Available online at www.worldscientificnews.com

World Scientific News

An International Scientific Journal

WSN 192 (2024) 76-87 EISSN 2392-2192

Prediction of Rainfall Data in the DKI Jakarta Area Using Cubic Spline Interpolation

Willen Vimelia*, Chusnul Chatimah Azis, Sri Purwani

Department of Mathematics, Universitas Padjadjaran, Sumedang 35463, Indonesia

*E-mail address: willen20001@mail.unpad.ac.id

ABSTRACT

No human can survive without water. Water holds numerous benefits for human life, encompassing essential needs such as drinking and bathing, as well as playing a crucial role in electricity generation. Rain serves as one of the primary sources of water. While rain brings various advantages to human existence, it is not uncommon for it to also bring about disasters. This is further compounded by climate change, which can lead to alterations in rainfall patterns and increases in air temperatures. Changes in rainfall patterns pose a potential risk of undesirable consequences. In urban areas, these consequences may manifest as floods, disruptions in transportation, and the outbreak of diseases. Therefore, it is imperative to meticulously record rainfall data for a given region. This recording serves as a foundation for implementing preventative measures to mitigate potential unwanted outcomes. Unfortunately, in some instances, rainfall data goes unrecorded, resulting in data gaps. This paper focuses on predicting rainfall data through the application of cubic spline interpolation. Cubic spline interpolation is a method employed to determine an interpolating function that accurately passes through a defined set of data points. In this research, the cubic spline interpolation method successfully predicted rainfall data with a minimal relative error.

***Keywords*:** Rainfall, Data Prediction, Cubic Spline Interpolation

(Received 24 February 2024; Accepted 19 March 2024; Date of Publication 21 March 2024)

1. INTRODUCTION

Water is one of the fundamental necessities for human survival. It is used for drinking, cooking, bathing, and even for electricity generation [1]. Rain is one of the primary sources of water for human life. According to Great Dictionary of the Indonesian Language (Kamus Besar Bahasa Indonesia or simply KBBI), rain is defined as water droplets that fall from the atmosphere due to the cooling process. Rain can have positive impacts, such as replenishing groundwater and preserving plants from wilting. However, excessive rainfall can also lead to adverse effects, such as floods and landslides [2].

Climate change is associated with long-term changes in the statistical distribution of weather patterns over significant periods [3]. These changes include temperature variations, extreme weather, sea level fluctuations, and wind patterns [4]. Climate change has currently become a widely discussed topic due to its felt impacts, influencing various aspects of human life. Rezanezhad et al. asserted that regions vulnerable to climate change include urban areas, one of which is Daerah Khusus Ibukota (DKI) Jakarta, the capital city of Indonesia [5]. This vulnerability is attributed to the high human activity and development in DKI Jakarta, contributing to the increase in greenhouse gas emissions. The potential impacts of climate change include alterations in rainfall patterns and rising air temperatures [6], [7]. Changes in extreme rainfall can lead to various consequences, such as floods, disease outbreaks, health disturbances, and disruptions in transportation, such as flight delays and train departure schedules being affected. People must be able to adapt to climate change and its impacts. However, they must also prepare to face the consequences of climate change [8]. Without action today, adapting to these impacts in the future will be more difficult and costly [9]. Therefore, it is essential to have the appropriate knowledge and methods to provide accurate information about these extreme events, aiming to mitigate the most severe impacts that may occur [10].

In the given circumstances, rainfall in DKI Jakarta is deemed essential to be regularly recorded to facilitate the planning and management of water resources, including river flow regulation and flood control. Additionally, rainfall data can support infrastructure planning in various DKI Jakarta regions, such as the construction and maintenance of water channels, dams, and drainage networks. However, in some cases, rainfall data may go unrecorded, leading to data gaps that can impact planning processes requiring such information. This challenge can be addressed through the utilization of interpolation methods.

Wen et al. conducted research with the objective of understanding investors perspectives on different risk criteria when making decisions related to blockchain technology. Investor preferences were to be extracted using personalized quantifiers. The cubic spline interpolation function was utilized to ensure that the quantifiers exhibited a smooth curve and accurate fitting. The results of the study indicated that quantifiers developed through cubic spline interpolation possessed favorable geometric features [11]. Azizan et al. utilized natural cubic spline and not-a-knot spline interpolation to visualize and predict rainfall data in Ipoh and Petaling, Malaysia [12]. The study demonstrated that both interpolation methods effectively performed interpolation and yielded reliable results. In a different context, Duan et al. employed spline interpolation to ensure smoothness and accuracy in preprocessing landslide displacement predictions with a multivariable Long Short-Term Memory (LSTM) model. This model integrated rainfall and reservoir water level into a traditional LSTM model [13].

