JEODEZİ VE JEOİNFORMASYON DERGİSİ
JOURNAL OF GEODESY AND GEOINFORMATION

ISSN : 2147-1339
e-ISSN: 2667-8519

Araştırma Makalesi / *Research Article*
Yıl / *Year:* 2025 **Cilt** / *Volume:* 12 **Sayı** / *Issue*: 1
Sayfa / *Page*: 42-57
Dergi No / *Journal No*: 121
Doi: 10.9733/JGG.2025R0004.E

Performance assessment of interpolation techniques for investigation Covid-19 spread in Türkiye

Duygu Arican[1*], Nursu Tunalioglu[1]

[1]*Yildiz Technical University, Campus of Davutpaşa, Faculty of Civil Engineering, Department of Geomatic Engineering, Esenler, Istanbul, Türkiye.*

Abstract: *Throughout history, viruses have posed significant threats to human life and health. In the context of the historical pandemics, Covid-19, rapidly spread across continents and was declared a pandemic by the World Health Organization on 11 March 2020. The first case in Türkiye was detected on the same date. Understanding the spatial distribution of the Covid-19 is crucial for effective public health planning and intervention. Geographic Information Systems (GIS) technology can be leveraged as a visualization aid to map the geographical distribution of the disease, the potential risk factors, and the resources available for treatment and prevention. To effectively map and analyze the spatial distributions, and local/global dynamics of the Covid-19 virus, various GIS-based interpolation methods were employed. To understand these dynamics, this study presents a detailed spatial analysis using interpolation methods to evaluate spatiotemporal changes on seasonal levels in the Covid-19 pandemic in Türkiye. Seasons investigated in a 1-year period were determined as follows: Spring, from 20 March 2021 to 18 June 2021; Summer, from 19 June 2021 to 17 September 2021; Autumn, from 18 September 2021 to 17 December 2021; and Winter, 18 December 2021 to 18 March 2022. Seasonal case distribution maps produced from city-level and district-level seasonal case data utilizing Inverse Distance Weighting (IDW), Radial Basis Function, Spline interpolation, and Empirical Bayesian Kriging (EBK) interpolation methods. Finally, the spread of Covid-19 in Türkiye was investigated on the seasonal scale, and interpolation results were assessed by standard deviation, mean absolute error, and root mean square error. The results of this study demonstrated that the period of highest incidence of cases of Covid-19 in Türkiye was winter. Overall, when considering error metrics, EBK and IDW generally proved to be the most reliable methods across different scales and conditions. In contrast, Spline interpolation's tendency to overfit the data made it less suitable for these datasets.*

Keywords: Covid, Pandemic, Data visualization, Spatial interpolation, Performance assessment

Türkiye'de Covid-19 yayılımının araştırılmasında enterpolasyon tekniklerinin performans değerlendirmesi

Öz: *Virüsler, tarih boyunca insan yaşamı ve sağlığı için önemli tehditler oluşturmuştur. Tarihsel pandemiler bağlamında Covid-19, kıtalar arasında hızla yayılmış ve 11 Mart 2020 tarihinde Dünya Sağlık Örgütü tarafından pandemi olarak ilan edilmiştir. Türkiye'deki ilk vaka da aynı tarihte tespit edilmiştir. Covid-19'un mekânsal dağılımının anlaşılması, etkili halk sağlığı planlaması ve müdahalesi için çok önemlidir. Coğrafi Bilgi Sistemleri (CBS) teknolojisi, hastalığın coğrafi dağılımını, potansiyel risk faktörlerini tedavi ve önleme için mevcut kaynakların haritalanması için bir görselleştirme aracı olarak kullanılabilmektedir. Covid-19 virüsünün mekânsal dağılımlarının ve yerel/küresel dinamiklerinin etkin bir şekilde haritalanması ve analiz edilmesi için çeşitli CBS tabanlı enterpolasyon yöntemleri kullanılmaktadır. Bu dinamiklerin anlaşılabilmesi için bu çalışma, Türkiye'deki Covid-19 pandemisindeki mekânsal-zamansal değişikliklerinin değerlendirilmesi üzerine enterpolasyon yöntemlerini kullanarak mevsimsel düzeyde ayrıntılı bir mekânsal analiz sunmaktadır. Bir yıllık dönemde incelenen mevsimler: İlkbahar, 20 Mart 2021 - 18 Haziran 2021 tarihleri arası; Yaz, 19 Haziran 2021 - 17 Eylül 2021 tarihleri arası; Sonbahar, 18 Eylül 2021 - 17 Aralık 2021 arası; ve Kış, 18 Aralık 2021 - 18 Mart 2022 tarihleri arası olarak belirlenmiştir. Mevsimsel dağılım haritaları, Ters Mesafe Ağırlıklandırma (IDW), Radyal Temelli Fonksiyon, Spline enterpolasyonu ve Ampirik Bayesian Kriging (EBK) enterpolasyon yöntemleri kullanılarak şehir ve ilçe düzeyinde mevsimsel vaka verilerinden üretilmiştir. Son olarak, Covid-19'un Türkiye'deki yayılımı mevsimsel ölçekte incelenmiş ve enterpolasyon sonuçları standart sapma, ortalama mutlak hata ve kök ortalama kare hatası ile değerlendirilmiştir. Bu çalışmanın sonuçları, Türkiye'de Covid-19 vakalarının en sık görüldüğü dönemin kış mevsimi olduğunu göstermiştir. Genel olarak, hata ölçütleri dikkate alındığında, EBK ve IDW'nin farklı ölçekler ve koşullar arasında en güvenilir yöntemler olduğu kanıtlanmıştır. Buna karşılık, Spline'ın verilere aşırı uyum sağlama eğilimi, onu bu veri kümeleri için daha az uygun hale getirmiştir.*

Anahtar Sözcükler: Covid, Pandemi, Veri görselleştirme, Mekânsal enterpolasyon, Performans değerlendirmesi

* Sorumlu Yazar/Corresponding Author: Tel: +90 212 383 5298

Geliş Tarihi/Received: 03.09.2024
Kabul Tarihi/Accepted: 18.11.2024

0000-0002-4618-4357, duyguaricann@gmail.com (Arican D.)*
0000-0001-9345-5220, ntunali@yildiz.edu.tr (Tunalioglu N.)

1. Introduction

Throughout history, viruses have posed significant threats to human life and health. From the Spanish Flu of 1918 to the more recent H1N1 pandemic, viral outbreaks have repeatedly demonstrated their potential for widespread morbidity and mortality (Johnson & Mueller, 2002; Taubenberger & Morens, 2006). The global incidence of outbreaks, defined as both pandemics and epidemics, has been considerable throughout history. While epidemics affect a substantial proportion of individuals within a given society, pandemics, in contrast to epidemics, have a significant impact on economic and social order across a wider geographical area (Jedwab et al., 2021).

In the context of historical pandemics, the current SARS-CoV-2 (Covid-19) virus-induced pandemic represents a modern-day crisis with an unforeseen global impact. In late December 2019, the initial cases of the novel coronavirus were identified in Wuhan, China, presenting with atypical pneumonia (Zhu et al., 2020). The initial situation report prepared by the World Health Organization (WHO) indicated that, as of the 20th of January 2020, there had been a total of 282 confirmed cases of the disease. Of these, 278 were in China, 2 were in Thailand, 1 was in Japan and 1 was in Korea. Subsequently, the virus spread rapidly across Asia, Europe, America and Africa, leading to its declaration as a pandemic by the WHO on 11 March 2020. On the same date, the first case of Covid-19 was detected in Türkiye and subsequently announced by the Minister of Health, Fahrettin Koca (URL-1). On 5 May 2023, the WHO announced that the status of the Covid-19 pandemic had been revised from a public health emergency of international importance to a pandemic due to the ongoing infectiousness and disease (URL-2). The rapid spread and high transmissibility of Covid-19 have underscored the critical need for effective public health strategies and interventions.

The trajectory of the pandemic in Türkiye mirrored global trends, with waves of infections prompting a series of public health interventions, including lockdowns, travel restrictions, and vaccination campaigns (URL-3). The impact on Türkiye's healthcare system and economy has been profound, highlighting the importance of understanding the spatial and temporal dynamics of the virus.

