

Data Conversion of GIS to Urban Area Model for Natural Disaster Simulation

Kengo Tanaka*, Sobhaninejad Gholamreza**, Muneo Hori***,
Tsuyoshi Ichimuara****, and Kenji Oguni†

*Student, Earthquake Research Institute, University of Tokyo (1-1-1 Yayoi, Bunkyo, Tokyo, 113-0023)

**Student, Earthquake Research Institute, University of Tokyo (1-1-1 Yayoi, Bunkyo, Tokyo, 113-0023)

***Ph.D., Earthquake Research Institute, University of Tokyo (1-1-1 Yayoi, Bunkyo, Tokyo, 113-0023)

****Ph.D., Department of Civil Engineering, Tokyo Institute of Technology (2-12-1 O-okayama, Meguro, Tokyo, 152-8552)

†Ph.D., Earthquake Research Institute, University of Tokyo (1-1-1 Yayoi, Bunkyo, Tokyo, 113-0023)

This paper presents a methodology of data conversion from GIS to urban area models which are used for the numerical simulations of natural disasters. The key point of the methodology is the use of intermediate data, i.e., GIS is converted to the intermediate data from which the models are constructed. Two kinds of intermediate data are developed, and they are converted to models for non-linear seismic response analysis and for flood and volcano eruption simulations. The robustness of the data conversion is examined, and commercial GIS is successfully converted to urban area models for earthquake, flood and volcano eruption simulations.

Key Words : *natural disaster simulation, geographic information system, data conversion, modeling*

1. Introduction

It is of primary importance to improve the preparedness for natural disasters in urban areas, since these areas are changing rapidly; for instance, declining birth rate alters the demographic composition, information revolution results in greater interdependencies among infrastructures and industries. Most of conventional prediction methods for natural disasters are based on static analysis of past event data, and hence the prediction of natural disasters which can take place in such rapidly changing urban areas may be beyond their limit. Therefore, more rational methods are needed. A candidate of such methods is the use of numerical simulation¹⁾. There are assets of numerical analysis methods which have been developed for engineering design purpose or scientific researches, and the assets can be used for predicting physical processes of natural disasters. Geographical information system (GIS) is being developed, so that more detailed data are available for urban areas. If a suitable analysis model for the urban areas is constructed from GIS, it is not impossible to use the assets of the numerical analysis methods to carry out natural disaster

simulations.

Constructing an urban area model is regarded as data conversion; data which are stored in GIS are converted to data of an analysis model. Developing a methodology of such data conversion is an interesting research topic²⁾ to explore new use of GIS; new information will be generated from GIS with the help of numerical analysis methods. We should mention that while GIS data has been analyzed from the beginning of GIS, these methods of analysis are statistical analysis or data mining. There are limited cases where more sophisticated analysis methods are used, but a hand-made model is constructed from GIS. As far as the authors have surveyed, no systematic methods are developed to automatically convert GIS to analysis models for numerical simulation of physical processes such as natural disasters.

Highest robustness is essential for this data conversion, since GIS stores data of numerous items. For instance, the number of structures in one urban area sometimes exceeds 1,000,000. Even if 0.1% of the GIS data is mistreated, it turns out that more than 1,000 models are wrongly constructed. If a data conversion method with highest robustness is developed, a wider

class of numerical analysis methods will be used and the simulation based prediction for natural disasters will be shared by an increasing number of domestic or international urban areas.

In this paper, we study the data conversion of GIS to urban area models in order to carry out numerical simulation of natural disasters, such as earthquake, flood and volcano eruption. The earthquake simulation needs an analysis model for each structure, and the flood and volcano eruption simulations need a model for spaces in which water or lava flows. To develop a robust method for the data conversion, we introduce a methodology which uses intermediate data, i.e., data stored in GIS are converted to the intermediate data, from which analysis models for a particular simulation are generated.

The contents of this paper are as follows: First, the data conversion methodology using the intermediate data is explained in Section 2. Actual intermediate data for the earthquake simulation and the flood or volcano eruption simulation are developed in Section 3 and 4, respectively. Examples of the data conversion are presented in Section 5, and it is shown that numerical analysis methods are applicable to the models which are constructed from GIS. Some concluding remarks are made in Section 6.

