Integrating GIS and 3D geostatistical methods for geotechnical characterization of soil properties

YOSOON CHOI[1] & HYEONG-DONG PARK[2]

[1] *School of Civil, Urban and Geosystem Engineering, Seoul National University. (e-mail: yspower7@snu.ac.kr)*
[2] *School of Civil, Urban and Geosystem Engineering, Seoul National University. (e-mail: hpark@snu.ac.kr)*

Abstract: Traditional approaches using Geographical Information System (GIS) based sample data have had to assume that all data are arranged in 2D layers. These approaches facilitate integrating GIS and geostatistical methods, but give difficulties when estimating 3D spatial variables owing to loss of the z-coordinate information. This study presents a new model which can integrate GIS and 3D geostatistical methods to predict variation of soil properties. Within the model, geotechnical variables are generated using 3D geostatistical methods and can update the GIS database. This provides a useful tool for estimating 3D geotechnical variables with a measure of reliability, making it possible to use the estimated results more efficiently within the GIS database. It is expected that the practical model suggested in this study can be used widely to develop a more effective GIS basis for construction projects in the urban environment.

Résumé: Les approches traditionnelles utilisant le Système Informatique Géographique (SIG) basé sur les données d'échantillon ont supposé que toutes données sont organisées dans 2D couches. Bien que ces approches aient grand avantage de s'intégrer facilement au SIG et aux méthodes géostatistiques, ils se donnent des difficultés pour estimer en 3D variables spatiales dû à la perte d'information coordonnée Z des données d'échantillon. Cette étude présente un modèle pratique pour intégrer dans SIG et dans 3D méthodes géostatiques pour prédire les variés de propriétés de sol précisément. Dans le modèle, les variables géotechniques sont établis par 3D méthodes géostatistiques et peuvent être mises à jour à la base de données de SIG. L'intégration de SIG et 3D méthodes géostatistiques fournissent un outil utile pour estimer les variables géotechniques de 3D avec plus de fiabilité, en utilisant efficacement les résultats estimés dans la base de données de SIG. Il est prévu que le modèle pratique suggéré dans cette étude sera utilisé largement pour développer plus efficace des applications basé sur SIG pour les projets de construction urbaine.

Keywords: 3D models, Data analysis, Data visualization, Database system, Geodata, Hydraulic conductivity.

INTRODUCTION

Geostatistics provides a tools for analysing spatial data and estimating unknown values (Issaks & Srivastava 1989). Geostatistics was first used for estimating the quality of mineral deposits, but has been extended to areas such as hydrogeology and engineering geology (Houlding 1994). Geostatistics is now offered as a tool within GIS such as ArcGIS (ESRI, USA), IDRISI (Clake Lab, USA), and GRASS (CERL, USA). ArcGIS provides a 'Geostatistical Analyst' tool which incorporates a framework for integration with a GIS database (Johnston *et al.* 2003). IDRISI and GRASS also have data conversion tools for coupling independent geostatistical software (e.g. Gstat) with GIS (Eastman 1999; Lennert 2004). Such GIS software has the limitation of using only two-dimensional (2D) spatial data, so cannot be used for estimation of spatial variables when data are distributed three-dimensionally.

Many researchers have tried to integrate geostatistical methods with GIS. Rosenbaum & Nathanail (1996) suggested a novel strategy for integration of geostatistical methods with GIS to characterize the underground rock properties using insufficient *in situ* data. Another example is the assessment of contaminated soil (Zhu, Charlet & Poffijn 2001; Zhang & McGrath 2004).

The main objective of this study is to present an approach for integrating GIS with 3D geostatistical methods to predict the variation of soil properties. The 2-dimensional approach inevitably leads to information loss when the data are distributed three-dimensionally. Figure 1 demonstrates the problems of the 2D approach for estimation of 3D spatial variables. The 2D approach may project data onto a surface, neglecting the information in the vertical dimension (Fig. 1a), or may decrease the number of available data by using the data located in the same depth (Fig. 1b). In order to overcome such limitations, 3D geostatistical methods have been applied GIS. Figure 2 shows the difference between the 2D and 3D approach. In the 3D approach, all data distributed in 3D space can be used for the estimation of 3D spatial variables. The problem is how best to couple GIS-based data with 3D geostatistics. Hydraulic conductivity on the Ilwon-Dong in Seoul, South Korea is used as an example.

