Available online at www.worldscientificnews.com

World Scientific News

An International Scientific Journal

WSN 153(2) (2021) 124-141 EISSN 2392-2192

Application of Interpolation and Extrapolation of Newton and Cubic Splines to Estimate and Predict the Gas Content of Hydrogen and Iodine in the Formation of Iodic Acid Reactions

Ati Maryati*, Naomi Pandiangan, Sri Purwani

Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Jl.Raya Bandung-Sumedang KM. 21, Jatinangor, Sumedang, West Java 45363, Indonesia

*E-mail address: ati19002@mail.unpad.ac.id

ABSTRACT

The problem that is mostly related to the pattern of experimental time series data is the function that involves the data. Experimental data in the field of exact sciences is very important to conclude a problem. Existing data can form certain functions. In this research, we are looking for a function that represents the gas content of hydrogen and iodine in the reaction of acid iodide formation. This is achieved by using interpolation in which the function interpolates a given group of data points. Interpolation can also be used to evaluate the function at points different from the group. In addition to constructing and evaluating a functions by interpolation, we can also predict experimental data outside the given group of data points by using extrapolation. The results of data extrapolation can be used as an alternative to experimental data, thereby saving time and cost. This research will also compare interpolation and extrapolation of both Newton method and cubic splines, which one better interpolates and extrapolates data on hydrogen and iodine gas content in the reaction of acid iodide formation. The research results show that the cubic spline method is better than Newton method at approaching data, in terms of interpolation, as well as extrapolation.

***Keywords*:** Newtom Interpolation, Cubic Spline Interpolation, Extrapolation and Time series data

(Received 16 December 2020; Accepted 05 January 2021; Date of Publication 06 January 2021)

1. INTRODUCTION

Performing experiments is very important in the field of exact sciences such as in chemistry, where the basis result from the experiment is referred to obtain accurate data so that a problem can be concluded or solved. In fact, existing experimental data are often incompletely presented for certain needs. Or often there are also missing data values. There are many causes of this condition, both due to human errors and limited measurement tools. Another condition that may arises from the data we have is the presence of outliers or values that are very different from the majority of the data we have. This value will have effect to the results of our analysis or conclusions. There are many ways to treat these conditions. A number of researchers choose to delete this data.

This can be done if the amount of data we have is large enough. What if we have little data and re-measurement is expensive or difficult to do? One way that can be done is to interpolate and extrapolate the data. Interpolation and extrapolation are the process of "guessing" the value of data by considering other data we have. Interpolation is the process of finding and calculating a function whose graph goes through a given set of points [1]. Interpolation can also be used to evaluate a fuction at points different from a given set of points. Whereas extrapolation is a technique used to evaluate a fuction at points outside a given set of points or outside known data range.

The interpolation rule can reach optimal levels for nonparametric estimation and prediction problems with quadra losses [2]. One of the implementations of basic Mathematics and Computer Science is polynomial interpolation. Polynomial interpolation basically uses polynomial as a base of interpolating function. Some examples of polynomial interpolation methods are the Newton and cubic spline methods, which will be used to estimate and predict the gas content of hydrogen and iodine in the reaction of the formation of iodic acid, at certain seconds and in subsequent seconds. Prediction is a process of estimating future value using past data. Predictions show what will happen in a certain situation and are input to the planning process and decision making.

The following is a description of previous research regarding spline and newton interpolation. Use a spline to get a resolution error correction in small angle [2], find a function with a cubic spline to measure the stem volume [3], prove cubic spline interpolation works much better than linear interpolation in some case [4]. end conditions is derived for cubic spline functions by use of integration [5], Newton-type method for solving linear complementarity problems [6], Linear method for interpolation [7], Multivariate polynomials with Newton approach [8], prove local quadratic convergence by viewing it as a semismooth Newton method [9], present a higher order predictor method with cubic spline [10], and Comparation between polynomial interpolation with Numerical vs Statistical techniques [11].

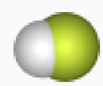
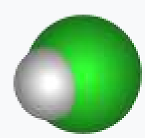
In this research, the polynomial interpolation methods used are Newton and Cubic Spline methods. With these two methods we look for a function that represents the gas content of hydrogen and iodine in the reaction of the formation of iodic acid. The results can be used to estimate and predict the gaseous content at points different from a given set of points and outside the set respectively. In addition, we also try to find out which method is better and more appropriate to solve this problem.

2. MATERIALS AND METHODS

In this research we use two interpolation and extrapolation methods to predict the gas content of hydrogen and iodine in the reaction of iodic acid formation.The two methods are Interpolation and Extrapolation of both Newton and Cubic Spline. The data is taken from the research conducted in [12]

2. 1. Iodic acid (HI)

Table 1. List of Chemical Compounds

Compound	Chemical formula	Bond length *d*(H−X) / pm (gas phase)	Model	Dipol µ / D	Solution phase (asam)
hydrogen fluoride (fluorana)	HF	H—F 91.7 pm		1.86	hydrofluoric acid
hydrogen chloride (klorana)	HCl	H—Cl 127.4 pm		1.11	hydrochloric acid
hydrogen bromide (bromana)	HBr	H—Br 141.4 pm		0.788	hydrobromic acid
hydrogen iodide (iodana)	HI	H—I 160.9 pm		0.382	iodic acid
hydrogen astatida astatine hydride (astatane)	HAt	H—At 172.0 pm		−0.06	astatida acid

HI is a colorless gas that reacts with oxygen to form water and iodine. HI can be used in organic synthesis and is better known as hydriodic acid. HI is very soluble in water. One liter of water can dissolve 425 liters of HI. HI is the Strongest Halide Acid. Acid Halide is the result of the reaction of hydrogen halide gas dissolved with water. Hydrogen halides are diatomic inorganic compounds with the general formula HX where X is one of the halogens: fluorine, chlorine, bromine, iodine, or astatine.

