This database includes six variables that are used to calculate coastal sensitivity on a national scale: sea-level change, decadal mean wave heights, ground ice, coastal material, backshore slope, and tidal range (Table 1, with details of scoring in Table 2). Further information on data sources and formats can be found in Manson, et al. (2019).

We classified results into five categories: "Very Low Sensitivity", "Low Sensitivity", "Medium Sensitivity", "High Sensitivity", and "Very High Sensitivity" using the Natural Jenks method (Jenks, 1967). This classification method relies on inter-class variance and means, which might be seen as introducing the assumptions that were avoided by using the μ -statistics method. The Natural Jenks method, however, is appropriate for use in this context because although the individual indicators are ordinal, the μ -score that is the result of the method represents the subtraction of two count variables (count of superior rows and count of inferior rows). Therefore, means and variances calculated on count data are valid and do not introduce invalid assumptions.

4.0 Results

4.1 Uncertainty

Using the synthetic dataset, we applied the three CSI calculation methods to each of the 30 runs. The mean percent errors for each method are 270% for CSI_c , 83% for CSI_g , and 27% for CSI_μ . The relative error in the CSI_μ is 190% less than that in the CSI_c and 50% less than the CSI_g . The μ -statistics method is less sensitive to uncertainty than either of the methods to which it is compared.

4.2 Application to CanCoast database

Three distinct statistical distributions of CSI values result from the application of the three aggregation methods to the national CanCoast data. CSI_c increases the range of values, skews towards lower sensitivity, and accentuates positive outliers (Fig. 3a). This last feature may be desirable for managing higher sensitivity coastlines, as they are easily captured by the method. The CSI_g distribution has a narrower range and is the most normal in shape (Fig. 3b), which shows the homogenizing influence of the geometric mean, as reported in Cogswell (2018). The CSI_μ distribution is tightly clustered around the mean and has a larger number of values on the negative tail of the distribution (Fig. 3c).

CSI_c identifies the eastern side of the Arctic Archipelago, the Beaufort Sea coast, and the east and west coasts of Hudson Bay as areas of high sensitivity. Much of the Canadian coastline is classified as having very low or low sensitivity. In some areas, such as Nova Scotia, small reaches of coastline of higher sensitivity are adjacent to reaches with highly variable sensitivity (Fig. 4a).

CSI_g identifies the same sensitivity hotspots as CSI_c , but includes areas of high sensitivity that are not highlighted by CSI_c . These areas include the mouth of Frobisher Bay, the coasts of the Bay of Fundy, New Brunswick, and northern British Columbia (Fig. 4b). Overall, CSI_g shows more areas classified as “high sensitivity” than CSI_c .

In general, CSI_μ shows higher sensitivity than the other two methods, particularly on the coasts of Hudson Bay, southern Coronation Gulf, Queen Maud Gulf, and much of the

western Arctic Archipelago (Fig. 4c). There are notable areas of very low values, such as around the northern end of Baffin Island, western Ellesmere, southern British Columbia, and the St. Lawrence Estuary.

Comparisons of the tested methods show that they are highly correlated at the extremes, but far less correlated near the centre (Fig. 5). Minimal variation in the lower extreme values shows that the μ -statistics method retains much of the variation between results that is minimized in the classic method. The same is true, but to a lesser extent, for the geometric mean (Fig. 5). The correlation diagrams show that mid-range variations of the classic or geometric mean are discounted by the μ -statistics method in place of increased variation at the extremes.

4.3 Extreme Value Detection

To investigate extreme values, we identified the coastal types above the 90th percentile (positive extremes) and below the 10th percentile (negative extremes) for each method. CSI_{μ} identified 31 positive extremes and 51 negative extremes. These coastal types are not extremes in the CSI_g or CSI_c distributions. These types are identical for both the classic method and the geometric mean. The μ -statistics method identifies two coastal types above the 90th percentile as a lower sensitivity class, and 20 coastal types below the 10th percentile identified by the classic and geometric mean method as a lower sensitivity class. If we subtract those missed from the added types, we see that the μ -statistics method includes 2.7% more negative extreme values and 2.5% more positive extreme values than either the classic or geometric mean method. Despite differences in the overall shapes of the distributions, the coastal types identified using the 90th percentile and 10th percentile overlap among the three methods, and there are 39 coastal types identified as either positive or negative extremes exclusively by the μ -statistics method.