Understanding the spatial distribution of the virus is crucial for effective public health planning and intervention. At this point, Geographic Information System (GIS) has proven invaluable in mapping and analyzing the spatial spread of infectious diseases, offering insights that are critical for targeted responses. For instance, GIS technology can be leveraged as a visualization aid to map the geographical distribution of the disease, the potential risk factors, and the resources available for treatment and prevention (Jia et al., 2023; Kang et al., 2020; Kumar et al., 2020; Murugesan et al., 2020).

A variety of GIS applications, including interpolation methods, have been employed to enhance comprehension of the spatial distribution and local and global dynamics of the Covid-19 disease (Franch-Pardo et al., 2020). Ibarra-Bonilla et al. (2023) used the Inverse Distance Weighting (IDW) method to analyze the spatial distribution Covid-19 cases and deaths in the state of Chihuahua, Mexico, from Winter 2019-2020 to Fall 2021. The study emphasized the significance of human mobility and socio-economic elements in the propagation of the disease and concluded that the IDW method was a valuable tool for elucidating the spatial dissemination of the virus and could potentially contribute to its management. Murugesan et al. (2020), employed the use of IDW interpolation to predict the spread of the novel coronavirus (Covid-19) in India. This approach allowed for the identification of high-risk areas, providing valuable information for government monitoring and response strategies. The study's findings suggest that IDW interpolation is an effective method for predicting the spread of the disease in India. Furthermore, the study proposes that this method can be applied to other countries with similar distance and density characteristics. Cong Nhut (2023) compares the Kriging methodology with traditional techniques for predicting the number of deaths resulting from Covid-19 in Vietnam. The findings demonstrate that the Kriging approach results in a smaller forecast error and is therefore a suitable method for developing a predictive model.

Despite the extensive literature on Covid-19, there is a notable absence of studies focusing on the geographic distribution of cases, particularly in Türkiye. Some studies have focused on the relationship between vehicle traffic and air quality during the pandemic in Türkiye (Alemdar et al., 2021; Bugdayci et al., 2023; Kotan & Erener, 2023). Kırlangıçoğlu (2022) investigated the impact of regional characteristics on the spatial distribution of the pandemic in Türkiye. The study employed the use of IDW interpolation and multiple linear regression analysis to examine the relationship between provincial incidence rates and 18 explanatory variables.

Although there are studies in the literature comparing interpolation methods (Ikechukwu et al., 2017; Li & Heap, 2008), there is a gap in the existing literature regarding the investigation of the performance of interpolation methods utilizing data regarding Covid-19. This is particularly regarding the visualization of the spatial distribution of the virus, which has not been adequately addressed in previous studies. This study aims to evaluate the performance of the selected interpolation methods regarding the spatial distribution of Covid-19 in Türkiye. For this purpose, the city-level data on the number of cases of Covid-19 in Türkiye was employed to derive interpolated maps. To assess the impact of the pandemic, the data were arranged on a seasonal timescale, and a district-level dataset was produced by weighting the population of each district. Both city-level and district-level data were randomly divided as 90% for interpolation models and 10% for the accuracy assessment of the interpolation. Interpolated maps produced from city-level and district-level seasonal case data utilizing IDW, Radial Basis Function (RBF), Spline interpolation, and Empirical Bayesian Kriging (EBK) interpolation methods. Finally, the spread of Covid-19 in Türkiye was investigated on the seasonal scale, and interpolation results were assessed by standard deviation (SD), mean absolute error (MAE), and root mean square error (RMSE).

2. Methodology

The analyses in this study are implemented as follows: first, the data sets were collected and organized, statistical quantities were calculated, and interpolation methods were applied to the data sets. Finally, accuracy was assessed by statistical results, and the spread of the pandemic was evaluated.

2.1 Study Area

This study has been realized using the seasonal coronavirus case numbers in Türkiye. The study area, Türkiye, is located at the intersection of two continents, centered on the Anatolian peninsula in Western Asia and within a minor portion of Southeast Europe. Türkiye comprises 81 cities, and 922 districts.

The data set comprised 90% of cities and districts, which were randomly selected for use in the construction of interpolated maps. The remaining 10% of cities and districts were reserved for the purpose of map control. In other words, the accuracy of the interpolated maps was assessed using 92 districts and 8 cities. Figure 1 provides a visual representation of the geographical scope of the study, illustrating the borders of cities and districts and the cities and districts utilized as a control.

2.2 Materials and Methods

The Covid-19 case number data in this study were provided by the Turcovid19 (https://turcovid19.com) open data platform. The Turcovid19 platform presents Türkiye's Covid-19 pandemic data in xls and csv file extensions, ready for analysis (Uçar et al., 2020). The Python programming language was used to organize and analyze the data for this investigation. The spatiotemporal studies were performed using the ArcGIS Pro 3.3.1 software.

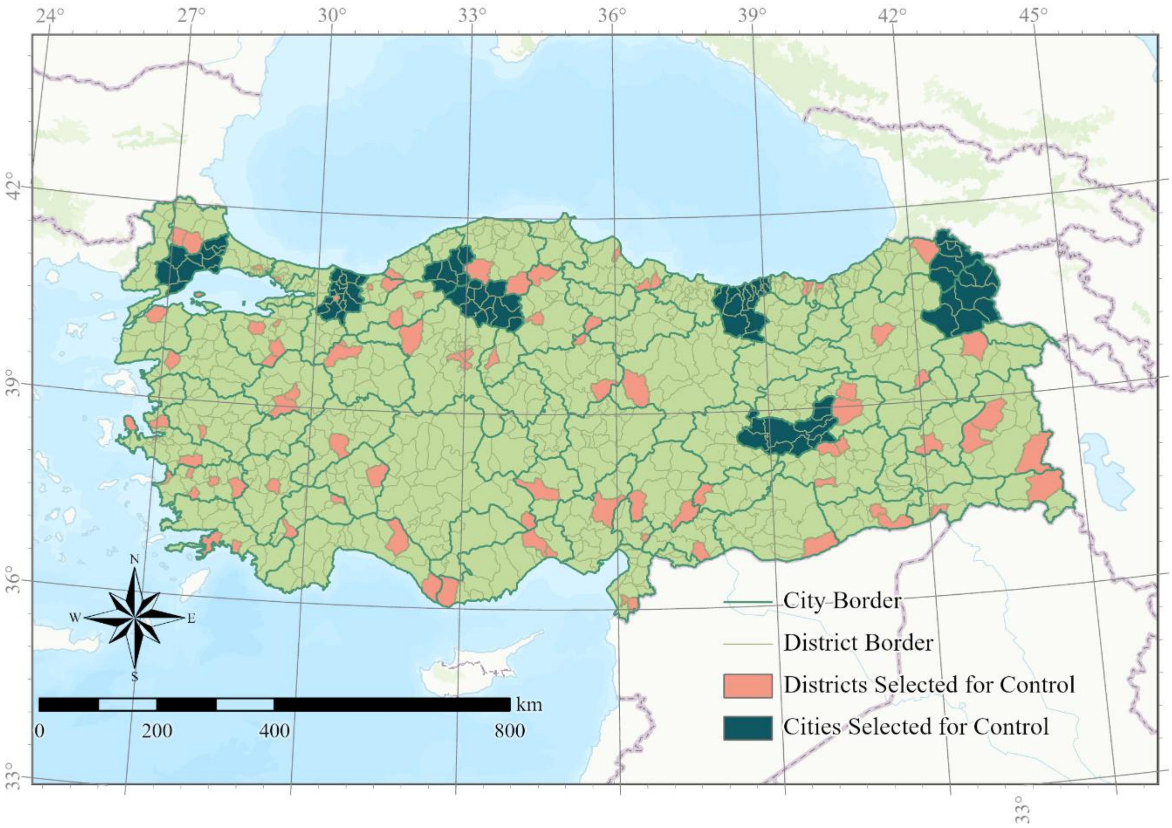

Figure 1: *Map of the study area, Türkiye*

Country-level data is available from the first case of Covid-19 in Türkiye, which was announced by the Turkish Ministry of Health on March 11, 2020, to May 31, 2022, on the Turcovid19 platform. In addition, city-level data is available as the number of cases per 100 000 people per week between February 8, 2021, and March 25, 2022. To examine the study on a seasonal scale, approximated equinox dates were determined as the start dates of the season. Consequently, the study encompasses one year following the onset of the pandemic in Türkiye, encompassing the dates between March 20, 2021, and March 18, 2022.