2. Common Modeling Data

In order to explain the data conversion, we regard GIS and an analysis model as lower and upper data, respectively. These upper and lower data have many kinds; although protocols are established, there are wide varieties for contents of data which are stored in GIS; and different models need to be constructed for one structure, depending on the numerical analysis method. It is not realistic to make a data conversion for each pair of lower and upper data. Instead, we introduce an intermediate data, called common modeling data (CMD)³⁾, so that lower data are converted to CMD, from which upper data are converted.

CMD are designed so that the conversion from the lower data to the upper data via CMD is made with minimum efforts. The contents of the lower data are configuration of structures; GIS has data for structure geometry which is expressed as a set of polygons. The upper data are for a numerical analysis method of solving a particular initial-boundary value problem of

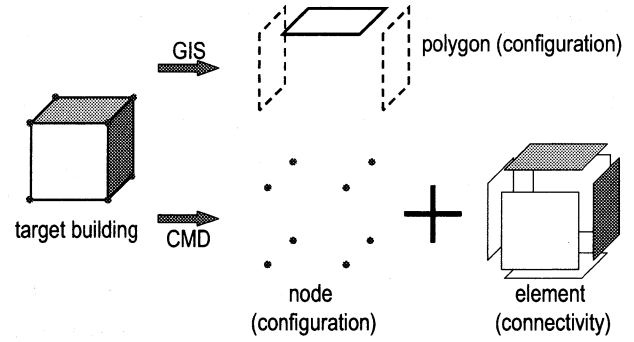


Fig. 1 Comparison of data contents of GIS and CMD for structure.

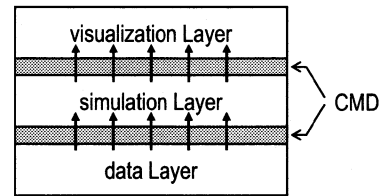


Fig. 2 Layer structure of GIS with simulation and visualization.

a natural disaster. Thus, we borrow input data of a finite element method (FEM) to design the data structure of CMD; they consist of two data, nodes and elements, so that a structure is described as a set of connected polygons or polyhedrons; see Fig. 1.

With the help of CMD, various numerical simulations of natural disasters are made by using GIS. We may regard a set of numerical analysis methods for the simulations as one layer of GIS^{4),5)}; GIS has a layer structure, i.e., data for particular items are stored in one layer. This layer is called a *simulation layer*; see Fig. 2. As new information is generated by combining plural data layers of GIS, new information related to natural disasters is generated by processing the data layers through the simulation layer^{6),7),8)}.

Another kind of CMD is developed for visualization of the results of the numerical simulations; the simulation results are converted to the new CMD from which data input to various visualization tools are converted. Like the simulation layer, a set of visualization tools are called a *visualization layer* of GIS; see Fig. 2.

In developing a GIS based simulation system, We used to use a *mediator* which plays a role of an interpreter for a particular simulation method¹⁾. A mediator, which must be made for each numerical analysis

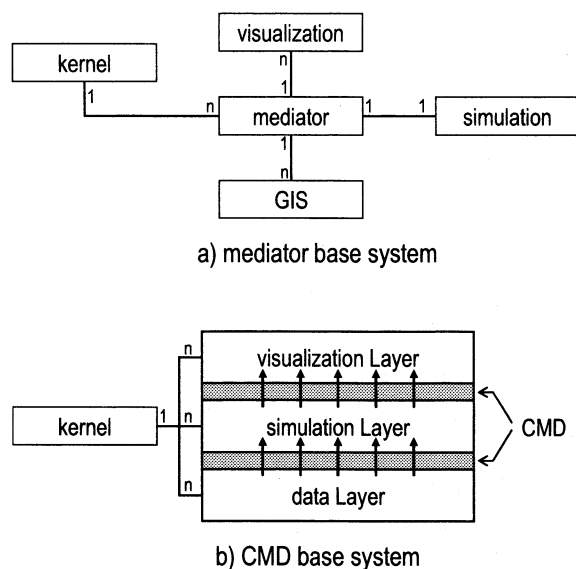


Fig. 3 Comparison of mediator and CMD base simulation systems.

method, directly converts lower data to upper data and produces data for visualization. In Fig. 3, we present a schematic view of the simulation systems based on the mediator and CMD; kernel in this figure is a controller of the system. As is seen, coding a mediator becomes more difficult as the number of GIS and numerical simulation increases. The coding task is decreased if CMD is used. Thus, the methodology of using CMD is a feasible choice for more general use of the numerical simulations in GIS.