Research on predicting rainfall data is essential as a preventive measure against the adverse impacts of extreme rainfall, particularly in urban areas. This prediction can be accomplished using cubic spline interpolation, which has the capability to generate a smooth and accurate curve. Therefore, this paper will delve into the prediction of rainfall data in the DKI Jakarta region utilizing cubic spline interpolation.

2. MATERIALS AND METHODS

2. 1. Rainfall

Within the framework of hydrological dynamics, rainfall is considered one of the primary sources of water. Additionally, rainfall can also affect various other aspects such as the economy, agriculture, and local regulations [14]. The natural process of rainfall formation involves the condensation of water vapor in the air, which then forms clouds. Under specific conditions, these clouds will precipitate rainfall. The phenomenon of rainfall is highly dependent on the prevailing weather conditions [15-19]. Rainfall measurement is crucial for analyzing and mitigating potential impacts from extreme changes in rainfaall patterns. Accurate rainfall measurements are important for various aspects of city life, such as urban planning and flood management [20].

Quoted from bmkg.go.id, the threshold values used to determine rainfall intensity are provided in the table below.

Table 1. Rain Categories Based on Rainfall Intensity.

Rainfall	Rain Categories
0 mm/day	Cloudy
0.5 – 20 mm/day	Light
20 – 50 mm/day	Moderate
50 – 100 mm/day	Heavy
100 – 150 mm/day	Very Heavy
>150 mm/day	Extreme

2. 2. Cubic Spline Interpolation

Interpolation is a method used to estimate the value of the dependent variable between two extreme values of the independent variable, based on the given values of both variables [21]. The term "spline" refers to craft tools, such as flexible strips of wood or thin metal, used to form smooth curves. Various loads are applied at different positions, causing the strip to bend according to the quantity and positions. These strips are forced to pass through fixed points, such as metal pins or boat ribs. On a flat surface, these points often take the form of loads with attached hooks, facilitating manipulation. The bending of naturally shaped material will form a spline curve [22]. Similarly, splines are used in statistics to mathematically represent flexible shapes. Nodes are placed at various locations within the data range to mark the points where adjacent parts of the function join together. As an alternative to metal or wood lines, smooth function pieces (usually low-degree polynomials) are selected to fit the data between two consecutive nodes. The type of polynomial, the number of nodes, and their placement then determine the type of spline produced [22].

Spline interpolation is an alternative approach and the development of polynomial interpolation. The advantage of spline interpolation lies in its ability to reduce interpolation errors, even when using low-degree polynomials. Spline interpolation is considered more advantageous as it can avoid the Runge phenomenon issues that arise when interpolating using high-degree polynomials [23].

Cubic spline interpolation aims to obtain a continuous interpolation function, not only on the intervals but also at the interpolation points, for the function itself, the first and second derivatives. This is intended to produce an interpolation function with better smoothness. The continuity of the first derivative indicates that the graph $y = S(x)$ does not have sharp corners or discontinuities. Meanwhile, the continuity of the second derivative signifies that the radii of curvature are well-regulated at each interpolation point [23]. Another advantage of cubic splines is their ability to preserve the original shape or pattern of the data. Cubic splines can accurately model data effectively with minimal deviation from the actual values [24].

For n data points $(x_1, y_1), \dots, (x_n, y_n)$, where x_i are distinct and $i = 1, .., n$. A cubic spline $S(x)$ passing through the data points $(x_1, y_1), \dots, (x_n, y_n)$ is a set of cubic polynomials:

$$S_1(x) = y_1 + b_1(x - x_1) + c_1(x - x_1)^2 + d_1(x - x_1)^3, \textbf{ on } [\mathrm{x}_1, \mathrm{x}_2]$$

$$S_2(x) = y_2 + b_2(x - x_2) + c_2(x - x_2)^2 + d_2(x - x_2)^3, \textbf{ on } [x_2, x_3]$$

$$\vdots$$

$$S_{n-1}(x) = y_{n-1} + b_{n-1}(x - x_{n-1}) + c_{n-1}(x - x_{n-1})^2 + d_{n-1}(x - x_{n-1})^3, \textbf{ on } [x_{n-1}, x_n]$$