Figure 2 presents the weekly new cases and cumulative cases of the Covid-19 outbreak in Türkiye between March 20, 2021, and March 18, 2022. The study revealed that the cumulative number of cases, 3 149 094 at the beginning, reached 14 663 508 by the end of the one-year study period. Upon examination of the number of new cases on the graph, it becomes apparent that there is a discernible pattern of peaks in the increase in the number of cases. The maximum weekly number of new cases was observed between January 29, 2022 - February 4, 2022, with 708 159 new cases detected. In contrast, the minimum weekly number of new cases was observed in the week of July 3, 2021 - July 9, 2021, with 34 933 new cases detected.

The analyzed seasons comprise equal time intervals, each of which is 13 weeks long, equivalent to 91 days. Table 1 presents the temporal intervals demarcating the seasons investigated in this research.

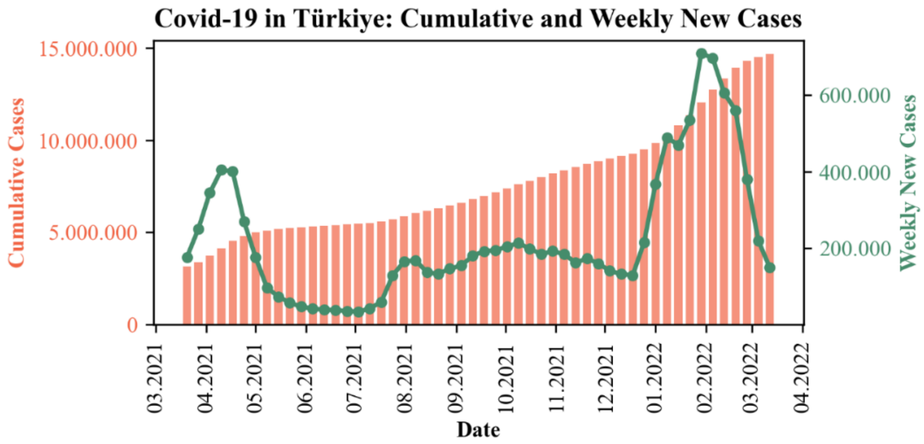

Figure 2: *Weekly new case and cumulative case numbers of the Covid-19 outbreak in Türkiye (March 20, 2021- March 18, 2022)*

Table 1: *Time intervals of the seasons analyzed*

#	Season	Time interval	Weeks	Days
1	Spring	20 March 2021 – 18 June 2021	13	91
2	Summer	19 June 2021 – 17 September 2021	13	91
3	Autumn	18 September 2021 – 17 December 2021	13	91
4	Winter	18 December 2021 – 18 March 2022	13	91
		Overall	52	364

The seasonal and cumulative case numbers of the Covid-19 outbreak in Türkiye from March 20, 2021, to March 18, 2022, are presented in Figure 3. The graph demonstrates the cumulative and seasonal incidence of novel coronavirus infections in Türkiye. The cumulative number of cases, represented by the red line, exhibits a consistent increase from approximately 5 million at the end of the spring season to nearly 15 million during the winter months. The green bars illustrate the seasonal total number of new cases. In the one-year data analyzed in this study, the proportion of new cases recorded in spring was 20.4%, which decreased to 12.3% in summer and then increased slightly to 20% in autumn. The highest proportion, 47.3%, was observed in winter. This seasonal distribution indicates a substantial increase in new cases during the winter months, contributing significantly to the cumulative increase.

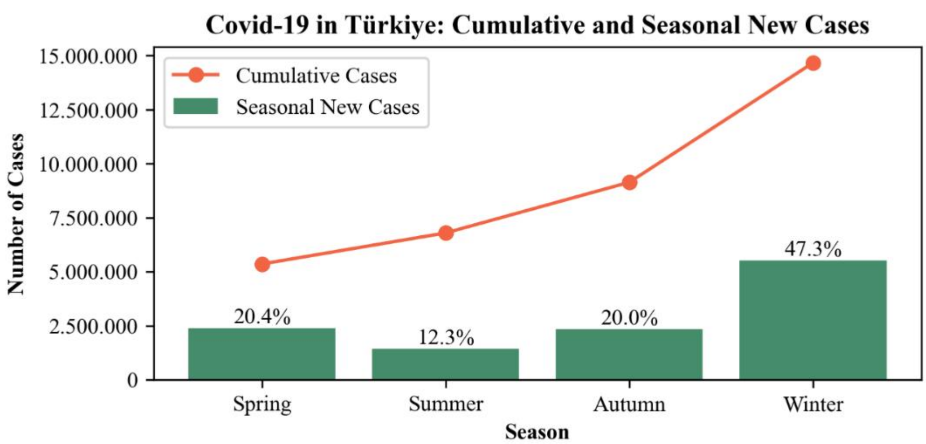

Figure 3: *Seasonal case and cumulative case numbers of the Covid-19 outbreak in Türkiye (March 20, 2021- March 18, 2022)*

Population data for Türkiye was collected from the Turkish Statistical Institute (TurkStat) at the city and district levels. According to the data, the country-level population is 84 680 273 in 2021 and 85 279 553 in 2022. The maximum population lives in İstanbul, and the minimum population lives in Ardahan both in 2021 and 2022. While İstanbul's population is 15 840 900 in 2021, it is increasing to 15 907 951 in 2022. Ardahan's population is 39 130 in 2021, slightly changed in 2022 as it is 39 123. City populations are illustrated in Figure 4 in alphabetical order.

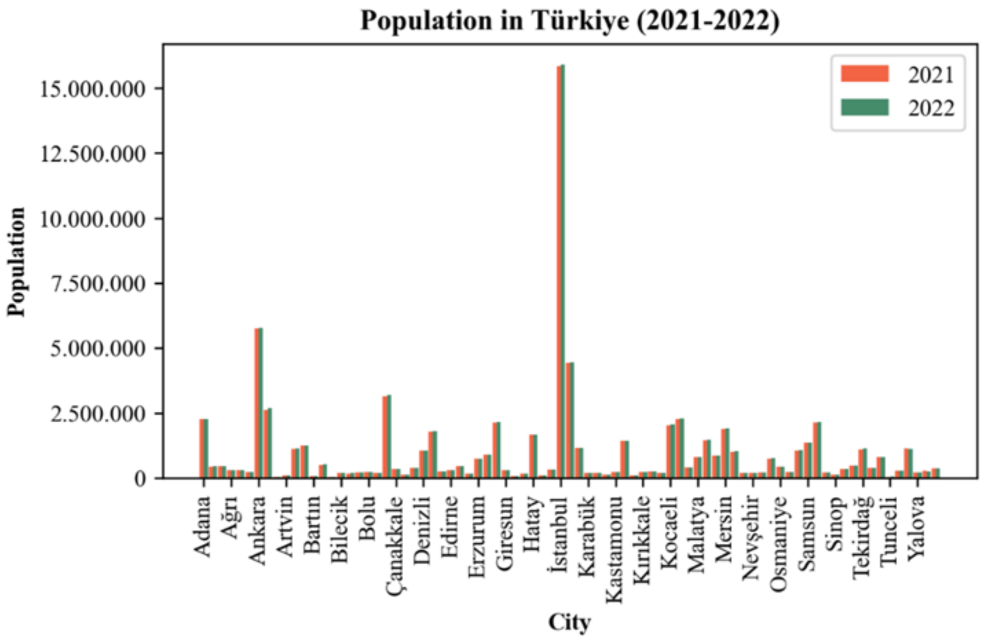

Figure 4: *Population Statistics of Turkish Cities (2021-2022)*

The average district population in Türkiye is 87 030 in 2021 and 87 646 in 2022. The lowest population is recorded in Konya, Yalıhüyük as 1532 and 1710 in the years 2021 and 2022, respectively. In addition, the highest population is recorded in İstanbul, Esenyurt as 977 489 and 983 571, in the years 2021 and 2022, respectively.