3. CMD for Earthquake Simulation

Most of GIS data for structure are of DXF or SHAPE format, which is actually used in CAD. Since the primary objective of CAD is drawing, data of DXF or SHAPE format have a set of configuration data of *structure pieces*, i.e., a structure is divided into several pieces and each piece is expressed as a polygon. Structure pieces are not related to their mother structure; the structure is drawn if its pieces are correctly located. Thus, linking the structure pieces to their mother structure is needed in the data conversion.

Structure CMD are thus constructed, following the style of FEM input data; structure CMD consist of a set of nodes and elements which are determined by analyzing polygons for the structure pieces. The constructed structure CMD are called *shape*. It seems easy to convert polygon data to *shape*. However, it

is not true. This is because GIS data for structure are not complete; structure pieces stored in GIS are measured by air- or satellite-borne laser profiling, and structure pieces which are missed by profiling are not stored. It is thus important to analyze the polygons so that a whole structure configuration is put into *shape*, with *guessing* missed structure pieces.

3.1 Shape

Structure pieces measured by laser profiling are the one which can be seen from the top; the polygon for the piece is on a horizontal plane with common height. The configuration of a base and vertical walls for the structure must be guessed in making *shape*. Guessing the configuration of unmeasured structure pieces is not trivial when a structure has complicated geometry. We have developed the following procedures of converting GIS data to *shape*:

1. gather polygons which are horizontally overlapped with each other.
2. order the gathered polygons according to their height.
3. find a polygon that covers other polygons on it.
4. make a polygon for the structure base using the polygon found above.
5. make vertical rectangles for walls connecting points of the gathered polygons.

The whole configuration is completed by adding the polygons guessed to the original polygons stored in the GIS. For each structure, *shape* is constructed by setting polygon points and polygons (or polyhedrons) as nodes and elements, respectively.

Beside for configuration, data for material and structure form are needed in order to construct a structure model. However, there are no GIS which store such data. We thus need to *guess* these data, too; for building structures, structure form can be guessed from the height and floor area; more details can be guessed if the construction year of the structure is known.

We construct another kind of structure CMD, called *data*, between the simulation layer and the visualization layer. As *shape* corresponds to FEM input, *data* mimics FEM output, i.e., data for nodal or element responses are stored in *data*. Thus, when C++ is used, we code *data* as a sub-class of *shape*.

3.2 Construction of standard model

Shape is converted to an analysis model for a particular numerical simulation method. For a building structure, we convert *shape* to a multi-degree-of-freedom (MDOF) model. This data conversion is not new to the authors, who have developed a mediator for the MDOF simulation. We briefly summarize the procedures of the data conversion. First, the number of floors is calculated from the building height using a typical value of floor height. For each floor, the weight is calculated from the floor area; typical values of the floor thickness and density are used. Finally, spring constants between the mass are determined so that the natural frequency and damping constant of the first and second modes coincide with the empirical formulae.

We seek to make the data conversion for a non-linear seismic response analysis method, called RESP, which is a non-linear analysis of MDOF and has been used for the purpose of design; Fig. 5 illustrates an example of the non-linear spring used by RESP. The structure configuration is determined by taking the same procedures as shown above. The classification of the structure form and the determination of the non-linear spring parameters are made according to *standard model*, i.e., a model which is guessed by experienced building designers, and a structure form is chosen from the building height and floor area and suitable spring parameters for the chosen structure form are calculated. Table 1 presents the parameters used in a RESP model and the classification with which a standard model is constructed for a building. We should mention that the standard model converted from *shape* is not always the most suitable. Instead of determining a unique set of non-linear spring parameters, we may try several sets to clarify a range of possible seismic responses which are predicted from limited GIS data. Such trial, however, is not made in this paper.

