STEM Education
STEM Education, 3 (3): 171–189
DOI: 10.3934/steme.2023011
Received: 06 July 2023
Revised: 27 July 2023
Accepted: 09 August 2023
Published: 17 August 2023

https://www.aimspress.com/journal/steme

Technical note

Solving problems involving numerical integration (II): Modified Simpson's methods for equal intervals of odd numbers

William Guo*

School of Engineering and Technology, Central Queensland University, Bruce Highway, North Rockhampton, QLD 4702, Australia

* **Correspondence:** Email: w.guo@cqu.edu.au; Tel: +61-7-49309686.

Academic Editor: Dianshun Hu

Abstract: The trapezium and Simpson's methods are widely used for numerical integration. In most circumstances, Simpson's method is more accurate than the trapezium method but only applicable to cases with equal intervals of even numbers. This technical note reports the formulation of two modified Simpson's methods, the trapezium-corrected Simpson's method (TCSM) and cubic-corrected Simpson's method (CCSM), as general-purpose symmetric formulas to solve numerical integrations with equal intervals of odd numbers ($n \geq 5$) with the same level of accuracy as that of Simpson's method applied to the even number near n. Error analysis in terms of the order of error bound and case studies in this note demonstrate and validate the usefulness of the proposed formulas for solving different types of theoretical problems and real-world applications. In terms of accuracy of approximation for cases with equal intervals of odd numbers, CCSM performs better than TCSM by at least one order in error bound whereas TCSM performs better than the trapezium method by at least one order in error bound.

Keywords: numerical integration, Simpson's method, trapezium method, order of error bound, trapezium-corrected Simpson's method (TCSM), cubic-corrected Simpson's method (CCSM)

1. Introduction

Some definite integrals cannot be solved by analytical means, for example $\int_0^1 \frac{\sin x}{x} dx$. As the definite integral of $y = f(x)$ in $[a, b]$ in geometry equals the area enclosed between $f(x)$ and the x-axis within $[a, b]$, one simple strategy to approximate this area is to divide the range $b–a$ into n equal vertical strips; then calculate the subarea of each of these vertical strips using a known method; finally add all these subareas together as an approximate to the total area as the integral. Such strategy has brought many widely used methods, such as the trapezium rule and Simpson's rules [1–8]. As an extension, numerical integration would be unique in estimating the 'area' under a set of discrete data points in sequence which are supposed to be on the curve of an unknown integrand $f(x)$ [7–9].

The trapezium rule is a general-purpose method applicable to all cases where the range can be divided into equal intervals of any number. Simpson's 1/3 rule, also referred to as Simpson's method, is only applicable to cases where the range is divided into equal intervals of even numbers. Under the same condition where both methods can be applied, Simpson's method is more accurate than the trapezium method if the integrand is continuous and changes smoothly. However, the trapezium method would perform better if dealing with integrals for very narrow peaks or zigzag curves [9,10].

In real applications where the set of known data has a smooth trend but with an odd number, the trapezium method becomes the simple choice for the corresponding numerical integration because Simpson's method is not applicable to sequential data points of odd numbers. If Simpson's method were directly applied to cases of odd numbers, larger errors than that of applying the trapezium method would occur [9]. There exist methods of third-order Simpson's rules for equal intervals of odd numbers as a multiple of 3 [5–7], but they do not cover data sets of many other odd numbers, particularly the prime numbers, such as 5, 7, 11, 13 and so forth. In practice, engineers and other practitioners may have a tendency towards choosing one of Simpson's formulas to approach numerical integrations by overlooking the requirements on using Simpson's methods as demonstrated by some engineering students [9] whereas other engineering students also demonstrated how to combine Simpson's method with the trapezium method to estimate the land area with survey datasets of odd number.

It is not new to combine two existing methods to deal with numerical integrations for cases with equal intervals of odd numbers. A few unpublished lecture notes or public websites have indicated the usefulness of some combinations [11,12]. However, these efforts are not enough in terms of firstly deriving a general-purpose symmetric formula for users to choose for solving appropriate numerical integrations, just like choosing a Simpson's rule; secondly providing a logical analysis on the error bound for the combined method so as to give users confidence in using the proposed formula; thirdly validating the usefulness of the proposed formula for solving different types of theoretical problems and real-world applications. This technical note aims to fulfil these three tasks.

Since this technical note follows the classroom note on numerical integration published previously [9], fundamentals of the trapezium and Simpson's methods are not repeated in this second part on numerical integration, but only a brief summary. Section 2 is to derive the proposed formulas based on the existing methods. Theoretical problems that have exact solutions are solved numerically by the existing and the proposed methods to assess the accuracy of these methods in Section 3. Section 4 presents more cases of solving real-world problems in science and engineering. The major points of this technical note are summarised in Section 5.

2. Modification of Simpson's method

2.1. Summary of the trapezium method

By dividing the range $[a, b]$ into n equal intervals, the subarea (or integral) of any single segment as a trapezoid can be estimated by

$$A_i = \frac{1}{2}(y_{i-1} + y_i)h \longleftarrow h = \frac{b-a}{n}. \tag{1}$$

Add the n subareas together to obtain an estimate to the integral as

$$\int_a^b f(x)dx = \int_a^b ydx = \frac{h}{2}[y_0 + 2(y_1 + y_2 + y_3 + ... + y_{n-1}) + y_n] = \frac{h}{2}[y_0 + 2\sum_{i=1}^{n-1} y_i + y_n]. \tag{2}$$

Assuming the integrand $f(x)$ is continuous and its second-order derivative exists in $[a, b]$, the error of a single strip and the total error of the n strips together using the trapezium method have been proven to be bounded by [13–16]

$$\begin{cases} E_i \le \dfrac{(b-a)^3 M}{12} \to O(h^3) \\ E_n \le \dfrac{(b-a)^3 M}{12n^2} \to O(h^2) \end{cases}, \tag{3}$$

where M is the maximum absolute value of the second-order derivative of $f(x)$ within $[a, b]$, i.e., $|f''(\xi)| \le M,\ \xi \in [a,b]$. This indicates that the error of the composite trapezium method is about the order of $O(h^2)$ even though the error for a single strip is in the order of $O(h^3)$. If the single interval is around 1/10 = 0.1, the maximum error would be around the order of $O(10^{-3})$ for a single strip and $O(10^{-2})$ for all n strips together.

