Computer Networks and Communications
https://ojs.wiserpub.com/index.php/CNC

UNIVERSAL WISER PUBLISHER

Research Article

Exploring Signal Interpolation: A Comparative Study of Convolution, Regression, and SVM Methods

Remus Stanca[1], Annamaria Sârbu[2*]

[1]"Lucian Blaga" University of Sibiu, Sibiu, Romania
[2]Department of Communications, IT and Cyber Defense, "Nicolae Bălcescu" Land Forces Academy, Sibiu, Romania
E-mail: paljanosanna@yahoo.com

Received: 16 January 2025; **Revised:** 18 March 2025; **Accepted:** 21 March 2025

Abstract: Signal interpolation plays a critical role in various signal processing applications, including wireless communications, image processing, and radar systems. Accurately reconstructing signals from decimated samples is essential for maintaining data integrity and improving transmission efficiency. To this extent, this paper presents a comparative study of five signal interpolation methods: convolution with three types of deterministic signals (triangular, rectangular and sinc signal), statistical linear regression and support vector machine (SVM). All these methods were applied on a sinusoidal signal corrupted by noise at different signal to noise ratio (SNR) values and on a QPSK (Quadrature Phase Shift Keying) modulated signal with 25 different decimation factors. The comparison between the investigated methods was made based on the inter-correlation coefficient, Euclidean distance and determinism for sinusoidal signal corrupted by noise. Two additional parameters, namely Euclidean letter distance and Bit Error Rate (BER), were defined and used for the QPSK modulated signal. Our findings suggest that for the sinusoidal signal corrupted by noise convolution with sinc function outperforms the other methods in terms of Euclidian distance in at least 98.57% of the cases and at least 95.71% of the cases in terms of inter-correlation coefficient. In the case of QPSK modulated signal it is the SVM method which surpasses all the other methods in terms of intercorrelation coefficient and Euclidean distance, in 80% and 88% of the cases respectively. If the Euclidean letter distance and the Bit Error Rate are considered for comparison, in the case of the QPSK modulated signal, convolution with sinc function was found to outperform the other investigated methods for at least 80% and 60% of the decimation factors respectively.

Keywords: interpolation, convolution, regression, support vector machine

1. Introduction

Interpolation is the process of estimating values within a range based on known data points. It involves filling in the gaps or missing values between the given data points to create a continuous representation of the data [1]. Interpolation methods are used in various fields involving different types of data from economic [2], financial [3], chemical, physical, and environmental [4, 5] domains.

Interpolation methods can be classified into two categories: deterministic and stochastic methods, based on their underlying principles and approaches [1]. Deterministic interpolation methods rely solely on the given data points and do

DOI: https://doi.org/10.37256/cnc.3120256443

not account for any uncertainty or variability [6]. The choice of interpolation method depends on the characteristics of the data, the level of uncertainty, and the specific requirements of the application.

Interpolation finds extensive usage in digital signal processing applications [7], where its primary objective is to approximate intermediate sample values based on a given set of existing samples. Interpolation enables the reconstruction of continuous signals from discrete samples, facilitates data recovery in the presence of missing or lost information, and improves the quality and accuracy of digital signals in various applications [8].

In wireless communication systems, interpolation is essential for digital signal processing, including sample rate conversion for signal synchronization [9, 10]. With their recent development, software-defined-radio platforms represent an agile system capable of conveniently implementing sample rate conversion, an important functionality to provide synchronization between the analog to digital converter and baseband signal processor [9]. The sampling rate converter finds utility in numerous communication and signal processing scenarios, enabling the interconnection of two signals or systems with different sampling rates [10].

Interpolation can be applied to estimate sensor readings at non-sampled time instances, providing a continuous and more accurate representation of the data [11]. The increase in the number of signal samples is essential for analyzing signals accurately, especially when dealing with high-frequency components, fast-changing dynamics, or precise frequency and time-domain analysis. Having more signal samples offers flexibility in signal processing operations such as filtering, modulation, or demodulation.

Signal interpolation finds applications in various fields involving incomplete measurement data sets. In [12] authors consider the problem of signal interpolation on graphs, by recovering one or multiple graph signal values from incomplete measurements. Research is carried out for developing efficient interpolation methods for wireless communications and signal processing applications [13, 14]. The method presented in [13] is based on cubic spline interpolation and improves the frequency-domain properties significantly while maintaining low complexity. The authors of [14] propose an interpolation method for wideband signal reconstruction, demonstrating its applicability for real-time blade damage diagnosis.

The authors of [8] propose a mixed signal interpolation method based on sinc and linear interpolation to improve the accuracy of trigger resampling. Their proposed method achieves better interpolation results with less resource consumption.

Comparative studies of interpolation methods are also available in public scientific literature. In [15], authors present a comparative study of the least-squares, linear nearest neighbor method and piecewise cubic Hermite polynomials interpolation methods applied on wideband, mm-wave communication signals. Their findings suggest that the choice of the interpolation method plays a role in defining the accuracy of the interpolated signal only for large time stamps, corresponding to a lower initial sampling frequency. In their study the least-square linear fit method applied on a 44 GHz wideband signal was found to achieve the best outcome in terms of error-vector magnitude minimization. In [16] bicubic interpolation is used to enhance satellite image resolution. This approach has been shown to outperform existing methods, providing superior resolution enhancement applicable to various types of images. Several studies investigating the performances of different interpolation methods for obtaining rainfall/precipitation patterns [17, 18]. The authors of [18] compare six interpolation methods in predicting the spatial distribution pattern of precipitation in a geographical area, proving that the Kernel interpolation with barrier method has the highest accuracy.

From the published literature we observed that different interpolation methods work well with various types of signals. The use of a customized interpolation technique tailored to different signal types offers significant advantages in signal processing applications [19]. Signals can vary in their characteristics, such as frequency content, dynamics, and noise characteristics. By selecting a specific interpolation technique, it becomes possible to exploit the unique properties of each signal and enhance the accuracy of the interpolation process [20].

The objective of this study is to expand upon previously published research by conducting a comparative analysis of five signal interpolation methods-convolution with three types of deterministic signals (triangular, rectangular, and sinc), statistical linear regression, and SVM-applied to signals commonly encountered in wireless communication environments. The analysis begins with a sinusoidal signal corrupted by noise and extends to a more complex case, a QPSK-modulated signal. Using Python, the original signals are generated, decimated, and then reconstructed using the selected interpolation methods. The study evaluates the performance of each method based on five key metrics: inter-correlation coefficient, Euclidean distance, determinism (for the sinusoidal signal), Euclidean letter distance, and BER (for the QPSK-modulated signal). By assessing the effectiveness of these interpolation techniques under different signal conditions, the study aims to identify the most suitable method for preserving signal integrity in wireless communication applications.