The number of Covid-19 cases at the district level was not recorded by the authorities. Therefore, assuming it changes directly to the district population data, the seasonal number of Covid-19 cases was calculated at the district-level from city-level data by weighting it according to the population data.

Based on the seasonal Covid-19 case data at the city and district level and population data size, it is inconvenient to illustrate as a graph. Consequently, GIS tools are beneficial for analyzing spatial data, improving decision-making processes, and enhancing mapping and geographic analysis efficiency and accuracy. Additionally, interpolation techniques are invaluable in this context. These methods allow estimating Covid-19 case numbers in areas where data might be sparse or unavailable by using known data points to predict unknown values. Interpolation techniques can provide a more detailed and continuous surface of infection rates, helping public health officials identify potential hotspots, allocate resources more effectively, and implement targeted interventions. By leveraging these interpolation methods within GIS platforms, deeper insights are gained, and more precise spatial representations of the pandemic's impact are created.

In this study, seasonal changes in the spreading of Covid-19 are evaluated using IDW, RBF, Spline interpolation, and EBK interpolation techniques for city-level and district-level data. Employing multiple interpolation methods aims to assess their comparative effectiveness in accurately modeling the spatial distribution of Covid-19 cases across different geographic scales. Each interpolation method has unique characteristics and assumptions that can influence the results. For instance, IDW emphasizes closer data points more heavily, making it useful for areas with dense data. RBF can model smooth variations over space, while Spline interpolation can capture local trends and variations. EBK considers spatial autocorrelation and can provide statistically optimal predictions (Ikechukwu et al., 2017). By applying a range of techniques, it is aimed to identify the most suitable methods for different scales of analysis and data distributions, thereby enhancing the robustness and reliability of our spatial analysis. This comprehensive approach allows for a better understanding of the seasonal patterns of Covid-19 and supports more targeted and effective public health interventions at city and district levels.

Table 2: *Exposure risk classifications are according to seasonal Covid-19 case numbers at city and district levels*

#	Number of seasonal Covid-19 cases (100.000)		Exposure risk
	City-level	District-level	
1	≤ 750	≤ 400	Very Low
2	≤ 1500	≤ 800	Low
3	≤ 2250	≤ 1200	Medium-Low
4	≤ 3000	≤ 1600	Medium
5	≤ 3750	≤ 2000	Medium-High
6	≤ 4500	≤ 2400	Moderate
7	≤ 5250	≤ 2800	Moderate-High
8	≤ 6000	≤ 3200	High
9	≤ 6750	≤ 3600	Very High

Seasonal Covid-19 cases data at the city and district level were organized as point data centered on the city and district locations. Then the raster data produced due to interpolation techniques were classified according to the determined statistical quantities proposed in Table 2.

The data distribution for each season was considered in determining the threshold values presented in Table 2. The data sets at the City and District levels were evaluated separately for this purpose. The average range was calculated using the range values (Table 3) of the seasonally separated data sets. The average range was then divided by nine, as the visualization will employ nine classes. The values were rounded to multiples of 250 for the city level and multiples of 100 for the district level to establish the threshold values.

The interpolated maps were evaluated through a comparison with the data allocated for control purposes. The seasonal data on cases of coronavirus disease 2019 (Covid-19) in provinces and districts determined for control purposes were compared with the values obtained by interpolations. The SD, MAE and RMSE were calculated.

3. Results and Discussion

In this study, interpolation methods were applied to evaluate spatiotemporal changes in the Covid-19 pandemic in Türkiye. Investigated seasons in 1 year period determined as follows: Spring, from 20 March 2021 to 18 June 2021; Summer, from 19 June 2021 to 17 September 2021; Autumn, from 18 September 2021 to 17 December 2021; and Winter, 18 December 2021 to 18 March 2022. The separation of these seasonal periods was conducted to capture variations in the number of cases in different periods, which can be influenced by factors such as weather, social behavior and public health interventions. Similarly, the objective of seasonal evaluation of Covid-19 data is also addressed by Ibarra-Bonilla (2023). All interpolation methods were applied using by seasonal number of Covid-19 cases per 100 000 people data on both city-level and district-level.

Table 3: *Descriptive statistics of Covid-19 seasonal number of cases data*

Level	Season	Average	Median	Minimum	Maximum	Range
City	Summer	2410.8	2425.4	410.2	5034.4	4624.2
City	Spring	1818.4	1687.6	411.2	4839.7	4428.5
City	Autumn	2753.2	2629.4	284.9	5394.8	5110.0
City	Winter	5497.0	5692.7	542.6	10723.1	10180.5
District	Summer	200.7	86.9	1.7	2527.8	2526.1
District	Spring	151.4	58.1	0.7	2545.2	2544.5
District	Autumn	229.2	87.8	0.8	3137.7	3136.9
District	Winter	457.6	191.8	1.6	5548.1	5546.6

Table 3 represents the descriptive statistical results of the data set used in this study. When one year of data was analyzed, it was already determined that most of the cases occurred in the winter season. Following the winter season, the distribution of the number of Covid-19 cases orders as spring, autumn, and summer, respectively. The observed increase in cases during the winter season can be attributed to the fact that cold weather conditions lead to an increased spread of respiratory tract infections. This increase is corroborated by McClymont and Hu (2021), who suggest that temperature is an important factor in the transmission of Covid-19.

Firstly, the IDW interpolation method was employed to process the city-level seasonal Covid-19 data for each season. The maps were reclassified according to the specified threshold values, thus providing a more straightforward interpretation. The maps produced with the IDW method and subsequently reclassified are presented in Figure 5.

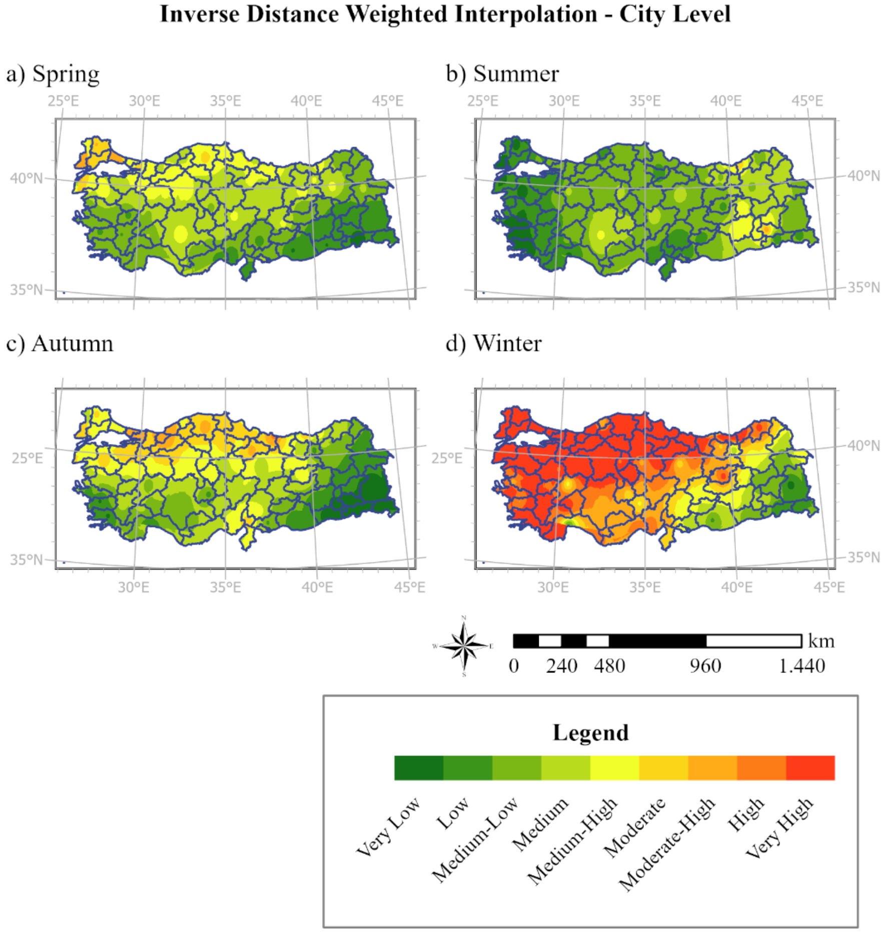

Figure 5: *Covid-19 city-level seasonal number of cases distribution maps using IDW method: a) Spring, b) Summer, c) Autumn, d) Winter*

The RBF interpolation method was applied to the city-level seasonal data set for each season relating to the incidence of Covid-19. The maps were reclassified according to the specified threshold values, thus enabling a more straightforward interpretation of the data. The maps produced with the RBF method and subsequently reclassified are presented in Figure 6.

