

Surface Estimation of an Ancient City Using Spatial Interpolation: The Case of the Capital of Silla, an Ancient State in Korea

Dongseok Kang¹, Jongyun Won¹

¹ Dept. of Art History & Archaeology, University Dongguk, Gyeongju, Korea – kydhsh1214@gmail.com, axa12345@naver.com

Keywords: Korea, Silla Capital, Interpolation, IDW, Kriging, Surface.

Abstract

Silla was an ancient state in Korea, whose capital was Gyeongju. The city began to develop in earnest around 6C A.D. and was the leading city in Northeast Asia until 935. More than 600 archaeological excavations have revealed the city's existence. In this study, we use the survey information from these sites to estimate the surface that was once inhabited by its citizens. This has the advantage of reconstructing the topographic landscape of the ancient city, identifying the land use patterns of the time, and predicting the elevation of the ruins below the current surface. The IDW and Kriging interpolation methods were utilized to estimate the historical land surface. The data used here basically uses spatial information provided by WebGIS, which was built and operated by the Korean government since 2009 to catalog and digitally document all archaeological survey information. Based on this, the elevation of features identified in the excavation reports was extracted to create a database. The total number of sites is 292, and the elevation of 985 points during the Silla period was obtained. This is a statistically reliable sample size. IDW and Kriging interpolation was performed by dividing the searching neighborhood type into standard and smooth, and cross-validation was performed for a total of 18 models. The most accurate model was Kriging interpolation, and the model with circular semi-variogram with smooth factor set to 1 had the lowest RMSE. Based on the cross-validation results of these models, the DEM of the ancient city of Silla was finally created. Comparing it to the current surface, it was found that the surface was highly curved in areas close to rivers at an elevation of less than 40 meters above sea level. These topographical conditions likely prevented a grid-like road network from being laid out here. Furthermore, the comparison of the present and past topographical cross-sections showed that the surface had been cut and embanked in the past. As more information on excavated sites is added in the future, the accuracy of the interpolation model is expected to improve. This will be a fundamental study for reconstructing the landscape of the Silla capital.

1. Introduction

Silla, the ancient state of Korea, is known as the Golden Kingdom and the eastern terminus of the Silk Road (Lee *et al.*, 2013). For over a thousand years, Gyeongju was the capital of Silla and served as the political, economic, and religious center of the country. The capital began to develop in earnest around 6C AD and was the leading city in Northeast Asia until the fall of Silla in 935.

The true nature of the Silla Kingdom is being revealed by more than 600 archaeological excavations. The results of these archaeological investigations are constructed and managed in the 'National Heritage GIS Integrated Intranet System' and 'Ancient City GIS' of the Korea Heritage Service. The excavation areas data are managed as spatial databases, and the research reports associated with them provide information on the past surface where the remains, such as building sites, have been identified. This paper attempts to estimate the capital land surface of the Silla Kingdom by integrating archaeological spatial information systematically managed at the national level in Korea. This study has the following implications. First, the overall topographic landscape of the city can be reconstructed. Second, it allows us to explore the way the developers of Silla capital understood the landscape at the time. Third, it can reveal the causal relationship between the arrangement of various functional spaces and topographic conditions. In other words, the estimation of the land surface of the Silla Kingdom capital is important research for understanding the development history of ancient cities, the landscape perception of urban agents, and establishing preservation measures.

This surface estimation of Shilla capital utilizes the spatial interpolation method of GIS. To this end, we construct an

integrated spatial database of the ancient city sites, create DEMs using IDW and Kriging interpolation methods, and select the most appropriate DEM through model validation. By comparing and examining the distribution of the archaeological sites, we aim to interpret the space utilization patterns of the Silla capital.

2. Study Area

The study area is the urban center of Gyeongju, South Korea. Gyeongju was the capital of Silla, and the Gyeongju Historic Areas are listed as a World Heritage Site because of the high concentration of archaeological sites and artifacts that not only preserve the history and cultural heritage of the ancient nation, but also trace its interactions with China and Japan. These Silla capitals are not surrounded by castles. Instead, high mountains surround the city, forming a basin with rivers flowing through it to the north, south, and west.



Figure 1. The Royal Tomb Area and Urban Center of Gyeongju, the Silla Capital (Copyright Gyeongju City).

Excavations have revealed that the entire city center is a densely packed urban site of mounds, temples, buildings, and roads. The royal castle and royal palace are located in the southern part of the city, with a large royal necropolis in the adjacent north. These tombs were in the center of the city until A.D. 5C, after which they were placed on the outskirts of the city. Fifteen Buddhist temples are known to have been located within a 5-kilometer diameter area, and urban development began in earnest after A.D. 6C, when buildings were constructed around these temples and streets were laid out.

The population of the city is estimated to have been around 170,000 based on historical records (Jeon, 2005; Kang, 2008; Lee, 2006; Min, 1986). The spatial scope of the city's inhabitants would have corresponded to the actual area of the capital. Based on the archaeological excavations conducted in the urban area of Gyeongju, a distribution map was created, and kernel density estimation was performed (Figure 2). The sites show high density within the spatial range surrounded by rivers. These analysis results allow for the identification of the city's central area. In this study, we aim to estimate the living surface from the 6th to 10th centuries A.D. based on the excavation data accumulated in this central zone.