The proposed study builds upon previous research by expanding the comparative evaluation of interpolation methods to signals commonly found in wireless communication environments. While earlier studies have primarily focused on interpolation techniques applied to wideband signals, satellite imagery, or precipitation patterns, this work extends the analysis to both a sinusoidal signal corrupted by noise and a QPSK modulated signal. In completing prior research that relies on conventional error metrics or frequency-domain properties, this study introduces original comparison indicators based on the time-domain representation of the signal. These indicators are important in applications where time-domain accuracy directly impacts system performance like network synchronization in 5G networks [21] or quantum control systems [22]. The introduction of Euclidean letter distance and BER as evaluation criteria for the QPSK modulated signal offers a novel perspective on how interpolation affects communication signal integrity. Moreover, the statistical assessment of interpolation performance across different SNR levels enables a deeper understanding of each method's robustness in noisy environments. This work, therefore, not only reinforces existing findings on interpolation performance but also provides new insights into method suitability for real-world applications, particularly in software-defined radio and high-precision digital signal processing.

The remainder of this paper is organized as follows: The materials and methods section describing the experimental framework for signal generation and processing (Section 2), the interpolation methods (Section 3) and the definition of comparison indicators (Section 4). The results and discussion are presented in Section 5 for both sinusoidal and QPSK modulated signals. The conclusions of the paper are synthetized in Section 6.

2. Experimental framework for signal generation and processing

The signal generation and processing (decimation and interpolation) were performed in Python, following the general framework presented in Figure 1.

Two types of signals were considered: sinusoidal signal corrupted by noise and a QPSK modulated signal. A sinusoidal signal with amplitude 1, initial phase 0, 1 kHz frequency and an initial sample rate of f_{s1} = 48,000 Hz with N = 960 samples was considered. The sinusoidal signal is corrupted with a noise signal with different amplitude to obtain 70 different values for signal to noise ratio (SNR). The obtained SNR values ranged between 14 and 25 dB to cover typical transmission conditions encountered in real-world scenarios. In practical wireless communication systems, signals often experience varying levels of noise due to channel impairments such as multipath fading, interference, and thermal noise. An SNR of 14 dB represents relatively challenging conditions with noticeable noise, which can impact demodulation and signal recovery. On the other hand, an SNR of 25 dB reflects a more favorable scenario where noise is lower, allowing for improved signal clarity and accurate data transmission. By covering this range, the study ensured that the interpolation methods were evaluated under conditions that realistically simulate different transmission environments, from moderate to good signal quality. The resulting signal, named $x_1[n]$ was then decimated by a factor of 10 resulting in a new sample frequency f_{s2} = 4,800 Hz and N_1 = 96 sample points.

For the QPSK signal, the modulating signal $i[n]$ was set as a binary signal composed of the ASCII code representation of a text string composed of 100 characters of "1". The QPSK carrier was a sinusoidal signal with frequency $f_c = 1{,}000$ Hz, amplitude 1 and a sample frequency $f_{s1} = 750{,}000$ Hz. The resulting QPSK signal was decimated by using 25 different values for the decimation factor representing the first 25 divisors for 750,000 Hz (1, 5, 10, 15, 20, 25, …, 100). All interpolation methods investigated were applied for each of the 25 signals which resulted after decimation.

The choice of decimation factors was made to systematically analyze the effects of signal down sampling on the interpolation performance of QPSK signals. Initially, experiments were conducted using a sinusoidal signal to establish a foundational framework for interpolation techniques before applying them to the more complex QPSK signal. This approach ensured that the interpolation methods were first tested in a controlled setting with a predictable signal structure. The decimation factors represent the degree of signal down sampling, where a factor of 1 implies no decimation, while increasing values (5, 10, 15, etc.) progressively reduce the number of available samples. Higher decimation factors introduce greater information loss, increasing the challenge for interpolation methods to accurately reconstruct the original signal. By varying the decimation factor, the study aimed to examine how different levels of down sampling impact the interpolation quality and performance, particularly for QPSK signals, which have distinct phase transitions that make accurate reconstruction crucial for signal integrity.

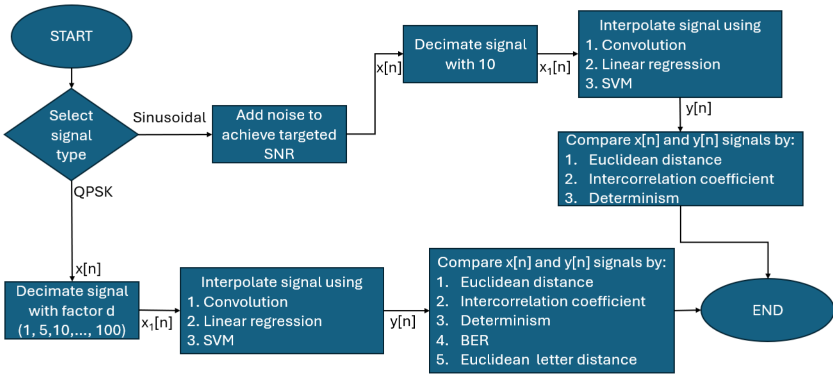

Figure 1. Experimental framework

Let $x[n]$ be the initial digital signal with sample frequency $f_{s1} = 4,800/750,000$ Hz, $x_1[n]$ the signal which results after decimating the initial signal with a factor of 10/d and $y[n]$ the signal which results after interpolating $x_1[n]$ back to the initial sample frequency.

3. Interpolation methods

In this paper we have compared five different interpolation methods: convolution with triangular, rectangular and sinc signal, statistical linear regression and SVM. Each method is mathematically described in the following subsection

3.1 *Convolution method*

Convolution method implies finding a general mathematical formula $x(t)$ for the signal by computing convolution between the digital signal and the time domain transfer characteristic for a low pass interpolation filter which can have a deterministic shape. Instead of the continuous variable t it can be written nT_s value, where T_s is the sampling period which will be considered to have a smaller value to induce an increased sample frequency.