5.0 Discussion

The shapes of the CSI_c , CSI_g , and CSI_μ differ (Fig. 3). The CSI_c distribution exhibits a positive tail that suggests a few coastal types are anomalously sensitive; CSI_g method predicts both negative and positive tails that are equally extreme. These two methods have fairly homogeneous sensitivities with minor variation above and below a dominant central tendency. The zero mean of the CSI_μ distribution indicates that there are an equal number of coastal types that are either more or less sensitive than the average. A positive median, however, suggests a bias in coastal types towards less sensitivity. This distribution suggests that many coastal types are in the mid-range of sensitivity, although there are a small group of diverse (wide- ranging) types that have either extremely low or extremely high sensitivity.

5.1 The μ -statistics method

Because the CSI_μ is rank-based, the addition or subtraction of a specific amount of uncertainty has little effect on the results. The CSI_μ method has no inherent assumptions about the magnitude difference between variables, so the effect of changing the score of a single indicator by 1 makes less difference than it does with the CSI_c and CSI_g methods.

The CSI_μ method shows that Canada's coastline has a significant cluster of similar mid-range sensitivity types, but there are a number of types that are far removed from this mid-range cluster. The CSI_μ method favours ambiguous types with outliers. We can't currently quantify their relative importance for sensitivity. This ambiguity is dealt with by assumption when using when using the classic method (e.g. Gornitz, 1991; Shaw et al., 1998) and the geometric mean method (Cogswell et al., 2018). Despite identical underlying data, the three aggregation methods produce three different interpretations. The differences in the conclusions from the three methods indicates the importance of the aggregation method in the assessment of coastal sensitivity based on these three methods. However, without empirical validation we do not know whether the increased precision offered by the CSI_g and CSI_c methods is warranted.

The compression of variation around the mean of the CSI_{μ} method is due to the high ambiguity of pairwise comparisons that are close, with a corresponding expansion of variation in CSI_{μ} at the positive and negative extremes of the distribution. Because no assumptions are made about the relative weights of each indicator, or their linear nature, the CSI_{μ} method accentuates coastal types at either end of the distribution, but diminishes the variation around the mean. This is not the case for the CSI_c and CSI_g distributions as these methods incorporate assumptions about the magnitude of the scores and uncertainty propagates through them with increasing magnitude.

5.2 The classic and μ -statistics methods

An argument against adopting the CSI_{μ} method in place of the CSI_c method is that it would eliminate cross-study comparisons that are possible using the latter method, which has been applied around the world (Abuodha and Woodroffe, 2006). The potential to unify cross-national assessments to broaden our understanding of global coastal sensitivity in a changing climate would justify the continued use of the CSI_c method. Unfortunately, despite a consistent method of aggregation, there is little consensus among researchers on how to select, bin, and rank underlying variables. For example, there is no consensus on how to score tidal range. Investigators typically decide on scores based on whether they are more concerned with inundation or erosion (see, for example, the difference between Shaw et al. (1998) and Abuodha & Woodroffe, (2006)). Variables are binned according to local conditions, so, for example, a study in Greece might place the limits of tidal range sensitivity between 0.2 m and 0.8 m (Karymbalis et al., 2012), whereas a Canadian assessment will bin values between 0.2 m and 16 m (Shaw et al., 1998). This precludes meaningful comparison of two countries even if the CSI_c method was used in both cases.

5.3 The geometric mean and μ -statistics methods

As Cogswell et al. (2018) have shown, the geometric mean maintains a consistent range in resulting distributions for small numbers of variables, while retaining explanatory power. This justifies its adoption in studies where all underlying variables are on an interval scale. However, when aggregating ordinal variables, the μ -statistics method is the

most mathematically justified method. For regional or local assessments of coastal sensitivity, in which all input variables are on ratio or interval scales that are accurately known, the geometric mean is mathematically justified and the most appropriate choice. In this sense, the two methods are complementary and facilitate assessments of coastal sensitivity at any scale and level of precision.