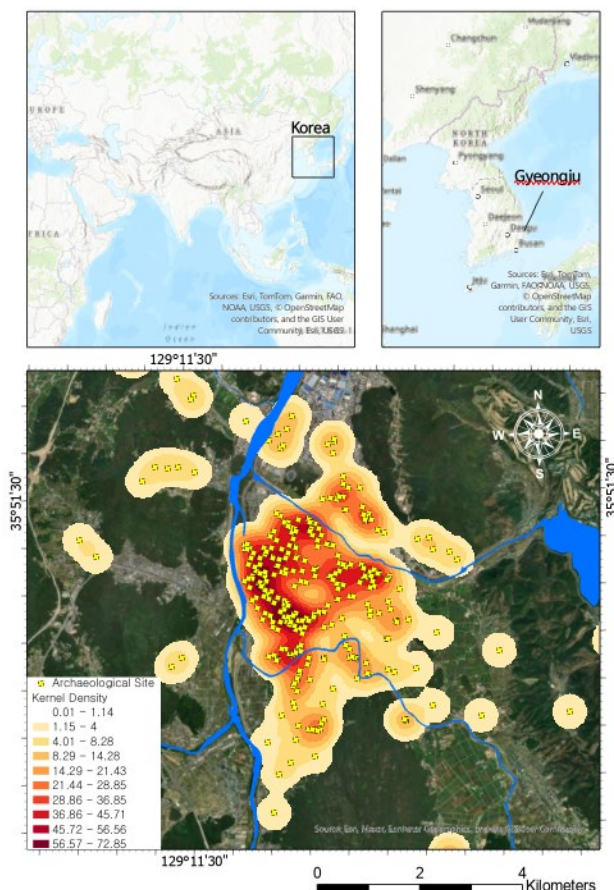


Figure 2. Geographical Location of the Study Area and Kernel Density Estimation Results of Archaeological Sites.

3. Method and Materials

3.1 Research Methodology

The overall research procedure for estimating the land surface of the Silla capital can be roughly divided into two parts (Figure 3): the construction of a spatial database for spatial

statistical analysis, and the creation and validation of a model by spatial interpolation. First, the basic data required to construct the spatial database is collected from the 'National Heritage GIS' (<https://intranet.gis-heritage.go.kr/>), which has been operated by the Korea Heritage Service since 2009, and the Ancient City GIS (<https://dosi.nrich.go.kr/index.do>), which was recently opened by the National Research Institute of Cultural Heritage. 'The National Heritage GIS' documents all archaeological excavations throughout Korea and is updated frequently. By registering for a member account, you can access the website, which contains information on more than 30,000 excavations conducted in Korea to date. 'Ancient City GIS' provides information on the layout of ancient city excavations sites. This data needs to be georeferenced to obtain coordinates. Through this process, the elevation of the identified sites related to the Silla capital is constructed into a spatial database. Based on this data, spatial statistical analysis using IDW and Kriging interpolation is performed to generate digital elevation models (DEM), and cross-validation of each model is performed. Through this process, an optimal DEM is created to estimate the surface during Silla's most prosperous period. Finally, through the comparison of this DEM with the current land surface, we predict the distribution of underground ruins and examine the terrain utilization patterns related to urban development during the Silla capital.

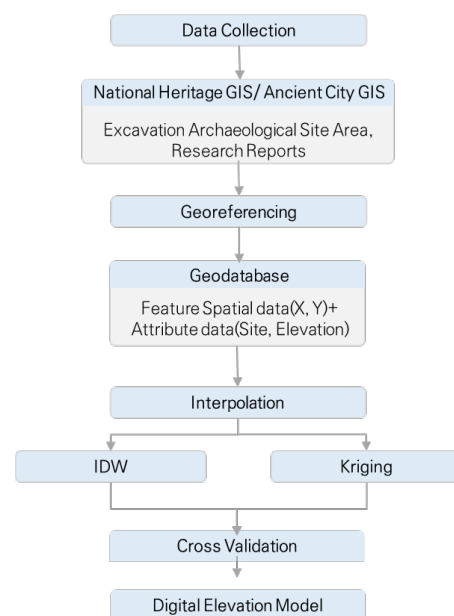
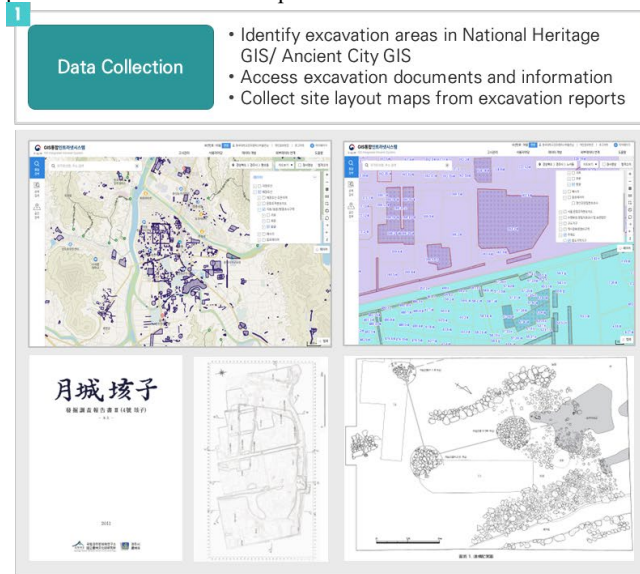


