

INTERPOLATION AND SPATIAL MATCHING METHOD OF VARIOUS PUBLIC DATA FOR BUILDING AN INTEGRATED DATABASE

INTAEK JUNG & KYUSOO CHONG

ICT Convergence and Integration Research Institute, Korea
Institute of Civil Engineering and Building Technology, South Korea

ABSTRACT

The Korea Institute of Construction Technology (KICT) is currently developing a platform system for the analysis of driving environment on the road using vehicle sensing data and various public data. The public data in the KICT platform have different collection cycles and space units to provide real-time road weather and traffic information to users. So, it is difficult to build a database in the form of a relational database to use as explanatory variables for driving environment prediction. This study aims to suggest effective time-series data interpolation and spatial data matching to build an integrated database using different types of public data. We suggested three interpolation methods (piecewise constant interpolation, linear interpolation, nonlinear interpolation) using time-series data and the spatial data matching using weighted average based on each administrative boundary. The quadratic and cubic spline interpolation was applied as the nonlinear interpolation method in this study. As a result of the case study, the linear spline interpolation was selected as the best method and the spatial matching method also showed good results. Although the linear spline interpolation is not so different from other interpolation methods in terms of estimation error, the linear spline interpolation is more effective than other interpolation methods in terms of algorithm implementation and computation speed to be applied to the system. We expect that the built database will be widely used as an explanatory variable for developing various driving environment prediction models.

Keywords: interpolation, spatial matching, database, public data, vehicle sensing data.

1 INTRODUCTION

In accordance with the Government 3.0 policy declared in 2013, all public data have been opened to the private sectors in various fields. According to National Information Society Agency (NISA) in South Korea, the number of open public data increased by from 5,272 in 2013 to 15,912 in 2015. The utilization of open public data also increased by from 13,923 in 2013 to 783,773 in 2015. This trend means that the utilization of open data will continuously increase for many years to come. A data analysis platform for utilizing open public data has been developing by Korea Institute of Civil Engineering and Building Technology (KICT) since 2015. This platform aims to provide real-time road weather and traffic information using individual vehicle sensing data and different types of public data related to the weather observation and road transport fields. The public data in this platform were collected and stored in real time using the Open API (Application Program Interface) provided by various public portal services [1], [2]. However, it is difficult to build an integrated database because the public data have different time intervals and space units to provide real-time information. For example, Korea Meteorological Administration (KMA) offers weather data in real time every 1 hour or 3 hours with the administrative district unit. National Transport Information Center (NTIC), on the other hand, offers traffic data in real time every 5 minutes with Intelligent Transport System (ITS) standard links. Because of the different forms of data collection, these public data are difficult to use as various variables to develop a prediction model. Therefore, interpolation and spatial matching are needed to convert public data with different space-time units into the same space-time unit. The



interpolation method can be classified into two types according to whether the spatial factors are considered [6]–[20] or not [3]–[5]. When interpolation is performed considering spatial factors, interpolation method are divided in two typologies (deterministic/stochastic) [6]. In this study, we also considered temporal and spatial factors, but spatial matching was not an interpolation problem of spatial data as in previous studies. Interpolation of time-series data was performed by applying several methods that do not consider spatial factors. Therefore, this study aimed to suggest effective interpolation and spatial matching method to build an integrated database using different types of public data. Through the built database, it is expected that the analysis and prediction of road driving environment can be carried out in the KICT platform.

2 BACKGROUND AND LITERATURE REVIEW

2.1 Public data in the KICT platform

KICT is currently developing a platform system to analyse the road driving environment using the sensing data observed from individual vehicles and different types of public data. The road driving environment means the weather and traffic conditions of the road that the driver is currently driving. Vehicle sensing data includes road surface temperature, precipitation, and traffic density. Public data includes traffic speed, traffic time, weather information, and incident and event information. As shown in Fig. 1, the KICT platform includes information infrastructure technologies (H/W, S/W) for real-time collection, storage, processing, processing, analysis, and information visualization of vehicle sensing data and public data. In the future, the KICT platform will provide more realistic road weather information and road traffic information (road icing, bad weather, traffic congestion, incident, etc.) through web service [1], [2]. Public data in the KICT platform were collected in real time using open APIs from various public institutions.