5.4 Theoretical considerations

If the goal of a sensitivity index is to communicate complex phenomena as simply as possible, it is advantageous to establish consistency and validity in the underlying methods such that there is confidence in the interpretations. McLaughlin (2001) concluded that "in the absence of a sound mathematical argument for a decision, the best course is to do the simplest calculation, i.e. addition." We suggest that the best method to use is one that is mathematically sound and that is practical. Based on the results of this study, the CSI_{μ} method is a practical and theoretically justified method for widespread use.

The two main drawbacks to wide-scale adoption of the CSI_{μ} method are (1) increased computational load, and (2) counterintuitive interpretation of results. The CSI_g and CSI_c methods have the advantage of simple mean statistics calculations, something that is intuitive for most users. The CSI_{μ} method, however, represents not a simple statistical summary, but instead a rank order whose magnitude depends on the sample population being studied. We would argue, however, that once this initial hurdle is overcome then the method provides a richer basis for interpretation of relative coastal sensitivity. Additionally, though a significant problem when thirty years ago, the increased computational load of the CSI_{μ} method no longer represents a challenge to its adoption. In this study we calculated results for the entire Canadian marine coastline on standard laptop hardware.

The use of the CSI_c has led to situations where small differences in scores are attributed significance that may be over-interpreted. A CSI_c value of 3.1 for one section of coast and 3.3 for another may not be statistically different, but this discrepancy is often not

addressed. Instead, interpretation of the results is based on any difference in values. The use of μ -statistics removes this requirement because the establishment of meaningful difference is a fundamental part of the method and is implicit in the partial ordering of results (Wittkowski et al., 2004). No superfluous assumptions are introduced, and the need for validation is reduced. The validity of the μ -statistics method is determined at the scoring stage, because the aggregation method adds no additional assumptions beyond the one that states that sensitivity increases with score.

It is important to note that in situations where underlying variables can be aggregated without introducing rankings to the underlying variables, the CSI_g method remains the most suitable for the reasons discussed in Cogswell *et al.*, (2018).

This study, which uses a mix of real and simulated data, shows appropriate uses of each method. The CSI_c , by far the most widely-used method, is not desirable for comparative purposes when there is a lack of consensus on the ranking of underlying variables. The CSI_μ method offers advantages in its ability to enhance extreme values, which are commonly of most interest in sensitivity and vulnerability studies, as well as its inherent insensitivity to expected ranges of uncertainty in the underlying variables. Furthermore, because it is not reliant on erroneous assumptions about the underlying variables, it is at a theoretical advantage over other methods. For these reasons we recommend the use of the CSI_μ in calculating coastal sensitivity indices.

6.0 Conclusions

Comparison of three coastal sensitivity index methods using identical data shows large difference among the resulting distributions. A commonly used method (CSI_c) increases the range of values and favours lower sensitivity. A second method using the geometric mean (CSI_g), yields generally higher sensitivity than the CSI_c method. It gives results similar to a third μ -statistics method (CSI_μ). The CSI_c and CSI_g methods make assumptions that are not required with the CSI_μ method. Because of its method of aggregation, the CSI_μ method is less sensitive to the propagation of error than the other two.

434
435 Our proposed method of CSI calculation does not involve assumptions about
436 the underlying variables, is less sensitive to uncertainty in the underlying
437 variables, and is easily interpretable to coastal managers. We hope to improve
438 coastal management practice by providing a sound mathematical method for
439 assessing coastal sensitivity using μ -statistics.

440

441

442 **Acknowledgments**

443 This research was funded by the Climate Change Geoscience Program, Lands and
444 Minerals Sector (LMS), of Natural Resources Canada and the Climate Change Impacts
445 and Adaptation Division (CCIAD) of LMS. Don Lemmen provided early encouragement
446 and funding for a preliminary version of CanCoast, and May-Ann Wilson with CCIAD
447 funded development of CanCoast 2.0. We gratefully acknowledge the work of Sheila
448 Hynes, Paul Fraser, Kimberley Jenner, Tracy Lynds and Barbe Szlavko in helping
449 coordinating the work of Eric Patton, Chelsea Smith, Nicholas Hosek and Elliot Wand
450 (Geological Survey of Canada-Atlantic), who provided valued assistance throughout the
451 study. We acknowledge the early work of John Shaw, Donald Forbes, Steven Solomon,
452 Robert Taylor, and David Frobel, without whom we would not have had the starting point
453 to conduct this research. This manuscript benefitted greatly from reviews by Thomas
454 James and Vladimir Kostylev (GSC) and X anonymous reviewers. This is Natural
455 Resources Canada contribution number 20210067. Data are publicly available under the
456 name CanCoast on the NRCan portal Geoscan.