Figure 3. Scheme of research methodology

3.2 Data Collection and Geodatabase Construction

The procedure for data collection and spatial database construction for estimating the ground surface of the Silla Capital is shown in <Figure 4>. The inventory and documentation of all archaeological excavation zones in Korea are integrated and managed as GIS data in 'National Heritage GIS' and 'Ancient City GIS'. In the first step of data collection, we checked the information of excavation in Gyeongju and collected the research reports. This WebGIS provides vector data in polygon type representing the excavation zone and attribute data on the kinds and ages of the identified features, and it is possible to download the research report, which is the raw data of these data. However, this GIS platform does not

provide information on the placement of the identified features



in the research zone. Using this national GIS, we collected

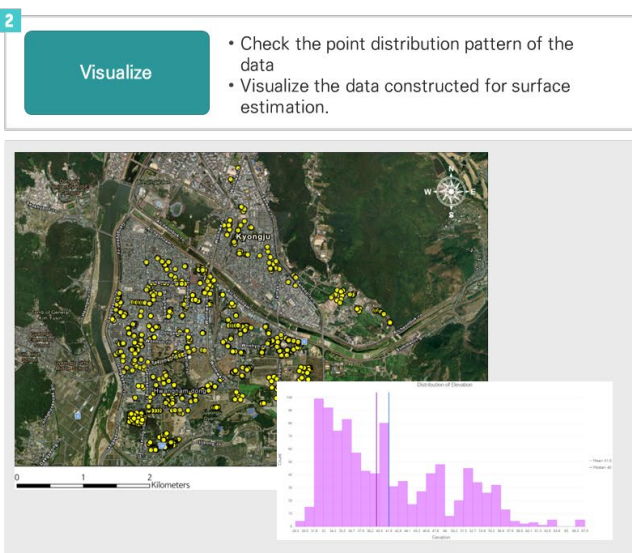
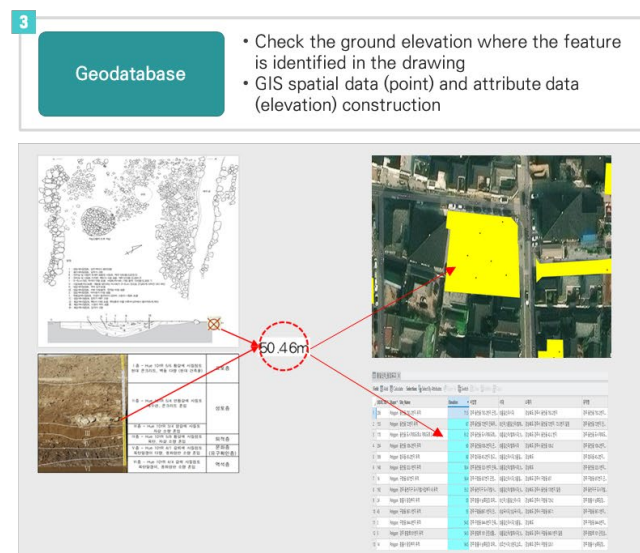
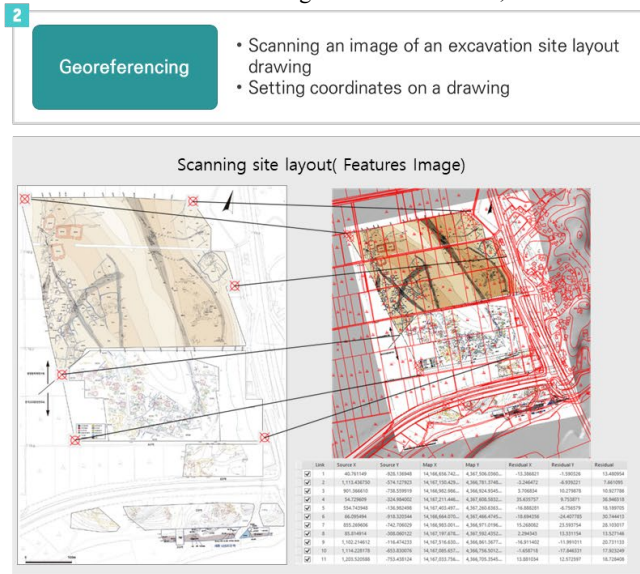


Figure 4. Process of Data Collection and Geodatabase Construction.

polygon data of the excavation zone in Gyeongju, and selected polygons corresponding to the Unified Silla period, the period when the Silla capital was the most prosperous. Then, we scanned the layout map that shows the placement of features in the associated report.

This scanning image does not have coordinate information, so spatial analysis cannot be performed on it. Here, we obtain spatial coordinates through georeferencing, which involves linking the image to a coordinate system to give it a location on Earth. The first step is to georeference the image to polygon data downloaded from a national GIS and determine the location of features such as buildings, fences, wells, roads, etc. that are located within the excavation zone.

The geodatabase required to estimate the land surface of the Silla capital should consist of point geodata representing the spots where features are identified and attribute data that can be connected to them to check the elevation of each point. After that, the elevation recorded in the cross-sectional plan of the

excavation site layout or the geological stratigraphy of the site is checked, and this is input as attribute data to finally complete the geodatabase.

3.3 Interpolations

In archaeological research, DEMs are often used to effectively represent the undulations of the terrain visually (Herzog and Yopez, 2015). Many GIS-based studies, such as archaeological predictive modelling, line-of-sight analysis, and least-cost path analysis, are possible using the elevations input into this raster data, and there is ongoing debate about their use (Conolly and Lake, 2006). As such, in archaeology, where research on the past land surface is conducted, modelling using separate geostatistical methods is necessary to produce a useful DEM. In this study, we use such a methodology to reconstruct the land surface of Gyeongju when it was the capital of Silla, based on a DEM based on the elevation of the excavated site.