There are several conditions or properties that must be satisfied [19]:

a) $S_i(x_i) = y_i$ **for** $i = 1, \dots, n-1$

b) $S_{i-1}(x_i) = S_i(x_i)$ **for** $i = 2, \dots, n-1$

c) $S'_{i-1}(x_i) = S'_i(x_i)$ **for** $i = 2, \dots, n-1$

d) $S''_{i-1}(x_i) = S''_i(x_i)$ **for** $i = 2, \dots, n-1$

3. DATA

This study utilizes monthly rainfall data for the DKI Jakarta region from January 2021 to December 2022, obtained from the site https://power.larc.nasa.gov. The data is compiled into a single table as shown below.

In Table 2, it can be observed that the rainfall data obtained in several regions has the same values. This occurs in the rainfall data for the South Jakarta region compared to the East Jakarta region, as well as in those data for the North Jakarta, West Jakarta, and Central Jakarta regions. Only Kepulauan Seribu has different values of the rainfall data.

Table 2. Monthly Rainfall Data for the Regions of DKI Jakarta.

Month	Kepulauan Seribu	South Jakarta	North Jakarta	East Jakarta	West Jakarta	Central Jakarta
Jan-21	379,69	321,68	348,05	321,68	348,05	348,05
Feb-21	627,54	564,26	696,09	564,26	696,09	696,09
Mar-21	400,78	485,16	411,33	485,16	411,33	411,33
Apr-21	269,64	206,77	205,34	206,77	205,34	205,34
May-21	190,15	168,97	141,31	168,97	141,31	141,31
Jun-21	129,06	157,14	146,19	157,14	146,19	146,19
Jul-21	70,12	60,87	53,97	60,87	53,97	53,97
Aug-21	107,38	84,7	66,49	84,7	66,49	66,49
Sep-21	169,25	147,35	118,02	147,35	118,02	118,02
Oct-21	209,8	210,12	170,08	210,12	170,08	170,08
Nov-21	239,67	321,58	294,41	321,58	294,41	294,41
Dec-21	341,01	306,26	302,43	306,26	302,43	302,43
Jan-22	284,36	220,38	249,95	220,38	249,95	249,95
Feb-22	269,35	230,21	223,88	230,21	223,88	223,88
Mar-22	227,19	253,55	259,86	253,55	259,86	259,86
Apr-22	256,41	323,17	286,27	323,17	286,27	286,27
May-22	236,34	186,48	144,77	186,48	144,77	144,77
Jun-22	139,59	159,57	135,94	159,57	135,94	135,94
Jul-22	114,1	124,29	128,64	124,29	128,64	128,64
Aug-22	123,59	105,22	92,04	105,22	92,04	92,04
Sep-22	134,02	173,86	141,13	173,86	141,13	141,13
Oct-22	236,47	296,8	235,17	296,8	235,17	235,17
Nov-22	185,94	263,27	218,05	263,27	218,05	218,05
Dec-22	247,65	265,53	239,71	265,53	239,71	239,71

4. RESULT AND DISCUSSION

In this study, the rainfall data in each region of DKI Jakarta for August 2021 and May 2022 will be used as sample data, assuming that rainfall is not recorded in those months. In the calculations, the sample data will not be included as interpolation points and will be compared to the corresponding result of interpolation. The obtained results are as follows:

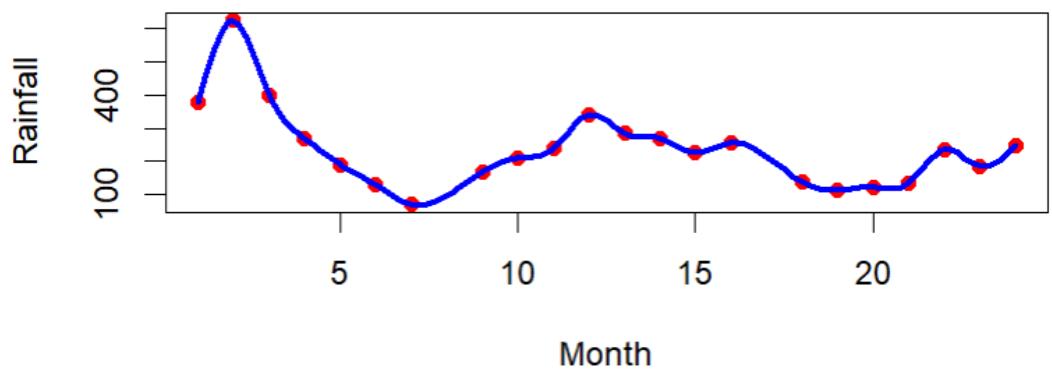

Figure 1. Cubic spline curve for Kepulauan Seribu

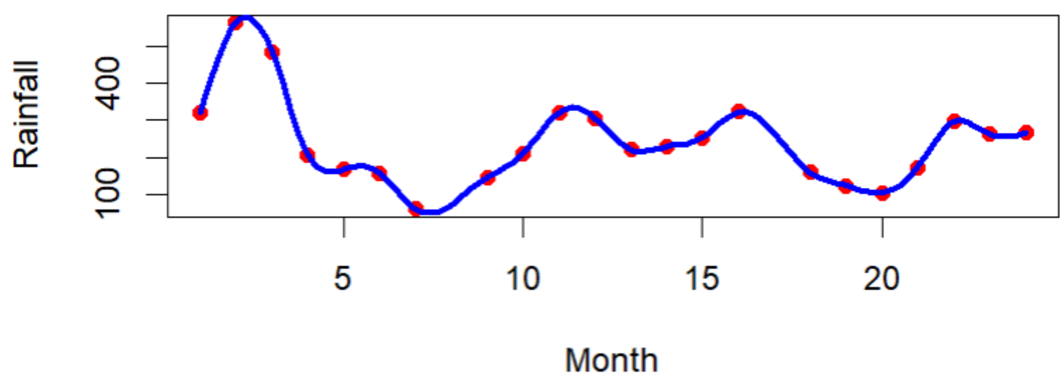

Figure 2. Cubic spline curve for South and East Jakarta

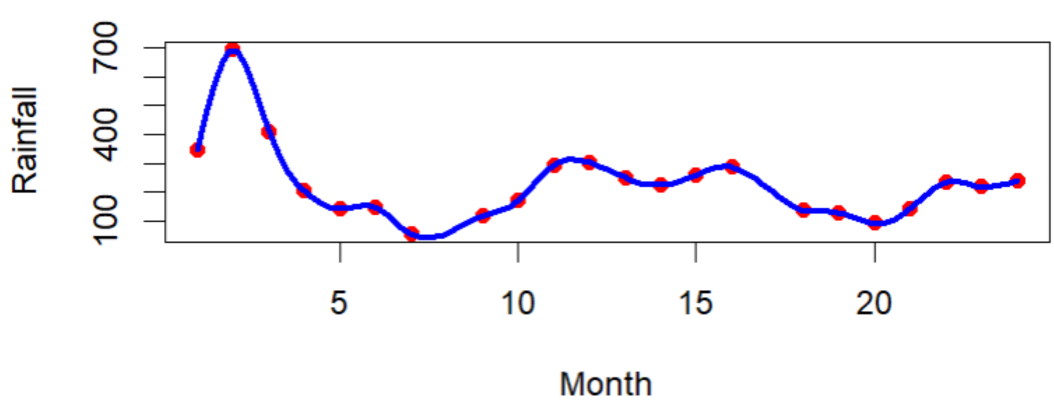

Figure 3. Cubic spline curve for North, West, and Central Jakarta

Applying a cubic spline to the data sample produces rainfall in the Kepulauan Seribu region of 104.61677 in August 2021 and 215.36346 in May 2022. The cubic spline results for rainfall data in South Jakarta and East Jakarta are 87.43487 in August 2021 and 185.69940 in May 2022. Lastly, for the North Jakarta, West Jakarta, and Central Jakarta regions, the results obtained are 69.09901 in August 2021 and 145.23249 in May 2022. These results can be seen in Table 3.

Table 3. Monthly Rainfall Data for the DKI Jakarta Regions with Interpolation Results.