References

- Abuodha, P. 2009. Application and evaluation of shoreline segmentation mapping approaches to assessing response to climate change on the Illawarra Coast, South East Australia (PhD thesis). University of Wollongong, Wollongong, Australia.
- Abuodha, P., & Woodroffe, C. 2006. Assessing vulnerability of coasts to climate change: A review of approaches and their application to the Australian coast. *In* Woodroffe C.D., Bruce, E., Puotinen, M. and Furness, R.A. (Eds), GIS for the Coastal Zone: A Selection of Papers from CoastGIS 2006. Australian National Centre for Ocean Resources and Security University of Wollongong, Wollongong, Australia, 2007.
- Cogswell, A., Greenan, B. J. W., and Greyson, P. 2018. Evaluation of two common vulnerability index calculation methods. *Ocean & Coastal Management*, 160: 46–51. <https://doi.org/10.1016/j.ocecoaman.2018.03.041>
- Cooper, J. A. G., and McLaughlin, S. 1998. Contemporary multidisciplinary approaches to coastal classification and environmental risk analysis. *Journal of Coastal Research*, 14(2): 512–524.
- Dupont, F., Hannah, C.G., Greenberg, D.A., Cherniawsky, J.Y. and Naimie, C.E. 2002. Modelling System for Tides for the Northwest Atlantic Coastal Ocean. Canadian Technical Report of Hydrography and Ocean Sciences 221, Fisheries and Oceans Canada, Ottawa.
- ESRI 2021. ArcGIS Pro – Overview. <https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview>
- Fulton, R.J. 1995. Surficial Materials of Canada, Geological Survey of Canada, Map 1880A. <https://doi.org/10.4095/205040>.
- GEBCO Compilation Group. 2021. GEBCO 2020 Grid (doi:10.5285/a29c5465-b138-234d-e053-6c86abc040b9).
- Gibb, J. 1992. A Standardised Coastal Sensitivity Index based on an Initial Framework for Physical Coastal Hazards Information. New Zealand Department of Conservation, Science & Research Series no. 55. Wellington, New Zealand.
- Gornitz, V. 1991. Global coastal hazards from future sea level rise. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 89. pp. 379–398.
- Gornitz, V., & Kanciruk, P. 1989. Assessment of Global Coastal Hazards from Sea Level Rise. Oak Ridge National Lab., TN (USA).
- Heginbottom, J.A., M.A. Dubreuil, and P.T. Harker. 1995. Canada: Permafrost. *In* National Atlas of Canada Fifth Edition. Natural Resources Canada, MCR 4177. <https://doi.org/10.4095/294672>.
- Hoeffding, W. 1948. A Class of Statistics with Asymptotically Normal Distribution. *Annals of Mathematical Statistics*, 19: 293–325.
- James, T.S., Robin, C., Henton, J.A., and Craymer, M. *in preparation*. Relative Sea-level Projections for Canada's Coastlines. Geological Survey of Canada, Open File, Natural Resources Canada, Ottawa.
- Jenks, G.F. 1967. "The Data Model Concept in Statistical Mapping", *International Yearbook of Cartography* 7: 186-190.
- Karymbalis, E., Chalkias, C., Chalkias, G., Grigoropoulou, E., Manthos, G., and Ferentinou, M. 2012. Assessment of the sensitivity of the southern coast of the Gulf of Corinth (Peloponnese, Greece) to sea-level rise. *Central European Journal of Geosciences*, 4.