The data used to construct this DEM does not represent all the land surface information of the Silla capital. It is just a sample of data from a few points in the ancient city. Although this is not many data, it is possible to estimate the historical land surface through spatial interpolation. IDW (Willmott *et al.*, 1985) and Kriging (Dirk *et al.*, 1998; Lynch, 2001; Zhao *et al.*, 2005) are the most used spatial interpolation methods to estimate the values of unobserved points that exist between these known data. In this study, we use these two interpolation methods to make DEMs and measure their precision to reconstruct the historical land surface of the capital city as precisely as possible.

IDW is a method that values unobserved points using a linear weighted combination that gives more weight to observation points in proximity (Park and Kim, 2013). The weights are inversely proportional to distance, and the surface is the dependent variable. This interpolation only considers the distance between observations, which is useful when you want to emphasize the influence of observations. The formula is:

$$y_p = \sum_{i=1}^n y_i w_i / \sum_{i=1}^n w_i, w = 1/d_i^k$$

Where y_p is the predicted value, y_i is the observed value, d_i is the distance from the observation point to the predicted point, n is the total number of observations used in the interpolation, and w_i is the weight on the distance of the known observations used in the interpolation used to determine the value of the point to be interpolated. As w_i increases with k , the closer observations have more influence.

Kriging is a method that considers trends in observations and uses spatial autocorrelation between measurements to construct a statistical model. Like IDW, Kriging weights the measurements around the point you want to interpolate, but unlike IDW, which uses an algorithm based simply on distance, Kriging uses a semi-variogram to capture the spatial structure of the data. A semi-variogram is the square of the difference between the attribute values of points located in space, and Gaussian, Exponential, and Spherical models are often used. The values calculated through these semi-variograms can identify the overall spatial distribution structure of the usage data and provide an optimal model for the distribution of spatial measurements. The variogram $2\gamma(h)$ between two points separated by a distance h is the covariance of the data values as a function of distance, expressed as $2\gamma(h) = \text{VAR} [y(h + h) - y(h)]$, and if there are n pairs of such data points, the semi-variogram $\gamma(h)$ is calculated as follows (Kim *et al.*, 2020).

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^n [y(x_i) - y(x_i + h)]^2$$

However, it is important to note that when creating a DEM using spatial interpolation, the accuracy of the data model must be evaluated. Cross validation is used to evaluate the model. This validation method shows how well the model derived from the interpolation predicts the unobserved values. This is often done using the leave-one-out cross validation (LOOCV) method, where for each point, one point is sequentially excluded, its value is predicted using the remaining data, and then the predicted value is compared to the measured value. The calculated statistics serve as a diagnostic indicator of whether the model is good for mapping.

The purpose of cross-validation is to aid in making informed decisions about which model provides the most accurate

predictions. The indicators used in this cross-validation are Mean Error (ME), Root Mean Squared Error (RMSE), and Root-Mean Squared Standardized Error (RMSSE) of the predicted data (ESRI 2003). For a model that provides accurate predictions, if the predictions are unbiased, the ME should be close to 0; if the standard error is accurate, RMSSE should be close to 1; and if the predictions are close to the measured values, RMSE should be small. The formula for this is as follows:

$$\text{ME} = \frac{\sum_{i=1}^n (\hat{z}(s_i) - z(s_i))}{n}, \text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\hat{z}(s_i) - z(s_i))^2}{n}}$$

$$\text{RMSSE} = \sqrt{\frac{\sum_{i=1}^n [(\hat{z}(s_i) - z(s_i)) / \hat{\sigma}(s_i)]^2}{n}}$$

$\hat{z}(s_i)$ is the predicted value from cross-validation, $z(s_i)$ is the observed value, and (s_i) is the prediction standard error for the location of s_i .

Since the kriging applied here does not know the average elevation value of the entire excavation site, ordinary kriging was applied to estimate it.

3.4 Materials

The data used for spatial interpolation is based on the excavation information of the Silla capital site, and the elevation values of 292 sites and 985 points were extracted (Figure 5). This data is based on the ground surface where features of the Unified Silla period were investigated, and as many sites as possible were selected in consideration of the density of sites.

Sample size is important when conducting spatial statistical analysis. If the sample size is not sufficient, the analysis results will be less reliable, accurate, and interpretable. Given the distance between the excavated sites (median 56 meters), we want to create a DEM with a cell size of 50 meters. Based on this, the population size would be 5,337 cells and based on commonly used values (95% confidence level, $\pm 5\%$ margin of error, population ratio of 0.5), a minimum of 360 samples would be required. In this study, the data used for interpolation is distributed in a total of 469 cells, so we can say that we have a sufficient sample size for the analysis.



Figure 5. Distribution of data used for interpolation of the surface of the Silla capital.

4. Results

4.1 IDW Interpolation

IDW relies on a value representing the inverse of the distance as a mathematical power. This power allows you to control the significance of known points in the interpolated value based on their distance from the output point. This parameter is a positive real number, and by defining a higher power value, more emphasis can be put on the nearest points. Here, we chose the optimized power, which provides the model with the lowest error, considering the RMSE of the model generated by the interpolation method. The searching neighborhood type was divided into standard and smooth.

The models generated by the standard neighborhood type were cross validated against the models generated by the following neighborhood search area types: 1sector, 4sector, 4sector with 45° offset, and 8sector. The validation results are shown in Table 1 below. Among them, when looking at the precision of the prediction results based on RMSE, the 4sector with 45° offset zone type was the best. However, the RMSE of the models with other sector types does not show much difference.

Model Name	Sector Type	ME	RMSE
IDW-ST1	1 Sector	0.017	0.450
IDW-ST2	4 Sector	0.019	0.449
IDW-ST3	4 Sector 45° offset	0.018	0.448
IDW-ST4	8 Sector	0.020	0.450

Table 1. Cross-validation results for interpolation models by standard neighborhood type.

