IRSTI 68.31.02 **DOI** https://doi.org/10.37884/4-2024/42

A. Bakirova[1,2], H. Arslan[1]*

[1]Ondokuz Mayis University, Samsun, Tukiye, 20281336@stu.omu.edu.tr, hakan.arslan@omu.edu.tr*
[2]Zhetysu University named after I.Zhansugurov, Taldykorgan, Kazakhstan

ANALYZING THE SPATIAL DISTRIBUTION OF ANNUAL PRECIPITATION ACROSS KAZAKHSTAN THROUGH DIFFERENT INTERPOLATION TECHNIQUES

Abstract

Various climatic factors, including temperature, precipitation, humidity, and wind, play a crucial role in plant growth, yield, and quality. Therefore, understanding a region's climate characteristics is essential before beginning agricultural activities to determine suitable crops for cultivation.

This study aims to map the spatial distribution of annual total precipitation in Kazakhstan by applying different interpolation methods to long-term (20-year) precipitation data from 80 meteorological stations. Precipitation data from 2000 to 2019 were obtained from the MERRA-2 Meteorological Re-Analysis database.

To identify the most accurate interpolation method, seven techniques were tested: Inverse Distance Weighting (IDW), Radial Basis Functions (RBF), Local Polynomial Interpolation (LPI), Ordinary Kriging (OK), Bayesian Empirical Kriging (BEK), CoKriging (COK), and Diffusion Kernel. Correlation coefficients and Root Mean Square Error (RMSE) values were used as criteria to evaluate each method's accuracy.

Results showed RMSE values of 35.568 for IDW, 30.777 for RBF, 30.031 for LPI, 25.751 for OK, 28.767 for BEK, 45.9 for COK, and 25.379 for Diffusion Kernel. The Diffusion Kernel method, with the lowest RMSE and highest correlation coefficient, was found to be the most suitable for creating the areal distribution map of precipitation in Kazakhstan.

Selecting the best interpolation method is essential as variations among methods can lead to different spatial distributions. Using the most appropriate interpolation method for each point dataset allows for more accurate mapping of water, soil, and climate data, leading to improved resource management.

Key words: *Kazakhstan, Precipitation, Interpolation methods, RMSE*

Introduction

Climate features such as precipitation and temperature are among the primary factors influencing agricultural production [1]. The spatial distribution of climate significantly impacts water resources [2]. In agriculture, irrigation plays a crucial role in production; however, in regions without access to irrigation, precipitation becomes the key factor for crop growth.

Traditionally, many climate data are obtained from meteorological or precipitation stations. However, limitations in the positioning of these measurement points can lead to inaccuracies in determining climate values in unmeasured areas. Therefore, data from radars or satellites have increasingly been used in hydrological studies over recent years [3]. Various climate parameters are employed across studies in multiple disciplines [4].

MERRA-2, developed by NASA's Global Modeling and Assimilation Office, is an analysis program widely and successfully used in current studies [5,6]. In a 2020 study conducted in Nepal, Hamal et al. compared precipitation data from 141 ground observation stations with precipitation data generated from MERRA and found a very high correlation between measured and MERRA-generated data.

Today, the need and interest in producing spatial climate data layers from point observations are growing. Geographic Information Systems (GIS), an essential component of spatial databases, have become indispensable in climate studies [7]. Modeling the spatial variation of precipitation, producing precipitation distribution maps, and preparing climate classifications are critical [8,9]. The preparation of spatial distribution maps of climate data holds great importance for meteorologists and water resource managers [10,11,12].

Different interpolation methods are frequently used to prepare spatial distribution maps. These methods are categorized as deterministic or stochastic. Among deterministic methods, Inverse Distance Weighted Interpolation (IDW) and Radial Basis Functions (RBF) are the most commonly applied, while the Kriging family of methods is the most widely used among stochastic techniques. In numerous studies on the spatial distribution of precipitation, interpolation methods have been employed [13,14,15,16].

In 2021, Antal et al. conducted a study in Portugal comparing four deterministic methods—Inverse Distance Weighted (IDW), Radial Basis Function (RBF), Local Polynomial Interpolation (LPI), and Global Polynomial—with three geostatistical methods: Ordinary Kriging (OK), Bayesian Kriging Regression (EBKR), and CoKriging (COK) for spatial precipitation distribution. The study concluded that the EBKR method provided more accurate results than the others.