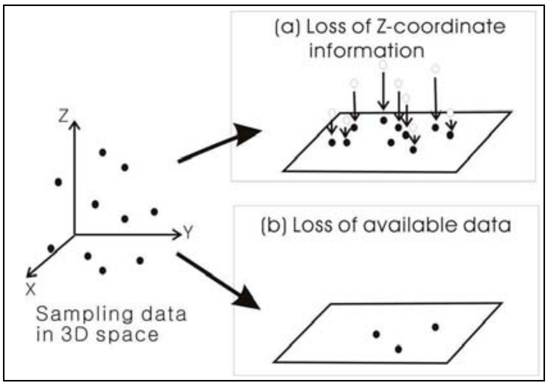

Figure 1. Limitations of the 2D approach in geostatistical estimation.

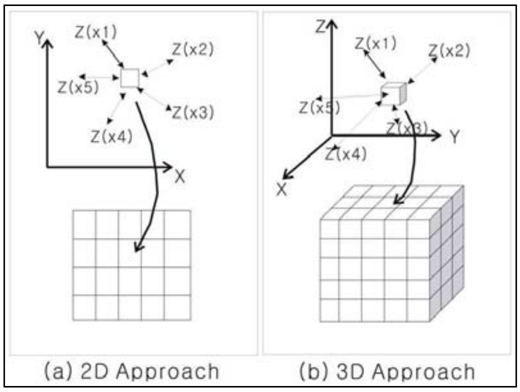

Figure 2. 2D vs. 3D in geostatistical estimation.

COUPLING 3D GEOSTATISTICS WITH GIS

Analysis of three-dimensional data structure

Although much recent research has continued to focus on 2D integration of geostatistics with GIS (Rosenbaum & Nathanail 1996, Zhu et al. 2001, Rosenbaum & Turner 2003, Zhang & McGrath 2004) because of the efficiency and speed of analysis, preservation of 3D information remains a goal.

Figure 3 illustrates some data structures for managing 3D geoscience data. Hole data clearly includes 3D information and corresponding attributes such as lithology, stratigraphy, and rock quality (Choi & Park 2005). The TIN data structure generally represents topological surfaces in vector GIS ("2.5"D volumes) and is popular in CAD and surveying packages. The 3D point data structure is suitable for managing both geometry and attributes of sampling data collected from field investigations. For a solid model, a geometric boundary is set up first, and then a 3D grid data structure divides the database into a finite grid system and allocates attribute values to each grid cell. This creates a kind of 3D raster data structure with four variables (x, y, z coordinates & attribute value).

Such a 3D grid is very effective for representing 3D objects. Its potential coupling with 3D geostatistics is now considered.

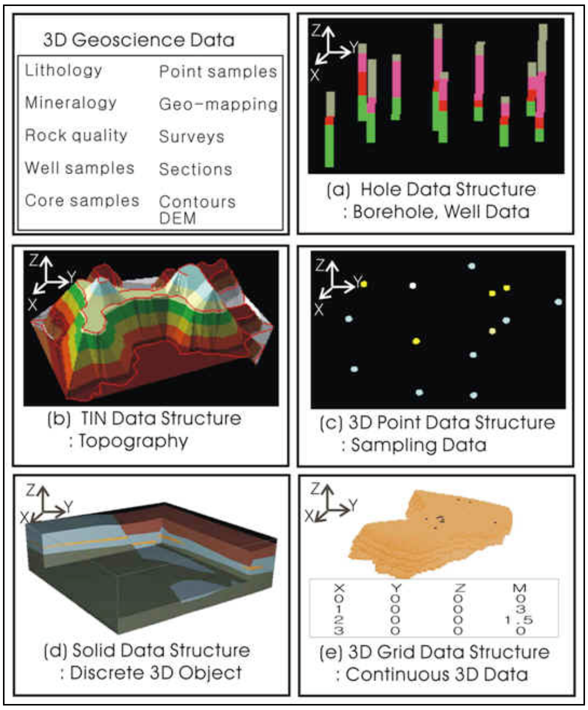

Figure 3. 3D data structures.