Next, let's look at the model with the smooth neighborhood type. Here, the smooth factor varies the search zone, and we interpolate all the point data by quartering the 0.2 to 1 factor. The results are shown in Table 2. Since the RMSEs are all similar, there is no difference between the spatial interpolation models.

Model Name	Smoothing Factor	ME	RMSE
IDW-SM1	0.25	0.020	0.452
IDW-SM2	0.50	0.020	0.452
IDW-SM3	0.75	0.020	0.452
IDW-SM4	1	0.019	0.452

Table 2. Cross-validation results for interpolation models by smooth neighborhood type.

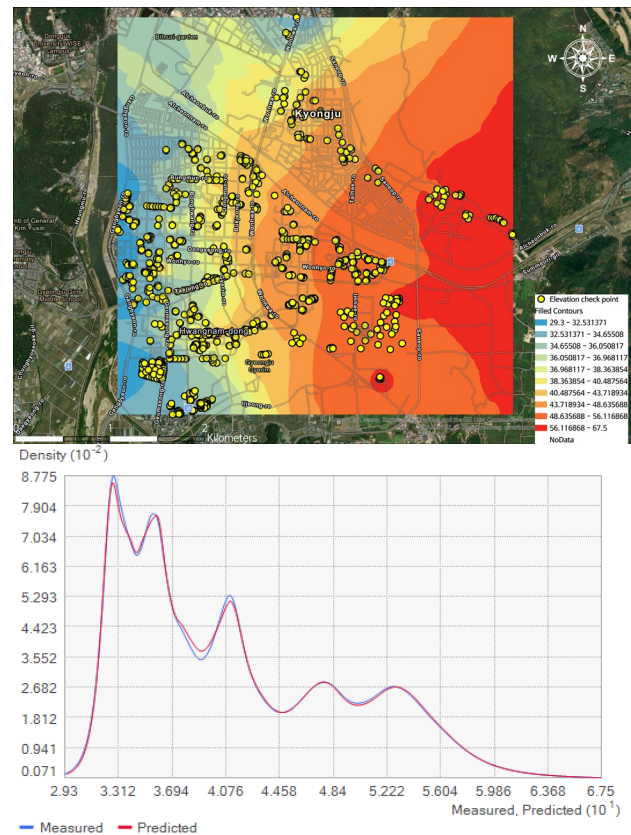


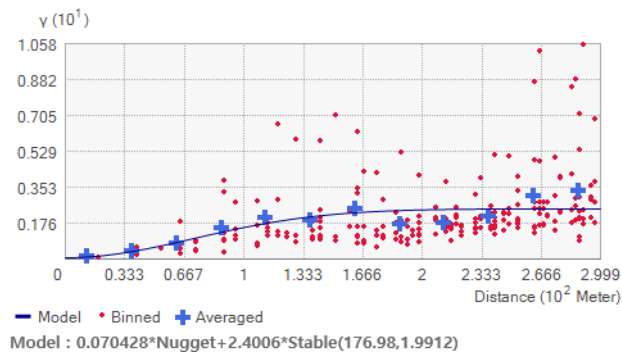
Figure 6. Error graph of a prediction model resulting from applying a searching neighborhood of 4sector 45° with offset of standard neighborhood type and the DEM.

The spatial interpolation results using IDW above show that the ME is close to 0 and the RMSE is low. This suggests that the DEM created by IDW is a model that can be assured of precision. The error graph and DEM of the prediction model generated by applying the searching neighborhood of 4sector 45° with offset of the standard neighborhood type, which is the most precise among them, are shown in Figure 6.

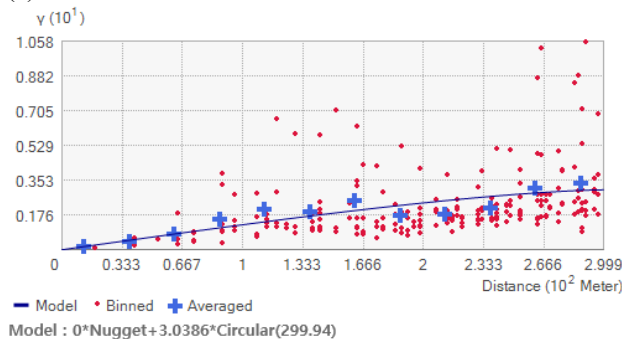
4.2 Kriging Interpolation

For interpolation by Kriging, linear, circular, spherical, exponential, and gaussian semi-variograms were applied to select the appropriate semi-variogram. The semi-variogram graphs of each functional model and the model equations representing the nugget effect are shown in Figure 7. The nugget values, which show the variability of the data applied to the analysis, show that the circular and spherical semi-variograms are more accurate.