In this study, the spatial distribution of the annual average precipitation in Kazakhstan is modeled using long-term (20-year) precipitation data from 80 city centers. Seven different interpolation methods (Inverse Distance Weighting (IDW), Radial Basis Functions (RBF), Local Polynomial Interpolation (LPI), Ordinary Kriging (OK), Bayesian Empirical Kriging (BEK), CoKriging (COK), and Diffusion Kernel) were compared to determine the most suitable interpolation method for the region.

Materials and methods

Study Area and Data

Kazakhstan's climate is primarily characterized by aridity. The climate map reveals that precipitation across the country is minimal and irregularly distributed, a pattern largely due to its central location in Eurasia and its distance from oceans. The annual distribution of atmospheric precipitation in Kazakhstan is irregular and highly seasonal. In the northern part of the country, 70-80% of annual precipitation falls during the warm season, with the majority occurring in July. In contrast, minimal summer rainfall is observed in the southern desert belt and in the foothills of the mountains in the east and southeast.

In southern regions, there are years when no rainfall occurs at all for 2-3 months during summer. In such times, "dry" rains are observed, where raindrops evaporate before reaching the ground. Precipitation is less frequent during the cold season: in the north, it makes up 20-30% of the annual norm, while in the south, it accounts for 50-60%, mainly due to cyclones passing through the southern regions.

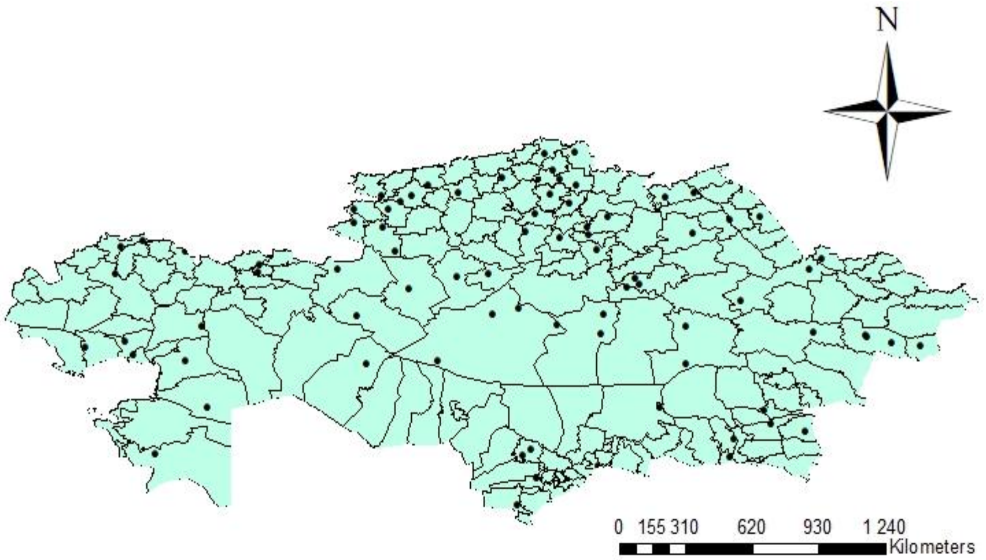

Figure 1. Study Area

Obtaining climate data for every region of Kazakhstan is challenging. Therefore, for this study, precipitation data covering the years 2000-2019 for 80 different locations in Kazakhstan were obtained from the MERRA-2 METEOROLOGICAL RE-ANALYSIS database.

Interpolation Methods

Many data related to climate, soil, and water quality are easily analyzed using Geographic Information Systems (GIS). GIS is fundamentally an information system designed to store data in flexible formats for real-time use to derive information. In the GIS environment, there are various methods for generating raster surfaces through interpolation.

Stochastic methods, also known as geostatistical methods, are primarily referred to as Kriging methods. Kriging includes several sub-methods, such as Simple Kriging, Ordinary Kriging, Universal Kriging, Indicator Kriging, Probability Kriging, Disjunctive Kriging, and CoKriging. While each method has specific applications, only CoKriging allows the inclusion of a secondary variable in the model. Deterministic methods include Inverse Distance Weighting (IDW), Local Polynomial Interpolation, Radial Polynomial, and Radial Basis Function (RBF).