Commercial HI available in the market usually contains about 48% - 57% HI which has formed an azeotropic solution with water. This azeotropic solution is a solution that is very difficult to separate up to 100% because in certain compositions it has the same boiling point. HI can be synthesized from hydrogen gas and iodine with several benefits.

The benefits of HI include:

1. Converts alcohols to alkyl halides
2. Cleaving ether into alkyl halides and alcohols
3. Can be used as a reducing agent for producing drugs

The above explanation illustrates the role of hydrogen in the reaction of the formation of halide acid types, especially HI type acid halides is very important. This is shown by the results of successful research showing that the role of hydrogen is not as a catalyst but as a reagent [13].

2. 2. Polynomial Interpolation

Polynomials are commonly used as an approximation function in most numerical analysis problems because of their simple structure, so that they can be used effectively [1]. Interpolation is the process of finding and evaluating a function whose grahp goes through a given set of points. Polynomial interpolation is an interpolation technique by assuming known data patterns form polynomials of degree less than or equal to n [1]. Polynomial interpolation is the job of interpolating points using curves whose representations are polynomials [2]. Interpolation with this method is done by first forming a polynomial equation. Types of polynomial interpolation include Lagrange, Newton and Cubic Spline polynomials. Lagrange polynomials are less preferred in practice for the reasons, such as a lower degree Lagrange interpolation is not included in the formulation of higher one [7]. Hence it is computaionaly expensive compared to Newton polynomial interpolation.

2. 3. Newton's Polynomial Interpolation.

Newton's interpolation is applied to get the polynomial function P(x) which has a certain degree across a number of data points. For polynomials of degree n, given $n+1$ distinct points, that is $x_0, x_1, x_2, ..., x_n$. Newton's polynomial equation is [1]:

$$P(x) = a_0 + a_1 x + a_2 x^2 + ... + a_{n-1} x^{n-1} + a_n x^n \qquad (1)$$

In Newton's polynomials, the preconceived polynomials can be used to create polynomials of higher degree. Since Newton's polynomials are formed by adding a single term with a polynomial of lower degrees, this makes it easier to calculate polynomials of higher degrees in the same program. For this reason, Newton's polynomials are often used especially in cases where the degree of the polynomial is not known in advance. In addition, it can be used to determine whether adding the degrees will increase or decrease the accuracy of the interpolation value [7].

2. 3. 1. Newton's Interpolation Calculation Formulas

Given the set $(n+1)$ of pairs of points x and y, that is $(x_0, y_0), (x_1, y_1), ..., (x_n, y_n)$, or (x_i, y_i), where $i = 0,1, \dots, n$, and x_i are different points. From this value it will be searched $p_n(x)$ that is, a polynomial of degree n [1]. For example, suppose we want to obtain a polynomial function of degree one passing through two points, namely (x_0, y_0) and (x_1, y_1). Newton's first degree polynomial has a form:

$$p_1(x) = f(x_0) + (x - x_0) f[x_1, x_0] \tag{2}$$

The general form of Newton's polynomials of degree n has a form [10]:

$$P_n(x) = \sum_{i=0}^{n} f[x_0, x_1, \dots, x_i] \prod_{j=0}^{i-1}(x - x_j) \tag{3}$$

The divided difference table on Newton's polynomials can be used repeatedly with different starting point values to estimate the value of the function at the interpolated value. Analytically, the approximation result will be more accurate when using a polynomial of higher degree, however, it contains / involves rounding / truncation errors so that the result is not always better. In addition polynomial interpolation of degree more than 5 generally results in significant errors near the ends of the interval, as well as producing high oscilation rate. A part from this issue we increase the degree of Newton's polynomial interpolation to reach a certain accuracy, starting from p_0 and ends at p_k where the error at iteration k meets the required accuracy.

2. 3. 2. Constructing Newton's Polynomial Functions

The steps to take:

1. Plot the points obtained
   At this stage, a curve matching from points in a Cartesian coordinate is carried out over a finite set of pairs of points x and y, namely $(x_0, y_0), (x_1, y_1), ..., (x_n, y_n)$.
   In this study, we took 11 data points to be matched with the polynomial function.
2. Determine the general equation of the interpolated polynomial according to the number of data points obtained. In general $(n + 1)$ data points can be matched to a polynomial of degree n, which has the general form:

   $$P_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \cdots + a_n(x - x_0)(x - x_2) \dots (x - x_{n-1}) \tag{4}$$

   In this study, with 11 points we constructed the polynomial function of degree $n = 10$.
   Determine the coefficients of the polynomial by calculating the divided differences over the data points. The equations used to calculate the coefficients are:

   $a_0 = f(x_0) = y_0$

$$a_1 = \frac{y_1-y_0}{x_1-x_0} = f[x_1, x_0]$$
$$a_2 = \frac{\frac{y_2-y_1}{x_2-x_1}-\frac{y_1-y_0}{x_1-x_0}}{x_2-x_0} = \frac{f[x_2,x_1]-f[x_1,x_0]}{x_2-x_0} = f[x_2, x_1, x_0] \quad (5)$$
$$\vdots$$
$$a_n = \frac{f[x_n, x_{n-1}, .., x_2, x_1] - f[x_{n-1}, x_{n-2}, \dots, x_1, x_0]}{x_n - x_0} = f[x_n, x_{n-1}, \dots x_1, x_0]$$

The value of the brackets function is called finite divide difference and is defined as $\frac{y_i-y_j}{x_i-x_j} = f[x_i, x_j]$ (the first order divided difference of *f*(*x*))

$\frac{y[x_i,x_j]-y[x_j,x_k]}{x_i-x_k} = f[x_i, x_j, x_k]$ (the second order divided difference of *f*(*x*))

$\vdots$

$\frac{f[x_n,x_{n-1},\dots,x_1]-f[x_{n-1},\dots,x_1,x_0]}{x_n-x_0} = f[x_n, x_{n-1}, \dots, x_1, x_0]$(the-*n* order divided difference of *f*(*x*))

3. Enter the coefficients obtained into the form of the polynomial equation.
4. Determines the value of x to be sought.
5. Draw the curves obtained.