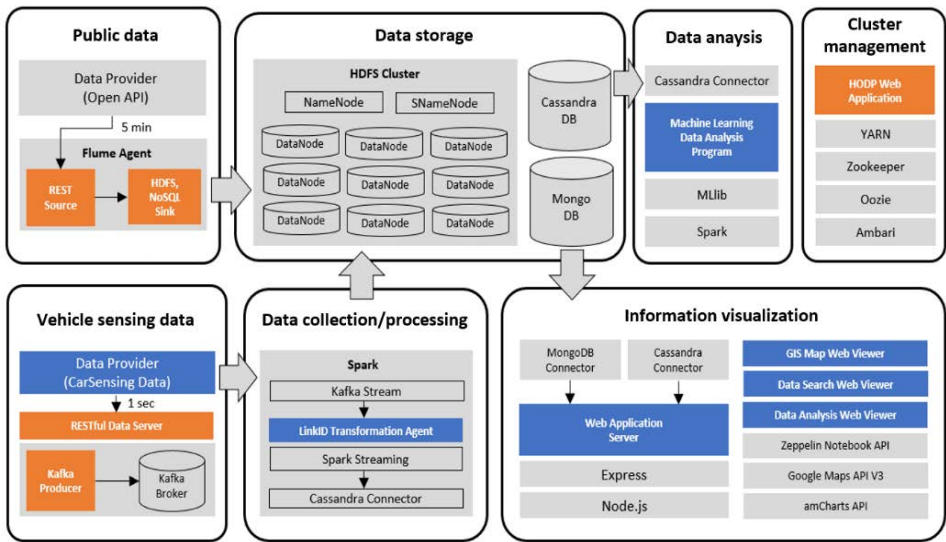


Figure 1: Structure diagram of KICT platform.



Table 1: Collected public data in the KICT platform.

Public institution	Public data	Provision cycle	Space unit
National transport information center (NTIS)	Travel speed	5 minutes	GIS link (ITS standard link)
	Travel time		
	⋮		
Korea meteorological administration (KMA)	Temperature	1 hour, 3 hours	Administrative district
	Humidity		
	Precipitation		
	Wind direction		
	Wind speed		
	⋮		
Urban traffic information system (UTIS)	Incident/Event	5 minutes	GIS link (ITS standard link)
	⋮		

Public data were fused and analysed with vehicle sensing data using the web-based analysis tool. As shown in Table 1, these data included travel speed, travel time, weather data, incident/accident, and so on [2].

2.2 Literature review

Interpolation is a method of constructing new data points within the range of a discrete set of known data points. This method can be divided into two cases: the case where the spatial factor is not considered and the case where it is considered. First, we examined interpolation methods that do not take into account the spatiality of the data. These methods are used to input the missing values of the time-series data in a specific space unit without considering spatial factors. This means that the discrete data points given in time-series data are used to average the values between points. There are basically three kinds of interpolation methods: piecewise constant interpolation, linear interpolation, and nonlinear interpolation. There are several methods for nonlinear interpolation such as polynomial interpolation, spline interpolation, and so on [3]–[5].

Second, we have investigated interpolation methods considering the space of data. These methods are used to estimate the values of an unknown point using the given actual values nearby in the specific space. When interpolation is performed considering spatial factors, interpolation method are divided in two typologies (deterministic/stochastic). There are classification, trend surfaces, Thiessen polygons, linear, moving averages, and inverse distance weighting in deterministic interpolation methods, there are regression models, thin plate splines, kriging, and conditional simulation in stochastic interpolation methods. Three interpolation techniques were commonly used in interpolating climate data such as inverse distance weighted averaging (IDWA), thin plate smoothing spline (TSP) and kriging [6]. IDWA combines the idea of proximity created by the Thiessen polygon with the gradual change of the trend line [7]. Observations closer to the predicted location have more impact on the predicted value than values farther away. This distance-weighted approach has been widely applied to interpolate climate data [8]. If the weight function is selected according to the IDWA equation, the interpolation result can be greatly affected. Various weighting functions are compared with each other and the analysed results are presented [9]. Splining refers to a local deterministic interpolation that represents a two-dimensional curve on a

three-dimensional surface [10]. It is conceptually similar to fitting rubber membranes through measured sample values while minimizing the total curvature of the surface. A special kind of TSP was developed principally [11]. TSP is a relatively simple calculation and is often used to interpolate elevations to create a digital elevation model. After further methodological developments, this method was applied for climate interpolation [12]. Kriging is a stochastic technique similar to IDWA in that it uses a linear combination of weights at known points to estimate the value of an unknown point. Kriging differs from deterministic methods like IDWA because it considers spatial correlations [13]. Kriging is divided into Simple Kriging, Ordinary Kriging, and Universal Kriging depending on how weights are calculated. In addition, kriging is divided into univariate kriging (simple kriging, general kriging, general kriging) and co-kriging, depending on whether secondary variables are used or not. Detailed description of each kriging method can be found in this study [14]. When interpolating with co-kriging, the variogram models have to fit the linear model of co-regionalization [15].