2.2. Summary of Simpson's methods

If dividing the range $[a, b]$ into equal strips of an *even number*, a local quadratic interpolation can be created using the three known boundary points of any two adjunct strips to replace $f(x)$ within the sub-range $[x_{i-1}, x_{i+1}]$. The area under the interpolation in $[x_{i-1}, x_{i+1}]$ is regarded as an approximate to the area under $f(x)$ in this segment:

$$A_i = \frac{h}{3}(y_{i-1} + 4y_i + y_{i+1}). \tag{4}$$

As the process is repeated by a sliding window of two adjacent strips over the whole range, the total area of all strips of *even-number* can be obtained by adding all $n/2$ subareas together as follows:

$$\int_a^b ydx = \frac{h}{3}[y_0 + 4(y_1 + y_3 + ...) + 2(y_2 + y_4 + ...) + y_n] = \frac{h}{3}[y_0 + 4\sum_{i=1}^{n-1} y_i(odd) + 2\sum_{i=2}^{n-2} y_i(even) + y_n]. \tag{5}$$

Assuming the integrand $f(x)$ is continuous and its fourth-order derivative exists in $[a, b]$, the maximum errors of a single subarea and the total area approximated by Simpson's formulas (4-5) have been proven to be bounded by [13–15]

$$\begin{cases} E_i \le \dfrac{(b-a)^5 M}{2880} \to O(h^5) \\ E_n \le \dfrac{(b-a)^5 M}{180n^4} \to O(h^4) \end{cases}, \tag{6}$$

where M is the maximum absolute value of the fourth-order derivative of $f(x)$ within the range, i.e., $\left|f^{(4)}(\xi)\right| \le M,\ \xi \in [a,b]$. Formula (6) indicates that the maximum error of the composite Simpson's method is about the order of $O(h^4)$. If the single interval is around 1/10 = 0.1, the error would be in the order of $O(10^{-4})$ for all n strips together, much more accurate than that of the trapezium method in general cases.

Simpson's rule (5) only applies to integrals with equal intervals of even numbers. If the range is divided into equal intervals by a number that is a multiple of 3, a section of three consecutive strips can form a cubic polynomial that can be integrated analytically to obtain the area of this segment. Moving this integral window across the whole range would form a composite formula expressed as follows, which is commonly referred to as Simpson's 3/8 rule [6],

$$\int_a^b y dx = \frac{3h}{8}[y_0 + 3(y_1 + y_2 + y_4 + ...) + 2(y_3 + y_6 + ...) + y_n] = \frac{3h}{8}\left[y_0 + 3\sum_{i=1, i\neq 3k}^{n-1} y_i + 2\sum_{i=1}^{n/3-1} y_{3i} + y_n\right]. \quad (7)$$

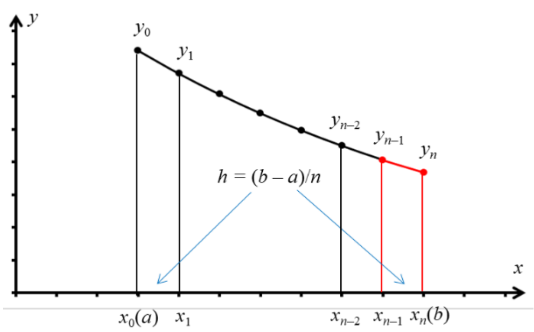
The maximum error for a single section of three strips by Simpson's 3/8 rule is

$$E_i \le \frac{(b-a)^5 M}{6480} \to O(h^5), \quad (8)$$

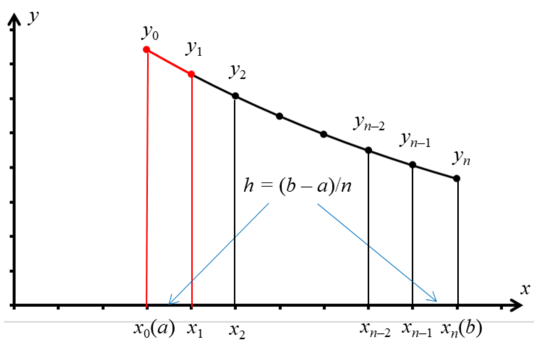
where M is the maximum absolute value of the fourth-order derivative of $f(x)$ within the range, i.e., $\left|f^{(4)}(\xi)\right| \le M,\ \xi \in [a,b]$. Formula (8) indicates that the maximum error of Simpson's 3/8 rule shares the same order of error bound in $O(h^5)$ for a single segment of three consecutive strips, the same order as Simpson's 1/3 rule for two adjacent strips, but it is about three times smaller than that of Simpson's method, so is true for the error bound in the order of $O(h^4)$ for the composite methods.

2.3. The trapezium-corrected Simpson's method (TCSM)

Where the range is divided into equal intervals of an odd number or the known datasets are with an odd number, if the users still prefer to use Simpson's method or alike rather than the trapezium method, one alternative could be incorporating the concept of quadratic interpolation behind Simpson's method and the concept for the trapezium method. As illustrated in Figure 1a, if the total strips come up with an odd number (n), we can apply Simpson's method to the first n–1 strips as an even number,



(a)



(b)

Figure 1. The trapezium-corrected Simpson's method (TCSM) with equal strips of odd numbers.

$$I_1 = \int_a^{b-h} ydx = \frac{h}{3}\left[y_0 + 4(y_1 + ... y_{n-2}) + 2(y_2 + ... y_{n-3}) + y_{n-1}\right] = \frac{h}{3}\left[y_0 + 4\sum_{i=1}^{n-2} y_i(odd) + 2\sum_{i=2}^{n-3} y_i(even) + y_{n-1}\right]. \quad (9)$$

The area of the last strip between x_{n-1} and x_n can be approximated by the trapezium method:

$$I_2 = \frac{h}{2}(y_{n-1} + y_n) = \frac{h}{6}(3y_{n-1} + 3y_n). \quad (10)$$

The total area or integral can be approximated by

$$\begin{aligned} I = I_1 + I_2 &= \frac{h}{3}\left[y_0 + 4(y_1 + ... y_{n-2}) + 2(y_2 + ... y_{n-3}) + y_{n-1}\right] + \frac{h}{2}(y_{n-1} + y_n) \\ &= \frac{h}{6}\left[2y_0 + 8(y_1 + ... y_{n-2}) + 4(y_2 + ... y_{n-3}) + 2y_{n-1}\right] + \frac{h}{6}(3y_{n-1} + 3y_n) \\ &= \frac{h}{6}\left[2y_0 + 8(y_1 + ... y_{n-2}) + 4(y_2 + ... y_{n-3}) + 2y_{n-1} + 3y_{n-1} + 3y_n\right] \\ &= \frac{h}{6}\left[2y_0 + 8(y_1 + ... y_{n-2}) + 4(y_2 + ... y_{n-3}) + 5y_{n-1} + 3y_n\right]. \end{aligned}$$

$$\int_a^b ydx = \frac{h}{6}\left[2y_0 + 8\sum_1^{n-2} y_i(odd) + 4\sum_2^{n-3} y_i(even) + 5y_{n-1} + 3y_n\right] \quad (n \text{ is odd}). \quad (11)$$

Alternatively, if we apply the same strategy to the strips of even number between x_1 and x_n, and the single strip between the first two points (x_0, y_0) and (x_1, y_1), as shown in Figure 1b, a new formula can be obtained as follows:

$$\int_a^b ydx = \frac{h}{6}\left[3y_0 + 5y_1 + 4\sum_3^{n-2} y_i(odd) + 8\sum_2^{n-1} y_i(even) + 2y_n\right] \quad (n \text{ is odd}). \quad (12)$$

A combined approximate value can be obtained by the average of formulas (11-12):

$$\int_a^b ydx = \frac{h}{12}\left[5y_0 + y_1 + 12\sum_{i=1}^{n-1} y_i + y_{n-1} + 5y_n\right] \quad (n \text{ is odd}). \quad (13)$$

Note that this is a symmetric formula that is easy to remember and implement. Since this modified approach is based on a correction with the trapezium method, this modified method is called the trapezium-corrected Simpson's method (TCSM). Hence, the order of error bound of TCSM would be the sum of that for the trapezium method to a single strip and that for the composite Simpson's method to other n-1 strips, i.e.,

$$E_n = E_{n-1}^S + E_1^T \approx O(h^4) + O(h^3) = O(h^3), \ \ h \ll 1 \ \& \ n \gg 5. \quad (14)$$

This estimate should hold true for cases where $h < 0.3$ and $n \geq 5$ in practice. It is a logical inference that the error bound of TCSM would be in the order of $O(h^3)$ in the worst case and $O(h^4)$ in the best case. Of course, the smaller the interval size h (or the more intervals divided), the more accurate the estimate.