Coupling 3D geostatistics with GIS by data conversion

The concept of data conversion method is shown in Fig. 4. The first step is to construct a GIS database using its 3D point data structure. The x, y, z coordinates and attribute values of the sample data (e.g. underground rock properties) are stored in a GIS table. In order to convert 3D points to a 3D grid data structure, the concept of moderated separation distance is applied. Moderated separation distance z-coordinates of 3D points are approximated by multiple 2D point vector layers (one of the main data structures in a GIS). The multiple 2D vector layers containing the sample data are then converted to multiple 2D raster layers.

After data conversion from 3D points to a 3D grid, 3D geostatistical estimation can be performed. The resulting estimates are recorded again using the 3D grid structure, as multiple 2D raster layers. Users can access this data in 2D form from the GIS database, and use them for reclassification or display.

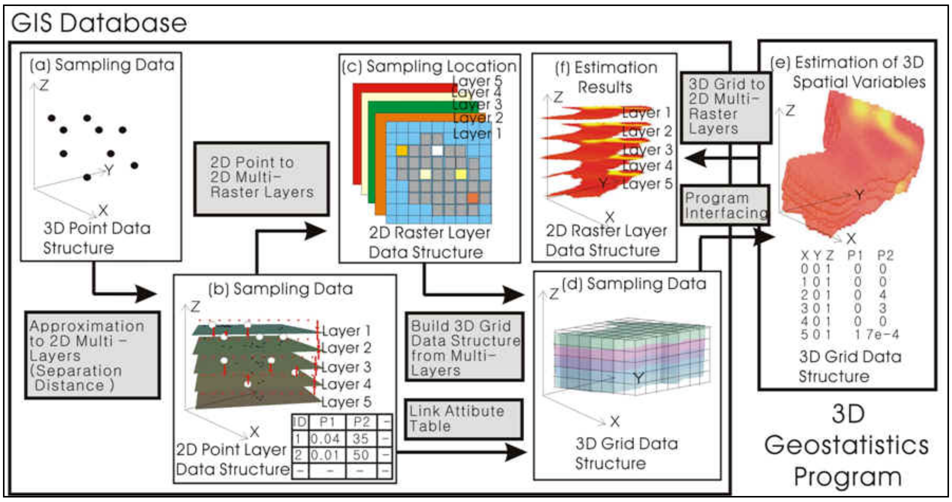

Figure 4. Concept of data conversion for integrating GIS with 3D geostatistical methods.

Program Development

Data conversion program

Data conversion comprises two procedures: (1) data conversion using GIS software and (2) coupling GIS database and external 3D geostatistics programs. For the coupling, a new program using Visual Basic 6.0 has been developed (Fig. 5). This can open and save an ASCII file of a 3D grid structure as shown in Fig. 6. The ASCII file consists of two parts: header and body. The header part includes the number of samples, the size and the number of grid system, and the referencing point coordinates for back-conversion to the GIS database. The body part contains the types and values of the variables. The first three columns are for coordinate information. The 'SampleT' variable gives information on the availability of the sample at the location determined by the first three columns, the 'Active' variable is for determining whether the point is located in the geometrical boundary or not, and the 'Value' variable is the data of sample value. Here, zero in the Value column means null, i.e. there is no value of spatial data within the cell. By 'Active' variable, an irregular grid system can be expressed and geostatistics applied.