4. CMD for Flood and Volcano Eruption Simulations

For flood and volcano eruption simulations, an analysis model is constructed for empty space, i.e., if a sufficiently large cubic is put to an urban area, the empty space is the remaining of the cubic from which all structures located in this area are excluded.

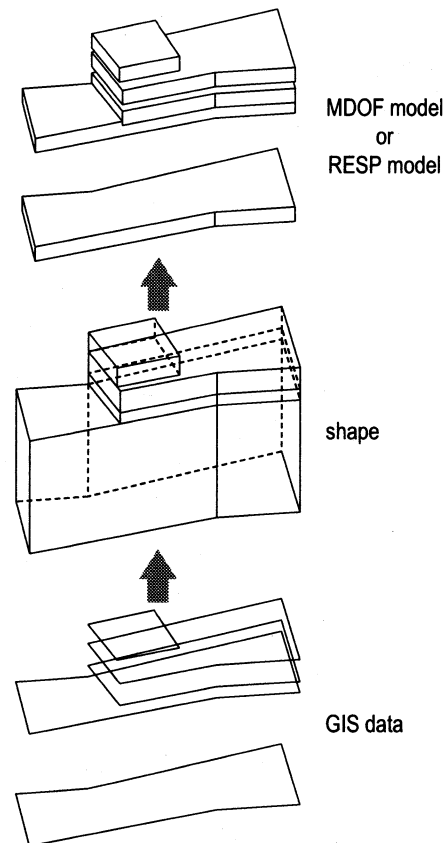


Fig. 4 Example of data conversion from original GIS data to MDOF model via *shape*.

Space CMD is constructed as intermediate data for the empty space.

We should mention that the flood and volcano eruption simulations are based on fluid mechanics. An analysis model used is often of a structured-element type, i.e., elements are an identical square or cubic. Thus, the space CMD is called *grid*.

4.1 Grid

Since *grid* is converted to a model with structured elements, grid data are used for *grid*. That is, a set of 0 or 1 value assigned for each grid; for a structure, one grid has a value of 1 if the grid is within the structure. By whipping the grid values, the grid data for a structure are converted to grid data for the empty space that surrounds the structure (*grid*); see Fig. 6 in which the configuration of a building is captured more clearly as finer grids are used. Note that *grid* has both nodes and elements to describe its square or cubic configuration, together with the binary grid value.

High robustness is expected in converting GIS data

a) parameters for RESP

parameter	remarks
floor number	basement floors are excluded
floor height	different height for some floors
floor area	some care for tower or open space
structure form	old/new S-rahmen, old/new S-brace, old/old modif./new RC-rahmen, old/new RC wall
ground type	1st/2nd/3rd type

b) classification of structure form

form	year	code	classification
steel rahmen	before 1980	codes before 1981	old S-rahmen
	after 1981	1981 code	new S-rahmen
steel brace	1980以前	codes before 1981	old S-brace
	after 1981	1981 code	new S-brace
RC/SRC rahmen	before 1970	codes before 1981	old RC-rahmen
	from 1971 to 1980	codes before 1981	old modif. RC-rahmen
	after 1981	1981 code	new RC-rahmen
RC/SRC wall	before 1970	codes before 1981	old RC wall
	after 1981	1981 code	new RC wall

Table 1 Summary of parameters of RESP model.

to **grid**, since the configuration of **grid** is simple (square or cubic) and it has binary data only. It will require smaller grids to make grid data if the structure geometry is complicated, and the data size and the computational effort for the data conversion will be increased accordingly.

In converting GIS data for a structure to grid data, we can take advantage of an algorithm of *painting*; the GIS data give outer boundary of the structure, and grids inside the boundary are painted or their grid values are changed to 1. This algorithm has been studied in the field of computational geometry and many codes are available. We make use of one of such codes.

In general, the numerical efficiency of the painting algorithm is improved by decomposing a target domain into subdomains of smaller size; see Fig. 7. Such domain decomposition is inevitable in converting GIS data for a large urban area. Once grid data are made for structures located in a subdomain, they are converted to **grid** for the subdomain. Thus, **grid** for the entire urban area is converted by assembling these **grid**'s.