We also investigated studies comparing various interpolation methods, including the three interpolation methods described above. Three interpolation approaches (IDWA, TSP, co-kriging) were evaluated for a 200,000 km^2 square area covering the state of Jalisco, Mexico [6]. This paper applied four different interpolation methods (kriging, IDWA, nearest neighbor, spatial averaging) to the monitoring data to derive air concentration levels [16]. A thin plate smoothing spline was used to interpolate daily climate variables, and ordinary kriging was used to interpolate daily and monthly rainfall [17]. Interpolation methods of spatial data are mainly Thiessen polygon, polynomial interpolation using the least-square, Lagrange approach, multi-quadric interpolation, kriging technique, and so on [18]. Four different methods were applied to several test cases, respectively: IDWA, kriging method, hardy's multi-quadric method, and tension finite-difference method [19]. GIS often requires the interpolation methods of an input dataset, and these methods also require the estimation of the unknown values at location-time pairs with a satisfactory level of accuracy. Using an actual real estate dataset, these methods were compared with other interpolation methods based on IDWA and kriging [20].

This study builds an integrated database to develop future driving environment prediction model. This database includes road traffic data and weather data, which are used as variables in the development model. However, these data are collected in different space-time units. So, we need to convert the collected data of different space-time units into the same unit. We have found through the literature review that there are various interpolation methods depending on whether data space is considered (IDWA, TSP, kriging, etc.) or not (piecewise constant interpolation, linear interpolation, and nonlinear interpolation). They also compared various interpolation methods according to application variables. In the case of spatial data of this study, it is used to convert the spatial data of administrative district unit to the spatial data of GIS link unit. This is not a spatial interpolation but a spatial matching problem, it was difficult to apply the existing interpolation methods using spatial data and another spatial matching method is needed for this study. But interpolation methods are used to convert the data collected every 1 hour to the 5 minutes data. It is necessary to compare three interpolation methods using time-series data that does not take spatiality of data into account.

3 METHODOLOGY

3.1 Concept of interpolation and spatial matching

This study deals with the problem of converting macroscopic space-time unit into microscopic space-time unit in order to build an integrated database using various kinds of



public data. In other words, the problem is to convert the space-time unit of KMA's data to the space-time unit of NTIS's data, which is the minimum unit of public data in the KICT platform. This problem can be solved by interpolating time series data and matching spatial data. The interpolation of the time-series data means that the collection data of 1 hour interval is converted every 5 minutes and The matching of the spatial data means that the spatial data of the administrative district unit is converted into the road segment unit (ITS standard links). First, we proposed three interpolation methods that do not consider the space of data for the interpolation of time-series data. Three interpolation methods for interpolating time-series data were piecewise constant interpolation, linear interpolation, and nonlinear interpolation. The linear interpolation method used the first-order form of the spline interpolation method. The nonlinear interpolation method used the second-order and the third-order form of the spline interpolation method (quadratic spline interpolation, cubic spline interpolation). Because the type of the collected data is time-series data and the data is interpolated continuously at intervals of 5 minutes according to the observed values. In this study, the best method is to minimize the estimation error among the three interpolation methods. Second, the spatial matching of this study deals with the problem of converting the spatial data of the administrative district unit into the road segment unit. It was not necessary to apply the existing spatial data interpolation methods and we proposed another spatial data matching method. If the links are included in a regional boundary, link data is applied to the public data of the administrative district. Otherwise, the link data are applied by weighted averaging of weather data through which the link passes.

By applying the proposed methods, we can build an integrated database of 5 minutes interval and road section unit. This database will be used as predictive variables for the development of driving environment prediction models using various public data. The Overall flow chart of this study is shown in Fig. 2.

3.2 Building input data

The input data of this study used public data collected in real time using Open API and GIS in the KICT platform. As shown in Table 1, the input data included public data collected from three public institutions. It can be seen that the collection unit of each public data is different. These public data should be converted to the same space-time unit to build an integrated database. Therefore, the input data was applied to weather data collected from KMA.

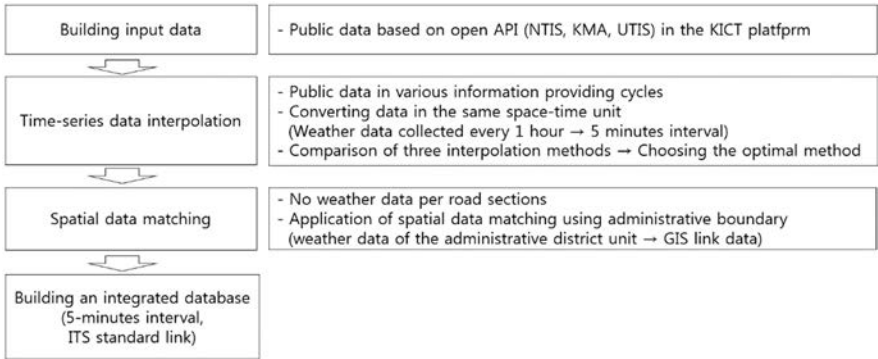


Figure 2: The flow chart of this study.