Geostatistics program

Figure 7 shows the 3D geostatistics program used for this study. After importing the ASCII file created above (Fig. 7a), the sample location and geometry boundary are visualised three-dimensionally (Fig. 7b). The program provides three estimation methods: (1) 3D polygon method, (2) 3D inverse distance weighted (IDW) method, (3) 3D kriging method. Details of the methods and their theoretical background are presented in Isaaks & Srivastava (1989). Each method has various options for variogram modelling (Fig. 7c). 3D spatial variables can thereby be estimated and visualised three-dimensionally (Fig. 7d). Moreover, the estimation result can be saved in ASCII format (Fig. 6) for back-conversion to the GIS database.

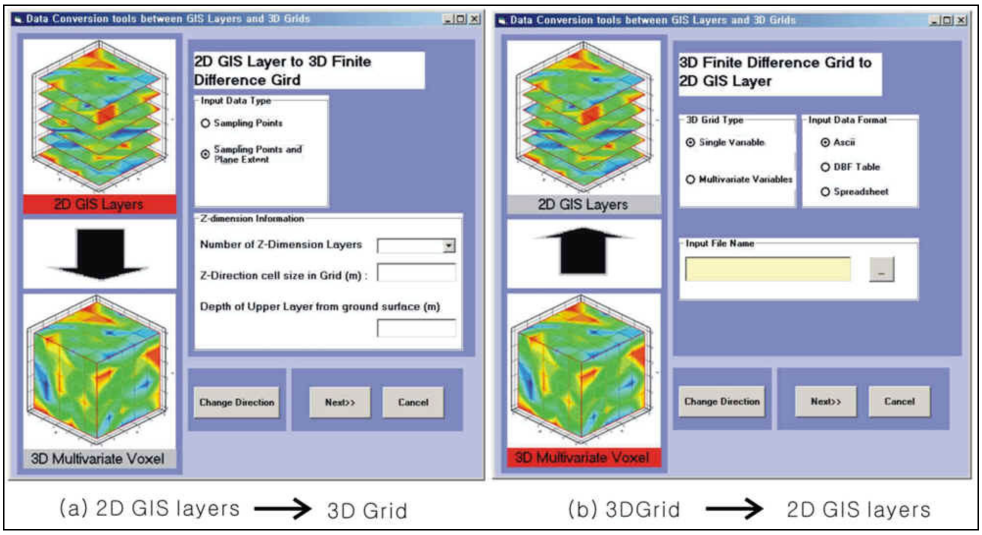

Figure 5. Data conversion program for coupling GIS database with a 3D geostatistics program

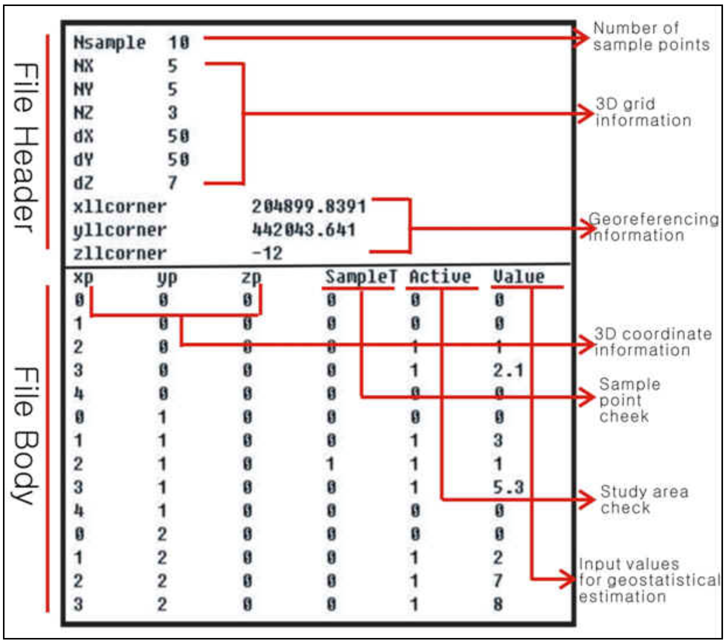

Figure 6. Data conversion program for coupling GIS database with a 3D geostatistics program