- 503 Lemmen, D., Warren, F., & Mercer Clarke, C. 2016. Introduction. *In* Canada's Marine
 504 Coasts in a Changing Climate. *Edited by* D. Lemmen, F. Warren, & C. Mercer
 505 Clarke. Government of Canada, Ottawa, Ontario. pp. 17–26. Ottawa, ON.
- 506 Manson, G. K., Couture, N., and James, T. 2019. CanCoast 2.0: Data and Indices to
 507 Describe the Sensitivity of Canada's Marine Coasts to Changing Climate. Geological
 508 Survey of Canada. Open File 8551. 1.zip file. <https://doi.org/10.4095/314669>.
- 509 McLaughlin, S. (2001). Assessment and Development of a Coastal Vulnerability Index
 510 for Northern Ireland Employing GIS Techniques Doctoral dissertation. University of
 511 Ulster. Ulster, Northern Ireland, United Kingdom.
- 512 Morales, J.F., Song, T., Auerbach, A.D., & Wittkowski, K.M. 2008. Phenotyping Genetic
 513 Diseases Using an Extension of μ -Scores for Multivariate Data. Statistical
 514 Applications in Genetics and Molecular Biology. 7 (1). [https://doi.org/10.2202/1544-](https://doi.org/10.2202/1544-6115.1372)
 515 [6115.1372](https://doi.org/10.2202/1544-6115.1372).
- 516 Natural Earth. 2021. 1:10M Large Scale Data. <https://www.naturalearthdata.com/>
 517 (Accessed 2021-06-27).
- 518 Natural Resources Canada. 2017. Canadian Digital Elevation Model Data.
 519 <ftp.geogratis.gc.ca/pub/nrcan-rncan/elevation/cdem-mnec> (Accessed 2021-06-27).
- 520 Natural Resources Canada. 2021. CanCoast - Coastal Materials Version 2.0.
 521 <https://open.canada.ca/data/en/dataset/73714ed4-a795-a7ae-7e93-36100ce7c242>
 522 (Accessed 2021-06-27).
- 523 Shaw, J., Taylor, R. B., Forbes, D. L., Ruz, M. H., and Solomon, S. 1998. Sensitivity of
 524 the Coasts of Canada to Sea-Level Rise. Geological Survey of Canada, Bulletin 505.
 525 <https://doi.org/10.4095/210075>
- 526 Thieler, E. R., and Hammar-Klose, E. S. 1999. National Assessment of Coastal
 527 Vulnerability to Sea-Level Rise; US Atlantic Coast. U.S. Geological Survey Open-
 528 File Report 99-593.
- 529 World Climate Research Program (WCRP), 2011. Coupled Model Intercomparison
 530 Project Phase 5. CLIVAR Exchanges Newsletter Special Issue 56
- 531 Wheeler, J.O., Hoffman, P.F., Card, K.D., Davidson, A., Sandford, B.V., Okulitch, A.V.,
 532 and Roest, W.R. 1996. Geological Map of Canada. Geological Survey of Canada Map
 533 1860A. <https://doi.org/10.4095/208175>.
- 534 Wittkowski, K. M. 1992. An Extension to Wittkowski. Journal of the American
 535 Statistical Association, 87(417), 258–258.
 536 <https://doi.org/10.1080/01621459.1992.10475202>
- 537 Wittkowski, K. M., Lee, E., Nussbaum, R., Chamian, F. N., & Krueger, J. G. 2004.
 538 Combining several ordinal measures in clinical studies. Statistics in Medicine, 23(10):
 539 1579–1592. <https://doi.org/10.1002/sim.1778>

Figures and Tables

Figure 1. Map showing the marine coasts of Canada with the 100 m and 1000 m bathymetric contours, with general place names, and specific place names referenced in the text. The map is in Lambert Conformal Conic projection, with the WGS 84 datum. Map created using ESRI ArcGIS 10.5 (ESRI, 2021). Land polygons for Canada from CanVec9.0 (Natural Resources Canada, 2017), and land polygons outside of Canada from Natural Earth 1:10M vector data (Natural Earth, 2021). Ocean contours taken from the GEBCO 2020 Gridded Bathymetry Dataset (GEBCO, 2020).