In this study, precipitation distribution maps for Kazakhstan will be generated using seven of the most commonly applied interpolation methods. The correlation coefficient and Root Mean Square Error (RMSE) values will be used to compare the methods. The method with the highest correlation coefficient and lowest RMSE value will be identified as the best approach. The RMSE value is calculated using the following equation:

$$RMSE = \sqrt{\frac{\sum(z_{i^*} - z_i)^2}{n}}$$

where; z_{i^*} represents the predicted value, z_i the observed value n the number of samples.

Findings and Discussion

Table 1 presents basic descriptive statistics for the annual average precipitation data. Based on a 20-year average for Kazakhstan, precipitation levels range from 112 mm to 419 mm. The average precipitation is 235 mm, indicating very low rainfall across much of the country. This underscores the significant importance of irrigation in agricultural production.

Table 1. Basic descriptive data of the average annual precipitation

	n	Min.	Max	Average	Standard deviation	Distortion	Kurtosis
Total Rainfall (mm)	80	112	419	235	58	0.43	3.62

Enterpolasyon yöntemi yapmadan önce veriler Kolmogrov Smirnov normallik testine tabi tutulmuş ve verilerin normal dağılıma uygun olduğu belirlenmiştir. Bu nedenle haritalama aşamasında normal veriler kullanılmıştır.

As part of the study, average total precipitation distribution maps were created using precipitation data from 2000–2019 obtained from the MERRA-2 METEOROLOGICAL RE-ANALYSIS database and applying seven different interpolation methods: IDW, RBF, LPI, O.K., BEK, COK, and Diffusion Kernel (Figure 2). Comparison results for these interpolation methods are presented in Table 2.

Table 2 shows that RMSE values for the interpolation methods range from 25.38 to 45.90 mm, with correlation coefficients (R) between 0.644 and 0.898. The specific RMSE values for each method are: 35.57 for IDW, 30.78 for RBF, 30.03 for LPI, 25.75 for OK, 28.77 for BEK, 45.90 for COK, and 25.38 for the Diffusion Kernel method. These findings reveal substantial variation in performance across the different methods.

Table 2. Comparison results of the methods

		R^2	RMSE
1	IDW	0.807	35.57
2	RBF	0.850	30.78
3	LPI	0.855	30.03
4	OK	0.895	25.75
5	BEK	0.867	28.77
6	COK	0.644	45.90
7	Kernel	0.898	25.38

The highest RMSE value and lowest R value were in the COK method, while the lowest RMSE and highest R value were in the Diffusion Kernel method. It was determined that the most suitable method for preparing the areal distribution map of precipitation in Kazakhstan was Diffusion Kernel, which gave the highest correlation coefficient and the lowest RMSE value. When the map prepared according to the best method was examined, it was seen that the annual precipitation total in the southern and central parts of the country was below 120 mm. The highest precipitation was determined to be Almaty City and the Northern Kazakhstan Region. The annual total precipitation in the region with the highest precipitation varied between only 280 mm and 420 mm. When these precipitation amounts were examined, it was determined that a large part of the country was located in the arid and semi-arid climate zone. These results show that irrigation is very important for agricultural production in the country. Therefore, it is of great importance to prefer irrigation methods that use water more economically in terms of agricultural production. In addition, the use of drought-resistant varieties is important for production in regions where irrigation is not possible.

Recommendations

In this study, 7 different interpolation methods (IDW, RBF, LPI, OK, BEK, COK and Diffusion Kernel) were compared to determine the areal distribution of precipitation using long-term precipitation data of Kazakhstan. Correlation coefficient and RMSE values were used to determine the best method. As a result of the study, it was determined that the best interpolation method for the distribution of annual precipitation was Diffusion Kernel. There were significant differences between the methods and it was determined that the method with the highest error for this region was COK.

It was determined that the areal distribution of precipitation in the country showed great change and that the annual precipitation total was very low, especially in the southern parts. According to these results, it was determined that many parts of the country were arid. This situation causes irrigation management to gain even more importance in terms of agricultural production. Growing drought-resistant agricultural products is of great importance in these regions.

Areal distribution maps provide great convenience in the management of water and soil resources. However, there are great differences between interpolation methods in the preparation of areal distribution maps. Therefore, in order to obtain more accurate results in planning and observation studies, the best interpolation method for the region and the parameter under investigation must first be determined and then other studies must be started.

Gratitude: Research works were carried out under the research grant project «IRN: AP23488859 Adding values to the Central Asian wheat through the means of whole genome scanning, automated phenotyping and intensive selection» for the period 2024-2026. Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan, the results of which are given in this article.