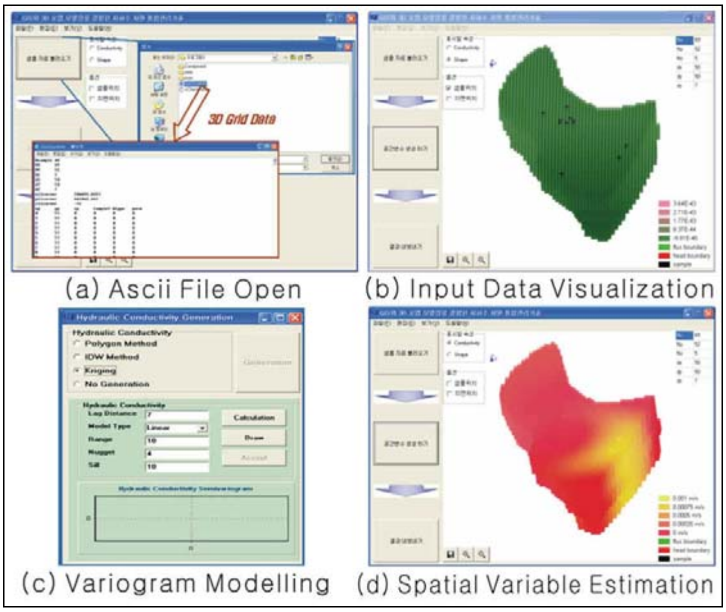

Figure 7. 3D geostatistics program for spatial data estimation

APPLICATIONS OF THE DEVELOPED COUPLING METHOD

Objectives

To validate the applicability of the newly developed method, the values of hydraulic conductivity of Ilwon-Dong, Seoul, South Korea were estimated. The x, y, z coordinates and attributes of the sample data were presented as shape files (ESRI 2004). Figure 8 shows the study area with a 3D visualisation created by ArcGIS 3D Analyst for the distribution of sample data. The area covers about 1 × 1 km, and the resolution has 10m pixels. The 63 original sample data (measurements) have been used to estimate values for the hydraulic conductivity in 3D.