It is necessary to convert space-time unit of public data collected from KMA to space-time unit of public data collected from NTIS. Because public data collected from NTIS has the minimum space-time unit among collected public data in the KICT platform. In the case of GIS data, we used the ITS standard node and link provided by the Ministry of Land, Infrastructure, and Transport (MOLIT) in South Korea. It is consisted of the node information, rotation information, link information, and link additional information.

3.3 Time-series data interpolation

The time-series data interpolation method is necessary to convert the weather data of 1-hour interval into weather data of 5 minutes interval in a given administrative district. We applied three interpolation methods like piecewise constant interpolation, linear interpolation, and nonlinear interpolation. All three methods were applied in case study, and the interpolation method which minimizes the estimation error was selected as the best method. First, piecewise constant interpolation is a method of equally applying the weather data of the current cycle every 5 minutes until the weather data of the next cycle is updated. This method assumes that there is no change in the 5 minutes interval of weather data every hour.

$$[w_{t+m_i}] = [w_{t+m_1}, w_{t+m_2}, \dots, w_{t+m_{12}}] = w_t, t = 1h, 2h, \dots, 24h, i = 1, 2, \dots, 12, \quad (1)$$

where, w_t : weather data at time t in an administrative district, m_i : i interval (every 5 minutes).

Second, linear interpolation is the first-order form of the spline interpolation method. This is a method of interpolating with weather data in 5 minutes interval between the previous time and the current time every 1 hour. As shown in Fig. 3, the given current time is x_t in an administrative district, and the previous time is x_{t-1} , the weather data of x_t is y_{x_t} , and the weather data of x_{t-1} is $y_{x_{t-1}}$. The estimated value $f(x_{t_m})$ of time x_{t_m} between the two observation points $(x_{t-1}, y_{x_{t-1}})$ and (x_t, y_{x_t}) can be calculated using eqn (2).

$$f(x_{t_m}) = y_{t-1} + \frac{y_{x_t} - y_{x_{t-1}}}{x_t - x_{t-1}} \times (x_{t_m} - x_{t-1}), x_{t-1} \leq x_{t_m} \leq x_t, x_t - x_{t-1} = 12 \quad (2)$$

Finally, the quadratic and cubic spline interpolation were applied to the nonlinear interpolation method. These methods were the most widely used, and it was algorithms that smoothly connect the given discrete points. The quadratic spline interpolation uses a quadratic equation as a curve connecting two points. The quadratic equation is shown in eqn (3).

$$f_k(x) = a_k x^2 + b_k x + c_k, \quad (3)$$

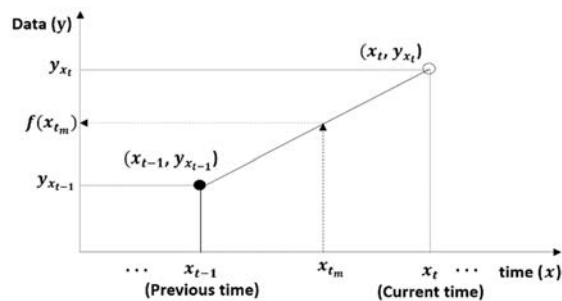


Figure 3: Concept of linear interpolation.

$f_k(x) = a_k x^2 + b_k x + c_k$ Known values ($n + 1$) must be determined for k data points ($k = 1, 2, \dots, n + 1$) in an administrative district. A quadratic spline equation ($n = 2$) that requires three unknown values are derived from eqn (4)–(7). Since we have obtained $3n$ conditions in total, we can calculate $n + 1$ unknown. As shown in Fig. 4, $f_1(x)$ and $f_2(x)$ can be calculated.

$$f(x_{k-1}) = a_{k-1}x_{k-1}^2 + b_{k-1}x_{k-1} + c_{k-1} = a_k x_{k-1}^2 + b_k x_{k-1} + c_k \quad (4)$$

$$f(x_0) = a_1 x_0^2 + b_1 x_0 + c_1, f(x_n) = a_n x_n^2 + b_n x_n + c_n \quad (5)$$

$$f'(x_{k-1}) = 2a_{k-1}x_{k-1} + b_{k-1} = 2a_k x_{k-1} + b_k \quad (6)$$

$$f''(x_0) = a_1 = 0 \quad (7)$$

The cubic spline interpolation uses a third-order polynomial equation (cubic curve) as a curve connecting two points. The cubic spline interpolation minimizes the estimation error in the function that changes locally rather than the quadratic spline interpolation in general. The cubic curve is shown in eqn (8).