2.4. The cubic-corrected Simpson's method (CCSM)

As illustrated in Figure 2a, if the number of the total strips is odd ($n \geq 5$), we can apply Simpson's method to the first n–3 strips of an even number as follows:

$$I_1 = \int_a^{b-3h} y dx = \frac{h}{3}[y_0 + 4(y_1 + ... y_{n-4}) + 2(y_2 + ... y_{n-5}) + y_{n-3}] = \frac{h}{3}\left[y_0 + 4\sum_{i=1}^{n-4} y_{odd} + 2\sum_{i=2}^{n-5} y_{even} + y_{n-3} \right]. \quad (15)$$

For the last three strips between x_{n-3} and x_n, a cubic polynomial can be created to represent that section of the integrand. This cubic polynomial is then integrated over the range x_{n-3} to x_n to estimate the area of these three strips, which is equivalent to applying Simpson's 3/8 rule (7) to the last three strips as

$$I_2 = \frac{3h}{8}(y_{n-3} + 3y_{n-2} + 3y_{n-1} + y_n). \quad (16)$$

The total area or integral can be approximated by

$$\int_a^b y dx I = I_1 + I_2 = \frac{h}{3}\left[y_0 + 4\sum_{i=1}^{n-4} y_{odd} + 2\sum_{i=2}^{n-5} y_{even} + y_{n-3} \right] + \frac{3h}{8}\left(y_{n-3} + 3y_{n-2} + 3y_{n-1} + y_n \right). \quad (17)$$

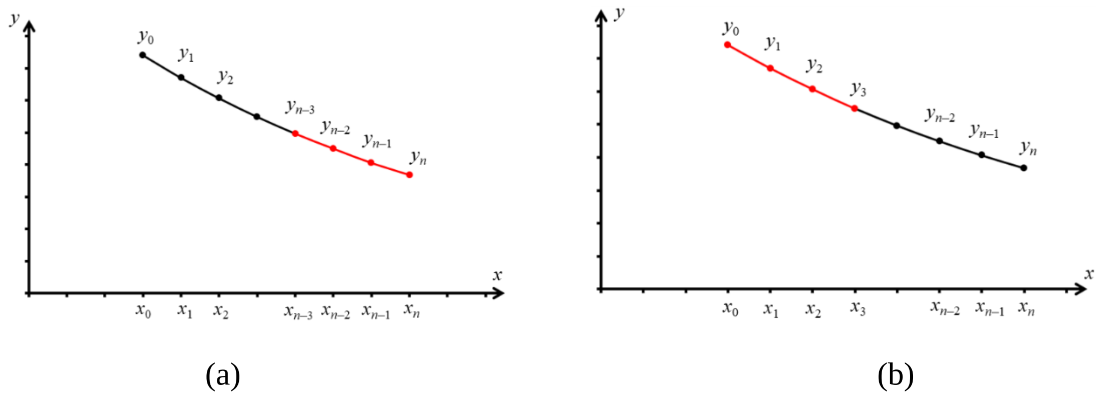


Figure 2. The cubic-corrected Simpson's method (CCSM) with equal strips of odd numbers.

Alternatively, as shown in Figure 2b, if we apply Simpson's method to the strips between x_3 and x_n, and the cubic approach to the first three strips between x_0 and x_3, a new formula can be obtained as follows:

$$\int_a^b y dx = \frac{3h}{8}\left(y_0 + 3y_1 + 3y_2 + y_3 \right) + \frac{h}{3}\left[y_3 + 4\sum_{i=4}^{n-1} y_{even} + 2\sum_{i=5}^{n-2} y_{odd} + y_n \right]. \quad (18)$$

Considering both the modified Simpson's methods (17-18) for datasets or divisions of odd numbers, their average can be regarded as the approximation to the integral as

$$\int_a^b y dx = \frac{h}{48}\left[17(y_0 + y_n) + 11(y_1 + y_{n-1}) - 5(y_2 + y_{n-2}) + (y_3 + y_{n-3}) + 48\sum_{i=1}^{n-1} y_i \right]. \quad (19)$$

This is also a symmetric formula applicable to datasets or divisions of odd numbers where $n \geq 5$. Since this modified approach is based on a correction to Simpson's method with a cubic interpolation, this modified method is called the cubic-corrected Simpson's method (CCSM). Hence, the order of error bound of CCSM would be the sum of that for the Simpson's 3/8 rule to one section of three consecutive strips and that for the composite Simpson's method to n-3 strips, i.e.,

$$E_n = E_{n-3}^S + E_1^{3/8} \approx O(h^4) + O(h^5) = O(h^4), \quad h \ll 1 \;\&\; n \gg 5. \tag{20}$$

This estimate should hold true for cases where $h < 0.3$ and $n \geq 5$ in practice. It is a logical inference that the error bound of CCSM would be in the order of $O(h^4)$ in the worst case and $O(h^5)$ in the best case. Again, the smaller the interval size h, the more accurate the estimate.

Considering both the odd and even numbers ($n \geq 2$) for datasets or divisions, the following set of Simpson's methods can be summarised:

$$\int_a^b y dx = \begin{cases} \frac{h}{3}\left[y_0 + 4\sum_{i=1}^{n-1} y_{odd} + 2\sum_{i=2}^{n-2} y_{even} + y_n \right] & n \text{ is even } \& \geq 2. \\ \frac{3h}{8}\left[y_0 + 3\sum_{i=1, i\neq 3k}^{n-1} y_i + 2\sum_{i=1}^{n/3-1} y_{3i} + y_n \right] & n = 3k,\ k = 1,2,3... \\ \frac{h}{48}\left[17(y_0 + y_n) + 11(y_1 + y_{n-1}) - 5(y_2 + y_{n-2}) + (y_3 + y_{n-3}) + 48\sum_{i=1}^{n-1} y_i \right] & n \text{ is odd } \& \geq 5. \end{cases} \tag{21}$$

3. Performances of the modified Simpson's method

In this section, six examples that have exact solutions are solved by applying both the trapezium and modified Simpson's methods to equal intervals of odd numbers deliberately to assess the accuracy or error bound of these methods with respect to the exact solutions respectively.

Example 1: Use the trapezium, TCSM and CCSM methods to approximate $\int_1^2 \frac{1}{1+x^2} dx$ by dividing the range into five equal intervals.

Solution

All the methods can be realised using Excel with five strips, or $h = 1/5 = 0.2$. The results are shown in Table 1. Note in the table, the subtotal of each row is italic and the total of all the subtotals is italic AND bold. The approximate solution is calculated based on this total in the middle part of the last row for each method.