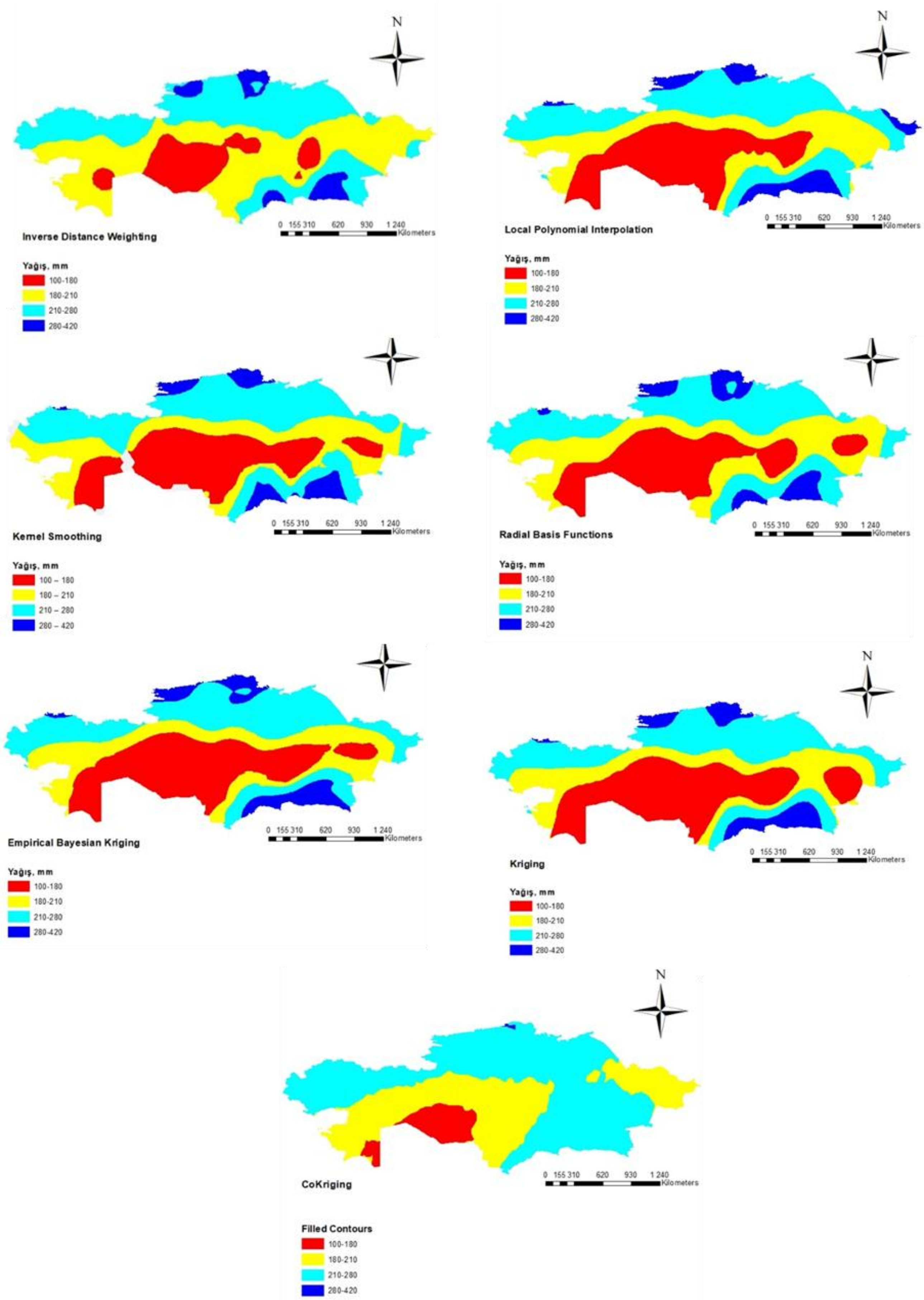

Figure 2. Areal distributions of annual total precipitation according to interpolation methods.

References

1. Borges, P. D. A., Franke, J., da Anunciação, Y. M. T., Weiss, H., Bernhofer, C., 2016. Comparison of spatial interpolation methods for the estimation of precipitation distribution in Distrito Federal, Brazil. Theoretical and applied climatology, 123(1): 335-348.