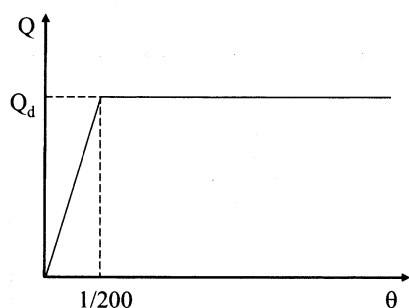
4.2 Robust data conversion from GIS to grid

While the domain decomposition contributes to improve the computational efficiency, there is a possibility that it leads to errors in the data conversion. If a structure of complicated geometry happens to be separated by the domain decomposition (i.e., one part is put in one subdomain and the remaining part is put in another subdomain), boundary of the separated structure crosses the boundary of the subdomain. There are many patterns for crossing of the separated structure boundary and the subdomain boundary; see Fig. 8 for examples.

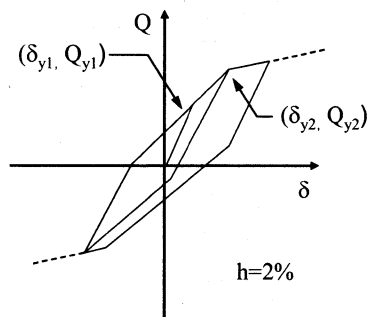
A sophisticated algorithm might be needed to treat each pattern of crossing. Instead of developing such an algorithm, we change the order of the domain decomposition and the painting. That is,

1. generate a small cubic for each structure so that the structure is fully included in the cubic.
2. convert GIS data for the structure to grid data, applying the painting algorithm.
3. generate **grid** for each subdomain by mapping a suitable part of the grid data to the subdomain.

The above procedures are shown in Fig. 9. There



a) drift angle-base shear relation



b) hysteresis curve

Fig. 5 Example of non-linear spring characteristics used in RESP (old S-rahmen).

are cases when the converted grid data of a structure are separated into some subdomains, but there is no room for an error to be made in converting the grid data to **grid** of the subdomain, since the grid data deliver only the grid value of 0 or 1. Therefore, high robustness is expected for this data conversion.

5. Example of Data Conversion of Commercial GIS

Data stored in a commercial GIS, TDM, are converted to structure or space CMD, i.e., **shape** or **grid**; this GIS is the product of Tokyo Digital Map Company. The data format of TDM is SHAPE. The data conversion from TDM to urban area models for natural disaster simulations via these CMD are examined in this section. We should mention that the objective is to confirm the completeness of the data conversion. No discussions are made for results of the numerical simulations which use the constructed models.

5.1 Earthquake simulation

RESP analysis models are constructed for around 1,400 buildings which are located in a certain part of

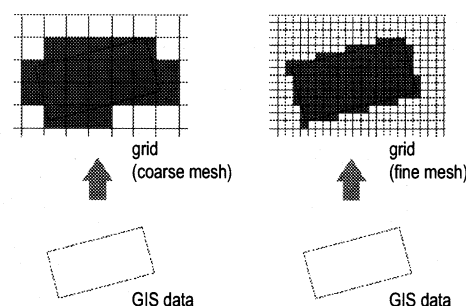


Fig. 6 Grid data for structure.

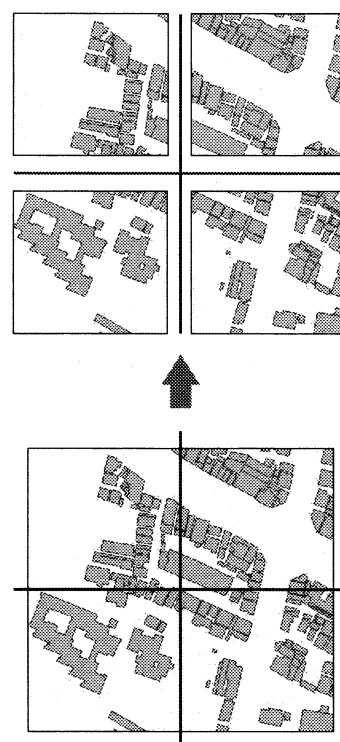


Fig. 7 Example of domain decomposition.