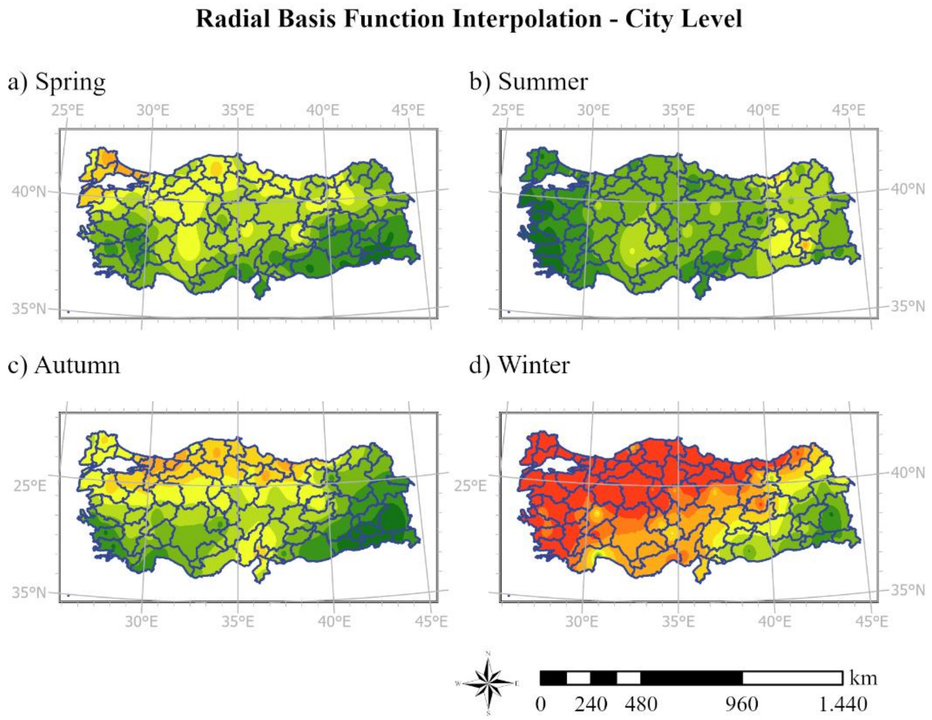

Figure 6: *Covid-19 city-level seasonal number of cases distribution maps using RBF method: a) Spring, b) Summer, c) Autumn, d) Winter*

The Spline interpolation method was utilized for the analysis of city-level seasonal data for Covid-19 for each season. All maps were reclassified according to the specified threshold values, thus providing a more straightforward interpretation. The maps produced with the Spline method and subsequently reclassified are presented in Figure 7.

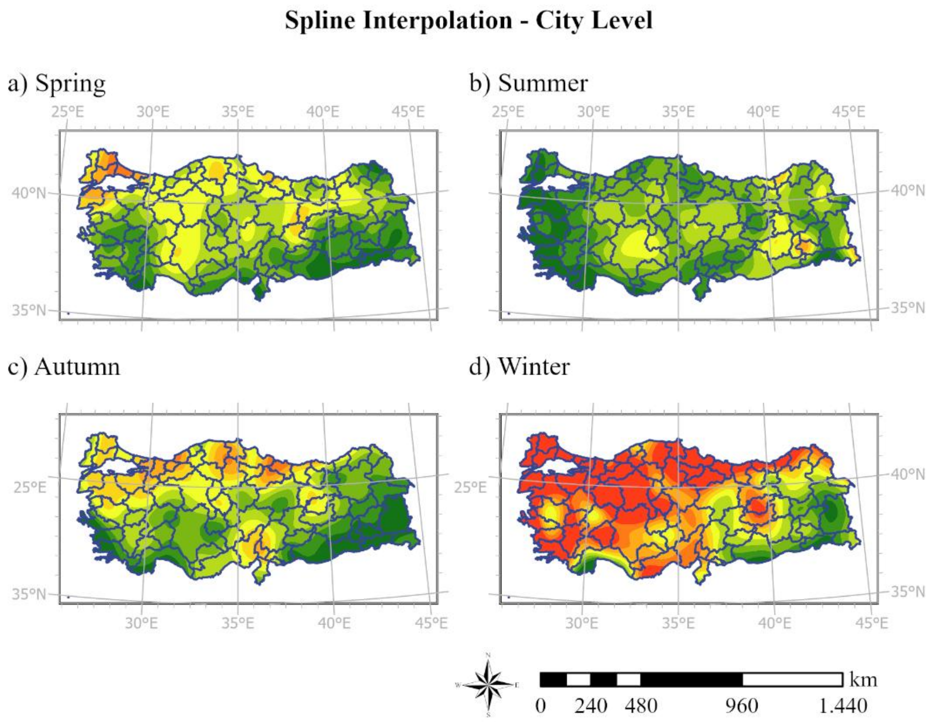

Figure 7: *Covid-19 city-level seasonal number of cases distribution maps using Spline interpolation method: a) Spring, b) Summer, c) Autumn, d) Winter*

EBK interpolation method was applied to the city-level seasonal Covid-19 data for each season. All the maps were reclassified by the specified threshold values for easier interpretation of maps. The produced with the EBK interpolation method and reclassified maps are represented in Figure 8.

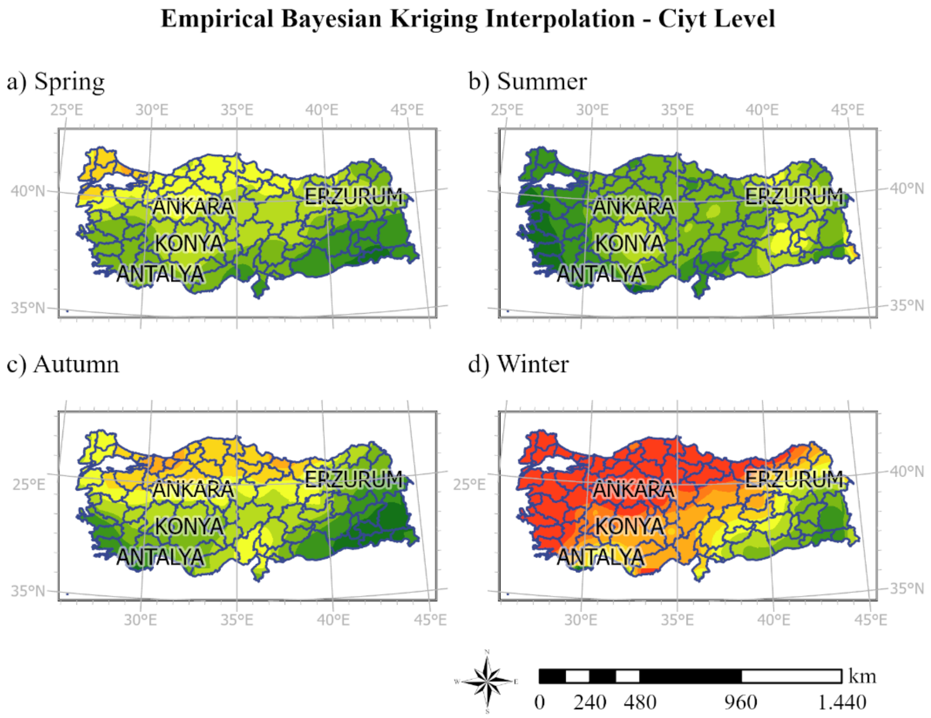

Figure 8: *Covid-19 city-level seasonal number of cases distribution maps using EBK interpolation method: a) Spring, b) Summer, c) Autumn, d) Winter*

Interpolation techniques have been utilized to generate continuous surfaces across the country, employing data on the number of cases of Covid-19 at the city level on a seasonal basis. This approach allows for the observation of geographical areas affected by the disease throughout the country and is employed for the visualization of spatial variations in case distribution. Notably, there was an increase in case densities during the winter and spring seasons in larger cities by area, such as Konya, Ankara, Antalya, and Erzurum, as can be observed on the maps.