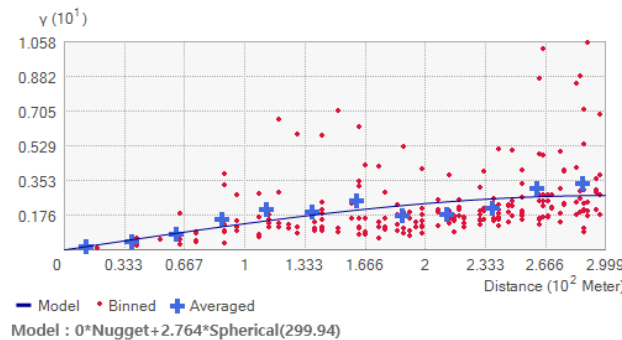
(1) Liner



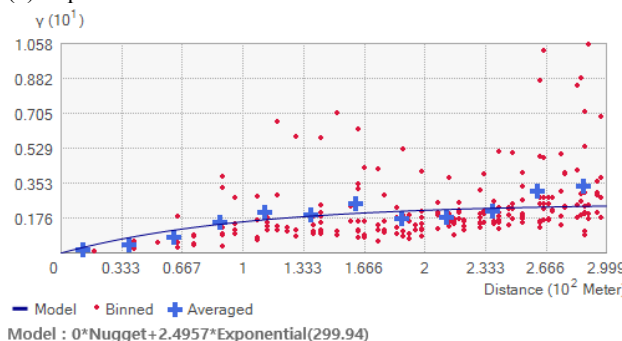
(2) Circular



(3) Spherical



(4) Exponential



(5) Gaussian

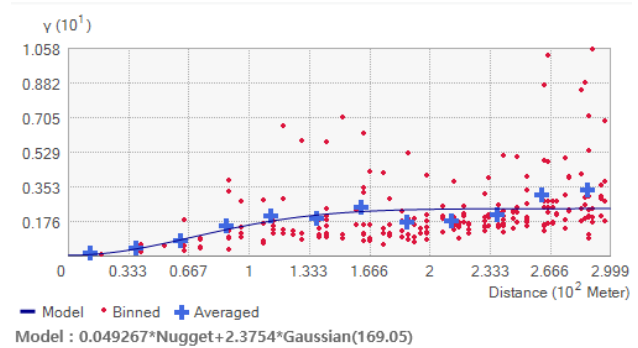


Figure 7. Semi variogram model graph.

The above seme-variogram was applied to generate models for standard and smooth searching neighborhood types. First, the cross-validation results of the models generated by the standard neighborhood type are shown in Table 3. Based on the RMSE, which can comprehensively judge the prediction accuracy of a model, the most accurate model among the five semi-variogram is the one with the circular function, which shows a low RMSE value close to zero. This model also has an RMSSE close to 1, indicating that the model has a good structural fit.

Model Name	Semi-variogram	ME	RMSE	RMSSE
KRI-M1	Linear	0.006	0.480	1.123
KRI-M2	Circular	0.012	0.449	1.076
KRI-M3	Spherical	0.012	0.451	1.038
KRI-M4	Exponential	0.016	0.454	0.786
KRI-M5	Gaussian	0.007	0.489	1.293

Table 3. Cross-validation results of models generated by the standard neighborhood type of kriging.

Model Name	Semi-variogram	Smoothing Factor				
		0.2	0.4	0.6	0.8	1
KRI-M1	Linear	0.481	0.480	0.475	0.468	0.458
KRI-M2	Circular	0.446	0.445	0.445	0.444	0.442
KRI-M3	Spherical	0.446	0.445	0.445	0.444	0.443
KRI-M4	Exponential	0.448	0.448	0.448	0.448	0.449
KRI-M5	Gaussian	0.482	0.485	0.479	0.470	0.457

Table 4. Cross-validation results for smooth neighborhood type of kriging model by RMSE.

The following are the cross-validation results of the models generated with the smooth neighborhood type (Table 4). The smoothing factor can be set between 0 and 1. Here, we evaluated the performance of the models by setting the smoothing factor in 0.2 intervals across the entire range and measured the accuracy of the models based on RMSE, which is the preferred metric when comparing the performance of the models. Overall, RMSE tends to be lower when the smoothing factor is 1. Among them, we can see that the model with circular semi-variogram has the best performance. This is the same as the result of applying the smooth neighborhood type. This model (KRI-M2) is also the most accurate among all models with IDW and Kriging based on RMSE (Figure 8).

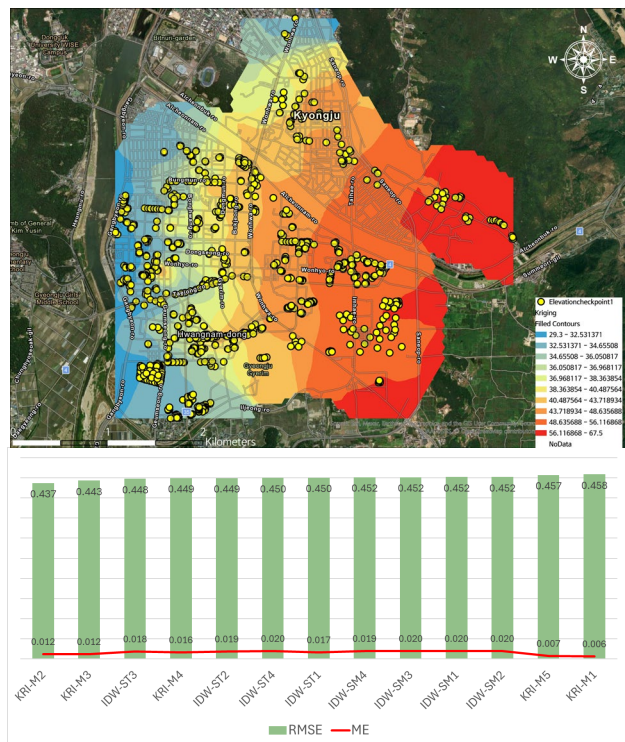
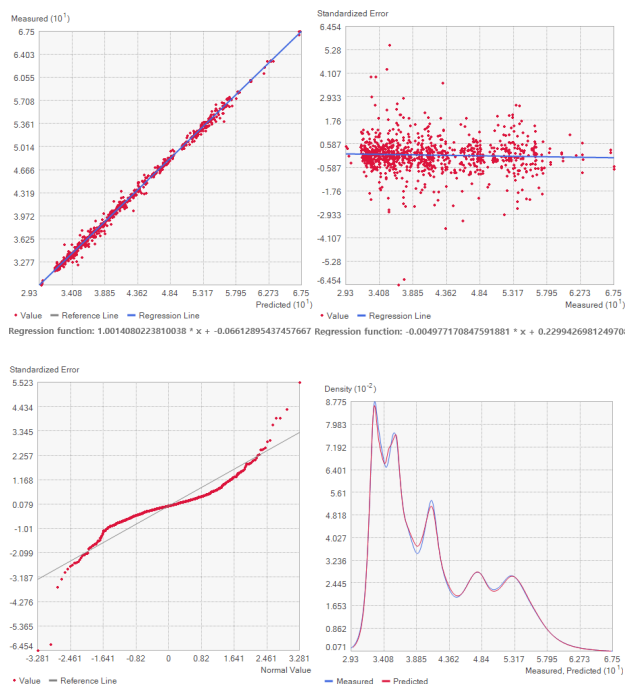