Tokyo. For each building, data of SHAPE format are converted to **shape**, and it is converted to the RESP analysis model. A schematic view of the model set is presented in Fig. 10. No error is found in the data conversion. Note that the RESP analysis models are located according to its target building, so that an urban area model for earthquake hazard simulation is constructed for this area.

A strong ground motion distribution is separately simulated, and RESP simulation is carried out for each RESP analysis model by inputting time series data of displacement at the model site. No error is found in the RESP simulation of all the models. The data conversion from TDM to the RESP analy-

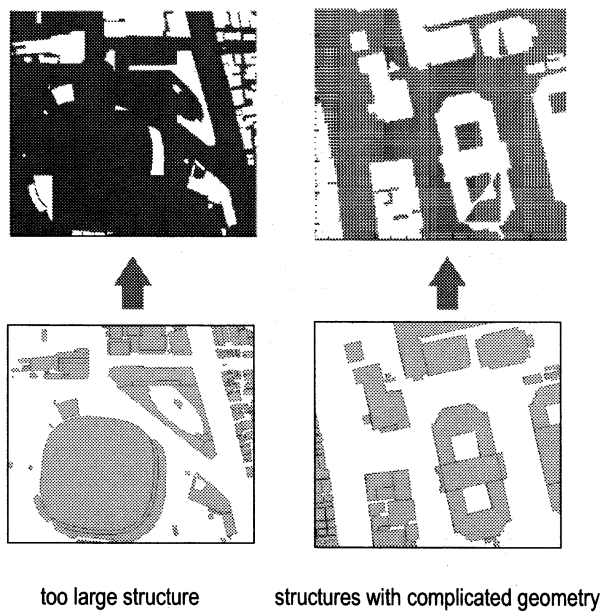


Fig. 8 Example of errors in converting GIS to grid.

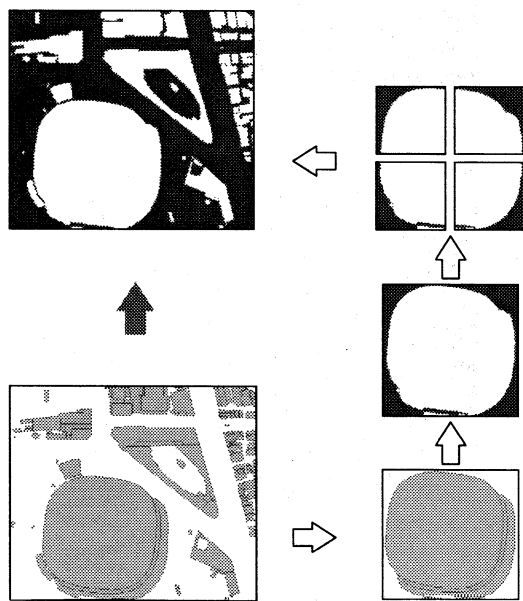
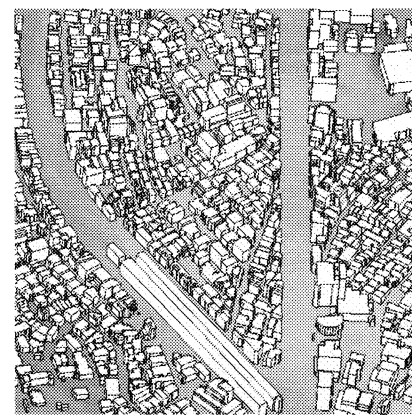
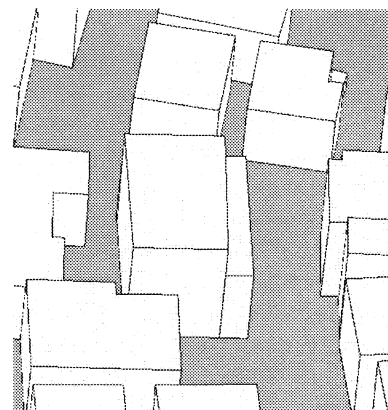


Fig. 9 Data conversion from GIS data to grid via grid data of larger structure.

sis models via **shape** is successfully done. We should mention that it is the building geometry that is transferred in the data conversion via **shape**; other input data such as structure form or non-linear spring parameters are determined by using the standard model which is explained in the preceding section. Thus, the quality of the RESP analysis model constructed in this manner should be discussed by comparing with the target building, although this comparison is out



a) wide view of RESP models



b) close view of RESP model

Fig. 10 RESP models constructed from TDM.

of scope of the present paper.