Similarly, district-level seasonal number of cases distribution maps for Covid-19 were produced at the district level by performing the same interpolation methods to assess interpolation performance on different resolutions of the data. Additionally, this approach has enabled the determination of which districts of cities are most influenced by the disease.

Figure 9 illustrates the mapping of districts where cases of Covid-19 are concentrated seasonally, as determined by the IDW method. A comparison of city-level and district-level maps reveals that, while the concentration is more widespread at the city level, at the district level, the density in cities covering larger areas, decreases under the distribution of population.

The RBF method employed in the production of the interpolation map of district-level Covid-19 data does not reveal a strong concentration of cases in specific areas, as observed in Figure 10. Figure 11 illustrates the seasonal distribution of cases of Covid-19 generated by Spline interpolation. As can be observed in the maps, the regions where cases are concentrated exhibit a similar pattern to that observed with the IDW method.

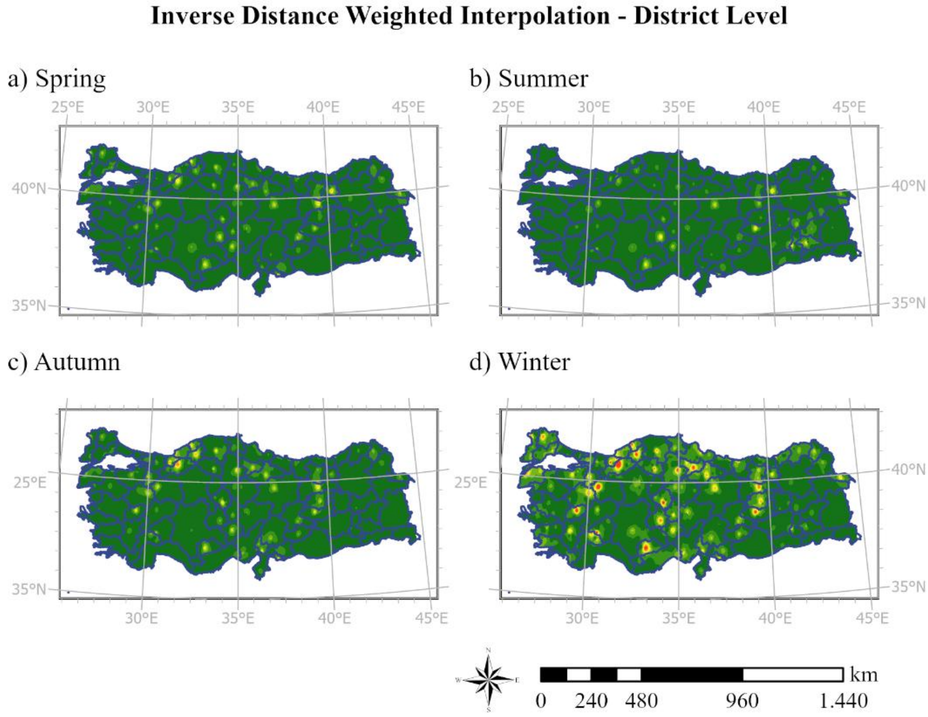

__Figure 9:__ Covid-19 district-level seasonal number of cases distribution maps using IDW interpolation method: a) Spring, b) Summer, c) Autumn, d) Winter

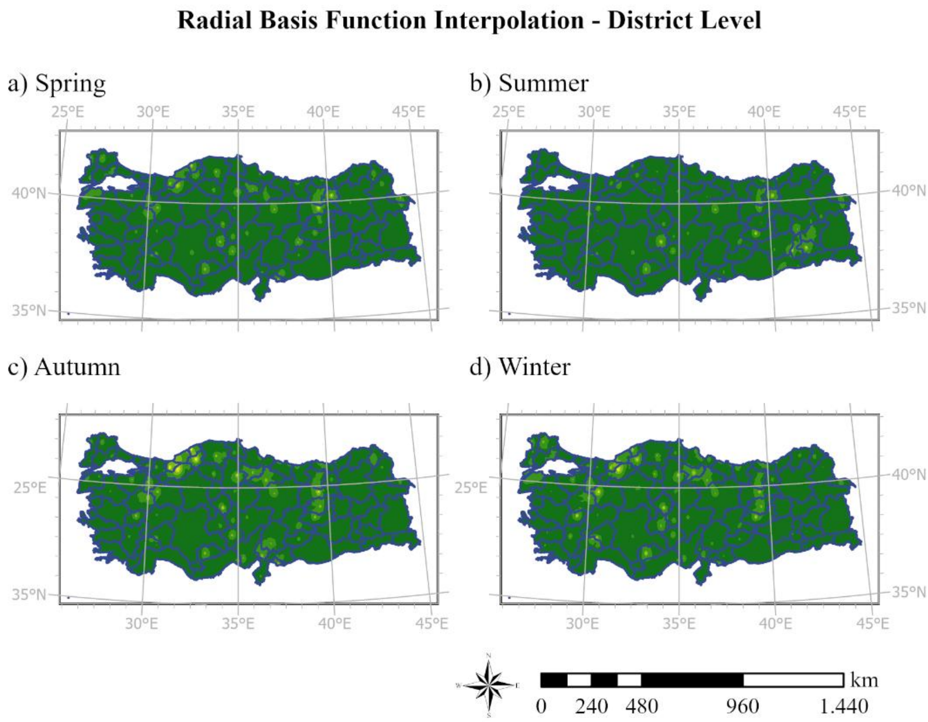

__Figure 10:__ Covid-19 district-level seasonal number of cases distribution maps using RBF interpolation method: a) Spring, b) Summer, c) Autumn, d) Winter

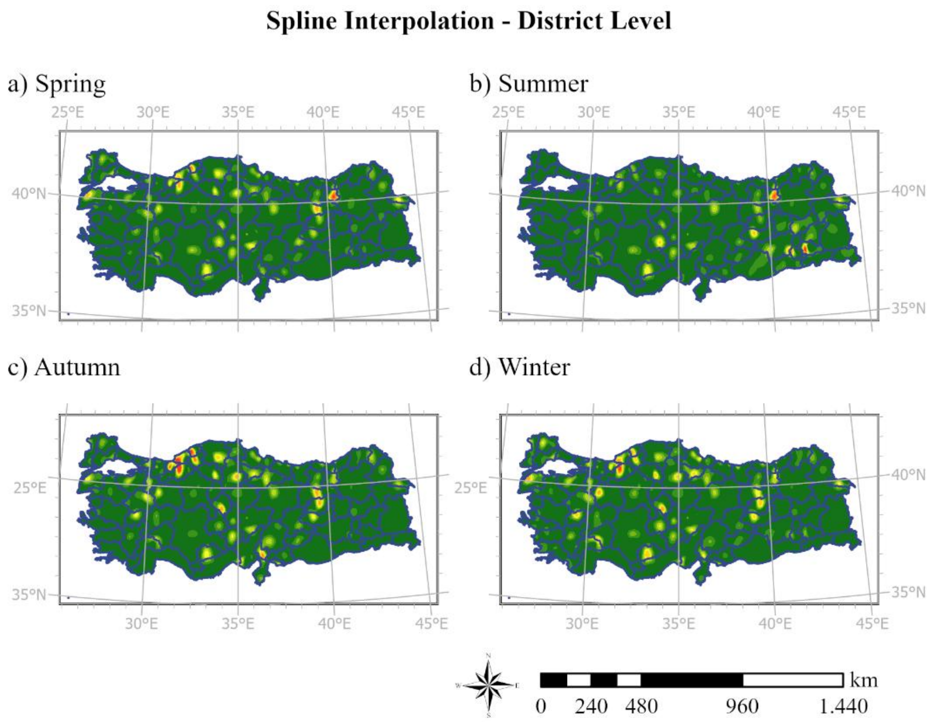

Figure 11: *Covid-19 district-level seasonal number of cases distribution maps using Spline interpolation method: a) Spring, b) Summer, c) Autumn, d) Winter*