Let $x[n]$ be a digital signal which comes from an analog signal $x_a(t)$ by using a sample frequency f_{s1}:

$$x[n] = x_a(n \cdot T_{s1}) = x_a\left(\frac{n}{f_{s1}}\right) \tag{1}$$

After decimating $x[n]$ by a decimation factor d, the sampling frequency will become $f_{s2} = \frac{f_{s1}}{d}$, and the sampling period will have the value $T_{s2} = \frac{d}{fs_1}$ which implies that the new decimated signal can be written in the following form:

$$x_1[n] = x_a\left(\frac{nd}{f_{s1}}\right) = x[nd] \tag{2}$$

The signal $x[n]$ will be deduced by performing convolution between the decimated signal $x_1[n]$ and the time domain transfer characteristic $h_{LP}(t)$ for a low pass interpolation filter [5]:

$$x_a(t) = \sum_{n=-\infty}^{\infty} x_1[n] \cdot h_{LP}(f_{s2}(t - nT_{s2})) \tag{3}$$

Instead of the continuous variable t will be used the value nT_{s1}, to obtain back the initial digital signal $x[n]$:

$$x[n] \cong \sum_{n=-\infty}^{\infty} x_1[n] \cdot h_{LP}(f_{s2}(nT_{s1} - nT_{s2})) \tag{4}$$

Because there is a finite number of sample points for the signal, equation (10) will be limited to a number of N sample points [23]:

$$x[n] \cong \sum_{n=0}^{N-1} x_1[n] \cdot h_{LP}(f_{s2}(nT_{s1} - nT_{s2})) \tag{5}$$

where $h_{LP}(t)$ is a deterministic signal which can have different time domain waveforms. In this article the following types of deterministic signals were used:

• Triangular signal defined as:

$$\mathrm{tri}(t) = (1 - |t|) \cdot (u_0(t+1) - u_0(t-1)) \tag{6}$$

where $u_0(t)$ is the unit-step function

• Sinc function:

$$\sin c(t) = \frac{\sin(\pi t)}{\pi t} \tag{7}$$

• Rectangular signal:

$$\mathrm{rect}(t - 0.5) = u_0(t) - u_0(t-1) \tag{8}$$

3.2 *Linear regression method*

Linear regression is based on finding a linear law which gives minimum values for differences between the actual values of the signal and the approximated values.

We consider a linear law $f(t) = at + b$ which approximates the relationship between the time values and the values of the signal samples. To obtain linear regression the following objective function must be minimized [24]:

$$F(a,b) = \sum_{i=0}^{M-1} (at_i + b - y_i)^2 \tag{9}$$

where y_i are the values for the signal samples, t_i is the time moment for which the signal has the value y_i and M is the number of samples on which it is made the linear approximation.

Let $x[n]$ be the initial signal which has a sample frequency f_{s1} and after decimation with a factor of d the signal becomes $y[n]$ with a new sample frequency $f_{s2} = \dfrac{f_{s1}}{d}$.

The signal $y[n]$ is split into sequences of 3 samples. On each sequence of three samples, an approximation in made by using a linear law $f(t)$, so that $M = 3$. The total number of obtained sequences is $\left[\dfrac{N}{3}\right] + C$, where N is the total number of samples for the signal $y[n]$.

$$C = \begin{cases} 0, \text{ for } N \bmod 3 = 0 \\ 1, \text{ for } N \bmod 3 \neq 0 \end{cases} \tag{10}$$

In $N \bmod 3 = 1$ then 2 values of “0” are added at the end of the signal and if $N \bmod 3 = 2$ then a single value of “0” is added to the end of the signal.

Let $S_i = \{y[3i], y[3i+1], y[3i+2]\}$ be the sequence number “i” with a number of 3 samples and let $T_i = \{t_{3i}, t_{3i+1}, t_{3i+2}\}$ be the time moments for which the signal takes values from the S_i sequence. For each S_i sequence a linear law $f(t) = at + b$ is found and this implies equating the first order partial derivatives of $F(a,b)$ with 0:

$$\begin{cases} \frac{\partial F}{\partial a} = 0 \\ \frac{\partial F}{\partial b} = 0 \end{cases} \tag{11}$$

$$\begin{cases} a\sum_{j=0}^{2} t_{3i+j}{}^2 + b\sum_{j=0}^{2} t_{3i+j} = \sum_{j=0}^{2} y[3i+j]t_{3i+j} \\ a\sum_{j=0}^{2} t_{3i+j} + 3b = \sum_{j=0}^{2} y[3i+j] \end{cases} \tag{12}$$

let $\alpha = \sum\limits_{j=0}^{2} t_{3i+j}{}^2$ and $\beta = \sum\limits_{j=0}^{2} t_{3i+j}$.

By using the notations presented above, the equations system from (12) becomes:

$$\begin{cases} a\alpha + b\beta = \sum_{j=0}^{2} y[3i+j]t_{3i+j} \\ a\beta + 3b = \sum_{j=0}^{2} y[3i+j] \end{cases} \tag{13}$$

By solving the system from (13) it can be found that:

$$a = \frac{3\sum_{j=0}^{2} y[3i+j]t_{3i+j} - \beta\sum_{j=0}^{2} y[3i+j]}{3\alpha - \beta^2} \text{ and } b = \frac{\alpha\sum_{j=0}^{2} y[3i+j] - \beta\sum_{j=0}^{2} y[3i+j]t_{3i+j}}{3\alpha - \beta^2} \tag{14}$$

In this way the sequence $S_i = \{y[3i], y[3i+1], y[3i+2]\}$ which has values that corresponds to the time moments $T_i = \{t_{3i}, t_{3i+1}, t_{3i+2}\}$ can be approximated with $f(t) = at + b$ by using the determined a and b values.

To determine the missing samples, the continuous time variable "t" is substituted by the value $\frac{n}{f_{s1}}$ and in this way the signal can be brought back to its initial sample frequency f_{s1}.

3.3 *SVM method*

SVM method consists in finding a law $f(t)$ which gives the minimum values between the approximated samples and the actual samples of the signal. $f(t)$ can be a non-linear law that uses a function $\Phi(t)$ which maps the one-dimensional vectors t into vectors with greater dimensions. To compute the dot product between the two vectors $\Phi(x)$ and $\Phi(y)$ from the new vector space, a "Kernel" function $K(x, y) = < \Phi(x) , \Phi(y) >$[25] can be used.

Starting with the MATLAB script presented in [26], the following mathematical algorithm was deduced:

The initial digital signal is denoted by $x[n], n = \overline{0, N-1}$, where N is the number of samples for this signal and let f_{s1} be the sample frequency of this signal. Let t_n be the time moments for which the signal has the values $x[n]$.

This signal is decimated by a factor of d with the sample frequency becoming $f_{s2} = \frac{f_{s1}}{d}$ and the obtained signal is denoted by $x_1[n], n = \overline{0, N_1 - 1}$, where N_1 is the number of samples obtained after decimation.

The following Kernel function $K(x, y) = e^{-\gamma(x-y)^2}$ is defined.