$$f_k(x) = a_k x^3 + b_k x^2 + c_k x + d_k \quad (8)$$

Known values ($n + 1$) must be determined for k data points ($k = 1, 2, \dots, n + 1$) in an administrative district. A cubic spline equation ($n = 3$) that requires four unknown values is derived from eqn (9) and eqn (10). The $n - 1$ second-derivative values are calculated using eqn (9) and eqn (10). The $n + 1$ unknown values are calculated by applying the condition that the second derivative is 0 on both sides like eqn (11). As shown in Fig. 4, $f_1(x)$, $f_2(x)$ and $f_3(x)$ can be calculated.

$$f_k(x) = \frac{f_k''(x_{k-1})}{6(x_k - x_{k-1})}(x_k - x)^3 + \frac{f_k''(x_k)}{6(x_k - x_{k-1})}(x - x_{k-1})^3 + \left[\frac{f(x_{k-1})}{(x_k - x_{k-1})} - \frac{f_k''(x_{k-1})(x_k - x_{k-1})}{6} \right](x_k - x) + \left[\frac{f(x_k)}{(x_k - x_{k-1})} - \frac{f_k''(x_k)(x_k - x_{k-1})}{6} \right](x - x_{k-1}), \quad (9)$$

$$(x_k - x_{k-1})f''(x_{k-1}) + 2(x_{k+1} - x_{k-1})f''(x_{k-1}) + (x_{k+1} - x_k)f''(x_{k+1}) = \frac{6}{(x_{k+1} - x_k)}[f(x_{k+1}) - f(x_k)] + \frac{6}{(x_k - x_{k-1})}[f(x_{k-1}) - f(x_k)], \quad (10)$$

$$f''(x_0) = 0, f''(x_3) = 0 \quad (11)$$

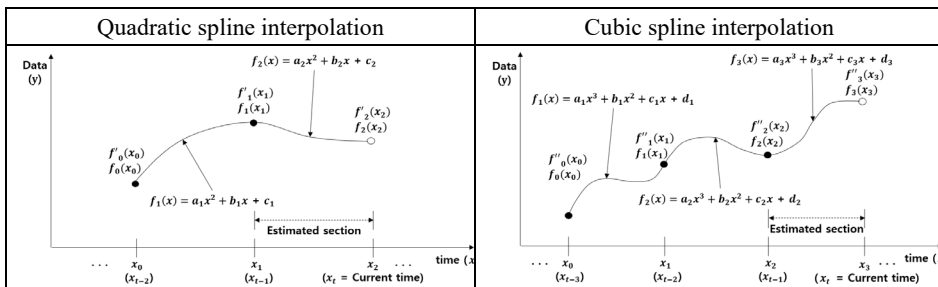


Figure 4: Concept of nonlinear interpolation.

3.4 Spatial data matching

The space unit for matching spatial data is the ITS standard link. Weather data for each administrative district should be converted into each link unit. In this study, we proposed the spatial matching method using administrative boundaries. This method can be applied in two ways, depending on whether the link is within the administrative district or is not. If the link is completely within the administrative district, the link data applies the weather data of the administrative district. Otherwise, it means that the link is passing through more than one administrative district. In this case, the link data is determined by calculating the weighted average of the weather data of each administrative district using the link length included in each administrative district using eqn (12).

$$l_w = \frac{a_{w,1} \times d_1 + a_{w,2} \times d_2 + \dots + a_{w,n} \times d_n}{d_1 + d_2 + \dots + d_n}, n \in \mathbb{N}, \tag{12}$$

where, l_w : link weather data, $a_{w,n}$: weather data of the n th administrative district, and d_n : n th link length divided.

As shown in Fig. 5, link 1, 2 belong to administrative district A, and link 5 belongs to district B. Link 7, 8 belong to district C. Therefore, the weather data of link 1, 2 become the weather data of district A, and the weather data of link 5 becomes the weather data of district B. The weather data of link 7, 8 become the weather data of district C. In the case of link 3, 4 and 6, each link is not completely contained within the administrative districts. The weather data of these links are calculated using eqn (12).