In addition to using Newton's interpolation, we will also use Cubic Spline interpolation, because the former shows the high oscillation rate when the degree is high.

2. 4. Cubik Spline Interpolation

The regression spline is special chunks of polynomial function defined in parts, which is widely used in interpolation problems requiring smoothing. In particular, for certain partitions $a = t_0 < t_1 < t_2 < \cdots < t_n = b$ of the interval [a, b], spline is a multi-polynomial function S (t) defined by [11]:

$$S(t) = S_1([a, t_1]) \cup S_2([t_1, t_2]) \cup \dots \cup S_{k-1}([t_{k-2}, t_{k-1}]) \cup S_k([t_{k-1}, b]) = \bigcup_{i=1}^{k} S_i([t_{i-1}, t_i]) \quad (6)$$

where k is the number of vertices $t = t_0, t_1, t_2, \dots, t_{k-1}$ which divides the interval [a, b] into $k-1$ subintervals. Each function $S_i(t)$, is a low degree (usually square) polynomial (sometimes it can be linear) corresponding to the appropriate interval $[t_{i-1}, t_i]$, where $i = 1,2,3, \dots, n$ hence, the spline is a continuous and smooth function.

A function $S(x)$ is called a cubic natural spline of degree k if

1. $S(x)$ is a polynomial function of degree $k \le 3$ for each subinterval $[x_{j-1}, x_j]$ $j = 2,3, \dots, n$
2. $S(x), S'(x), and\ S''(x)$ continuous on $[a, b]$.
3. $S''(x_1) = S''(x_n) = 0$

A function $f(x)$ delimited by the interval a and b, and has a number of data points $a = x_0 < x_1 < x_2 < \cdots < x_n = b$. The cubic spline interpolation $S(x)$ is a slice of a third degree (cubic) polynomial function that connects two adjacent points with the condition,

1. $S(x)$ is cubic at each interval $[x_{j-1}, x_j]$, hence its functions $S''(x)$ is linear on the interval. Linear functions are determined by their value at two points, using

$$S''(x_{j-1}) = M_{j-1},\ S''(x_j) = M_j \tag{7}$$

$$S''(x) = \frac{(x_j-x)M_{j-1}+(x-x_{j-1})M_j}{(x_j-x_{j-1})} \qquad x_{j-1} \leq x \leq x_j \tag{8}$$

2. Functions $S(x)$ is obtained by integrating $S''(x)$ twice, on the condition of interpolation $S(x_{j-1}) = y_{j-1}$, $S(x_j) = y_j$, and $\int_a^b [S''(x)]^2\, dx$ should be as small as possible.
3. To ensure $S'(x)$ is continuous at $[a, b]$, then $S'(x)$ must have the same value at the point of intersection between each interval $[x_{j-1}, x_j]$ and $[x_j, x_{j+1}]$, that is $x = x_j$ for $j = 2,3, \ldots, n-1$.
4. By applying all the terms and assumptions $S''(x_1) = S''(x_n) = 0$, the linear equation is obtained:

$$\frac{x_j-x_{j-1}}{6} M_{j-1} + \frac{x_{j+1}-x_{j-1}}{3} M_j + \frac{(x_{j+1}-x_j)}{6} M_{j+1} = \frac{y_{j+1}-y_j}{x_{j+1}-x_j} - \frac{y_j-y_{j-1}}{x_j-x_{j-1}} \tag{9}$$

By solving all equations (9), obtained M_j for $j = 1,2, \ldots, n$

with, $M_1 = M_n = 0$

The linear system of equations (9) is called the tridiagonal system.Then the value of M_j is substituted in equation (10) so that the function $S(x)$ is obtained.

5. After some manipulation, a cubic polynomial is produced:

$$S(x) = \frac{(x_j-x)^3 M_{j-1}+(x-x_{j-1})^3 M_j}{6(x_j-x_{j-1})} + \frac{(x_j-x)y_{j-1}+(x-x_{j-1})y_j}{(x_j-x_{j-1})} - \frac{1}{6}(x_j - x_{j-1})\left[(x_j - x)M_{j-1} + (x - x_{j-1})M_j\right] \tag{10}$$

For interval $x_{j-1} \leq x \leq x_j$. Equation (10) is applied to all intervals $[x_1, x_2], \ldots, [x_{n-1}, x_n]$, hence the function $S(x)$ is a continuous function.

2. 4. 1. Building a Cubic Spline Function

The steps taken in forming a cubic spline function by taking n data points:

1. Plot the points from the data.
Determine the number of splines that pass through these points according to the number of points, that is, form the general equation for interpolation of cubic sline according to the number of points

$$y = f(x) = S_j(x) = a_j + b_j(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3 \tag{11}$$

The cubic spine to be formed exists $(n-1)$ spline taken from n the data points to be matched. Equation (11) leads to 4 unknown values, namely $a_j, b_j. c_j, dan\ d_j$ in each equation (11). Therefore there are $4(n-1)$ unknown values.
Or in other words, it will form an equation (10).