Let $A \in M_{0 \le i, j \le N_1 - 1}[R]$ be a matrix with the following elements:

$$a_{ij} = \begin{cases} 0, \text{ for } i = j = 0 \\ 1, \text{ for } ij = 0 \text{ and } i \neq j \\ e^{-\gamma(t1_{i-1} - t1_{j-1})^2} + z, \text{otherwise} \end{cases} \tag{15}$$

where

$$z = \begin{cases} \frac{1}{C}, \text{for } i = j, \text{but } i \neq 0 \text{ and } j \neq 0 \\ 0, \text{otherwise} \end{cases} \tag{16}$$

In this article C is 100 and thus z can be written as:

$$z = \begin{cases} \frac{1}{100}, \text{for } i = j, \text{but } i \neq 0 \text{ and } j \neq 0 \\ 0, \text{otherwise} \end{cases} \tag{17}$$

The matrix c is defined as $c = (0, x_1[0], \ldots, x_1[N_1 - 2])$ and is used to find $b = A^{-1}c^T$.

Using the elements from the matrix b it can be deduced an approximated form for the analog signal from which comes the digital signal $x[n]$:

$$x(t) = \sum_{j=1}^{N_1 - 1} b[j][0] \cdot e^{-\gamma(t - t1_{j-1})^2} + b[0][0] \tag{18}$$

The continuous time domain variable t is replaced by the discrete values with the form $\frac{n}{f_{s1}}$ to obtain back the signal $x[n]$ with the sample frequency f_{s1}:

$$x[n] \cong \sum_{j=1}^{N_1-1} [j][0] \cdot e^{-\gamma\left(\frac{n}{f_{s1}} - t1_{j-1}\right)^2} b + b[0][0], n = \overline{0, N-1}. \tag{19}$$

The SVM method was applied for each *SNR* value using a range of values for the parameter γ of the Kernel function (γ varied in the range 30,000,000-39,000,000 with a step of 80). The γ value that yielded the highest inter-correlation coefficient and the γ value that resulted in the lowest Euclidean distance were determined for each *SNR* value. Finally, the average of these two γ values was computed. Subsequently, each γ value obtained for every *SNR* value using this method was exported to an ".xlsx" document and further processed using the designed Python application.

In the case of the QPSK modulated signal the inter-correlation coefficient was computed between the SVM interpolated signal and the original digital signal. The method was applied considering for γ parameter values ranging from 900,000 to 900,000,000,000 in a geometric progression with a common ratio of 10. For each decimation factor, the γ value which provides the greatest inter-correlation coefficient was determined and this value was exported to an ".xlsx" document and further processed using the designed Python application.

The computational complexity of signal interpolation methods varies significantly based on their mathematical foundations and implementation. Each method presents a trade-off between computational efficiency and accuracy, with simple convolution being faster and more complex methods like sinc interpolation and SVM providing greater precision but at a higher computational cost.

4. Definition of comparison indicators

Let $x[n]$ be the initial signal and $x_1[n]$ be the signal which comes from the initial one by decimation. After applying a certain interpolation method for $x_1[n]$ an approximated signal denoted by $y[n]$ will be obtained. In this paper the inter-correlation coefficient, Euclidian distance and determinism will be computed between $x[n]$ and $y[n]$ by using the following formulas:

- Inter-correlation coefficient in point $t = 0$:

$$\rho_{xy}(0) = \frac{\sum_{i=0}^{n-1} x[i]y[i]}{\sqrt{\left(\sum_{i=0}^{n-1} x[i]^2\right)\left(\sum_{i=0}^{n-1} y[i]^2\right)}} \tag{20}$$

- Euclidian distance between the two signals:

$$d(x,y) = \sqrt{\sum_{i=0}^{n-1} (x[i] - y[i])^2} \tag{21}$$

- Determinism computed for the recurrence matrix obtained between the two signals.

Considering the two signals $x[i]$ and $y[i]$ with $i = \overline{0, n-1}$ then the recurrence matrix for x and y is defined as $A = (a_{i,j})_{0 \le i \le j \le n-1}$ with $a_{i,j}$ defined by [27] as:

$$a_{i,j} = \begin{cases} 1, \ if \ \sqrt{(x[i]-y[j])^2 + (x[i+1]-y[j+1])^2 + (x[i+2]-y[j+2])^2} < \varepsilon \\ 0, \ \text{otherwise} \end{cases} \tag{22}$$

The determinism has the following formula [28]:

$$\text{determinism} = \frac{\sum_{i=2}^{n} l(i) \cdot i}{N} \tag{23}$$

where $l(i)$ is the number of diagonal structures which consists only of "1" values and have length i and N represents the total number of "1" values from the recurrence matrix.

Two additional parameters were used for the QPSK modulated signal:

• Bit error rate (BER)

$$\text{BER} = \frac{\text{number\_ of\_ error\_ bits}}{\text{total\_ number\_ of\_ bits}} \tag{24}$$

Considering i_1 to be the binary sequence which results after demodulating $x[n]$ and i_2 the binary sequence which results after demodulating $y[n]$. Then number_ of_ error_bits represents the total number of bits which differ between the two binary sequences and total_ number_ of_ bits represents the length of the two sequences.

• Euclidian letter distance:

By using the ASCII representation (on 7 bits), the binary sequences i_1 and i_2 introduced above can be transformed into two texts text_1 and text_2. Each symbol is assigned with its ASCII decimal value, obtaining two sequences: ASCII_1 and ASCII_2 .

The Euclidian letter distance is defined as follows:

$$d(x, y)_{\text{letter}} = \sqrt{\sum_{i=0}^{N-1} (\text{ASCII}_1[i] - \text{ASCII}_2[i])^2} \tag{25}$$

To compare among the applied methods we have calculated the percentage for which a method is better than another one and the average value by which it manages to outperform that specific method by a specific procedure defined in the following section.

For the sinusoidal signal the following metrics apply:

Let two methods of interpolation meth_1 and meth_2 and let CORR_{ij} and EUC_{ij} be the inter-correlation coefficient and Euclidian distance respectively, computed for meth_j with $\text{NR} = \text{SNR}_i$, $i = \overline{1, 70}$, $j = \overline{1, 2}$.

Let $\text{diff}_{\text{CORR}} = \text{CORR}_{i1} - \text{CORR}_{i2}$. If $\text{diff}_{\text{CORR}} > 0$ then meth_1 is better than meth_2 for $\text{SNR} = \text{SNR}_i$ in terms of correlation coefficient.

Let $\text{diff}_{\text{EUC}} = \text{EUC}_{\text{i1}} - \text{EUC}_{i2}$. If $\text{diff}_{\text{EUC}} < 0$ then meth_1 is better than meth_2 for $\text{SNR} = \text{SNR}_i$ in terms of Euclidian distance.

If $\text{diff}_{\text{CORRi}} > 0$ for $i = \overline{1, m}$ then meth_1 is better than meth_2 in $\frac{m}{70} \cdot 100\% = p_{\text{CORR}}\%$ of the cases. If $p_{\text{CORR}} > 50$ then it can be said that meth_1 is better than meth_2 in $p_{\text{CORR}}\%$ of the cases with an average value of $\text{CORR}_{\text{avg}} = \frac{\sum_{i=1}^{m} \text{diff}_{\text{CORRi}}}{m}$.