Figure 2. A) Example data set showing four "coastal types" (rows), each with three sensitivity indicators. B) Partial ordering of table in A. It is a partial ordering because there are only two groups of rows. Row 3 is in group 1, and the other rows are in group 2. Rows in each group cannot be sorted, but groups as a whole can be sorted. (C) Results of score calculation.

Table 1 : Description of the six datasets in the CanCoast database used in the calculation of CSIs (taken with permission from Manson et al., 2019).

Dataset	Description
Sea-level change 2006 to 2020	Amount of relative sea-level (RSL) change between 2006 and 2020 (James et al., in prep).
Decadal mean wave height including sea ice (2000s)	Mean significant wave height between 1995 and 2005 (WCRP, 2011). Derived from CMIP5 data.
Ground ice	Amount of ground ice in unconsolidated sediments in the form of pores, veins, lenses and massive bodies (Heginbottom et al., 1995).
Material	Combination of surficial (Fulton, 1995) and bedrock (Wheeler et al., 1996) geology. Bedrock geology is used except where surficial materials were mapped in which case the surficial geology was used.
Slope	Mean slope landward within 200 m of the shoreline. Derived using the Canadian Digital Elevation Model (CDEM) (NRCan, 2017).

560

Tidal range	The 30 year mean of hourly predicted water levels (Dupont et al., 2002).
-------------	--

Draft

Table 2: Classification of CanCoast variables into scores from one to five based on coastal sensitivity (taken with permission from Manson et al., 2019).

Variables Scores	1	2	3	4	5
Sea-level change 2006 to 2020	<= -0.33 m	-0.32 m to -0.20 m	-0.19 m to 0.20 m	0.21 m to 0.70 m	> 0.70 m
Decadal mean wave height including sea ice (2000s)	<= 0.25 m	0.26 m- 0.75 m	0.76 m to 1.50 m	1.51 m to 2.25 m	> 2.25 m
Ground Ice	Permafrost does not exist or ground ice is Nil	nil to Low and Low	Low to Moderate and Moderate	Moderate to High	High
Material	<ul style="list-style-type: none">– Intrusive rocks– Metamorphic rocks– Volcanic rocks	<ul style="list-style-type: none">– Sedimentary and volcanic rocks– Sedimentary rocks– Unknown bedrock	<ul style="list-style-type: none">– Blocks, and rubble with sand and silt– Rubble and silt– Sand and gravel– Sand and gravel and locally diamicton– Sand, gravel and pockets of finer sediment– Sand, silt and gravel	<ul style="list-style-type: none">– Thick and continuous till– Sand and locally gravel– Sand and minor silt– Stratified silt, sand, clay and gravel	<ul style="list-style-type: none">– Silt, and clay, locally containing stones– Silt, clay and fine sand– Peat, muck and minor inorganic sediments– Fluid silty clay and clayey silt– Ice and minor morainal debris
Slope	> 24°	12.1° to 24°	5.1° to 12.0°	1.1° to 5.0°	<= 1.0°
Tidal range	2.1 m to 4.0 m	4.1 m to 6.0 m	1.1 m to 2.0 m	> 6.0 m	<= 1.0 m

Draft

Figure 3. Histograms of the results of the three tested methods: (a) the classic method (CSI_c), (b) the geometric mean method (CSI_g), and (c) the μ -statistics method (CSI_u). The unit of measure for all three x axes is a unit-less CSI index score.

Figure 4. Results of the three CSI methods mapped to the Canadian marine coastline: (a) the classic method (CSI_c), (b) the geometric mean method (CSI_g), and (c) the μ -statistics method (CSI_u). All results have been classified using the Natural Jenks method with five classes ranging from Very Low Sensitivity to Very High Sensitivity. Map created using

ESRI ArcGIS Desktop 10.5 (ESRI, 2021). Coastline vector taken from the CanCoast Coastal Materials Version 2.0 database (Natural Resources Canada, 2021).

***PRINT FIGURE IN COLOUR**

Figure 5. Scatter plots comparing results of the new (CSI_u) method to results of (a) the classic (CSI_c), and (b) the geometric mean (CSI_g) methods. Both scatter plots exhibit an inflection point just below 0 on the x-axis, with a marked increase in slope in the positive values of CSI_u . All axes are in unit-less index scores.