As the exact solution to this problem is 0.321751 [2], the errors of all relevant methods for this question are shown in Table 2. The best result is from CCSM that is slightly better than that from Simpson's method applied to four equal intervals. However, both share actual errors in the order of $O(10^{-6})$, which is much more accurate than the maximum error bound of $O(10^{-3})$ for the interval h = 0.2. The result from TCSM for this case is the second best with an actual error in the order of $O(10^{-4})$. The trapezium method for this case is the worst with an error in the order of $O(10^{-3})$.

Note that it would make more sense to use relative errors to discuss the accuracy of approximation, but the relative error would be more dependent on the quantity of the exact solution.

Hence, absolute errors are used to make comparisons for simplicity in this work. This case also demonstrated that CCSM is applicable to cases with at least five intervals that make a combination of one single quadratic segment plus one single cubic segment.

Table 1. Results from the trapezium, TCSM, and CCSM methods for Example 1.

	i	0	1	2	3	4	5	Sum
	x_i	1	1.2	1.4	1.6	1.8	2	
Trapezium	y_0 or y_n	0.5					0.2	*0.700000*
	$2y_i$		0.819672	0.675676	0.561798	0.471698		*2.528844*
	Integral			h×Sum/2 = **0.322884**				***3.228844***
TCSM	$5y_0$ or $5y_n$	2.5					1	*3.500000*
	y_1 or y_{n-1}		0.409836			0.235849		*0.645685*
	$12y_i$		4.918033	4.054054	3.370787	2.830189		*15.173062*
	Integral			h×Sum/12 = **0.321979**				***19.318747***
CCSM	$17y_0$ or $17y_n$	8.5					3.4	*11.900000*
	$11y_1$ or $11y_{n-1}$		4.508197			2.594340		*7.102536*
	$-5y_2$ or $-5y_{n-2}$			-1.689189	-1.404494			*-3.093684*
	y_3 or y_{n-3}			0.337838	0.280899			*0.618737*
	$48y_i$		19.672131	16.216216	13.483146	11.320755		*60.692248*
	Integral			h×Sum/48 = **0.321749**				***77.219838***

Table 2. Solutions and errors from different numerical methods for Example 1.

Method	Approximate solution	Absolute error	Order of actual error $O(10^{-n})$	Order of error bound $O(h^n)$
Trapezium (n = 5)	0.322884	0.001133	10^{-3}	4.00×10^{-2}
TCSM (n = 5)	0.321979	0.000229	10^{-4}	8.00×10^{-3}
CCSM (n = 5)	***0.321749***	0.000001	10^{-6}	1.60×10^{-3}
Simpson (n = 4)	0.321748	0.000003	10^{-6}	3.91×10^{-3}
Exact solution		**0.321751**		

Example 2: Use the trapezium, TCSM and CCSM methods to approximate $\int_0^1 e^x \cos x dx$ by dividing the range into seven equal intervals.

<u>Solution</u>

Similar to Example 1, all the methods can be realised using Excel with seven strips, or h = 1/7 and the results are shown in Table 3.

As the exact solution to this problem is 1.378025 [17], the errors of all relevant methods for this question are shown in Table 4. The best result is from CCSM and slightly better than that from Simpson's method applied to six equal intervals. Both share actual errors in the order of $O(10^{-5})$, which is more accurate than the maximum error bound of $O(10^{-4})$ for h = 1/7. The result from TCSM for this case is the second best with an actual error in the order of $O(10^{-4})$. The trapezium method for this case is the worst with an error in the order of $O(10^{-3})$.

Table 3. Results from the trapezium, TCSM, and CCSM methods for Example 2.

	i	0	1	2	3	4	5	6	7	Sum
	x_i	0	0.142857	0.285714	0.428571	0.571429	0.714286	0.857143	1	
Trapezium	y_0 or y_n	1							1.468694	*2.468694*
	$2y_i$		2.283628	2.553532	2.792465	2.978935	3.086811	3.085023		*16.780394*
	Integral					h×Sum/2 = **1.374935**				***19.249088***
TCSM	$5y_0$ or $5y_n$	5							7.34347	*12.34347*
	y_1 or y_{n-1}		1.141814					1.542512		*2.684326*
	$12y_i$		13.701767	15.321191	16.754789	17.873608	18.520867	18.510140		*100.682363*
	Integral					h×Sum/12 = **1.377502**				***115.710158***
CCSM	$17y_0$ or $17y_n$	17							24.967797	*41.967797*
	$11y_1$ or $11y_{n-1}$		12.559953					16.967628		*29.527582*
	$-5y_2$ or $-5y_{n-2}$			-6.383829			-7.717028			*-14.100858*
	y_3 or y_{n-3}				1.396232	1.489467				*2.885700*
	$48y_i$		54.807069	61.284763	67.019158	71.494434	74.083470	74.040560		*402.729453*
	Integral					h×Sum/48 = **1.378005**				***463.009673***

Table 4. Solutions and errors from different numerical methods for Example 2.

Method	Approximate solution	Absolute error	Order of actual error $O(10^{-n})$	Order of error bound $O(h^n)$
Trapezium (n = 7)	1.374935	0.003090	10^{-3}	2.92×10^{-2}
TCSM (n = 7)	1.377502	0.000523	10^{-4}	4.16×10^{-3}
CCSM (n = 7)	***1.378005***	0.000020	10^{-5}	4.16×10^{-4}
Simpson (n = 6)	1.378001	0.000024	10^{-5}	4.16×10^{-4}
Exact solution		**1.378025**		

Example 3: Use the trapezium, TCSM and CCSM methods to approximate $\int_1^2 x(\ln x)^2\,dx$ by dividing the range into seven equal intervals.

Solution

Similar to Example 1, all the methods can be realised using Excel with seven strips, or h = 1/7 and the results are shown in Table 5.

This integral has an analytical solution 0.324612 [18]. With this exact solution as a reference, the errors of all relevant methods for this question are shown in Table 6. The best result is still from CCSM and also slightly better than that from Simpson's method applied to six equal intervals. Both share actual errors in the order of $O(10^{-6})$, which is much more accurate than the maximum error bound of $O(10^{-4})$ for h = 1/7. The result from TCSM in this case is the second best with an actual error in the order of $O(10^{-4})$. The trapezium method for this case is the worst with an error in the order of $O(10^{-3})$.

In normal circumstances, CCSM is expected to be with an error equal to or one-order better than the theoretical error bound. However, the actual error for this case is two orders better than the theoretical error bound, which is beyond the normal expectation. Such superior performance may be attributed to not only the method alone, but also the characteristics of the integrand in the given range. On the plot of the integrand in Figure 3, the function has a slow and smooth increase in [1, 2], which naturally minimises the approximation error by both quadratic and cubic fittings in any segment.