If $\text{diff}_{\text{EUCi}} < 0$ for $i = \overline{1, m}$ then meth_1 is better than meth_2 in $\frac{m}{70} \cdot 100\% = p_{\text{EUC}}$ of the cases. If $p_{\text{EUC}} > 50$ then it can be said that meth_1 is better than meth_2 in $p_{\text{EUC}}\%$ of the cases with an average value of $\text{EUC}_{\text{avg}} = \frac{\sum_{i=1}^{m} \text{diff}_{\text{EUCi}}}{m}$.

For the QPSK signal the calculated metrics are defined as:

Let two methods of interpolation meth_1 and meth_2 and let $\text{letter\_ distance}_{ij}$ and BER_{ij} be the letter distance and BER respectively, computed for meth_j with $\text{ECIM} = \text{DECIM}_i$, $i = \overline{1, 25}$, $j = \overline{1, 2}$.

Let $\text{diff}_{\text{letter\_ distance}} = \text{letter\_ distance}_{i1} - \text{letter\_ distnace}_{i2}$. If $\text{diff}_{\text{letter\_ distance}} < 0$ then meth_1 is better than meth_2 for $\text{DECIM} = \text{DECIM}_i$ in terms of letter distance.

Let $\text{diff}_{\text{BER}} = \text{BER}_{i1} - \text{BER}_{i2}$. If $\text{diff}_{\text{BER}} < 0$ then meth_1 is better than meth_2 for $\text{DECIM} = \text{DECIM}_i$ in terms of BER.

If $\text{diff}_{\text{letter\_distance}} < 0$ for $i = \overline{1,m}$ then meth_1 is better than meth_2 in $\frac{m}{25} \cdot 100\% = p_{\text{letter}}\%$ of the cases. If $p_{\text{letter}} > 50$ then it can be said that meth_1 is better than meth_2 in $p_{\text{letter}}\%$ of the cases with an average value of $\text{letter}_{\text{avg}} = \frac{\sum_{i=1}^{m} \text{diff}_{\text{letter\_distancei}}}{m}$.

If $\text{diff}_{\text{BERi}} < 0$ for $i = \overline{1,m}$ then meth_1 is better than meth_2 in $\frac{m}{25} \cdot 100\% = p_{\text{BER}}$ of the cases. If $p_{\text{BER}} > 50$ then it can be said that meth_1 is better than meth_2 in $p_{\text{BER}}\%$ of the cases with an average value of $\text{BER}_{\text{avg}} = \frac{\sum_{i=1}^{m} \text{diff}_{\text{BERi}}}{m}$.

5. Results and discussion

5.1 *Comparison between interpolation methods for sinusoidal signal corrupted by noise*

The inter-correlation coefficient and Euclidean distance parameters were computed between the signal resulting from the interpolation method and the original signal, for each SNR value in the case of the sinusoidal signal corrupted by noise.

The dependence between the Euclidean distance values obtained between the original signal, $x[n]$, and the interpolated signal, $y[n]$, for each interpolation method across 70 distinct SNR values is presented in Figure 2a and the corresponding boxplots are presented in Figure 3b. As theoretically expected, we observe that the Euclidean distance decreases with SNR increase. A small Euclidean distance translates to a better signal approximation, and all interpolation methods perform better for higher SNR conditions. By analyzing the distribution of the Euclidean distance values, we can deduce that convolution with sinc signal demonstrates superior performance compared to other interpolation methods. This is evident as it consistently produces the lowest values for this indicator, regardless of the SNR value. From Figure 2b, it can be seen that the convolution method with sinc signal provides the lowest median value for Euclidian distance compared to all the other methods. It is also clear that this interpolation method yields a very similar distribution with the SVM method. However, one can observe that for lower SNR values the SVM interpolation method can yield to better results than the sinc interpolation method.

The graph in Figure 4a presents the inter-correlation coefficient values obtained between the original signal, $x[n]$, and the interpolated signal, $y[n]$, for each interpolation method across 70 distinct SNR values and the corresponding boxplot values (b)

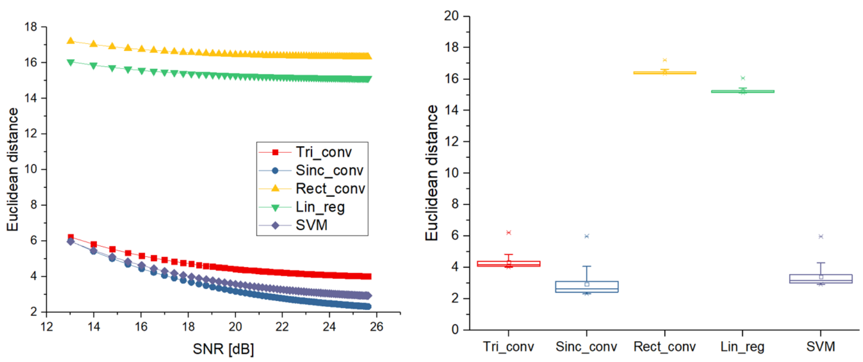

Figure 2. The dependence between Euclidian distance and SNR (dB) (a) and corresponding boxplots (b) for sinusoidal signal corrupted by noise

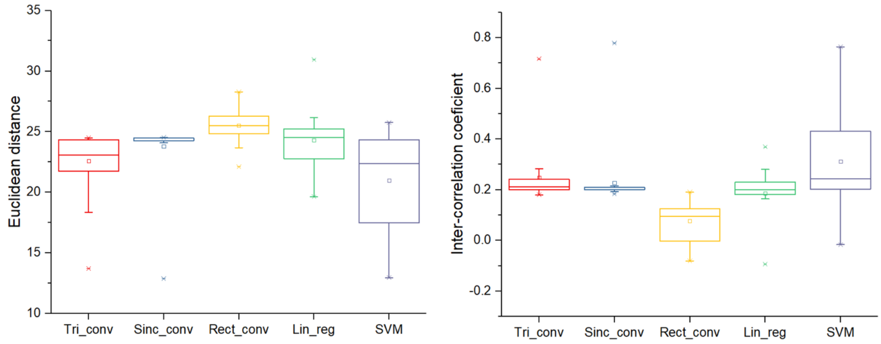

Figure 3. Boxplot graph for Euclidian distance (a) and inter-correlation coefficient (b) distribution in the case of the QPSK modulated signal