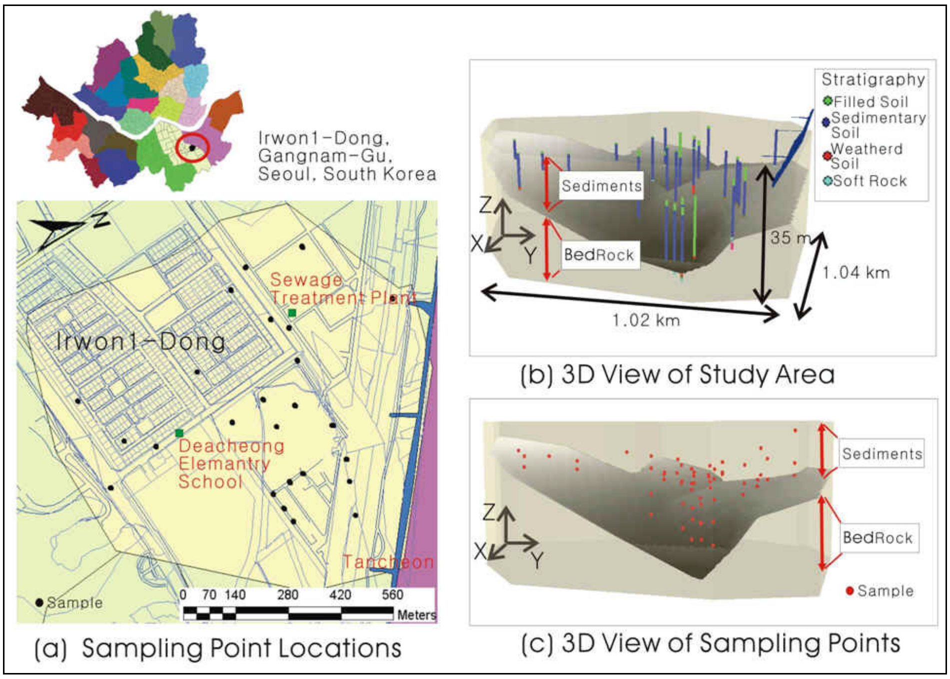

Figure 8. Study area: Ilwon-Dong, Seoul, South Korea

Estimation results

The estimation results for the 3D hydraulic conductivity are shown in Fig. 9. 3D kriging and the 3D IDW method have estimated a smooth distribution of hydraulic conductivity. However, estimation results using the 3D polygon method showed drastic changes in value in some areas. The reason may be that the local abnormal values have greater influence in the polygon method than the others.

Figure 10 shows the difference between 2D kriging and 3D kriging method. In 2D kriging, each sliced layer presents significantly different results. On the other hand, 3D kriging creates similarity between each layer. It is therefore believed that the estimated result using the 3D geostatistical approach helps overcome the limitations of the 2D approach.