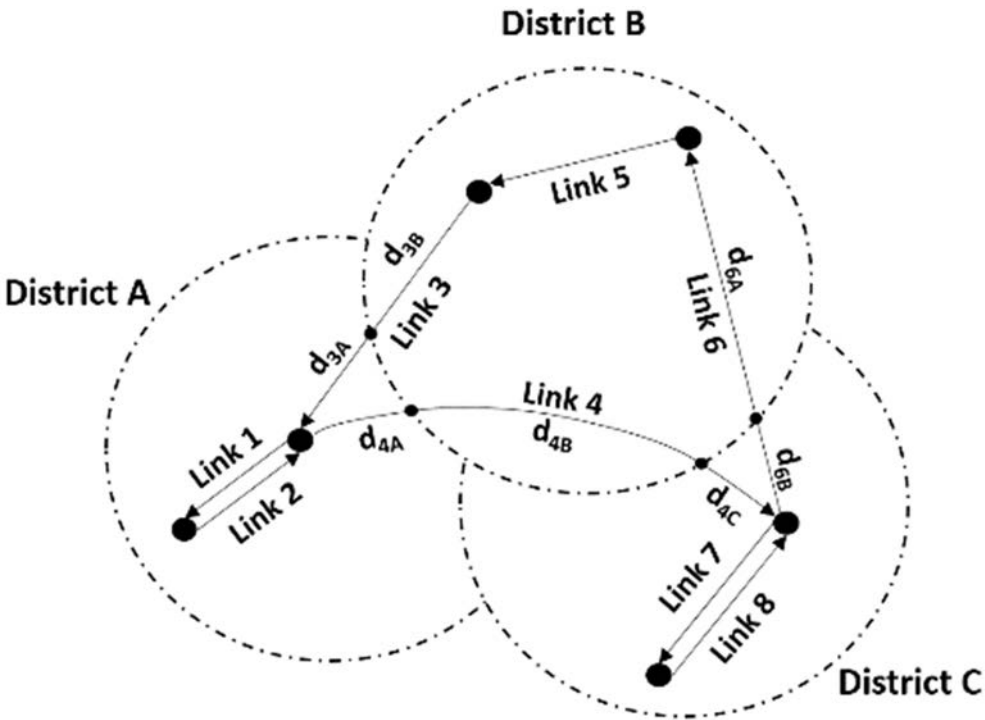


Figure 5: Concept of spatial data matching.

4 CASE STUDY AND EVALUATION

The KICT is currently developing a platform system to analyse road driving environments using vehicle sensing data and various public data. Some road sections of Jayouro (Seongdong IC – Isanpo IC) and Seoul Beltway (Ilsan IC – Tongilro IC) around the cities of Goyang and Paju in South Korea were set up as the case study sections to apply the proposed methods in this study. These road sections are the same as the road sections for testing the KICT platform. GIS data was based on ITS standard links provided by the MOLIT. The total number of road sections is 58 link (Jayouro: 44 links, Seoul Beltway: 14 links). As shown in Table 1, input data were based on the real-time weather data provided by KMA among the collected public data. These data are the weather data of the administrative district unit, which are collected in real time every 1 hour from various Open APIs. The observation data were received from KMA (Paju meteorological observatory) between October 1 and 3, 2016, and these data were used for validation and evaluation.

In the case of time-series data interpolation, all four methods proposed in this study were applied. The method of minimizing the evaluation index was selected as the optimal method. The evaluation index is mean absolute percent error (MAPE, %), which is the error rate of the observed and estimated values. The weather data for evaluation were used air temperature (°C) and humidity (%). Four methods were applied to compare the estimated error using the estimated value and the observed value. As shown in Table 2, the best method was determined to be the linear spline interpolation, but MAPE was calculated as 1.14–2.82%, and overall error rate is low and there was almost no difference among the three methods except for the piecewise constant interpolation. As shown in Fig. 6 and Fig. 7, this is because the temperature and humidity fluctuations at intervals of 5 minutes are not large. Therefore, choosing any of the three methods does not seem to be a big problem. But the linear spline interpolation is more effective than the quadratic and cubic spline interpolation in terms of algorithm implementation and computation speed to be applied to the KICT platform.

Spatial matching of weather data converted weather data of four administrative districts including test road sections into weather data of links. When the whole link was included in the administrative district, the weather data of the link applied the weather data of the administrative area as it is. When one link passed through more than one administrative district, the weather data of the link was calculated by weighted average using the link lengths divided by the boundary of each administrative district. Spatial data matching was performed only on the case analysis sections. So, we thought that the accuracy evaluation of spatial data matching was not meaningful. If the spatial extent extends to the whole country, it is necessary to establish a systematic evaluation method for evaluating the accuracy of spatial data.

Table 2: Types of comparison of error rate.