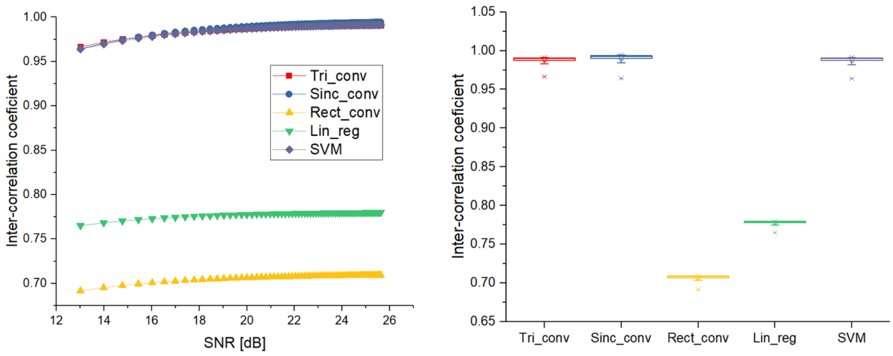

Figure 4. The dependence between the inter-correlation coefficient and SNR (dB) (a) and the corresponding boxplots (b) for sinusoidal signal corrupted by noise

From Figure 4 and Table 1 (right) it is clear the fact that sinc convolution method provides the best performances compared to all the other 4 methods in terms of inter-correlation coefficient, demonstrating outstanding performances in terms of this indicator. It can also be seen that it has quite similar distribution like triangular convolution method and SVM interpolation method. Table 1 presents the calculated p_{EUC} (left) and p_{CORR} (right) percentages.

From analyzing the two indicators presented in Table 1 we can observe that for the Euclidian distance convolution with sinc outperforms all other convolution methods (with triangular and rectangular signal) as well as the linear regression model for all SNR values (100). The same observation is valid when analyzing the p_{CORR} values when convolution with sinc is also ranked first followed by convolution with the triangular signal.

Table 1. p_{EUC} elements (left) and p_{CORR} (right) for sinusoidal signal corrupted by noise

	p_{EUC} (%)						p_{CORR}(%)				
	Tri_conv	**Sinc_conv**	**Rect_conv**	**Lin_reg**	**SVM**		**Tri_conv**	**Sinc_conv**	**Rect_conv**	**Lin_reg**	**SVM**
Tri_conv	0	0	100	100	0	**Tri_conv**	0	4.28	100	100	60
Sinc_conv	100	0	100	100	98.57	**Sinc_conv**	95.71	0	100	100	100
Rect_conv	0	0	0	0	0	**Rect_conv**	0	0	0	0	0
Lin_reg	0	0	100	0	0	**Lin_reg**	0	0	100	0	0
SVM	100	1.42	100	100	0	**SVM**	40	0	100	100	0

The values of the determinism indicator were computed for the lowest SNR (13.01 dB) and the highest SNR (25.62 dB) values, considering all five interpolation methods used in the article. The results are presented in Table 2.

Table 2. Determinism values for two SNR values (13.01 dB and 25.62 dB) computed for each of the interpolation methods for $\varepsilon = 0.2$

Method used	SNR = 13.01 dB	SNR = 25.62 dB
tri_conv	0.06183	0.90779
sinc_conv	0	0.95091
rect_conv	0.03636	0.80316
lin_reg	0.01943	0.83656
SVM	0	0.94802

From Table 2 it can be inferred that for SNR = 13.01 dB the method which implies performing convolution between the sampled signal and a triangular signal provides the greatest value for determinism. This observation suggests that there are more sequences inside the approximated signal using triangular convolution which are similar to certain sequences from the sinusoidal signal corrupted by noise. Thus, for noisy signals/communication channels triangular convolution outperforms the other studied methods in terms of determinism.

For SNR = 25.62 dB the method for interpolation which consists in performing convolution between the sampled signal and sinc signal provides greatest value for determinism compared to all the other 4 methods. This sugests that sinc convolution outperforms all the other methods in terms of determinism for large values of SNR/good signal conditions.

5.2 Comparison between interpolation methods for QPSK modulated signal

In Figure 3 we present boxplot graphs for Euclidian distance and inter-correlation coefficient to highlight the distribution of this values for all the five interpolation methods.

From Figure 3a it can be deduced that SVM interpolation method has the best performances in terms of Euclidian distance compared to all the other methods as it yields a median value which is smaller than all the median values which corresponds to the other 4 methods. It can also be said that the range which corresponds for the values which are lower than the median value is wider and extends to even lower possible values for Euclidian distance compared to all the other methods.

Table 3. p_{EUC} elements (left) and p_{CORR} (right) for QPSK modulated signal

	p_{EUC} (%)						p_{CORR}(%)				
	Tri_conv	**Sinc_conv**	**Rect_conv**	**Lin_reg**	**SVM**		**Tri_conv**	**Sinc_conv**	**Rect_conv**	**Lin_reg**	**SVM**
Tri_conv	0	96	100	88	12	**Tri_conv**	0	68	100	80	16
Sinc_conv	4	0	88	64	12	**Sinc_conv**	32	0	100	64	20
Rect_conv	0	12	0	12	4	**Rect_conv**	0	0	0	4	0
Lin_reg	12	36	88	0	4	**Lin_reg**	20	36	96	0	8
SVM	88	88	100	96	0	**SVM**	84	80	100	92	0

The superiority of the SVM method is also supported by the values of the inter-corellation coefficient presented in Figure 5b, where one can observe that the that SVM methods provides the best performances for the QPSK modulated signal. In Table 3, we present the results for the computed p_{EUC} (left) and p_{CORR} (right) percentages.

From analyzing the two indicators presented in Table 3 we can observe that for the QPSK modulated signal, in terms of Euclidean distance, SVM outperforms all other convolution methods (with triangular, sinc and rectangular signal) for at least 88% of the decimation factors. The SVM also outperforms the linear regression model for 96% of the decimation factors. The same observation is valid when analyzing the p_{CORR} values when SVM method is also ranked first followed by convolution with the triangular signal.

In Figure 5 we present boxplot graphs for Euclidian letter distance and BER variations for the QPSK modulated signal.

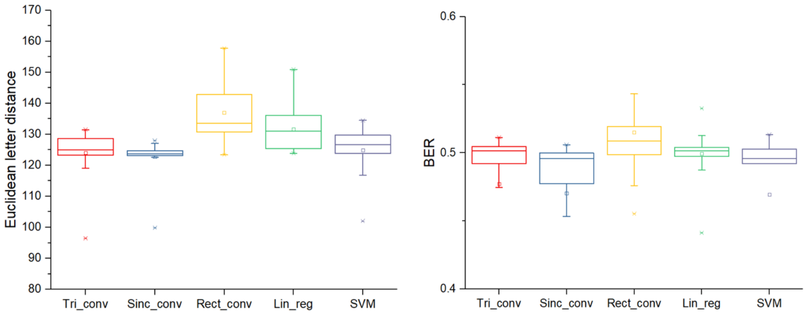

Figure 5. Boxplot graph for letter Euclidian distance (a) and BER (b) distribution in the case of the QPSK modulated signal

By analyzing Figure 5 it can be stated that sinc convolution method has lower median values and wider ranges for values which are lower than the median value for Euclidian letter distance and BER indicators compared to the other methods implying that it has the best performances when these two indicators are used. In Table 4, we present the results for the computed p_{letter} (left) and p_{BER} (right) percentages.