2. Determine the conditions formed from $4(n-1)$ unknown parameter.
   a. $S_j(x_j) = y_j \quad j = 1,2,3,\dots,n-1$

$$S_j(x_{j+1}) = y_{j+1} \quad j = 1,2,\dots,n-1,$$
$$S_j{}'(x_{j+1}) = (S_{j+1})'(x_j) \quad j = 1,2,3,\dots,n-2 \tag{12}$$
$$S_j{}''(x_{j+1}) = (S_{j+1})''(x_j) \quad j = 1,2,3,\dots,n-2$$
$$S''(x_1) = S''(x_n) = 0 \ ,$$
$$M_j \text{ for } j = 1,2,\dots,n \quad , M_1 = M_n = 0$$

   b. Solve the triagonal system (9), by: forming a matrix in accordance with the equation obtained, after the matrix is formed, solve the linear equation system formed using the Gauss method.

$$A.C = V \leftrightarrow \begin{bmatrix} m_1 & u_1 & 0 & 0 & 0 \\ l_1 & m_2 & u_2 & 0 & 0 \\ 0 & l_2 & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & u_{n-1} \\ 0 & 0 & 0 & l_{n-1} & m_n \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ \vdots \\ v_n \end{bmatrix} \tag{13}$$

where: A = tridiagonal matrix with the order $3x3$
C = coefficient matrix with order $3x1$
V = matrix constants with the order $3x1$

Solve $A.C = V$ for the coefficient c_j

There are 3 groups of formulas:

Group 1:

$$u_j = l_j = x_{j+1} - x_j \ , \quad j = 1,2,\dots,n-1 \tag{14}$$

Group 2: The formula for the main diagonal

$$m_j = 2\,(x_j - x_{j-1} + x_{j+1} - x_j), \qquad j = 2,3,\dots,n \tag{15}$$

Group 3: The formula for the coefficient matrix

$$v_j = 6\,\left(\frac{y_{j+1}-y_j}{x_{j+1}-x_j} + \frac{y_j-y_{j-1}}{x_j-x_{j-1}}\right) \qquad j = 2,3,\dots,n \tag{16}$$

Find the coefficient b_j

$$b_j = \frac{y_{j+1}-y_j}{x_{j+1}-x_j} - \frac{x_{j+1}-x_j}{3}(2c_j + c_{j+1}) \quad j = 1,2,3,\dots,n-1 \tag{17}$$

Find the coefficient d_j

$$d_j = \frac{c_{j+1}-c_j}{3(x_{j+1}-x_j)} \qquad j = 1,2,3,\dots,n-1 \tag{18}$$

3. Enter the coefficients obtained into the cubic spline equation. A cubic spline is obtained through which the data points are given.
4. Determines the value of x to be sought. To extrapolate with observation points outside the known range of points, use the polynomial equation that is at the end closest to the x value for which the y value will be sought.

2. 5. Exstrapolation

Extrapolation is the estimate of prices outside the limits of the observed data. Ekstrapolation is estimating the attribute values of locations outside the range of available data using known data values [13]. The equations used to determine the function of numerical data using interpolation are the same as those of extrapolation. Linear Ekstrapolation means creating a tangent line at the end of the known data and extending it beyond that limit, Linear extrapolation will provide good results only when used to extend the graph of an approximately linear function or not too far beyond the known data. If the two data points nearest to the point x^* to be extrapolated are (x_k, y_k) and (x_{k-1}, y_{k-1}), linear extrapolation gives the function [13]

$$f(x^*) = y_{k-1} + \frac{x^*-x_{k-1}}{x_k-x_{k-1}}(y_k - y_{k-1}) \tag{19}$$

with $f(x^*)$ is the value to be sought, and x^* is an independent variable, y_k, y_{k-1} , x_k and x_{k-1} represent numerical data.

The error on linear interpolation is:

$$eT = (\frac{f''(c)}{2})((x^* - x_k)(x^* - x_{k-1});\ x_{k-1} \leq c \leq x_k \tag{20}$$

A polynomial curve can be created through the entire known data or just near the end. The resulting curve can then be extended beyond the end of the known data. Polynomial extrapolation is typically done by means of Lagrange interpolation or using Newton's method of finite differences to create a Newton series that fits the data. The resulting polynomial may be used to extrapolate the data [13].

2. 6. Research Flow

The research flow diagram can be seen in **Figure 1** below.

We use Maple and Matlab applications to implement the interpolation and extrapolation of both methods. The estimate result of hydrogen and iodine content is then compared with the actual content of each of these gases, by using the following error formula:

$$E = \frac{|x_i-x_j|}{x_j} \tag{21}$$

where: x_i : interpolation value

x_j : true value

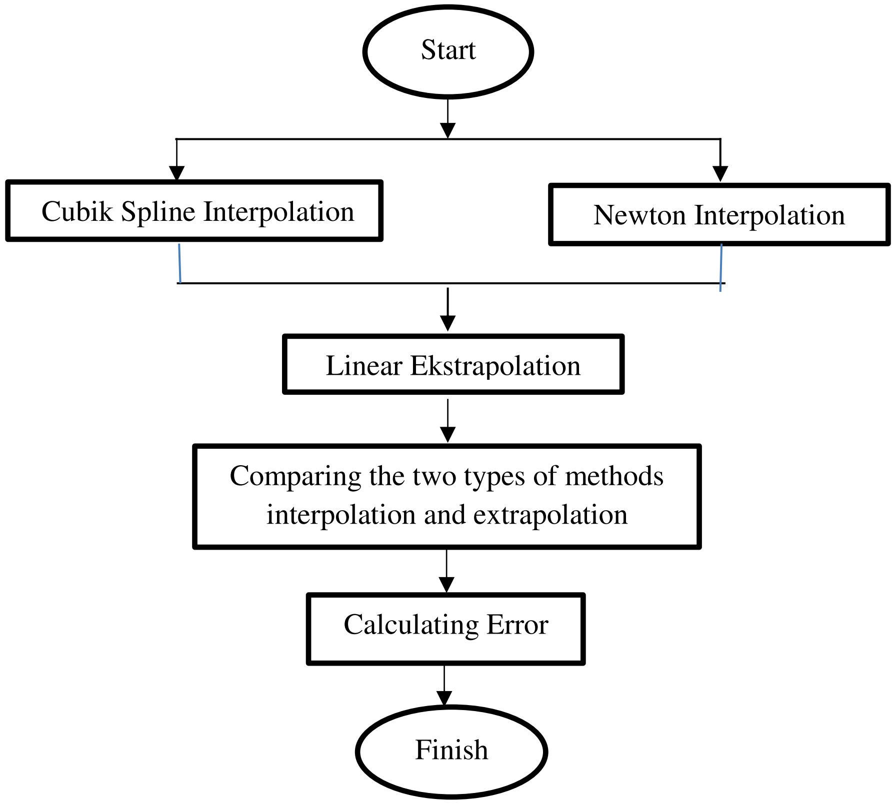


Figure 1. Research Flowchart

The absolute value will be obtained from the difference between the interpolated results and the actual price. Then the accuracy of the estimate is measured using the standard deviation. The formula for finding the standard deviation used is:

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{22}$$

where,

n : the amount of data

x_i : the amount of hydrogen and iodine order i

$\bar{x}$: the mean or average of the sample data

Statistically, the ideal standard deviation is a standard deviation that is close to unity. With a standard deviation of one, it means that the variance of the data is also one. With that, the goodness of the test model can be accounted for. Then to compare the accuracy between the two algorithms, the calculation of the standard deviation relative to the other algorithms is used. Another measure of fitability is the root mean square deviation or error (RMSD or RMSE), which calculates the square root of the expected difference between the predicted value and the observed value [14]. The calculation formula for Root Mean Squared Error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}[y_i - \hat{y}_i]^2}{n}} \tag{23}$$

where,

y_i : true value

$\hat{y}_i$: interpolation value

n : the amount of data

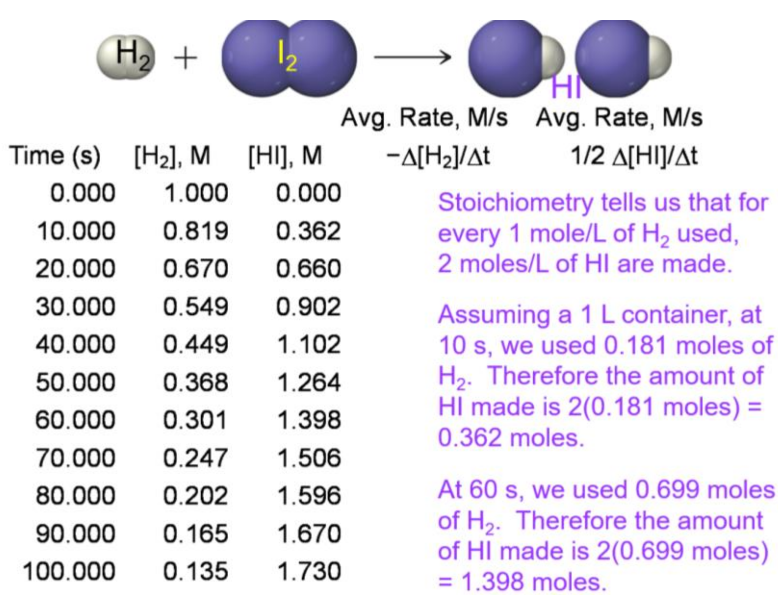
The final measure of the fitting ability used in the analysis is relative absolute error (RAE), which calculates the relative value of RMSE according to the expected observed value, as shown by the formula [11]:

$$RAE = \frac{\sqrt{\frac{\sum_{i=1}^{n}[y_i - \hat{y}_i]^2}{n}}}{\sqrt{\frac{\sum_{i=1}^{n}[y_i]^2}{n}}} \tag{24}$$

RAE is widely used in machine learning, data mining, and operations management applications, and it represents the RMSE analogy relative to the expected value of the observed value.

3. RESULT AND DISCUSSION

The data used are research data available in [12]



Time (s)	$[H_2]$, M	[HI], M	Avg. Rate, M/s $-\Delta[H_2]/\Delta t$	Avg. Rate, M/s 1/2 $\Delta[HI]/\Delta t$
0.000	1.000	0.000		
10.000	0.819	0.362		
20.000	0.670	0.660		
30.000	0.549	0.902		
40.000	0.449	1.102		
50.000	0.368	1.264		
60.000	0.301	1.398		
70.000	0.247	1.506		
80.000	0.202	1.596		
90.000	0.165	1.670		
100.000	0.135	1.730		

Figure 2. Research data

Plotting the data will produce the following graph:

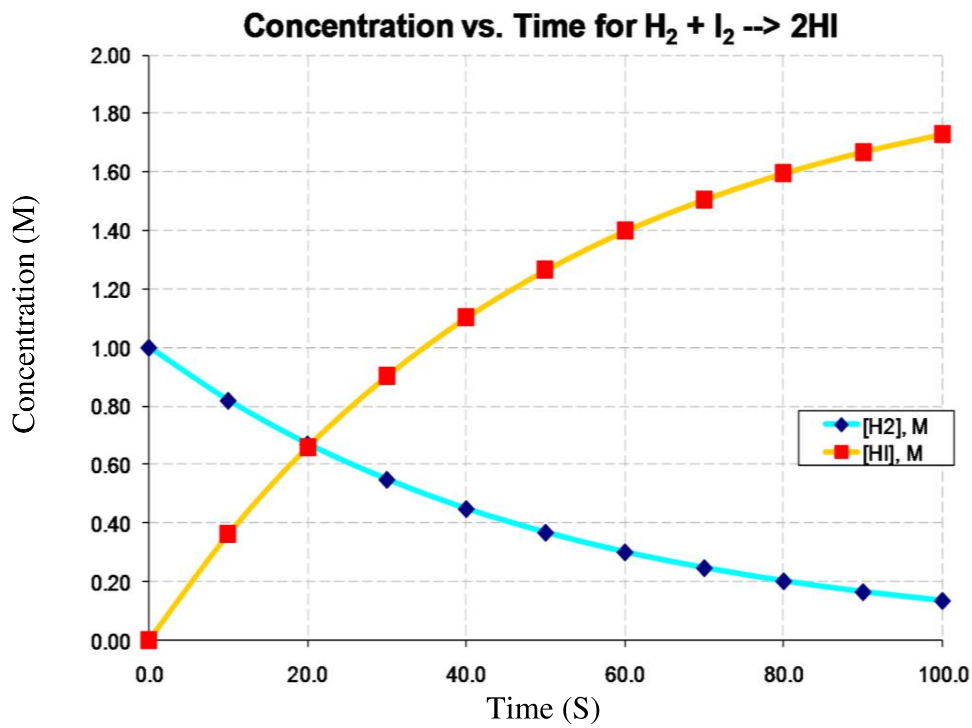


Figure 3. Research Data Graph

Once plotted, it can be assumed that the data spread H_2 has the same pattern as the function $f(x) = e^{-0,02x}$. From content data H_2 , the function will be searched numerically by interpolating newtons.

To form a polynomial function from 11 data points, search using the Maple application. The polynomials produced by the Newton interpolation method from the 11 data are:

$$P_{10}(x) = 0,99999 - 0,01463x - 0,00121x^2 + 0,00014x^3 - 0,00001x^4 + 2,79715.10^{-7}x^5 - 5,93398.10^{-9}x^6 + 7,97875.10^{-11}x^7 - 6.59804.10^{-13}x^8 + 3,06024.10^{-15}x^9 - 6,090167.10^{-18}x^{10} \quad (25)$$

with the following graph,

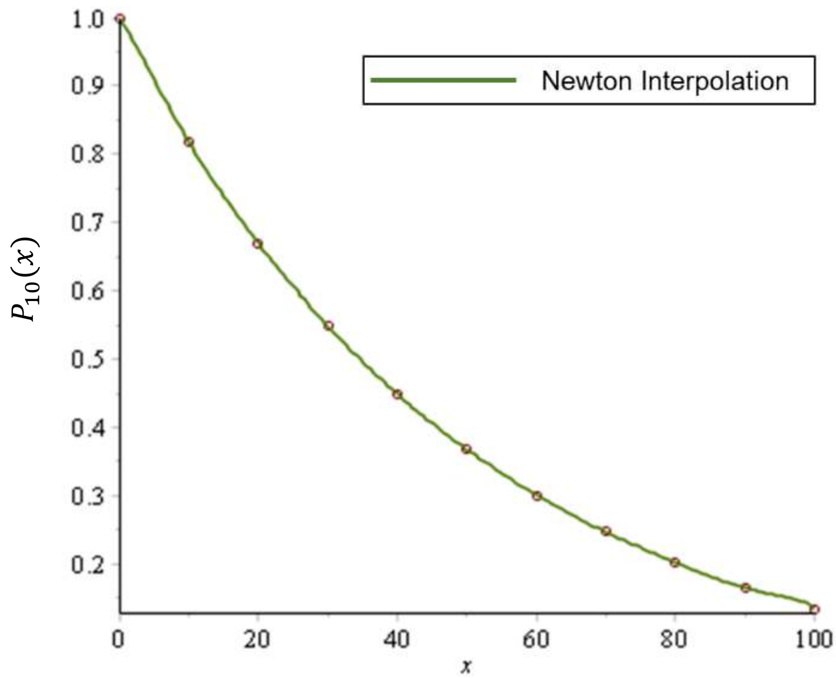


Figure 4. Graph of Newton's Interpolation

The use of polynomials with a very high degree does not always give a more accurate approximation. This is because the higher the degree of the polynomial used will result in more calculations so that rounding errors will significantly affect the result.

So, in addition to newton interpolation, a functional approximation of the data will be sought using the Spline cubic method, using the Maple application the following results will be obtained,

$$g(x) = \begin{cases} 1 - 0{,}018788x + 0{,}000006x^3, untuk\ x \in [0,10] \\ 1{,}009312 - 0{,}021582x + 0.000279x^2 - 0{,}00000242x^3, untuk\ x \in [10,20] \\ 0{,}999320 - 0{,}020083x + 0{,}000204x^2 - 0{,}0000011x^3, untuk\ x \in [20,30] \\ 0{,}963775 - 0{,}016528x + 0{,}000085x^2 + 1{,}38564.10^{-7}x^3, untuk\ x \in [30,40] \\ 1.060728 - 0{,}023800x + 0{,}000267x^2 - 0.00000137x^3, untuk\ x \in [40,50] \\ 0.84284211 - 0.01072725x + 0.0000062x^2 + 3.667622.10^{-7}, untuk\ x \in [50,30] \\ 1.1576577 - 0.0264680x + 0.0002686x^2 - 0.00000109x^3, untuk\ x \in [60,70] \\ 0.7848767 - 0.0104917x + 0.0000403x^2 - 3.8925.10^{-9}x^3, untuk\ x \in [70,80] \\ 0.7284646 - 0.0083762x + 0.0000139x^2 + 1.062875.10^{-7}x^3, untuk\ x \in [80,90] \\ 1.842044 - 0.0454955x + 0.0004263x^2 - 0.00000142x^3, untuk\ x \in [90,100] \end{cases} \tag{26}$$

The graph of the results of cubic spline interpolation is presented in Figure 5.