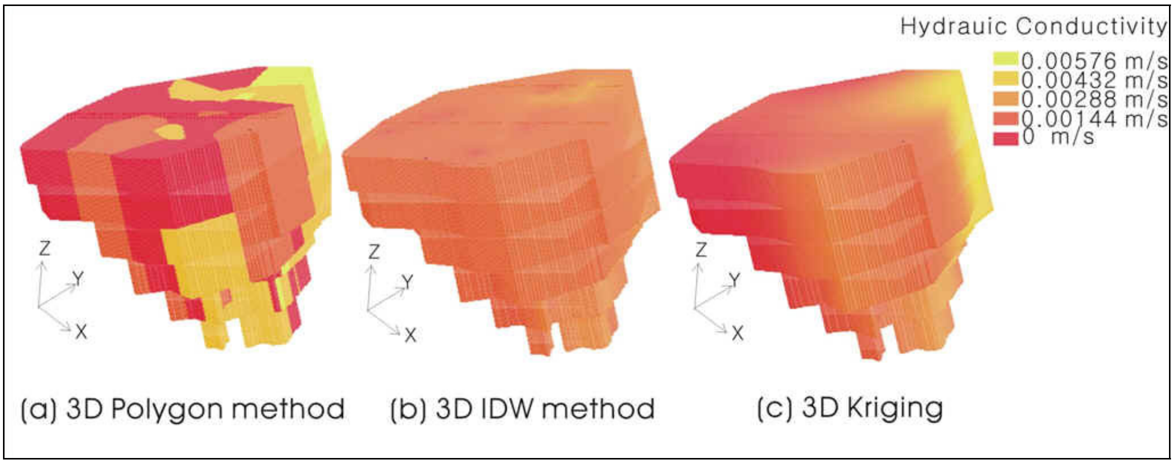

Figure 9. Estimates for hydraulic conductivity using 3D geostatistical methods

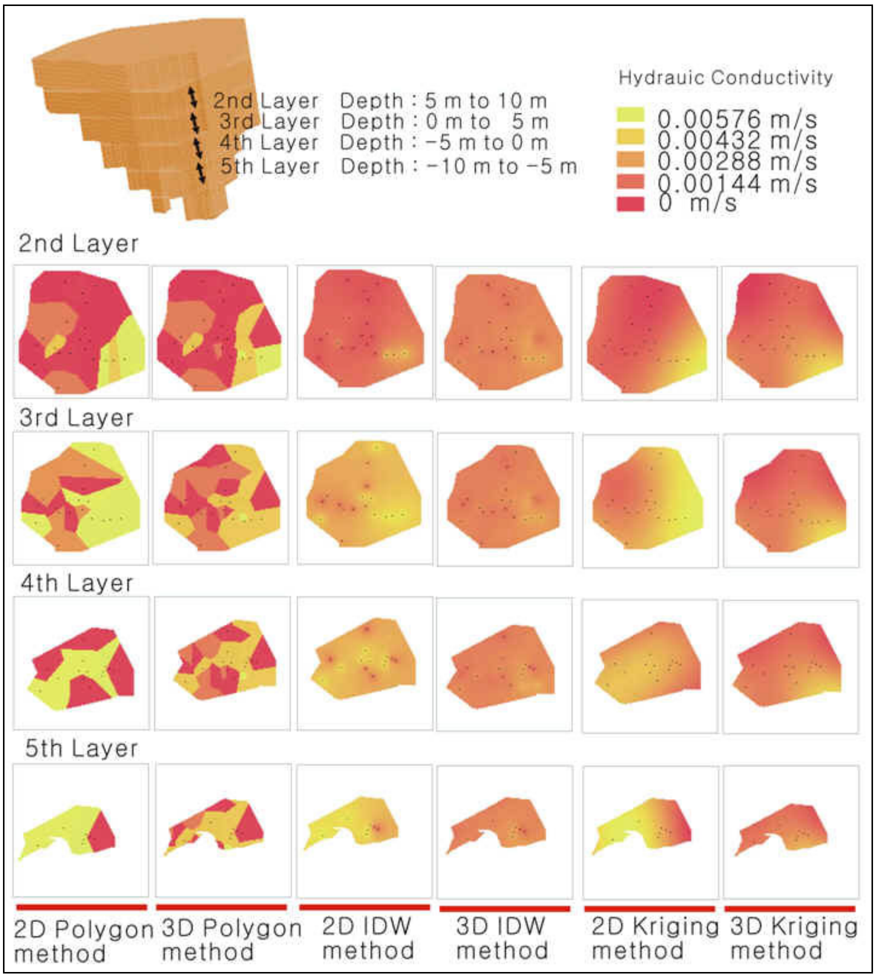

Figure 10. Comparing 2D with 3D estimation

Combining estimated results within the GIS database

The estimated results obtained by 3D geostatistics can be saved as an ASCII file and transferred back to the GIS database. Figure 11 shows the result of converting such an ASCII file to its 2D raster data structure. Through this procedure, estimated values can be reclassified if needed, and combined with other GIS layers as a part of the modelling phase. Combined with existing GIS database information, raster data generated by 3D geostatistical methods can be used together with other functions to present a powerful visualisation tool (Nathanail & Rosenbaum 1998; Lee et al. 2004), functions for integrating various data (Choi et al. 2003), and analysis and overlay of multiple layers represented by raster data (Zhang & McGrath 2004).