Table 4. p_{letter} elements (left) and p_{BER} (right) for QPSK modulated signal

	p_{letter} (%)						p_{BER}(%)				
	Tri_conv	**Sinc_conv**	**Rect_conv**	**Lin_reg**	**SVM**		**Tri_conv**	**Sinc_conv**	**Rect_conv**	**Lin_reg**	**SVM**
Tri_conv	0	16	80	88	60	**Tri_conv**	0	12	72	68	60
Sinc_conv	84	0	88	100	84	**Sinc_conv**	88	0	88	88	60
Rect_conv	20	12	0	28	20	**Rect_conv**	28	12	0	44	24
Lin_reg	12	0	72	0	24	**Lin_reg**	32	12	56	0	32
SVM	40	16	80	76	0	**SVM**	40	40	76	68	0

However, from analyzing the two indicators presented in Table 4 we can observe that for the QPSK modulated signal, in terms p_{letter}, convolution with sinc outperforms all other convolution methods (with triangular and rectangular signal) for at least 84% of the decimation factors. The convolution with sinc also outperforms the linear regression model for 100% of the decimation factors. The same observation is valid when analyzing the p_{BER} values when convolution with sinc method is also ranked first followed by convolution with the triangular signal.

The values of the determinism indicator were computed for the lowest decimation factor (2) and the highest decimation factor (200) values, considering all five interpolation methods used in the article:

From Table 5, it can be inferred that for d = 2 and d = 200, the SVM method, when applied to the sampled signal, provides the greatest values for determinism. This implies that the SVM interpolation method exhibits more sequences within the approximated signal that resemble certain sequences from the QPSK modulated signal, indicating that the SVM method outperforms all other methods for both decimation factor values.

Table 5. Determinism values for d = 2/200 computed for each of the interpolation methods for $\varepsilon = 0.5$

Method used	d = 2	d = 200
tri_conv	0.42729	0.23813
sinc_conv	0.50271	0.26237
rect_conv	0.39582	0.00776
lin_reg	0.49991	0
SVM	0.50958	0.47709

The computational complexity of the investigated interpolation methods significantly impacts their practical usability in real-world applications, particularly in wireless communication systems where real-time processing is essential. The five interpolation methods analyzed in this study each exhibit different trade-offs between computational efficiency and interpolation accuracy.

The rectangular and triangular convolution methods are computationally efficient, typically operating in time domain, as they involve straightforward averaging or linear weight application to neighboring samples. The sinc convolution method, while offering superior frequency domain performance, has a higher complexity due to its need for longer kernel support and truncation, making it less practical for real-time applications with large datasets.

The SVM-based interpolation method, while yielding good accuracy in Euclidean distance and inter-correlation metrics, is computationally expensive. SVM interpolation is suitable for offline processing but might pose problems for real-time applications without sufficient computational resources.

In real-world applications, the choice of interpolation method depends on the trade-off between computational resources and accuracy requirements. While SVM-based interpolation provides the highest fidelity, its computational cost may be prohibitive for real-time embedded systems or battery-powered devices. In contrast, sinc convolution offers a practical balance, especially in applications where real-time performance is critical, such as in adaptive equalization and signal restoration in communication networks.

6. Conclusions

This article presents a comparative study on the efficiency of five interpolation methods applied for two types of signals: sinusoidal signal corrupted by noise (Section 5.1) and QPSK modulated signal (Section 5.2). The comparative study is performed with statistics enabled for a wide range of SNR values for the sinusoidal signal and decimation factors for the QPSK modulated signal.

The comparative study was performed based on five indicators, namely: Euclidean distance, inter-correlation coefficient, Euclidean letter distance, BER and determinism. To compare between the interpolation methods for the whole SNR or decimation factors we have proposed the use of $p_{\text{indicator}}$ parameter to quantify the percentage of cases one method outperforms another one.

The main findings of the study can be summarized as follows:

• For the sinusoidal signal corrupted by noise the method which outperforms the other methods in terms of Euclidian distance and inter-correlation coefficient is convolution with sinc (Section 5.1).

• For the QPSK modulated signal it is the SVM method which outperforms all the other methods in terms of Euclidean distance and inter-correlation coefficient (Section 5.2).

• If Euclidean letter distance and BER indicators are considered, in the case of QPSK modulated signal, we have observed that convolution with sinc signal outperformes all other investigated methods (Section 5.2).

By conducting this study, we contribute to the improvement of research methods that investigate the usability of different interpolation methods for QPSK signals, a modulation type commonly present in the radiofrequency environment. Moreover, we present an original and proven-to-be reliable research methodology that can be used to compare the efficiency of different interpolation methods on the applied signal types (Section 2).

Building upon the findings of this paper and previous research, we emphasize the importance of tailoring interpolation methods to specific signal types and applications. The results highlight that there is no universally optimal interpolation method; rather, the best choice depends on the characteristics of the signal and the intended application. Future work should explore adaptive and hybrid interpolation strategies, optimizing performance based on real-time signal conditions and requirements.

Conflicts of interest

The authors declare no conflict of interest.

References

[1] M. Lepot, J.-B. Aubin, and F. H. L. R. Clemens, “Interpolation in time series: An introductive overview of existing methods, their performance criteria and uncertainty assessment,” *Water*, vol. 9, no. 10, p. 796, 2017. https://doi.org/10.3390/w9100796

[2] M. Gaß, K. Glau, M. Mahlstedt, and M. Mair, “Chebyshev interpolation for parametric option pricing,” *Finance and Stochastics*, vol. 22, no. 3, pp. 701-731, 2018. https://doi.org/10.1007/s00780-018-0361-y

[3] S. Fujimoto, T. Mizuno, and A. Ishikawa, “Interpolation of non-random missing values in financial statements’ big data using CatBoost,” *Journal of Computational Social Science*, vol. 5, no. 2, pp. 1281-1301, 2022. https://doi.org/10.1007/s42001-022-00165-9

[4] P. P. Amoroso, F. J. Aguilar, C. Parente, M. A. Aguilar, “Statistical assessment of some interpolation methods for building grid format digital bathymetric models,” *Remote Sensing*, vol. 15, no. 8, p. 2072, 2023. https://doi.org/10.3390/rs15082072