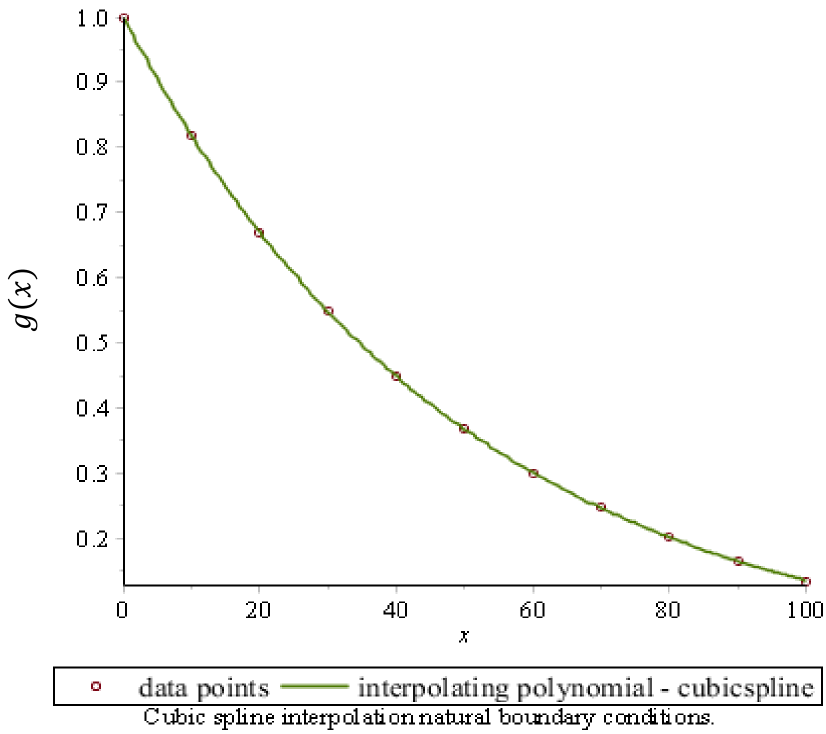


Figure 5. Graph of Spline Cubic Interpolation

If we combine the graph of the interpolated Newton (purple) and Cubic Spline (red) results, the following results are obtained,

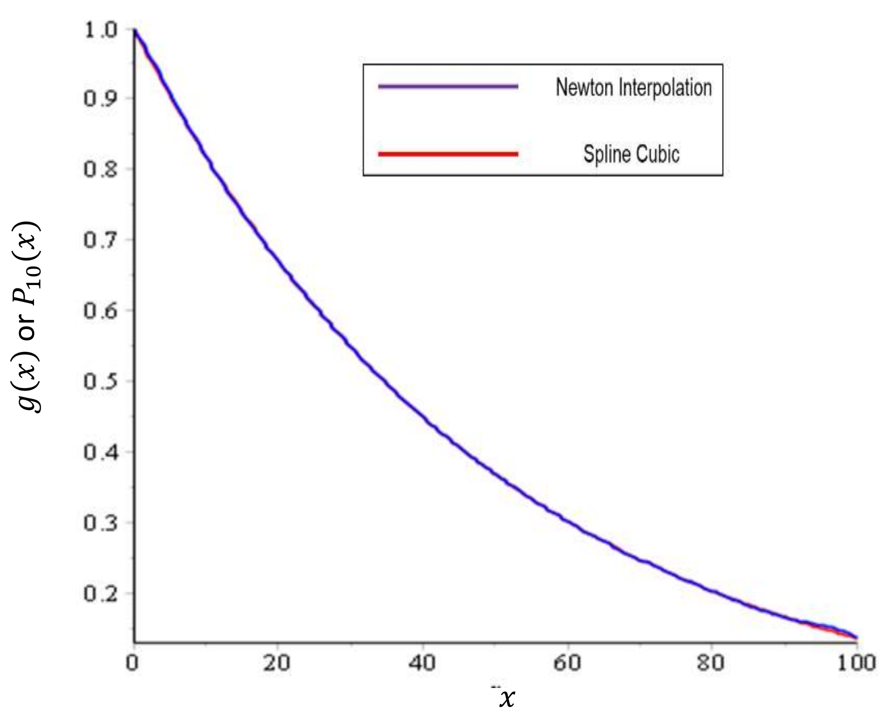


Figure 6. Combined Graph

The graph of the results of the two methods is not too much different, except at the end of the interval, it can be seen that Newton's interpolation has more angles, or in other words, the polynomial function of cubic spline interpolation is smoother than that of the newton interpolation polynomial function. This is in accordance with the theory, that Newton's interpolation will have difficulty when applied to higher degrees, compared to the cubic spline which is more stable since having the maximum degree of 3. After getting the function for each method, its values as estimates at several points will be calculated, their errors are calculated from the actual value of the function $f(x) = e^{-0,02x}$, as well as their RMSE values. The results of the calculations can be seen in Table 2 and Table 3.

Table 2. Calculation of Relative Error and Error.

No	xi	Original value	Newton's estimate	Cubic Spline Estimation	|x-xe|N	εN	|x-xe|SK	εSK
1	5	0.904837418	0.910514106	0.906917985	0.005676688	0.00627371	0.002080567	0.002299383
2	15	0.740818221	0.739544262	0.740246044	0.001273959	0.001719665	0.000572177	0.000772358
3	25	0.60653066	0.606914713	0.606597839	0.000384053	0.000633197	6.71794E-05	0.00011076
4	35	0.496585304	0.496332732	0.4964876	0.000252572	0.000508617	9.77042E-05	0.000196752
5	45	0.40656966	0.40648955	0.406451762	8.01097E-05	0.000197038	0.000117897	0.000289981
Σ					**0.007667382**	**0.001866**	**0.002935525**	**0.000796**

Table 3. Calculation MSE, RSME and RAE.