Month	Kepulauan Seribu	South Jakarta	North Jakarta	East Jakarta	West Jakarta	Central Jakarta
Jan-21	379,69	321,68	348,05	321,68	348,05	348,05
Feb-21	627,54	564,26	696,09	564,26	696,09	696,09
Mar-21	400,78	485,16	411,33	485,16	411,33	411,33
Apr-21	269,64	206,77	205,34	206,77	205,34	205,34
May-21	190,15	168,97	141,31	168,97	141,31	141,31
Jun-21	129,06	157,14	146,19	157,14	146,19	146,19
Jul-21	70,12	60,87	53,97	60,87	53,97	53,97
Aug-21	104,61677	84,43487	69,09901	84,43487	69,09901	69,09901
Sep-21	169,25	147,35	118,02	147,35	118,02	118,02
Oct-21	209,8	210,12	170,08	210,12	170,08	170,08
Nov-21	239,67	321,58	294,41	321,58	294,41	294,41
Dec-21	341,01	306,26	302,43	306,26	302,43	302,43
Jan-22	284,36	220,38	249,95	220,38	249,95	249,95
Feb-22	269,35	230,21	223,88	230,21	223,88	223,88
Mar-22	227,19	253,55	259,86	253,55	259,86	259,86
Apr-22	256,41	323,17	286,27	323,17	286,27	286,27
May-22	215,3635	185,6994	145,2325	185,6994	145,2325	145,2325
Jun-22	139,59	159,57	135,94	159,57	135,94	135,94
Jul-22	114,1	124,29	128,64	124,29	128,64	128,64
Aug-22	123,59	105,22	92,04	105,22	92,04	92,04
Sep-22	134,02	173,86	141,13	173,86	141,13	141,13
Oct-22	236,47	296,8	235,17	296,8	235,17	235,17
Nov-22	185,94	263,27	218,05	263,27	218,05	218,05
Dec-22	247,65	265,53	239,71	265,53	239,71	239,71

To assess the accuracy of the interpolation results using cubic spline, the error from the interpolation results against the actual data is calculated using relative error (ε) where y is the actual data and $\hat{y}$ is the interpolation result. Since the rainfall data for South Jakarta and East Jakarta are uniform, the interpolation results and error calculations are the same. Similarly, those interpolation results and errors for North Jakarta, West Jakarta, and Central Jakarta regions are the same. The error calculation for the interpolation results of each region can be seen below:

- Kepulauan Seribu
  a) August 2021
  $y = 107.38$
  $\hat{y} = 104.61677$

$$\varepsilon = \left|\frac{y - \hat{y}}{y}\right| = \left|\frac{107.38 - 104.61677}{107.38}\right| = 0.0257$$

Therefore, the relative error is 0.0257 or 2.57%.

  b) May 2022
  $y = 236.34$
  $\hat{y} = 215.36346$

$$\varepsilon = \left|\frac{y - \hat{y}}{y}\right| = \left|\frac{236.34 - 215.36346}{236.34}\right| = 0.0888$$

Therefore, the relative error is 0.0888 or 8.88%.

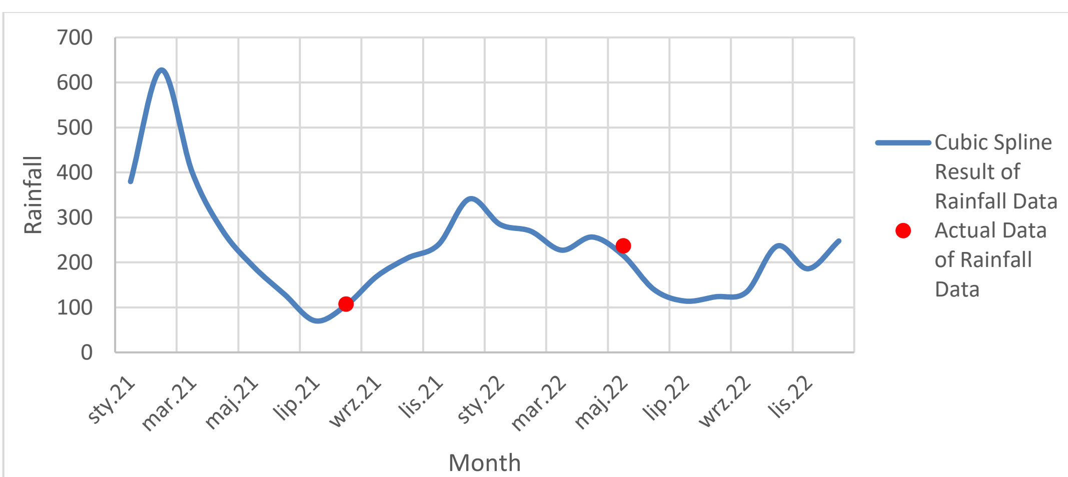

Figure 4. Comparison of the actual data and the predited data from Cubic Spline for Kepulauan Seribu