2. Wagner PD, Fiener P, Wilken F, Kumar S, Schneider K. 2012. Comparison and evaluation of spatial interpolation schemes for Daily rainfall in data scarce regions. J Hydrol 464: 388–400.

3. Cho, Y., 2020. Application of NEXRAD radar-based quantitative precipitation estimations for hydrologic simulation using ArcPy and HEC software. Water, 12(1): 273.

4. Irvem, A., Ozbuldu, M. 2019. Evaluation of satellite and reanalysis precipitation products using GIS for all basins in Turkey. Advances in Meteorology doi.org/10.1155/2019/4820136.

5. Gelaro, R., McCarty, W., Suarez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., 2017. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). J. Clim., 30: 5419–5454.

6. Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Zhao, B., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). Journal of climate, 30(14): 5419-5454.

7. Yıldırım, D., Arslan, H., Cemek, B., 2015. Assessment of spatial distribution of precipitation with different interpolation methods for Yeşilırmak catchment. 2nd International Conference on Sustainable Agriculture and Environment (2nd ICSAE), 30 september – 3 october, 2015, Konya, Turkey.

8. Fares, A., Bayabil, H. K., Zekri, M., Mattos-Jr, D., Awal, R., 2017. Potential climate change impacts on citrus water requirement across major producing areas in the world. Journal of Water and Climate Change, 8(4): 576-592.

9. Javari, M. 2017. Comparison of interpolation methods for modeling spatial variations of Precipitation in Iran. International Journal of Environmental and Science Education, 12(5): 1037-1054.

10. Benavides, R., Montes, F., Rubio, A., Osoro, K., 2007. Geostatistical modelling of air temperature in a mountainous region of Northern Spain. Agricultural and Forest Meteorology, 146(3-4): 173-188.

11. Hadi, S. J., Tombul, M. 2018. Comparison of spatial interpolation methods of precipitation and temperature using multiple integration periods. Journal of the Indian Society of Remote Sensing, 46(7): 1187-1199.

12. Pellicone, G., Caloiero, T., Modica, G., Guagliardi, I. 2018. Application of several spatial interpolation techniques to monthly rainfall data in the Calabria region (southern Italy). International Journal of Climatology, 38(9): 3651-3666.

13. Katipoğlu, O.M. 2022. Spatial analysis of seasonal precipitation using various interpolation methods in the Euphrates basin, Turkey. Acta Geophys.

14. Antal, A., Guerreiro, P.M.P., Cheval, S., 2021. Comparison of spatial interpolation methods for estimating the precipitation distribution in Portugal. Theor Appl Climatol, 145: 1193–1206.

15. Hamal, K., Sharma, S., Khadka, N., Baniya, B., Ali, M., Shrestha, M. S., Dawadi, B. 2020. Evaluation of MERRA-2 precipitation products using gauge observation in Nepal. Hydrology, 7(3): 40.

16. A.G. Rau1,Ye.M. Kalybekova1,I.S. Seitasanov1,S.B. Zulpibekova1,M.B. Arystanov (2024). The role of groundwater ın cotton crop rotatıon crops of the Shardara irrigation massif. *Izdenister Natigeler*, (1 (101), 210–220. https://doi.org/10.37884/1-2024/21

Ә. Бәкірова*[1,2], Х. Арслан[1]

[1]Университет Ондокуз Майыс, Самсун, Турция, 20281336@stu.omu.edu.tr,*
hakan.arslan@omu.edu.tr
[2] І.Жансүгіров атындағы Жетісу университеті, Талдықорган, Қазақстан

ҚАЗАҚСТАНДАҒЫ ЖЫЛДЫҚ ЖАУЫН-ШАШЫННЫҢ КЕҢІСТІК БОЙЫНША ТАРАЛУЫН ӘРТҮРЛІ ИНТЕРПОЛЯЦИЯ ӘДІСТЕРІ АРҚЫЛЫ ТАЛДАУ

Аңдатпа

Температура, жауын-шашын, ылғалдылық және жел сияқты әртүрлі климаттық факторлар өсімдіктердің өсуіне, өнімділігіне және сапасына үлкен әсер етеді. Сондықтан ауыл шаруашылығында аймақтың климаттық ерекшеліктерін түсіну егістіктерге қолайлы дақылдарды анықтау үшін маңызды.