No	xi	Original value	Newton's estimate	Cubic Spline Estimation	|x-xe|N	$(\|x-xe\|N)^2$	|x-xe|SK	$(\|x-xe\|SK)^2$
1	5	0.904837418	0.910514106	0.906917985	0.005676688	3.22248E-05	0.002080567	4.32876E-06
2	15	0.740818221	0.739544262	0.740246044	0.001273959	1.62297E-06	0.000572177	3.27386E-07

3	25	0.60653066	0.606914713	0.606597839	0.000384053	1.47497E-07	6.71794E-05	4.51307E-09
4	35	0.496585304	0.496332732	0.4964876	0.000252572	6.37925E-08	9.77042E-05	9.54611E-09
5	45	0.40656966	0.40648955	0.406451762	8.01097E-05	6.41757E-09	0.000117897	1.38998E-08
Σ						3.40655E-05		4.68411E-06
MSE						**6.81309E-06**		**9.36821E-07**
RMSE						**0.00261019**		**0.000967895**
RAE						**0.003982918**		**0.001485475**

The standard deviation for Newtonian interpolation is 0.199390102 and that for Cubic Spline interpolation is 0.198253376. While the relative standard deviation for Newton's interpolation is 1.0093855817 and that for Cubic Spline interpolation is 1.003631294.

Table 4. Comparison of Newton's Interpolation and Cubic Spline.

No	Interpolation Type	Absolute Error	Relative Error	RMSE	RAE	Standard Deviation	Relative Standard Deviation
1	Newton	0.007667382	0.001866	0.00261019	0.003982918	0.199390102	1.0093855817
2	Cubic Spline	0.002935525	0.000796	0.000967895	0.001485475	0.198253376	1.003631294

We can notice from Table 4, that cubic spline is better than Newton method in term of all measurement used. However, we still aim to obserbe the comparison of both methods in terms of extrapolation.

3. 1. Linear Esktrapolation

The results of each interpolation will be used separately for linear extrapolation, that is used to predict data outside the range of given set of data. A linear extrapolation of Newton method constructed by using the data (87.5, 0.17265184) and (97.5, 0.14732079) gives its value calculated at $x = 102.5$ of 0.134655265 with the absolute error value of 0.00592036. Whereas that of cubic spline constructed by using the data (87.5, 0.173173296) and (97.5, 0.142606813) gives its value calculated at x = 102.5 of 0.127323571 with the smaller absolute error value of 0.001411333.

4. CONCLUSIONS

In principle, the use of Newton and cubic spline interpolation methods is very good in finding the function of the Hydrogen data distribution in the acid formation reaction of Iodida. Because the distribution of data has the same pattern as the function $f(x) = e^{-0,02x}$, then it is assumed that the data has the same distribution as the function $f(x) = e^{-0,02x}$. Hence, for Newton's method, calculated at some points (see Table 2), this gives all results of measurements that are less than those of cubic spline.

It can be concluded that the Cubile Spline Interpolation method is more appropriate to be used as an estimate of. Similarly, from the estimation results of Newton and Cubic Spline, in terms of extrapolation. A better value of linear extrapolation calculated at $x = 102.5$ is given when using the estimation results of cubic spline than Newton.

From the 11 data, a polynomial interpolation of degree 10 is supposed to be obtained from the Newton method. However, we found that the data used has a pattern of function with the highest degree of 2. Hence, even though we use all the 11 data to construct a polinomial interpolation, the resulted interpolation will have a degree at most two.This is why, although the data taken to construct the interpolation function is up to 11 data points, Newton's polynomials are not highly oscillating.

References

[1] B. Das and D. Chakrabarty. Newton's forward interpolation: representation of numerical data by a polynomial curve. *Int. J. Math. Trends Technol* 34(2) (2000) 64-72

[2] J. Scelten, F. Hossfeld. Application of spline functions to the correction of resolution errors in small-angle scattering. *J. Appl. Cryst* 4(3) (1971) 210-223

[3] C. J. Goulding. Cubic spline curve and calculation of volume of sectionally measured trees. *J. of Forest Research* 9(1) (1977) 89-99

[4] A. James, Ligget, R James, Salmon. Cubic spline boundary elements. *Int. J. for Numerical Methods in Eng.* 17(4) (1981) 5-19

[5] G. Behforooz. End conditions for cubic spline interpolation derieved from integration. *Applied Math. And Computation* 29(3) (1989) 231-244

[6] A. Fischer. A Newton-type method for positive-semidefinite linear complementarity problems. *J. of Opt. Theory and Applications* 86 (1995) 585-608

[7] J. Hoschek, U. Schwanecke, Interpolation and approximation with ruled surfaces. *J. Information Geometers* 8 (1998) 213-231

[8] M. Gasca, T. Sauer. Polynomial interpolation in several variables. *Advances in Computational Mathematics* 12(4) (2000) 377-410

[9] L. Asen, Dontchev, Hou-Dou Qi, L. Qi, H. Yin. A Newton Method for Shape-Preserving Spline Interpolation. *SIAM Journal on Optimization* 13(2) (2002) 588-602

[10] M. I. Syam, Cubic spline interpolation predictors over implicitly defined curves. *J. of Computational and Applied Math.* 157(2) (2003) 283-295

[11] R. Goonatilake, V. Ruiz. On polynomial interpolation approximations: umerical vs. statistical techniques. *J. of Modern Methods in Numerical Math.* 5(1) (2014) 17-27

[12] Nivaldo J. Tro. Chemistry: A Molecular Approach. Pearson; 2nd edition (January 15, 2010). ISBN-13: 978-0321651785

[13] A. D. Rosalia, P. Patiha, and E. Heraldy. An empirical study on the hydrogen peroxide reaction with iodide in acid condition. *ALCHEMY J. Penelit. Kim* 11(1) (2015) 67-72

[14] A. P. de Camargo. On the numerical stability of newton's formula for lagrange interpolation. *J. Comput. Appl. Math* (365) (2020) 23-69

[15] K. Demertzis, D. Tsiotas, and L. Magafas, Modeling and forecasting the Covid-19 temporal spread in Greece: An exploratory approach based on complex network defined splines. *Int. J. Environ. Res. Public Health* 17(13) (2020) 1-18