- South Jakarta and East Jakarta
  a) August 2021
  $y = 84.7$

$\hat{y} = 87.43487$

$$\varepsilon = \left|\frac{y-\hat{y}}{y}\right| = \left|\frac{84.7 - 87.43487}{84.7}\right| = 0.0323$$

Therefore, the relative error is 0.0323 or 3.23%.

b) May 2022
$y = 186.48$
$\hat{y} = 185.69940$

$$\varepsilon = \left|\frac{y-\hat{y}}{y}\right| = \left|\frac{186.48 - 185.69940}{186.48}\right| = 0.0042$$

Therefore, the relative error is 0.0042 or 0.42%.

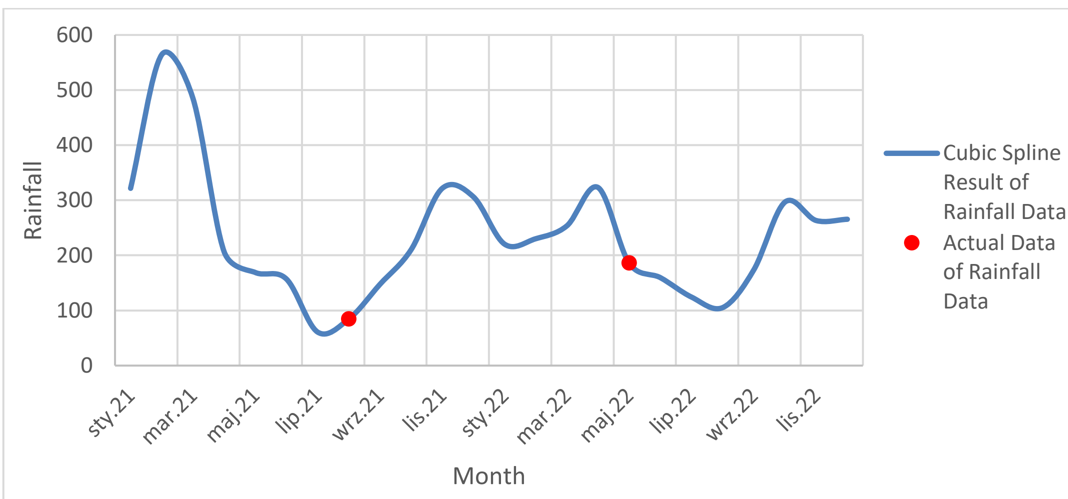

Figure 5. Comparison of the actual data and the predicted data from Cubic Spline for South and East Jakarta

- North Jakarta, West Jakarta, and Central Jakarta

a) August 2021
$y = 66.49$
$\hat{y} = 69.09901$

$$\varepsilon = \left|\frac{y-\hat{y}}{y}\right| = \left|\frac{66.49 - 69.09901}{66.49}\right| = 0.0392$$

Therefore, the relative error is 0.0392 or 3.92%.

b) May 2022
$y = 144.77$
$\hat{y} = 145.23249$

$$\varepsilon = \left|\frac{y - \hat{y}}{y}\right| = \left|\frac{144.77 - 145.23249}{144.77}\right| = 0.0032$$

Therefore, the relative error is 0.0032 or 0.32%.