Table 5. Results from the trapezium, TCSM, and CCSM methods for Example 3.

	i	0	1	2	3	4	5	6	7	Sum
	x_i	1	1.142857	1.285714	1.428571	1.571429	1.714286	1.857143	2	
Trapezium	y_0 or y_n	0							0.960906	*0.960906*
	$2y_i$		0.040756	0.162409	0.363477	0.642056	0.996059	1.423350		*3.628106*
	Integral					h×Sum/2 = **0.327787**				***4.589012***
TCSM	$5y_0$ or $5y_n$	0							4.804530	*4.804530*
	y_1 or y_{n-1}		0.020378					0.711675		*0.732053*
	$12y_i$		0.244534	0.974452	2.180863	3.852336	5.976354	8.540098		*21.768639*
	Integral					h×Sum/12 = **0.325062**				***27.305222***
CCSM	$17y_0$ or $17y_n$	0							16.335402	*16.335402*
	$11y_1$ or $11y_{n-1}$		0.224157					7.828423		*8.052580*
	$-5y_2$ or $-5y_{n-2}$			-0.406022			-2.490148			*-2.896169*
	y_3 or y_{n-3}				0.181739	0.321028				*0.502767*
	$48y_i$		0.978138	3.897809	8.723453	15.409345	23.905418	34.160393		*87.074555*
	Integral					h×Sum/48 = **0.324611**				***109.069134***

Table 6. Solutions and errors from different numerical methods for Example 3

Method	Approximate solution	Absolute error	Order of actual error $O(10^{-n})$	Order of error bound $O(h^n)$
Trapezium (n = 7)	0.327787	0.003175	10^{-3}	2.92×10^{-2}
TCSM (n = 7)	0.325062	0.000450	10^{-4}	4.16×10^{-3}
CCSM (n = 7)	***0.324611***	0.000001	10^{-6}	4.16×10^{-4}
Simpson (n = 6)	0.324610	0.000002	10^{-6}	4.16×10^{-4}
Exact solution		**0.324612**		

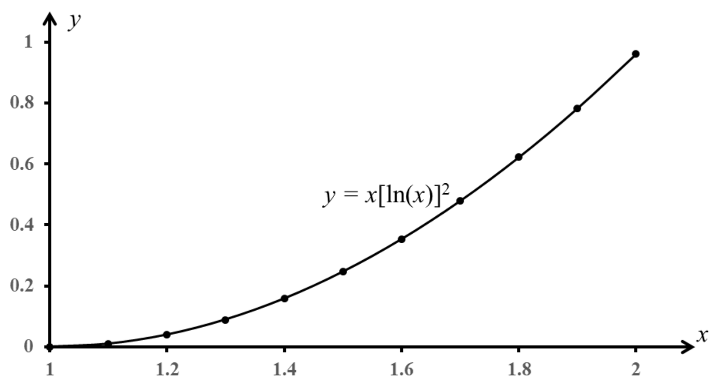


Figure 3. Plot of the integrand in [1, 2] in Example 3.

Example 4: Use trapezium, TCSM and CCSM methods to approximate $\int_1^2 \frac{\arctan x}{x^4} dx$ by dividing the range into seven equal intervals.

Solution

With seven strips, or $h = 1/7$, the results from all three methods are shown in Table 7.

This integral has an analytical solution 0.262334 [17]. With this exact solution as a reference, the errors of all relevant methods for this question are shown in Table 8. The best result is from CCSM and slightly better than that from Simpson's method applied to six equal intervals. This time, both share actual errors in the order of $O(10^{-4})$, same as the order of the maximum error bound for $h = 1/7$, but both are still smaller than this maximum error bound. The results from both TCSM and trapezium methods in this case are with actual errors in the order of $O(10^{-3})$, but still batter than the maximum error bound for both methods. TCSM is more accurate than the trapezium method for this case too.

Table 7. Results from the trapezium, TCSM, and CCSM methods for Example 4.

	i	0	1	2	3	4	5	6	7	Sum
Trapezium	x_i	1	1.142857	1.285714	1.428571	1.571429	1.714286	1.857143	2	
	y_0 or y_n	0.785398							0.069197	*0.854595*
	$2y_i$		0.998814	0.665849	0.461026	0.329317	0.241471	0.181053		*2.877530*
	Integral	h×Sum/2 = **0.266580**								***3.732125***
TCSM	$5y_0$ or $5y_n$	3.926990							0.345984	*4.272975*
	y_1 or y_{n-1}		0.499407					0.090527		*0.589934*
	$12y_i$		5.992884	3.995093	2.766155	1.975902	1.448828	1.086319		*17.265181*
	Integral	h×Sum/12 = **0.263430**								***22.128090***
CCSM	$17y_0$ or $17y_n$	13.351769							1.176346	14.528114
	$11y_1$ or $11y_{n-1}$		5.493477					0.995792		6.489269
	$-5y_2$ or $-5y_{n-2}$			-1.664622			-0.603678			-2.268301
	y_3 or y_{n-3}				0.230513	0.164659				0.395171
	$48y_i$		23.971537	15.980374	11.064619	7.903608	5.795313	4.345274		69.060725
	Integral	h×Sum/48 = **0.262515**								***88.204979***

Table 8. Solutions and errors from different numerical methods for Example 4.

Method	Approximate solution	Absolute error	Order of actual error $O(10^{-n})$	Order of error bound $O(h^n)$
Trapezium (n = 7)	0.266580	0.004246	10^{-3}	2.92×10^{-2}
TCSM (n = 7)	0.263430	0.001095	10^{-3}	4.16×10^{-3}
CCSM (n = 7)	0.262515	0.000181	10^{-4}	4.16×10^{-4}
Simpson (n = 6)	0.262549	0.000215	10^{-4}	4.16×10^{-4}
Exact solution		**0.262334**		

Example 5: Use the trapezium, TCSM and CCSM methods to approximate $\int_1^2 \frac{\ln x}{\sqrt{x}} dx$ by dividing the range into seven equal intervals.

Solution

This is the same question to Example 1 in [9]. With seven strips, or h = 1/7, the results from all three methods are shown in Table 9.

This integral has an analytical solution 0.303662 [9]. With this exact solution as a reference, the errors of all relevant methods for this question are shown in Table 10. The best result is from CCSM and slightly better than that from Simpson's method applied to six equal intervals. Both share actual errors in the order of $O(10^{-5})$, one order better than the maximum error bound of $O(10^{-4})$ for h = 1/7. The result from TCSM in this case is the second best with an actual error in the order of $O(10^{-4})$, also one order batter than the actual error for the trapezium method.

Table 9. Results from the trapezium, TCSM, and CCSM methods for Example 5.

	i	0	1	2	3	4	5	6	7	Sum
Trapezium	x_i	1	1.142857	1.285714	1.428571	1.571429	1.714286	1.857143	2	
	y_0 or y_n	0							0.490129	*0.490129*
	$2y_i$		0.249814	0.443277	0.596831	0.721119	0.823331	0.908501		*3.742873*
	Integral				h×Sum/2 = **0.302357**					***4.233002***
TCSM	$5y_0$ or $5y_n$	0							2.450645	*2.450645*
	y_1 or y_{n-1}		0.124907					0.454251		*0.579158*
	$12y_i$		1.498886	2.659662	3.580988	4.326713	4.939985	5.451007		*22.457240*
	Integral				h×Sum/12 = **0.303417**					***25.487044***
CCSM	$17y_0$ or $17y_n$	0							8.332194	8.332194
	$11y_1$ or $11y_{n-1}$		1.373979					4.996756		6.370735
	$-5y_2$ or $-5y_{n-2}$			-1.108192			-2.058327			-3.166519
	y_3 or y_{n-3}				0.298416	0.360559				0.658975
	$48y_i$		5.995545	10.638648	14.323952	17.306851	19.759938	21.804028		89.828962
	Integral				h×Sum/48 = **0.303644**					***102.024347***

Table 10. Solutions and errors from different numerical methods for Example 5.