Figure 8. Accuracy of IDW and Kriging models by RMSE and ME.



The prediction error graph and visualization of the KRI-M2 model is shown in Figure 9.

Figure 9. Prediction error graphs for the KRI-M2 model and visualization of the model.

4.3 Comparing the surface of the present-day and Silla capitals

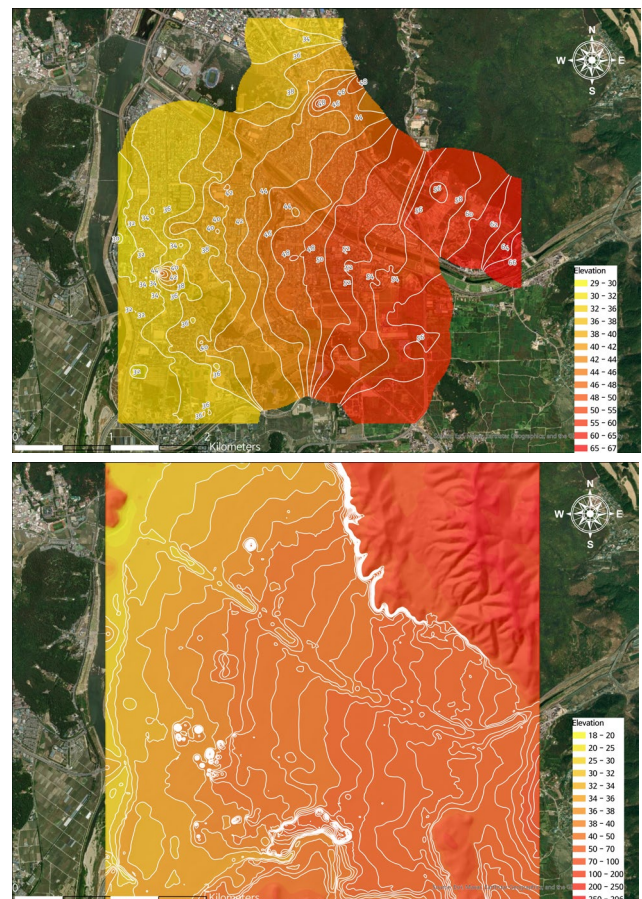


Figure 10. Comparison of KRI-M2 DEM (top) and current DEM (bottom).

Figure 10 shows a comparison of the land surface DEM of the Silla capital generated by the KRI-M2 model with the current surface DEM. In general, the trend of decreasing elevation from east to west is the same. However, when comparing the western area adjacent to the river, it can be seen that the surface of the Silla capital in the past was highly curved, while the current surface is gently sloping. This difference is especially evident in the area below 40 meters altitude. In this area, many excavations have been carried out, so there is ample information on the past landscape. This means that the KRI-M2 model is a good representation of the past land surface, and it tells us that the current land surface has changed a lot compared to the past. In other words, the topography of the Silla period was very different from the present. This can be further explored in Figure 11, which shows a cross-sectional comparison of the land surface in the past and present. In some cases, the land surface is much lower than it was during the Silla period, and in other cases, artificial soil has been piled up to create a higher land surface than in the past.

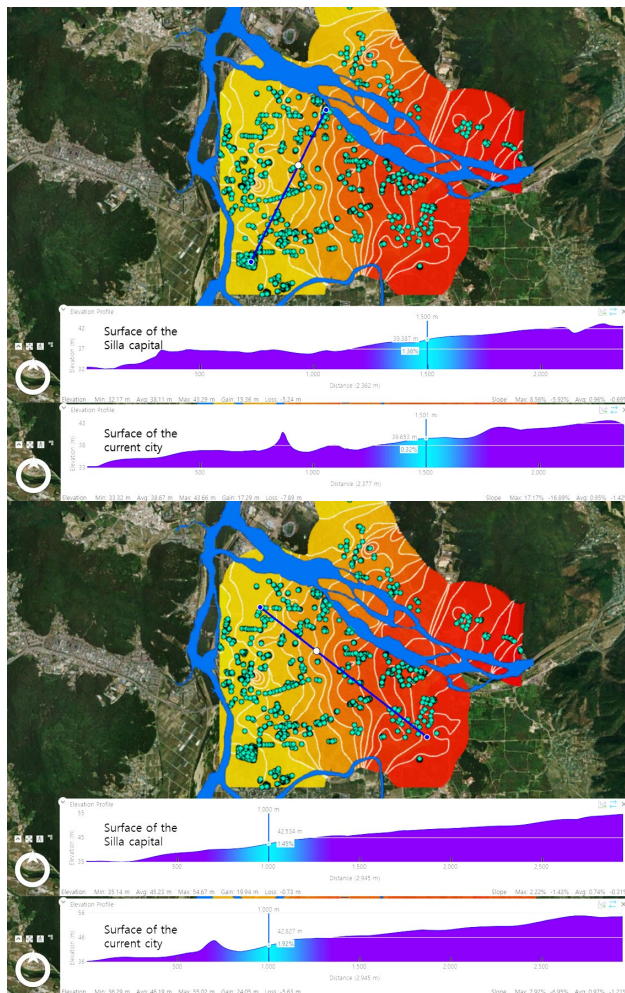


Figure 11. Comparison of land surface cross-sections from the Silla period and today.