[5] F. Turco, L. Azevedo, and D. Herold, “Geostatistical interpolation of non-stationary seismic data,” *Computational Geosciences*, vol. 23, no. 4, pp. 665-682, 2019. https://doi.org/10.1007/s10596-019-9812-6

[6] Z. Xu and T. Zhou, “On sparse interpolation and the design of deterministic interpolation points,” *SIAM Journal on Scientific Computing*, vol. 36, no. 4, pp. A1752-A1769, 2014. https://doi.org/10.1137/13094596x

[7] S. Marcus, M. Frigura-Iliasa, D. Vatau, and L. Matiu-Iovan, “New interpolation tools for digital signal processing,” In Proc. 2016 International Conference on Information and Digital Technologies, Rzeszow, Poland, Jul. 5-7, 2016, pp. 266-270. https://doi.org/10.1109/DT.2016.7557184

[8] M. Cao, F. Xu, H. Jia, L. Zhou, E. Ji, J. Wu, “A multiple interpolation algorithm to improve resampling accuracy in data triggers,” *Electronics*, vol. 12, no. 6, p. 1291, 2023. https://doi.org/10.3390/electronics12061291

[9] S. Pinjerla, S. Rao S, and C. P. Reddy, “Sampling rate conversion techniques-A review,” In Proc. 2021 4th International Conference on Recent Trends in Computer Science and Technology, Jamshedpur, India, Feb. 11-12, 2022, pp. 278-282. https://doi.org/10.1109/ICRTCST54752.2022.9781914

[10] N. Hussain and O. Singh, “A review on design of sampling rate converter using symmetric technique,” *International Journal of Engineering Science and Computing*, vol. 8, no. 4, pp. 17221-17225, 2018. https://doi.org/10.4010/2018.1970

[11] L. Chen, G. Li, G. Huang, P. Shi, “A missing type-aware adaptive interpolation framework for sensor data,” *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-15, 2021. https://doi.org/10.1109/TIM.2021.3089783

[12] A. Mazarguil, L. Oudre, and N. Vayatis, “Non-smooth interpolation of graph signals,” *Signal Processing*, vol. 196, p. 108480, 2022. https://doi.org/10.1016/j.sigpro.2022.108480

[13] S. C. Lin, K. Chuang, C. W. Chang, J. H. Chen, “Efficient interpolation method for wireless communications and signal processing applications,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 69, no. 5, pp. 2753-2761, 2021. https://doi.org/10.1109/TMTT.2021.3061563

[14] S. Chen, Y. Yang, H. Hu, F. Guan, G. Shen, Z. Bian, H. Guo, “Interpolation method for wideband signal reconstruction based on blade tip timing measurement,” *Measurement*, vol. 176, p. 109168, 2021. https://doi.org/10.1016/j.measurement.2021.109168

[15] R. D. Horansky, D. C. Ribeiro, K. A. Remley, P. D. Hale, C. M. Wang, D. F. Williams, N. B. Carvalho, “Comparison of timebase interpolation methods for traceable, wideband mm-wave communication signals,” In Proc. 2016 87th ARFTG Microwave Measurement Conference, San Francisco, CA, USA, Jun. 2-3, 2016, pp. 1-4. https://doi.org/10.1109/ARFTG.2016.7501941

[16] Anonymous, “Satellite image enhancement using optimized wavelet decomposition and bicubic interpolation,” *International Journal of Recent Technology and Engineering*, vol. 8, no. 5, pp. 6829-6834, 2020. https://doi.org/10.35940/ijrte.E6829.018520

[17] D. Liu, Q. Zhao, D. Fu, S. Guo, P. Liu, Y. Zeng, “Comparison of spatial interpolation methods for the estimation of precipitation patterns at different time scales to improve the accuracy of discharge simulations,” *Hydrology Research*, vol. 51, no. 4, pp. 583-601, 2020. https://doi.org/10.2166/nh.2020.146

[18] R. Yang and B. Xing, “A comparison of the performance of different interpolation methods in replicating rainfall magnitudes under different climatic conditions in Chongqing Province (China),” *Atmosphere*, vol. 12, no. 10, p. 1318, 2021. https://doi.org/10.3390/atmos12101318

[19] S.-C. Lin, K. Chuang, C.-W. Chang, and J.-H. Chen, “Efficient interpolation method for wireless communications and signal processing applications,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 69, no. 5, pp. 2753-2761, 2021. https://doi.org/10.1109/TMTT.2021.3061563

[20] Y. Du, X. Jin, H. Wang, and M. Lu, “An adaptive multipath linear interpolation method for sample optimization,” *Mathematics*, vol. 11, no. 3, p. 768, 2023. https://doi.org/10.3390/math11030768

[21] C. Shao, Z. Xu, and H. Ranganathan, “A novel resynchronization circuit to enable high-accuracy PTP timestamping,” *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1-8, 2024. https://doi.org/10.1109/TIM.2024.3420369

[22] W. Asavanant, B. Charoensombutamon, S. Yokoyama, T. Ebihara, T. Nakamura, R. N. Alexander, M. Endo, J.-i. Yoshikawa, N. C. Menicucci, H. Yonezawa, and A. Furusawa, “Time-domain-multiplexed measurement-based quantum operations with 25-MHz clock frequency,” *Physical Review Applied*, vol. 16, no. 3, p. 034005, 2021. https://doi.org/10.1103/PhysRevApplied.16.034005

[23] The University of Texas at Dallas. (2023). [online]. Available: https://personal.utdallas.edu/~torlak/courses/DSProject/lectures/lecture05/interpolation.pdf. [Accessed May 20, 2023]

[24] T. Daniya, M. Geetha, B. Santhosh Kumar, and R. Cristin, “Least square estimation of parameters for linear regression,” *International Journal of Control and Automation*, vol. 13, no. 2, pp. 447-452, 2020.

[25] G. Gordon, Support Vector Machines and Kernel Methods. [Online]. Available: https://www.cs.cmu.edu/~ggordon/SVMs/new-svms-and-kernels.pdf. [Accessed May 28, 2023]

[26] L. Cuoghi, “Support vector regression (LS-SVR) for non linear functions,” MathWorks. [Online]. Available: https://www.mathworks.com/matlabcentral/fileexchange/73706-support-vector-regression-ls-svr-for-non-linear-functions. [Accessed May 20, 2023]

[27] M. Fukino, Y. Hirata, and K. Aihara, “Coarse-graining time series data: Recurrence plot of recurrence plots and its application for music,” *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 26, no. 2, p. 023116, 2016. https://doi.org/10.1063/1.4941371

[28] V. Houshyarifar and M. C. Amirani, “Early detection of sudden cardiac death using Poincaré plots and recurrence plot-based features from HRV signals,” *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 25, no. 3, pp. 1541-1553, 2017. https://doi.org/10.3906/elk-1509-149