Method	Approximate solution	Absolute error	Order of actual error $O(10^{-n})$	Order of error bound $O(h^n)$
Trapezium (n = 7)	0.302357	0.001305	10^{-3}	2.92×10^{-2}
TCSM (n = 7)	0.303417	0.000245	10^{-4}	4.16×10^{-3}
CCSM (n = 7)	***0.303644***	0.000018	10^{-5}	4.16×10^{-4}
Simpson (n = 6)	0.303640	0.000022	10^{-5}	4.16×10^{-4}
Exact solution		**0.303662**		

Example 6: Use the trapezium, TCSM and CCSM methods to approximate $\int_1^2 x^3 \ln x dx$ by dividing the range into seven equal intervals.

Solution

This is the same question to Example 2 in [9]. With seven strips, or h = 1/7, the results from all three methods are shown in Table 11.

This integral has an analytical solution 1.835089 [9]. With this exact solution as a reference, the errors of all relevant methods for this question are shown in Table 12. The best result is from CCSM and slightly better than that from Simpson's method applied to six equal intervals. Both share actual errors in the order of $O(10^{-5})$, one order better than the maximum error bound of $O(10^{-4})$ for h = 1/7. The result from TCSM in this case is the second best with an actual error in the order of $O(10^{-3})$, one order better than that from the trapezium method in the actual error.

Table 11. Results from the trapezium, TCSM, and CCSM methods for Example 6.

	i	0	1	2	3	4	5	6	7	Sum
Trapezium	x_i	1	1.142857	1.285714	1.428571	1.571429	1.714286	1.857143	2	
	y_0 or y_n	0							5.545177	*5.545177*
	$2y_i$		0.398648	1.068269	2.079737	3.507826	5.430822	7.930199		*20.415502*
	Integral					h×Sum/2 = **1.854334**				***25.960679***
TCSM	$5y_0$ or $5y_n$	0							27.725887	*27.725887*
	y_1 or y_{n-1}		0.199324					3.965100		*4.164423*
	$12y_i$		2.391886	6.409617	12.478424	21.046957	32.584931	47.581194		*122.493010*
	Integral					h×Sum/12 = **1.837897**				***154.383320***
CCSM	$17y_0$ or $17y_n$	0							94.268017	*94.268017*
	$11y_1$ or $11y_{n-1}$		2.192562					43.616095		*45.808657*
	$-5y_2$ or $-5y_{n-2}$			-2.670674			-13.577055			*-16.247728*
	y_3 or y_{n-3}				1.039869	1.753913				*2.793782*
	$48y_i$		9.567544	25.638468	49.913695	84.187830	130.339725	190.324778		*489.972039*
	Integral					h×Sum/48 = **1.835103**				***616.594766***

Table 12. Solutions and errors from different numerical methods for Example 6.

Method	Approximate solution	Absolute error	Order of actual error $O(10^{-n})$	Order of error bound $O(h^n)$
Trapezium (n = 7)	1.854334	0.019245	10^{-2}	2.92×10^{-2}
TCSM (n = 7)	1.837897	0.002808	10^{-3}	4.16×10^{-3}
CCSM (n = 7)	***1.835103***	0.000015	10^{-5}	4.16×10^{-4}
Simpson (n = 6)	1.835106	0.000018	10^{-5}	4.16×10^{-4}
Exact solution		**1.835089**		

4. Solving real world problems using the modified Simpson's methods

Four more examples are presented in this section. The first one is a combination of theoretical derivation with numerical calculation to determine the surface area of a parabolic antenna. This problem has an analytical solution that can be used to compare the results from the numerical methods. Other three examples deal with sequential discrete datasets, which do not have analytical solutions. These problems together demonstrate how to solve a problem using different numerical methods.

Example 7: If a parabolic antenna is formed by rotating the function $f(x)=\frac{1}{2}\sqrt{x}$ in the section from

x = 0 to 1 m about the x-axis, estimate the surface area of revolution of this parabolic antenna, accurate to the third decimal place in square metres.

Solution

The surface area of revolution of the parabola $f(x) = \frac{1}{2}\sqrt{x}$ within [0, 1] about the x-axis can be calculated by

$$S = \int_0^1 2\pi f(x)\sqrt{1+(\frac{df}{dx})^2}\,dx = \int_0^1 \pi\sqrt{x}\sqrt{1+(\frac{1}{4\sqrt{x}})^2}\,dx = \int_0^1 \pi\sqrt{x}\sqrt{1+\frac{1}{16x}}\,dx$$

$$= \int_0^1 \frac{\pi}{4}\sqrt{16x+1}\,dx = \frac{\pi}{96}(16x+1)^{\frac{3}{2}}\Big|_0^1 = 2.2611\ m^2.$$

If we use numerical integration to estimate the surface area in section 0-1 m, Simpson's methods would be the first choice due to their higher accuracy. By dividing the range into 10 equal intervals of h = 0.1, the maximum error bound for Simpson's method should be $h^4 = 10^{-4}$ that is accurate enough to the third decimal place in square metres or in the order of $O(10^{-3})$. However, for the purpose of testing the usefulness of the modified Simpson's method CCSM for this case, the range can be divided into 9 equal intervals, which would lead to an estimate with a maximum error bounded by $(1/9)^4 = 1.5 \times 10^{-4}$ that is also accurate enough to the order of $O(10^{-3})$. The results estimated by Simpson's method and CCSM are shown in Table 13, along with the result from the trapezium and TCSM for comparison.

Table 13. Solutions and errors from different numerical methods for Example 7.

Method	Approximate	Final estimate (m^2)	Absolute error (to 10^{-3})
Trapezium (n = 9)	2.2563	2.256	0.005
TCSM (n = 9)	2.2596	2.260	0.001
CCSM (n = 9)	2.2606	2.261	0
Simpson's (n = 10)	2.2608	2.261	0
Analytical	2.2611	2.261	0

In terms of achieving an estimate accurate to the third decimal place in square metres, both Simpson's method for 10 equal intervals and CCSM for 9 equal intervals produced a surface area of 2.261 m^2, the same as the exact solution by analytical means. TCSM for 9 equal intervals is just short of the required standard by one-thousandth. The trapezium method for 9 equal intervals is further back with an error of five-thousandths to the required accuracy standard.

Example 8: A plot of land lies between a straight fence (x-axis) and a stream (northern bound) illustrated in Figure 4. Measured from the western end of the fence, the breadth of the plot (y-axis) was recorded in the table below. Choose an appropriate method to estimate the area of this plot of land. Keep 1 decimal place in the final result.

x (metre)	0	3	6	9	12	15	18	21	24	27
y (metre)	16.3	17.9	20.7	22.8	23.7	23.3	21.9	19.8	18.5	19.7

Solution

This problem is the same as Example 4 solved in [9] by the trapezium method with an estimated area of 559.8 m^2. Note that the Simpson's method cannot be directly applied to this problem as there are 9 equal-width strips over the length of 27 metres. However, Simpson's 3/8 rule can be applied to this case with an estimated area of 559.2 m^2 [9]. Both TCSM and CCSM can be directly applied to this case. The results from all these methods are shown in Table 14.