As such, areas bordering rivers during the Silla period had undulating ground surfaces, unlike today. The Silla Dynasty greatly improved the capital city, installing grid-like roads in various sections. However, this urban grid was not implemented in areas close to rivers. This is thought to be a reflection of the topographical conditions at the time. The land was highly utilized for residential purposes, as many sites have been discovered here. However, due to the curvature of the land surface, this area is not thought to have been subject to urbanization, which would have required large-scale civil works.

5. Conclusion

As the eastern terminus of the Silk Road and the site of several UNESCO World Heritage Sites, the Silla capital is being excavated for its many urban sites, including tombs, temples, buildings, and roads. These ruins are mostly found below the current surface and provide information about the surface where citizens lived in the past. In this study, we used IDW and Kriging interpolation, a spatial geostatistics technique, to estimate these surfaces. The data used here is the elevation of 985 points, a total of 292 sites, based on excavations of the Silla capital ruins. This data is basically spatial information downloaded from WebGIS operated by the Korean government, and the spatial database was built by checking the elevation of the points where Silla-age features were identified in the excavation report.

IDW interpolation was conducted by dividing the searching neighborhood type into standard and smooth. Four models were created for each type and cross-validated, and the model with 4 Sector 45 ° offset showed the lowest RMSE. All models with smooth factors showed the same performance. Kriging interpolation was also separated by searching neighborhood type, and prediction models were generated by applying linear, circular, spherical, exponential, and gaussian semi-variograms. The cross-validation results showed that the model with the circular semi-variogram with the smooth factor set to 1 had the lowest RMSE. In order to estimate the ground surface of ancient urban sites, it is necessary to create interpolation models by setting various parameters, and cross-validate them to ensure the accuracy of the model.

The DEM of the Silla capital was finally created through this verification process. By comparing it to the current land surface, we found that the terrain in areas close to rivers at an elevation of less than 40 meters above sea level was highly curved, unlike today. These topographic conditions probably prevented the grid-like road network from being installed here. Furthermore, the comparison of the present and past topographical cross-sections revealed that the ground surface had been cut and graded in the past. As such, the exposed surface of the features identified in the excavation provides important information to reconstruct the living surface of the citizens who created them. If additional spatial information on the excavated ruins in Gyeongju can be obtained in the future, the accuracy of the interpolation model is expected to be much better than it is now. This, in turn, will be fundamental to the reconstruction of the Silla capital's landscape.

We hope that with the accumulation of more historical ground surface reconstruction data than is currently available, we can reconstruct the topographic landscape of the Silla Capital at the time of its establishment. This could significantly contribute to the preservation and management of ancient urban sites by providing information on the probability of archaeological remains being identified at specific points beneath the current surface. Additionally, it could be used as crucial data to understand the topographic factors that influenced urban development and to examine the spatial context among urban components.

References

- Dirks, K.N., Hay, J.E., Stow, C.D. and Harris, D. 1998. High-resolution studies of rainfall on Norfolk Island Part II: Interpolation of rainfall data. *Journal of Hydrology* 208, pp.187-193.
- ESRI, 2003, *Using_ArcGIS_Geostatistical_Analyst*.
- Jeon, D.J., 2005, The Implementation and Characteristics of Silla's Bangri System, *The Structure and System of Silla's Capital*, Proceedings of the Silla Cultural Festival Academic Conference 26, pp.120, 130.
- Kang, B.W., 2008, A Historical and Archaeological Study on the Population of Silla's Capital, *Daegu History* 90, pp.1-37.
- Lee, E.S., 2004, The Establishment and Development of the Capital, *The 28th National Conference of Korean Archaeology: Archaeology of the Unified Silla Period*, Korean Archaeological Society, pp.9-25.

Lee, S.Y., Leidy, D.P., Rhi, J.Y, Lee, I.S., 2013, *Silla: Korea's Golden Kingdom*, The Metropolitan Museum of Art.

Lynch, S.D., 2001, *Converting point estimates of daily rainfall onto a rectangular grid*, Department of Agricultural Engineering, University of Natal, South Africa. <http://gis.esri.com/library/userconf/proc98/proceed/TO200/PAP196/P196.HTM>.

Min, D.S, 1986, A Study on the Urban Design and Management of Silla's Capital, *The Paek-San Hakpo* 33, pp.9-62.

Park, J. C. and Kim, M. K., 2013, Comparison of precipitation distributions in precipitation data sets representing 1km spatial resolution over South Korea produced by PRISM, IDW, and cokriging, *Journal of the Korean Association of Geographic Information Studies* 16, pp.147-163.

Zhao, C., Nan, Z. and Cheng, G. 2005. Methods for modelling of temporal and spatial distribution of air temperature at landscape scale in the southern Qilian Mountains, China. *Ecological Modelling* 189, pp.209-220.