	MAPE (%)			
	Piecewise constant interpolation	Spline interpolation		
		Linear	Nonlinear	
			Quadratic	Cubic
Temperature	2.46	1.14	1.22	1.49
Humidity	2.82	1.58	1.61	1.79

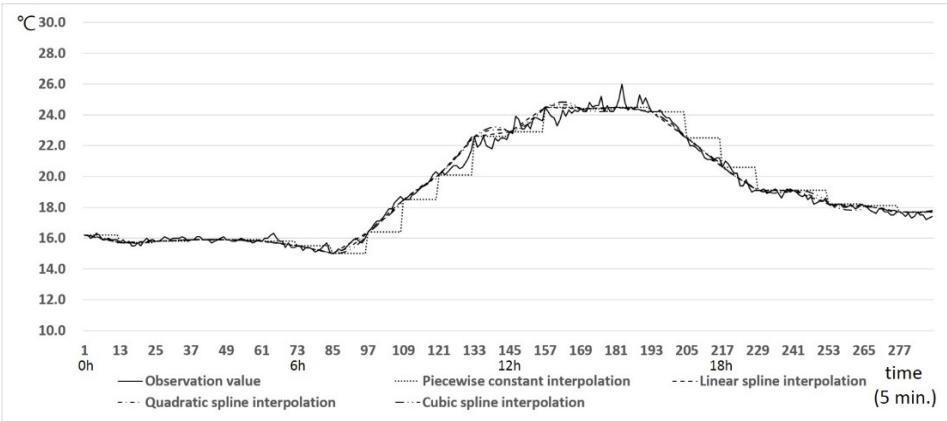


Figure 6: Comparison of temperature estimation results over time intervals (OCT 1, 2016).

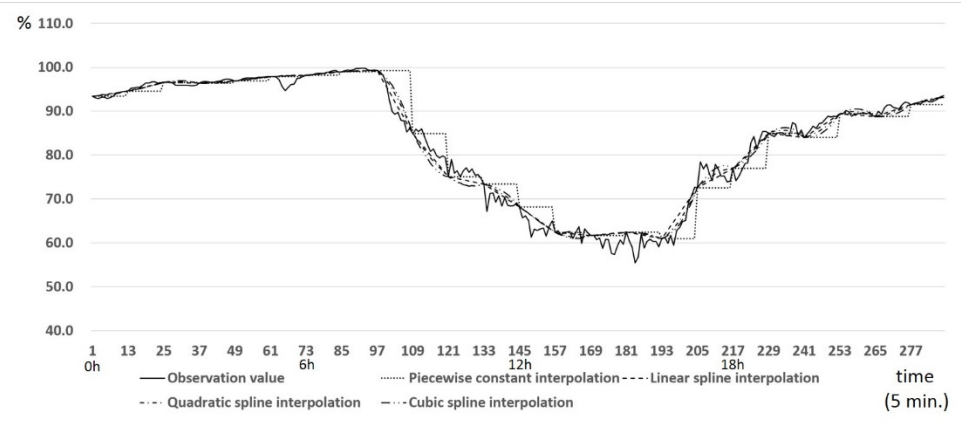


Figure 7: Comparison of humidity estimation results over time intervals (OCT 1, 2016).

As shown in Fig. 2, the integrated platform was constructed using spatial matching and linear spline interpolation, which have the lowest MAPE. The results of the database construction are shown in Table 3. The time unit of this database is 5 minutes, and the space unit is the ITS standard link.

5 CONCLUSION

This study suggested the time-series data interpolation and spatial data matching method to build an integrated database using various public data. Three methods were proposed for interpolating time series data such as piecewise constant interpolation, linear interpolation, and nonlinear interpolation. For the linear interpolation, we applied the spline interpolation method using the linear equation. For the nonlinear interpolation, we applied the quadratic spline interpolation and the cubic spline interpolation, which was generally used. The space matching method based on the administrative district boundaries was also applied because



Table 3: Result of integrated database construction.

Collection time	Link ID	Travel speed (kph)	Travel time (sec)	Temperature (°C)	Precipitation (mm)	Humidity (%)	Wind direction (degrees)
2016-09-01 0:05	218000 5300	80	48	17.1	0	93.4	197
2016-09-01 0:10	218000 5300	79	48	17.2	0	93.0	146.6
2016-09-01 0:15	218000 5300	80	48	17.0	0	92.9	291
2016-09-01 0:20	218000 5300	78	49	16.6	0	93.2	232.6
2016-09-01 0:25	218000 5300	80	48	16.4	0	93.2	178.4
2016-09-01 0:30	218000 5300	81	47	16.2	0	92.9	173.5
2016-09-01 0:35	218000 5300	81	47	16.3	0	93.1	180.6
2016-09-01 0:40	218000 5300	79	48	16.4	0	93.5	171.5
2016-09-01 0:45	218000 5300	78	49	16.2	0	94.0	193.5
2016-09-01 0:50	218000 5300	78	49	16.0	0	94.2	205.3
2016-09-01 0:55	218000 5300	80	48	16.0	0	94.4	237.3
2016-09-01 1:00	218000 5300	80	48	16.0	0	94.5	275.4