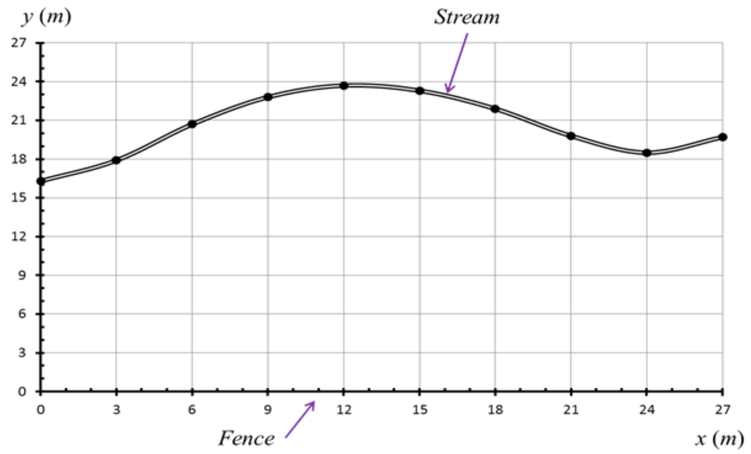


Figure 4. Plot of the measurements for the block of land in Example 8.

Table 14. Solutions from different numerical methods for Example 8.

	Trapezium	TCSM	CCSM	Simpson 3/8 rule
Area (m^2)	559.8	559.9	559.2	559.2

The results from the trapezium and TCSM are close to each other whereas both CCSM and Simpson's 3/8 rule produced the same result. The actual result does not matter too much as the maximum difference between any two estimated areas is very small for this case. However, in terms of mathematical reasoning, the estimated area of 559.2 m^2 by both CCSM and Simpson's 3/8 rule would be more accurate. In theory, both CCSM and Simpson's 3/8 rule perform at least two orders more accurate than the trapezium method and at least one order more accurate than TCSM.

Example 9: A speed gun was used to measure the speed variation of an airplane during taking-off in an airport. The measurement began at the fifth second and then took a reading every five seconds till the end of the first minute. The speed readings in metres per second (m/s) are listed in the table below. Choose an appropriate method to estimate the distance in km the airplane travelled during this period. Keep 3 decimal places in the final result.

t (second)	5	10	15	20	25	30	35	40	45	50	55	60
v (m/s)	10.56	19.44	30	48.72	85.44	149.04	179.04	198.48	213.36	227.52	241.92	256.08

Solution

When taking off, an airplane would initially start with a slow speed and then keep accelerating in the first 10-15 minutes until reaching the steady cruising status. The plot of these speed readings in the first minute during taking off shows a nonlinear variation in speed for the airplane in Figure 5. Note that Simpson's method cannot be directly applied to this problem as there are 11 equal time intervals over the 55 seconds of measurement.

As the distance travelled during the period is equivalent to the area under the curve connecting the speed readings in sequence in Figure 5, it can be estimated using appropriate methods for numerical integration. Since each measurement is made every 5 seconds between 5-60 seconds, the interval size is h = 5 s and the relative interval is 5/55 = 1/11 for the period. The speed reading is in metres per second (m/s) and accurate to centimetres. The accumulated distance travelled would be in metres accurate to centimetres initially by numerical integration, which is then converted to km by keeping three decimal places or accurate to metres. This means that the trapezium, TCSM and CCSM are all credible methods for this problem as any can produce an estimate accurate to at least the millimetre level. The distance estimated from these methods are listed in Table 15.

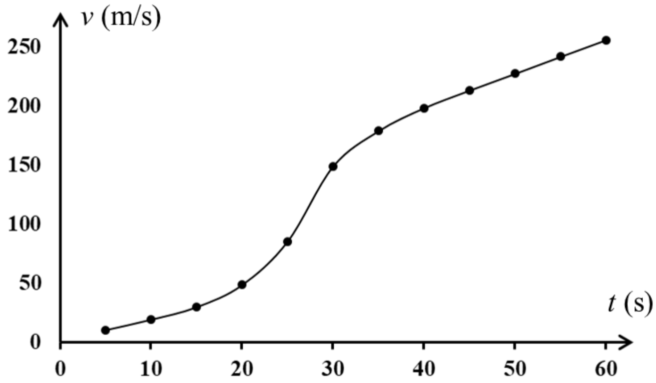


Figure 5. Plot of the speed readings of airplane in Example 9.

Table 15. Solutions from different numerical methods for Example 9.

	Trapezium	TCSM	CCSM
Distance (km)	7.6314	7.6292	7.6296

As analytical solution to this problem is not availability, the distance of 7.630 km estimated by CCSM is likely to be the most accurate result due to its higher accuracy, though the results of 7.631 from the trapezium method and of 7.629 from TCSM are different by only 1 metre, respectively.

Example 10: A lake with an irregular shape is illustrated in Figure 6. To estimate the area of the water surface of this lake, civil engineers obtained a set of boundary survey datasets for this lake, which are also shown in the figure. The measurements were made along the western-eastern direction with an interval of 100 metres. Use Simpson's methods to estimate the area of the water surface of this lake in square metres. Keep 2 decimal places in the final result.

Solution

This problem is reworked from an assessment to engineering students several years ago. As the final result is required to keep two decimal places in square metres, the intermediate results should keep at least three decimal places so that the final result can be properly rounded. As the shape of the lake is irregular, the area of water surface can be estimated by dividing the lake into three areas. Shown in Figure 6, the first area (A_1) is enclosed by the red pots and the black dots; the second area (A_2) is enclosed by the blue dots and the black dots; the third area (A_3) is enclosed by the green dots and the black dots. For each interval, the vertical values feeding into calculation should be the difference between the two measurements at the same horizontal point. Hence, the values feeding into calculation at the four vertexes at x = 0, 900, 1600, and 1900 m should be zero.

As each measurement was made every 100 metres over 900 metres in A_1, 1000 metres in A_2, and 700 metres in A_3, the largest relative interval among the three areas is h_r = 100/700 = 1/7. The largest error to use Simpson's methods for approximation would be bounded by $h_r^4 = 4.16 \times 10^{-4}$, sufficient for the accuracy requirement.