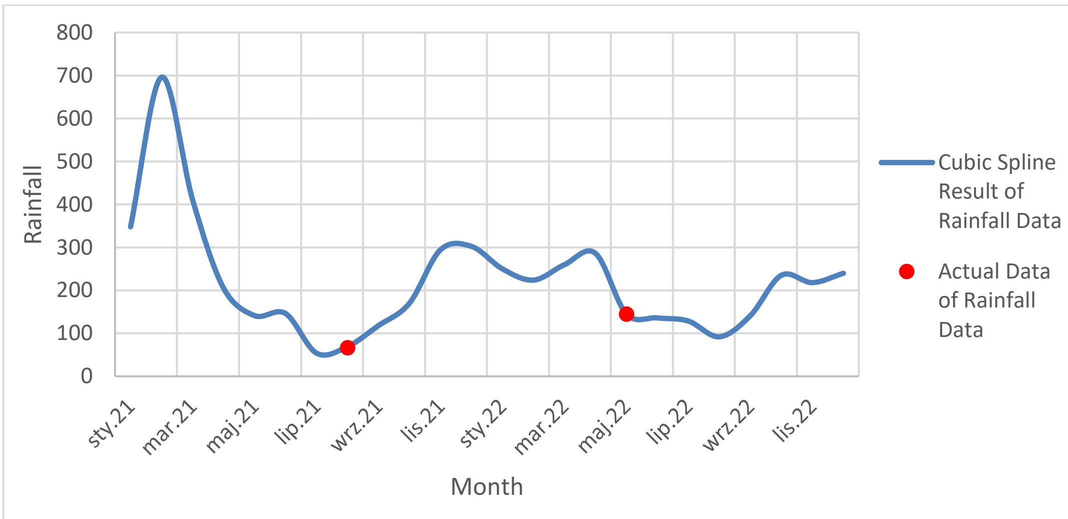

Figure 6. Comparison of the actual data and the prediction data from Cubic Spline for North, West, and Central Jakarta

Based on the relative error calculations and the graphs above, it can be observed that the rainfall data obtained from the cubic spline is not significantly different from the original rainfall data. This is indicated by the values of each relative error being less than 10% ($\varepsilon < 0.1$). Additionally, from Figures 4, 5, and 6, it can be seen that the actual data are closely concide with the cubic spline.

Based on Table 3, several pieces of information can be gleaned to assist the government in managing water resources and preventing floods. In February 2021, rainfall increased drastically in all regions of DKI Jakarta. In March 2021, rainfall decreased compared to February 2021 but still remained relatively high. Therefore, the government still needs to remain vigilant and pay close attention. Furthermore, high rainfall also occurred at the end of 2021. Hence, to prevent floods, the government should start conducting checks on drainage channels, cleaning waterways, and other measures starting towards the end of the year, around October to February. Additionally, based on Table 3, the government can also determine the priority scale regarding which areas should be given precedence, thus making the interventions more effective.

5. CONCLUSION

In this case, it can be said that cubic spline interpolation has proven to be capable of interpolating rainfall data effectively. This is evident from the relatively small values of the relative error and the closeness of the actual data to the cubic spline. Comprehensive rainfall data is highly beneficial and valuable for local governments, in this case, DKI Jakarta, to determine which areas need to be prioritized in order to prevent adverse impacts from changes in rainfall patterns. Therefore, cubic splines can be considered as an option for predicting rainfall data in the DKI Jakarta region if needed in the future.

References

[1] P. Tirkey, T. Bhattacharya, and S. Chakraborty, Water Quality Indices- Important Tools for Water Quality Assessment . *Int. J. Adv. Chem.* vol. 1, no. 1, pp. 15–28, 2013

[2] Q. xiang Huang, X. tao Xu, P. H. S. W. Kulatilake, and F. Lin, Formation mechanism of a rainfall triggered complex landslide in southwest China, *J. Mt. Sci.* vol. 17, no. 5, pp. 1128–1142, 2020, doi: 10.1007/s11629-019-5736-9

[3] R. J. H. Dunn, K. M. Willett, and D. E. Parker, Changes in statistical distributions of sub-daily surface temperatures and wind speed, *Earth Syst. Dyn.* vol. 10, no. 4, pp. 765–788, 2019, doi: 10.5194/esd-10-765-2019.

[4] P. Balamurugan, A. Ilanthirayan, M. S. Kaviraj, and S. Boobathiraja, GIS Application for Climatic Change Prediction Using Stepwise Regression International organization of Scientific Research International organization of Scientific Research, *IOSRJEN*, vol. 04, no. 02, pp. 9–17, 2014.

[5] F. Rezanezhad, M. Bieroza, A. Contosta, and P. Van Cappellen, Editorial: The cold regions in transition: Impacts on soil and groundwater biogeochemistry, *Front. Environ. Sci.* vol. 11, pp. 1–3, 2023, doi: 10.3389/fenvs.2023.1143388.