the weather data for each administrative district should be converted to the weather data for each link. As a result of case analysis and evaluation, the linear spline interpolation was the best among the four methods, and the spatial matching method was also excellent. There was almost no difference among the three methods except for the piecewise constant interpolation. But the linear spline interpolation is more effective than other methods in terms of algorithm implementation and computation speed to be applied to the system. Therefore, we constructed an integrated database of the KICT platform using linear spline interpolation with minimum MAPE and spatial data matching. It was judged that various models for predicting the road driving environment could be developed using the integrated database in the KICT platform. If the spatial coverage extends to the whole country, we need to compare with more interpolation methods, and it is also necessary to evaluate the accuracy of spatial data matching proposed in this study.

ACKNOWLEDGEMENT

This research was supported by a grant from a Strategic Research Project (*Development of Driving Environment Observation, Prediction and Safety Technology Based on Automotive Sensors*) funded by the Korea Institute of Civil Engineering and Building Technology.



REFERENCES

- [1] Ha, J. & Chong, K., Development of a road big-data storage platform for predicting the driving environment, *Journal of Emerging Trends in Computing and Information Sciences*, **7**(1), pp. 14–22, 2016.
- [2] Korea Institute of Civil Engineering and Building Technology, Development of driving environment prediction platform based on big data (2nd year), 2016.
- [3] Cheney, W. & Kincaid, D., Numerical Mathematics and Computing 6th Edition, Thomson Brooks/Cole, a part of The Thomson Corporation, 2008
- [4] Burden, R.L. & Faires, J.D., Numerical analysis 9th, BROOKS/COLE CENGAGE Learning, 2011.
- [5] Chapra, S.C. & Canale, R.P., Numerical methods for engineers, McGraw-Hill, 2012.
- [6] Hartkamp, A.D., De Beurs, K., Stein, A. & White, J.W., Interpolation techniques for climate variables, NRG-GIS Series 99-01, Mexico, D.F.: CIMMYT, 1999.
- [7] Thiessen, A.H., Precipitation averages for large areas, *Monthly Weather Review*, **39**, pp. 1082–1084, 1911.
- [8] Legates, D.R. & Willmott, C.J., Mean seasonal and spatial variability in global surface air temperature. *Theoretical Application in Climatology*, **41**, pp. 11–21, 1990.
- [9] Lancaster, P. & Salkauskas, K., *Curve and Surface Fitting: An Introduction*, Academic Press: London, 1986.
- [10] Hutchinson M.F. & Gessler, P.E., Splines—more than just a smooth interpolator *Geoderma*, **62**, pp. 45–67, 1994.
- [11] Wahba, G. & Wendelberger, J., Some new mathematical methods for variational objective analysis using splines and cross-validation, *Monthly Weather Review*, **108**, pp. 1122–1145, 1980.
- [12] Hutchinson, M.F., Interpolating mean rainfall using thin plate smoothing splines. *International journal of geographical information systems*, **9**(4), pp. 385–403, 1995.
- [13] Krige, D.G., Two-dimensional Weighted Average Trend Surfaces for Ore Evaluation. *Journal of the South African Institute of Mining and Metallurgy*, **66**, pp. 13–38, 1966.
- [14] Cressie, N., *Statistics for Spatial Data*. John Wiley and Sons: New York, 1993.
- [15] Goulard, M. & Voltz, M., Linear coregionalization model: tools for estimation and choice of cross-variogram matrix, *Mathematical Geology*, **24**(3), pp. 269–286, 1992.
- [16] Wong, D.W., Yuan, L. & Perlin, S.A., Comparison of spatial interpolation methods for the estimation of air quality data, *Journal of Exposure Analysis and Environment Epidemiology*, **14**(5), pp. 404–415, 2004.
- [17] Jeffrey, S.J., Carter, J.O., Moodie, K.B. & Beswick, A.R., Using spatial interpolation to construct a comprehensive archive of Australian climate data, *Environmental Modelling & Software*, **16**, pp. 309–330, 2001.
- [18] Guillermo, Q. & Jose, D., A comparative analysis of techniques for spatial interpolation of precipitation, *The Water Resources Bulletin American Water Resources Association*, **21**(3), pp. 365–380, 1985.
- [19] Caruso, C. & Quarta, F., Interpolation methods comparison, *Computers & Mathematics with Applications*, **35**(12), pp. 109–126, 1998.
- [20] Li, L. & Revesz, P., Interpolation methods for spatio-temporal geographic data, *Computers, Environment and Urban Systems*, **28**(3), pp. 201–227, 2004.