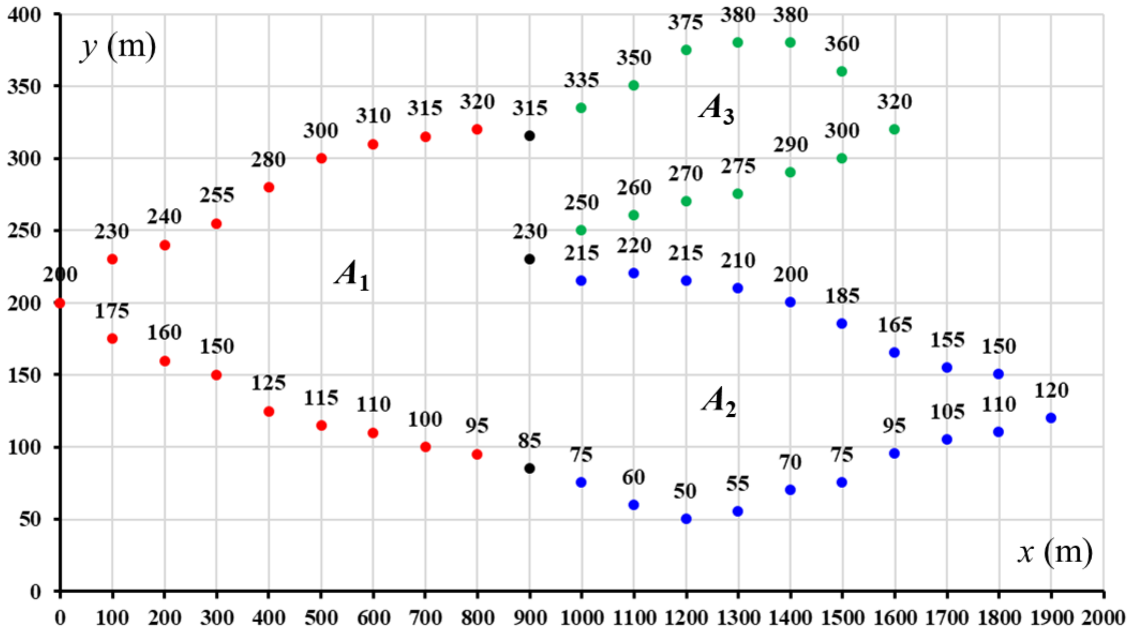


Figure 6. Plot of the measurements for the lake in Example 10.

Refer to formula (21), the first portion has 9 intervals so either CCSM or Simpson's 3/8 rule can be used to estimate the area. The second portion has 10 intervals, whose area can be estimated by Simpson's method. The third portion has 7 intervals, to which CCSM can be applied. The areas estimated from these methods are shown in Table 16.

Table 16. Areas estimated from different numerical methods for Example 10.

	A_1 (Simpson 3/8 rule)	A_2 (Simpson's method)	A_3 (CCSM)	Total
Area (m^2)	134437.50	109166.67	58395.83	302000.00 (= 0.302 km^2)

5. Discussion and conclusion

This technical note reported the formulation of two modified Simpson's methods applicable to subdivisions or datasets with equal intervals of odd numbers for numerical integration. TCSM is based on the average of the combinations of Simpson's method with corrections by the trapezium method at both ends of the given range. CCSM is based on the average of the combinations of Simpson's method with corrections by cubic interpolations at both ends of the given range. TCSM and CCSM are symmetric, hence, easy to remember and implement.

In terms of accuracy of approximation for equal intervals of odd numbers, CCSM performs with an error in the order of $O(h^4)$ where $h < 0.3$ and $n \geq 5$, the same level of accuracy as Simpson's method applied to the even number near the odd number. TCSM performs with an error in the order of $O(h^3)$, lower than that of CCSM but higher than that of the trapezium method.

The outcomes from the examples presented in this note not only verified above observations, but also showed that CCSM should be the first choice in numerical integration for cases with equal divisions or datasets of odd numbers once $n \geq 5$.

This note focuses only on modifying the Simpson's 1/3 rule that is applicable to equal intervals of even numbers so that the modified Simpson's method such as CCSM can be applied to equal intervals of odd numbers with similar accuracy of approximation. Such effort is driven by the fact that numerical integration has been commonly taught in universities through both the trapezium and Simpson's methods. In real applications, other numerical integration methods with even better accuracy are also available for selection with respect to different circumstances, for example, the Euler-Maclaurin rules reported in [10]. This may indicate a need to reconsider a new teaching plan on numerical integration in universities to include both the legacy methods and some more advanced methods in the near future.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The author would like to thank the constructive feedback provided by the reviewers. Special appreciation goes to Dr. Yuri Kalambet for his recommendations to include more advanced methods in numerical integration in tertiary mathematics teaching and learning in the future.

Conflict of interest

The authors declare there is no conflict of interest in any part of this article.

Ethics declaration

The author declared that the ethics committee approval was waived for the study.

References

1. Stewart, J., *Calculus: Concepts and Contexts*, 4th ed. 2019. Boston, USA: Cengage.
2. Guo, W.W., *Essentials and Examples of Applied Mathematics*, 2nd ed. 2020, Melbourne, Australia: Pearson.
3. Larson, R. and Edwards, B., *Calculus*, 12th ed. 2023, Boston, USA: Cengage.
4. Trim, D., *Calculus for Engineers*. 4th ed. 2008, Toronto, Canada: Pearson.

5. Sauer, T., *Numerical Analysis*, 2^{nd} ed. 2014, Harlow, UK: Pearson.
6. Chapra, S.C., *Applied Numerical Methods with MATLAB for Engineers and Scientists*. 2005, Boston, USA: McGraw-Hill Higher Education.
7. Wheatley, G., *Applied Numerical Analysis*, 7^{th} ed. 2004, Boston, USA: Pearson.
8. Stroud, K.A. and Booth, D.J., *Engineering Mathematics*. 7th ed. 2013, London, UK: Palgrave McMillian.
9. Guo, W., Solving problems involving numerical integration (I): Incorporating different techniques. *STEM Education*, 2023, 3(2): 130–147. https://doi.org/10.3934/steme.2023009
10. Kalambet, Y., Kozmin, Y. and Samokhin, A., Comparison of integration rules in the case of very narrow chromatographic peaks. *Chemometrics and Intelligent Laboratory Systems*, 2018, 179: 22–30. https://doi.org/10.1016/j.chemolab.2018.06.001
11. Wolfram MathWorld, *Newton-Cotes Formulas*. Retrieved from: https://mathworld.wolfram.com/Newton-CotesFormulas.html
12. Tebeest, K.G., *Combined Simpson's 1/3 and Simpson's 3/8 Examples*, 2000. Kettering University. Retrieved from: https://paws.kettering.edu/~ktebeest/math305/simp38b.pdf
13. Rozema, E., Estimating the error in the trapezoidal rule. *The American Mathematical Monthly*, 1980, 87(2): 124–128. https://www.jstor.org/stable/2321991
14. Fazekas E.C. and Mercer, P.R., Elementary proofs of error estimates for the midpoint and Simpson's rules. *Mathematics Magazine*, 2009, 82(5): 365–370, https://doi.org/10.4169/002557009X478418
15. Cruz-Uribe, D. and Neugebauer, C.J., Sharp error bounds for the trapezoidal rule and Simpson's rule. *Journal of Inequalities in Pure and Applied Mathematics*, 2002, 3(4): Article 49. http://eudml.org/doc/123201
16. Cruz-Uribe, D. and Neugebauer, C.J., An elementary proof of error estimates for the trapezoidal rule. *Mathematics Magazine*, 2003, 76(4): 303–306. https://doi.org/10.1080/0025570X.2003.11953199
17. Guo, W., A guide for using integration by parts: *Pet-LoPo-InPo*. *Electronic Research Archive*, 2022, 30(10): 3572–3585. https://doi.org/10.3934/era.2022182
18. Guo, W., Streamlining applications of integration by parts in teaching applied calculus. *STEM Education*, 2022, 2(1): 73–83. https://doi.org/10.3934/steme.2022005

Author's biography

Dr. William Guo is a professor in mathematics education with Central Queensland University, Australia. He is specialized in teaching applied mathematics for both engineering and education students. His research interests include mathematics education, applied computing, data analysis and numerical modeling. He is a member of IEEE and Australian Mathematical Society.