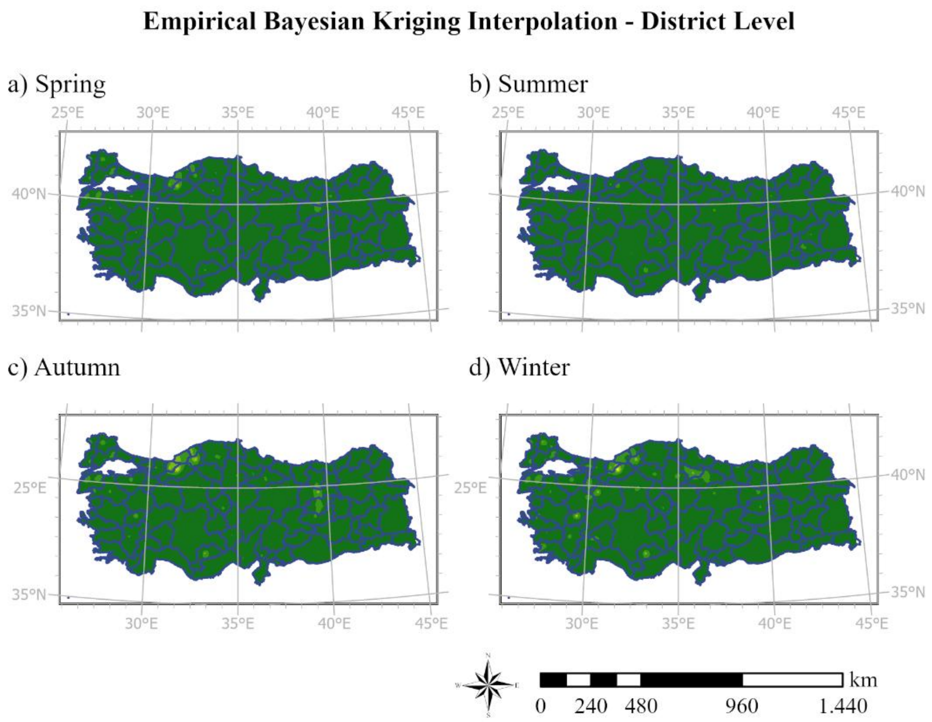

Figure 12: *Covid-19 district-level seasonal number of cases distribution maps using EBK interpolation method: a) Spring, b) Summer, c) Autumn, d) Winter*

The maps generated with the EBK interpolation method, as illustrated in Figure 12, exhibit similarities to those produced with the RBF interpolation model, as demonstrated in Figure 10. However, it is not possible to detect the presence of cases of Covid-19 at the district level in maps created with either the RBF or EBK methods.

Following the generation of the maps utilizing the presented interpolation techniques with seasonal case data for Covid-19, the validation of the resulting maps was carried out. The estimated values, which are extracted from the raster map correspond to 10% of the data in the dataset that was not employed in the model construction. Then estimated values and absolute values from control data were compared, and statistical calculations were performed. The findings are presented in Table 4.

Table 4 presents the statistical accuracy assessment results for the interpolation methods applied to both city-level and district-level seasonal Covid-19 case data. The evaluation metrics include the SD of absolute values, SD of estimated values, MAE, and RMSE.

Table 4: *Statical accuracy assessment results*

Level	Season	Method	SD (Absolute)	SD (Estimated)	MAE	RMSE
City	Spring	IDW	782.8	714.0	458.1	646.4
City	Summer	IDW	639.1	562.0	660.7	813.5
City	Autumn	IDW	1087.2	983.3	531.0	671.6
City	Winter	IDW	2028.7	2027.9	1438.9	1919.9
City	Spring	RBF	782.8	797.4	417.3	543.1
City	Summer	RBF	639.1	450.7	578.6	717.0
City	Autumn	RBF	1087.2	1043.1	522.1	655.5
City	Winter	RBF	2028.7	1779.7	1111.2	1238.2
City	Spring	Spline	782.8	1307.5	841.8	888.9
City	Summer	Spline	639.1	653.5	829.8	1009.6
City	Autumn	Spline	1087.2	1061.5	730.5	900.2
City	Winter	Spline	2028.7	2009.4	1598.1	1734.0
City	Spring	EBK	782.8	740.9	444.9	569.1
City	Summer	EBK	639.1	413.0	542.9	671.1
City	Autumn	EBK	1087.2	1067.0	494.5	607.7
City	Winter	EBK	2028.7	1778.8	1206.6	1286.7
District	Spring	IDW	323.5	206.9	161.1	300.3
District	Summer	IDW	314.9	155.1	121.9	278.6
District	Autumn	IDW	355.0	266.4	192.4	366.9
District	Winter	IDW	785.7	526.0	375.8	701.5
District	Spring	RBF	323.5	151.6	176.4	311.1
District	Summer	RBF	314.9	121.1	138.1	294.4
District	Autumn	RBF	355.0	220.5	203.6	371.9
District	Winter	RBF	785.7	384.6	408.8	738.7
District	Spring	Spline	323.5	429.4	300.8	477.1
District	Summer	Spline	314.9	322.9	216.8	364.3
District	Autumn	Spline	355.0	549.4	360.0	581.9
District	Winter	Spline	785.7	1147.6	726.4	1131.9
District	Spring	EBK	323.5	96.7	149.8	314.7
District	Summer	EBK	314.9	73.7	118.6	296.6
District	Autumn	EBK	354.9	164.7	167.4	361.3
District	Winter	EBK	785.7	269.8	352.9	736.9

3.1 City-Level Analysis

The IDW method shows relatively high MAE and RMSE values across all seasons, particularly during the winter season. This indicates that while IDW is effective in capturing local variations, it may struggle with accuracy in periods of high case density, such as winter. The method's performance is better during the spring and summer seasons, where case densities are lower, resulting in more accurate interpolations. Similarly, Murugesan et al. (2020) employed the Kriging and IDW methods to visualize the state-wise spatial distribution of the disease at the early stages of the pandemic in India, when the number of cases was low, and the density was relatively sparse. In this study, the authors demonstrated the efficacy of the IDW method in detecting and predicting potential disease risk.

RBF interpolation generally provides lower MAE and RMSE values compared to IDW, especially during the winter season. This suggests that RBF is more capable of smoothing spatial variations, making it a better choice when the data exhibits a strong gradient, as seen in the winter.

The Spline method consistently shows higher error metrics, particularly during the winter season, where both MAE and RMSE values are the highest among the interpolation methods. This result suggests that the Spline method may overfit the data, leading to less reliable predictions in cases of extreme variations.

EBK consistently outperforms the other methods, especially in the winter season, where it has the lowest MAE and RMSE values. This indicates that EBK's ability to account for spatial autocorrelation results in more accurate predictions, making it the most reliable method among those tested for high-density data scenarios.

3.2 District-Level Analysis

The IDW method produced the lowest MAE and RMSE values across all seasons at the district level. This indicates that IDW is highly effective for capturing local variations in data with lower spatial variability. Even during the winter season, where case variability tends to be higher, IDW outperformed the other methods with lower error values. This suggests that IDW is particularly well-suited for this scale and data set.

The RBF method yielded slightly higher MAE and RMSE values compared to IDW. However, it performed comparably well during the spring and summer seasons, producing results close to those of IDW. This suggests that RBF is a reliable method during seasons with less spatial variability, although it may not be as effective as IDW in more variable conditions.

The Spline method consistently showed the highest MAE and RMSE values across all seasons, particularly during the winter. This result indicates that the Spline method may be prone to overfitting, leading to less reliable predictions, especially in data sets with high spatial variability.

EBK produced the second-lowest error metrics, following IDW. While it performed well, particularly in the winter season, its MAE and RMSE values were slightly higher than those of IDW. This suggests that while EBK is a strong method, it may not be as well-suited as IDW for this specific district-level data set.