Бұл зерттеудің мақсаты Қазақстандағы жылдық жалпы жауын-шашынның кеңістік бойынша таралуын картографиялық көрсетілімін жасау, үшін 80 метеорологиялық станциядан алынған 20 жылдық (2000-2019) жауын-шашын деректерін пайдаланып, әртүрлі интерполяция әдістерін қолдану. Жауын-шашын деректері MERRA-2 метеорологиялық қайта талдау базасынан алынды.

Нақты интерполяция әдісін анықтау үшін жеті әдіс тексерілді: Инверс қашықтықпен салмақтау (IDW), Радиалды негіз функциялары (RBF), Жергілікті полиномды интерполяция (LPI), Қарапайым Крининг (OK), Байес әдісімен эмпирикалық Кринунг (BEK), Ко-кринунг (COK) және Диффузиялық ядро. Әрбір әдістің дәлдігін бағалау критерийі ретінде корреляциялық коэффициенттер мен Орташа квадраттық қате (RMSE) мәндері пайдаланылды.

Нәтижелер IDW үшін RMSE 35.568, RBF үшін 30.777, LPI үшін 30.031, OK үшін 25.751, BEK үшін 28.767, COK үшін 45.9 және Диффузиялық ядро үшін 25.379 болды. Ең төмен RMSE және ең жоғары корреляциялық коэффициентпен Диффузиялық ядро әдісі Қазақстандағы жауын-шашынның кеңістіктік таралу картасын жасау үшін ең қолайлы болып табылды.

Ең жақсы интерполяция әдісін таңдау маңызды, өйткені әдістер арасындағы айырмашылықтар әртүрлі кеңістіктік таралымдарға әкелуі мүмкін. Әрбір нүктелік дерек үшін ең қолайлы интерполяция әдісін пайдалану су, топырақ және климат деректерінің дәл карталарын жасауға мүмкіндік береді, бұл ресурстарды басқаруды жақсартады.

Кілт сөздер: Қазақстан, Жауын-шашын, Интерполяция әдістері, RMSE

А. Бакирова*[1,2], Х. Арслан[1]

[1]Университет Ондокуз Майыс, Самсун, Турция, 20281336@stu.omu.edu.tr, hakan.arslan@omu.edu.tr*

[2]Жетысуский университет имени И.Жансүгурова, Талдыкорган, Казахстан

АНАЛИЗ ПРОСТРАНСТВЕННОГО РАСПРЕДЕЛЕНИЯ ГОДОВЫХ ОСАДКОВ В КАЗАХСТАНЕ С ПОМОЩЬЮ РАЗЛИЧНЫХ ИНТЕРПОЛЯЦИОННЫХ ТЕХНИК

Аннотация

Различные климатические факторы, включая температуру, осадки, влажность и ветер, играют ключевую роль в росте растений, урожайности и качестве. Поэтому понимание климатических характеристик региона является важным перед началом сельскохозяйственной деятельности для определения подходящих культур для выращивания.

Целью данного исследования является картирование пространственного распределения годового общего количества осадков в Казахстане с применением различных методов интерполяции к долгосрочным (20-летним) данным об осадках с 80 метеорологических станций. Данные об осадках за период с 2000 по 2019 годы были получены из базы данных метеорологических переанализов MERRA-2.

Для определения наиболее точного метода интерполяции были протестированы семь техник: метод обратного расстояния (IDW), радиальные базисные функции (RBF), локальная полиномиальная интерполяция (LPI), обыкновенное кригинг (OK), байесовский эмпирический кригинг (BEK), кокригинг (COK) и метод диффузионного ядра. Коэффициенты корреляции и значения среднеквадратичной ошибки (RMSE) использовались в качестве критериев для оценки точности каждого метода.

Результаты показали значения RMSE: 35.568 для IDW, 30.777 для RBF, 30.031 для LPI, 25.751 для OK, 28.767 для BEK, 45.9 для COK и 25.379 для метода диффузионного ядра. Метод диффузионного ядра, обладающий наименьшим RMSE и наибольшим коэффициентом корреляции, оказался наиболее подходящим для создания карт распределения осадков в Казахстане.

Выбор наилучшего метода интерполяции является важным, поскольку различия между методами могут привести к различным пространственным распределениям. Использование наиболее подходящего метода интерполяции для каждого набора данных позволяет более точно картировать данные о воде, почве и климате, что ведет к улучшению управления ресурсами.

Ключевые слова: Казахстан, осадки, методы интерполяции, RMSE