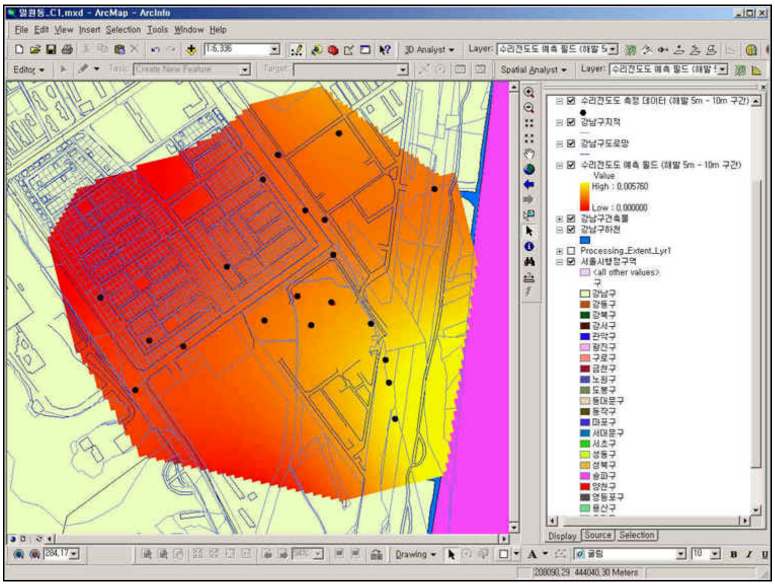

Figure 11. Integration of estimation result with GIS database

CONCLUSIONS

A model for integrating 3D geostatistics with GIS has been presented. Using a data conversion program and 3D geostatistics, 3D spatial variables have been estimated from sparse sample data. It is concluded that: (1) the limitation of two-dimensional approaches for integrating GIS and geostatistics can now be overcome, (2) by using 3D geostatistical method, estimated spatial data in a GIS database can be more reasonable than previous studies based on 2D approach, (3) 3D spatial variables such as hydraulic conductivity can be used and updated in a GIS database after back-conversion. It is expected that the practical model suggested in this study can be used widely to develop a more effective GIS-based application for construction projects in the urban environment.

Acknowledgements: This work was supported by a grant from the Construction Core Technology Research & Development Project of Ministry of Construction & Transportations in South Korea.

Corresponding author: Prof Hyeong-Dong Park, Seoul National University, School of Civil, Urban and Geosystem Engineering, Seoul, 151-744, South Korea. Tel: +82-2-880-8808. Email: hpark@snu.ac.kr.

REFERENCES

CHOI, S.Y., GO, W.R., YOON, W.J., HWANG, S.H. & KANG, M.K. 2003. Analysis of the Distribution Pattern of Seawater Intrusion in Coastal Area using the Geostatistics and GIS. *The Journal of GIS Association of Korea,* **11**, 251-260 (in Korean with English abstract).

CHOI, Y.S. & PARK, H.D. 2005. Development of an Intelligent Borehole Information System (IBIS) based on a Knowledge Approach. *GISPlanet 2005 Proceeding CD, Estoril, Protugal, May 30-June 2*.

EASTMAN, J.R. 1999. Idrisi 32: Guide to GIS and Image Processing. *Clark University: Worcester.*

ENVIRONMENTAL SYSTEMS RESEARCH INSTITUTE. 2004. What is GIS? : GIS by ESRI. *ESRI Press: Redlands, California.*

HOULDING, S.W. 1994. 3D Geoscience Modeling. *Springer: Berlin.*

ISSAKS, E.H. & SRIVASTAVA, R.M. 1989. Applied Geostatistics. *Oxford University Press: London.*

JOHNSTON, K., VER HOEF, J.M., KRIVORUCHKO, K., LUCAS, N. 2003. Using ArcGIS Geostatistical Analyst. *ESRI Press: Redlands, California.*

LEE, H.J., PARK, S.W., YOO, S.D. & KIM, H.T. 2004. An Estimation of Long-Term Settlements in the Large Reclamation Site and Determination of Additional Sampling Positions using Geostatistics and GIS. *Journal of the KGS,* **20,** 131-141 (in Korean with English abstract).

LENNERT, M. 2004. Grass Tutorial. *Grass Development Team (http://grass.itc.it/index.html)*, accessed (04-03-2004).

NATHANAIL, C.P. & ROSENBAUM, M.S. 1998. Spatial Management of Geotechnical Data for Site Selection. *Engineering Geology,* **50,** 347-356.

ROSENBAUM, M.S. AND NATHANAIL, C.P. 1996. Petrophysical Databases for Ground Characterization: Design Concepts and Considerations. *Marine and Petroleum Geology*, **13,** 427-435.

ROSENBAUM, M.S. & TURNER, A.K. (eds) 2003. New paradigms in subsurface prediction. Characterization of the shallow subsurface; implications for urban infrastructure and environmental assessment. Lecture Notes in Earth Sciences **99**. *Springer: Berlin.*

ZHANG, C., MCGRATH, D. 2004. Geostatistical and GIS Analyses on Soil Organic Carbon Concentrations in Grassland of Southeastern Ireland from Two Different Periods. *GEODERMA*, **119,** 261-275.

ZHU, H.C., CHARLET, J.M., POFFIJN, A., 2001. Radon Risk Mapping in Southern Belgium: An Application Of Geostatistical and GIS Techniques. *The Science of the Total Environment*, **272,** 203-210.