The analysis of interpolation methods across both city and district levels reveals distinct advantages and limitations for each approach. At the city level, EBK emerges as the most reliable method, particularly during periods of high case density, such as the winter season, due to its ability to effectively account for spatial autocorrelation. RBF also performs well, especially in the winter, showing a capacity for smoothing spatial variations. In contrast, IDW shows strong performance during periods of lower-case densities (spring and summer) but struggles with higher densities. The Spline method, however, exhibits the highest errors, indicating potential overfitting and reduced reliability, particularly during extreme case variations. At the district level, IDW consistently outperforms other methods, demonstrating its effectiveness in handling data with lower spatial variability. Although EBK and RBF also perform well, IDW's lower MAE and RMSE values suggest that it is the most suitable method for district-level analysis. The Spline method again shows the highest error rates, underscoring its limitations in handling data with varying spatial characteristics. Overall, while EBK and IDW are generally the most reliable methods across different scales and conditions, Spline's tendency to overfit makes it less appropriate for these datasets.

4. Conclusion

The results of this study demonstrated that the period of highest incidence of cases of Covid-19 in Türkiye was winter. This finding is consistent with the prevalence of respiratory tract infections during the colder months. At the district level, the population-weighted dataset displayed lower deviation values in comparison to the city level, indicating that more detailed data leads to more accurate interpolation results. This study illustrated the efficacy of interpolation methods in identifying high-risk areas on a seasonal basis and provided strategic guidance for public health interventions. The findings provide a useful basis for understanding the spread of the pandemic and optimizing public health strategies. The data obtained can be used to identify risk areas for future outbreaks in advance and to develop appropriate response plans. Finally, the accurate modeling of spreading pandemics is of great importance for potential future pandemics, so that they may be understood, surveilled, and prevented from spreading.

Acknowledgements

Duygu Arican is a Ph.D. scholarship holder from the Council of Higher Education (YÖK) in the field of "GIS and Informatics Applications", which is one of the 100 national priority areas determined by YÖK within the scope of the YÖK 100/2000 Doctorate Program.

Author Contribution

Duygu Arican: Design, Literature review, Data Collection, Analysis and interpretation, Writing, Review of article. **Nursu Tunalioglu:** Conception, Design, Literature review, Supervision, Writing, Review of article.

Declaration of Competing Interests

The authors declare that they have no known relevant competing financial or non-financial interests that could have appeared to influence the work reported in this paper.

References

Alemdar, K. D., Kaya, Ö., Canale, A., Çodur, M. Y., & Campisi, T. (2021). Evaluation of air quality index by spatial analysis depending on vehicle traffic during the COVID-19 outbreak in Turkey. *Energies, 14*(18), 5729.

Bugdayci, I., Ugurlu, O., & Kunt, F. (2023). Spatial Analysis of SO2, PM10, CO, NO2, and O3 Pollutants: The Case of Konya Province, Turkey. *Atmosphere*, *14*(3), 462.

Cong Nhut, N. (2023). Kriging interpolation model: The problem of predicting the number of deaths due to COVID-19 over time in Vietnam. *EAI Endorsed Transactions on Context-Aware Systems and Applications*, *9*(1).

Franch-Pardo, I., Napoletano, B. M., Rosete-Verges, F., & Billa, L. (2020). Spatial analysis and GIS in the study of COVID-19. A review. *Science of the total environment*, *739*, 140033.

Ibarra-Bonilla, J. S., Villarreal-Guerrero, F., Pinedo-Alvarez, A., & Prieto-Amparán, J. A. (2023). COVID-19 in Chihuahua, Mexico: Assessing its spatial behaviour through the inverse distance weighted interpolation technique. *The Canadian Geographies / Géographies Canadiennes*, *67*(4), 470-483.

Ikechukwu, M. N., Ebinne, E., Idorenyin, U., & Raphael, N. I. (2017). Accuracy Assessment and Comparative Analysis of IDW, Spline and Kriging in Spatial Interpolation of Landform (Topography): An Experimental Study. *Journal of Geographic Information System*, *09*(03), 354–371.

Jedwab, R., Khan, A. M., Russ, J., & Zaveri, E. D. (2021). Epidemics, pandemics, and social conflict: Lessons from the past and possible scenarios for COVID-19. *World Development*, *147*.

Jia, H., Zang, S., Zhang, L., Yakovleva, E., Sun, H., & Sun, L. (2023). Spatiotemporal characteristics and socioeconomic factors of PM2.5 heterogeneity in mainland China during the COVID-19 epidemic. *Chemosphere*, *331*.

Johnson, N. P. A. S., & Mueller, J. (2002). Updating the accounts: global mortality of the 1918-1920" Spanish" influenza pandemic. *Bulletin of the History of Medicine*, *76*(1), 105–115.

Kang, D., Choi, H., Kim, J. H., & Choi, J. (2020). Spatial epidemic dynamics of the COVID-19 outbreak in China. *International Journal of Infectious Diseases*, *94*, 96–102.

Kırlangıçoğlu, C. (2022). Investigating the effects of regional characteristics on the spatial distribution of COVID-19 pandemic: a case of Turkey. *Arabian Journal of Geosciences*, *15*(5).

Kotan, B., & Erener, A. (2023). Seasonal analysis and mapping of air pollution (PM10 and SO2) during Covid-19 lockdown in Kocaeli (Türkiye). *International Journal of Engineering and Geosciences*, *8*(2), 173–187.

Kumar, J., Sahoo, S., Bharti, B. K., & Walker, S. (2020). Spatial distribution and impact assessment of COVID-19 on human health using geospatial technologies in India. *International Journal of Multidisciplinary Research and Development*, *7*(5), 57–64.

Li, J., & Heap, A. D. (2008). *A review of spatial interpolation methods for environmental scientists*.

McClymont, H., & Hu, W. (2021). Weather variability and COVID-19 transmission: A review of recent research. *International Journal of Environmental Research and Public Health*, 18(2), 396.

Murugesan, B., Karuppannan, S., Mengistie, A. T., Ranganathan, M., & Gopalakrishnan, G. (2020). Distribution and Trend Analysis of COVID-19 in India: Geospatial Approach. *Journal of Geographical Studies*, *4*(1), 1–9.

Taubenberger, J. K., & Morens, D. M. (2006). 1918 Influenza: the mother of all pandemics. *Revista Biomedica*, *17*(1), 69–79.

Uçar, A., Arslan, Ş., Manap, H., Gürkan, T., Çalışkan, M., Dayıoğlu, A., Efe, H., Yılmaz, M., İbrahimoğlu, A., Gültekin, E., Durna, R., Başar, R., Osmanoğlu, F., & Ören, S. (2020). Türkiye'de Covid-19 Pandemisinin Monitörizasyonu Için Interaktif Ve Gerçek Zamanlı Bir Web Uygulaması: TURCOVID19. *Anatolian Clinic the Journal of Medical Sciences*, *25*(Special Issue on COVID 19), 154–155.

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., Zhao, X., Huang, B., Shi, W., Lu, R., Niu, P., Zhan, F., Ma, X., Wang, D., Xu, W., Wu, G., Gao, G. F., & Tan, W. (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine*, *382*(8), 727–733.

URL-1: Republic of Türkiye Ministry of Health. (2021, March 11). *Bakan Koca, Türkiye'nin Kovid-19'la 1 Yıllık Mücadele Sürecini Değerlendirdi*. https://www.saglik.gov.tr/TR,80604/bakan-koca-turkiyenin-kovid-19la-1-yillik-mucadele-surecini-degerlendirdi.html (Accessed: 1 September 2024).

URL-2: World Health Organization. (2023, May 5). *Virtual Press conference on COVID-19 and other global health issues transcript - 5 May 2023*. https://www.who.int/publications/m/item/virtual-press-conference-on-covid-19-and-other-global-health-issues-transcript---5-may-2023 (Accessed: 1 September 2024).

URL-3: Republic of Türkiye Ministry of Interior. (2020, December 1). *Koronavirüs ile Mücadele Kapsamında - Yeni Kısıtlama ve Tedbirler Genelgeleri*. https://www.icisleri.gov.tr/koronavirus-ile-mucadele-kapsaminda-sokaga-cikma-kisitlamalari---yeni-kisitlama-ve-tedbirler-genelgeleri (Accessed: 1 September 2024).