[6] M. Ellena, M. Breil, and S. Soriani, The heat-health nexus in the urban context: A systematic literature review exploring the socio-economic vulnerabilities and built environment characteristics, *Urban Clim.* vol. 34, p. 100676, 2020, doi: 10.1016/j.uclim.2020.100676

[7] C. He, L. Zhou, W. Ma, and Y. Wang, Spatial assessment of urban climate change vulnerability during different urbanization phases, *Sustain.* vol. 11, no. 8, 2019, doi: 10.3390/su11082406

[8] L. Berrang-Ford, J. D. Ford, and J. Paterson, Are we adapting to climate change? *Glob. Environ. Chang.* vol. 21, pp. 25–33, 2011, doi: 10.1016/j.gloenvcha.2010.09.012

[9] O. Adedeji, O. Reuben, and O. Olatoye, Global climate change, *J. Geosci. Environ. Prot.*, vol. 2, no. 2, pp. 114–122, 2014, doi: 10.1007/978-94-007-7353-0_3

[10] J. Yang, Y. Huang, G. Li, and Y. Li, Changes of extreme precipitation in the middle and lower reaches of the Yangtze River and their correlation with atmospheric circulation, *Front. Earth Sci.* vol. 11, pp. 1–12, 2023, doi: 10.3389/feart.2023.1162220

[11] Z. Wen, H. Liao, and A. Emrouznejad, Information representation of blockchain technology: Risk evaluation of investment by personalized quantifier with cubic spline interpolation, *Inf. Process. Manag.* vol. 58, no. 4, pp. 1–13, 2021, doi: 10.1016/j.ipm.2021.102571

[12] I. Azizan, S. A. B. A. Karim, and S. Suresh Kumar Raju, Fitting Rainfall Data by Using Cubic Spline Interpolation, *MATEC Web Conf.* vol. 225, pp. 1–9, 2018, doi: 10.1051/matecconf/201822505001

[13] G. Duan, Y. Su, and J. Fu, Landslide Displacement Prediction Based on Multivariate LSTM Model, *Int. J. Environ. Res. Public Health*, vol. 20, no. 2, 2023, doi: 10.3390/ijerph20021167

[14] S. Sangkhaphan and Y. Shu, The Effect of Rainfall on Economic Growth inThailand: A Blessing for Poor Provinces, *Economies*, vol. 8, no. 1 pp. 1–17, 2020

[15] R. Panditharathne, M. B. Gunathilake, I. M. Chathuranika, U. Rathnayake, M. S. Babel, and M. K. Jha, Trends and Variabilities in Rainfall and Streamflow: A Case Study of the Nilwala River Basin in Sri Lanka, *Hydrology*, vol. 10, no. 1, pp. 1–17, 2023, doi: 10.3390/hydrology10010008

[16] K. Rajendram, Rainfall Variability and Drought Occurrences in the Batticaloa District, Sri Lanka. *World News of Natural Sciences* 39 (2021) 30-45

[17] Wondimu Zeberie, Modelling of Rainfall-Runoff Relationship in Big-Akaki Watershed, Upper Awash Basin, Ethiopia. *World News of Natural Sciences* 27 (2019) 108-120

[18] S. Mathanraj, M. I. M. Kaleel, Rainfall Variability in the wet-dry seasons. An analysis in Batticaloa District, Sri Lanka. *World News of Natural Sciences* 9 (2017) 71-78

[19] M. I. M. Kaleel, Rainfall Variability of South-West Monsoon: A Special Study Based on Ratnapura District, Sri Lanka. *World News of Natural Sciences* 18(2) (2018) 195-202

[20] R. Janco, J. Ostrometzky, and H. Messer, In-City Rain Mapping from Commercial Microwave Links - Challenges and Opportunities, *Sensors*, vol. 23, no. 10, 2023, doi: 10.3390/s23104653

[21] B. Das and D. Chakrabarty, Lagrange's Interpolation Formula: Representation of Numerical Data by a Polynomial curve, *Int. J. Math. Trends Technol.*, vol. 34, no. 2, pp. 64–72, 2016, doi: 10.14445/22315373/ijmtt-v34p514

[22] A. Perperoglou, W. Sauerbrei, M. Abrahamowicz, and M. Schmid, A review of spline function procedures in R. *BMC Med. Res. Methodol.* vol. 19, no. 1, pp. 1–16, 2019, doi: 10.1186/s12874-019-0666-3

[23] K. Wang, A study of cubic spline interpolation. *In Sight RIVER Acad. J.* vol. 9, no. 2, pp. 1–15, 2013

[24] S. Purwani, R. A. Hidayana, V. P. Balqis, and Sukono, A Comparison of Newton's Divided Differences Interpolation and a Cubic Spline in Predicting the Poverty Rate of West Java, *Eng. Lett.* vol. 31, no. 4, pp. 1786–1791, 2023.