5.2 Flood and volcano eruption simulations

TDM consists of 273 data sets for distinct wards, and the data conversion from TDM to **grid** is made for each data set. There are 1,965,575 building data, and no errors are found for the data conversion of 273 grids. As an example, Fig. 11 presents the source TDM data set and the converted **grid** for a data set of ward near Tokyo Dome.

Numerical analysis methods for flood and volcano eruption simulations are provided by professional researchers. The flood simulation needs a two-dimensional model for the empty space, and this model is constructed from **grid**. The volcano eruption simulation is three-dimensional, and TDM is converted to three-dimensional **grid** of cubic elements, and a three-dimensional model is constructed; elevation data stored in DEM are added to this three-dimensional **grid**. Including DEM data to **grid** is important since TDM does not have data for rivers

and the sea, but they can be identified by using DEM (which distinguish them as domain of 0 elevation).

A snapshot at the initial step of the flood and volcano eruption simulations is presented in Fig. 12; a bulk volume of water and lava is put at the left edge of the urban area model. This problem setting is unrealistic since no river nor volcano is located there. However, we emphasize that that no error is made during the simulations and that the urban area models converted from TDM via `grid` are suitable for the third-party numerical simulation methods.

6. Concluding Remarks

Two kinds of intermediate data are developed so that urban area models for earthquake, flood or volcano eruption simulations are constructed from TDM. For another GIS (MapCube, made by PASCO Co. Ltd.), the data conversion to urban area models via CMD is successfully done. While further examination is definitely needed, we are expecting that the robust data conversion is made by the use of CMD.

Acknowledgements: The authors gratefully acknowledge Prof. Kashiyama, Chuo University, Dr. Ishimine and Fujita, NIDE, for their support of providing their programs of flood and volcano eruption simulation.

REFERENCES

- 1) M. Hori, *Introductio to computational earthquake engineering*, Imperial Collge Press, London, 2006.
- 2) K. Tanaka, Data conversion of GIS configuration data and natural disaster simulation, Master Thesis, Department of Civil Enginnering, University of Tokyo, 2008.
- 3) S. Nallathamby, M. Hori and Y. Ariga, Improvement of IES wrapper for plug-in of advanced simulation programs, *Journal of Applied Mechanics*, JSCE, **10**, 2007.
- 4) M.H. Serror, J. Inoue, M. Hori, and Y. Fujino, Distributed object-based software environment for urban system integrated simulation under urban-scale hazard – Part I: Infrastructure, *Earthquake Engng Struct. Dyn.*, **36**, 1545-1560, 2007.
- 5) M.H. Serror, J. Inoue, M. Hori, Y. Fujino and T. Ichimura, Distributed object-based software environment for urban system integrated simulation under urban-scale hazard – Part II: Application, *Earthquake Engng Struct. Dyn.*, **36**, 1561-1579, 2007.
- 6) N. Senin, D. R., Wallace, and N. Borland, Distributed object-based modeling in design simulation market-

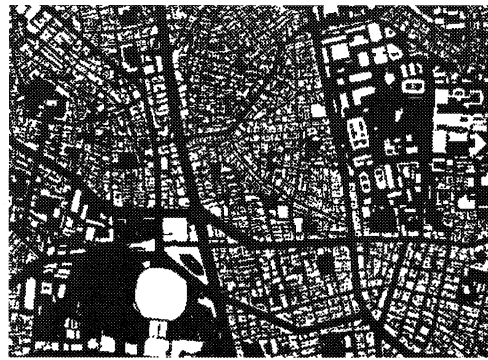
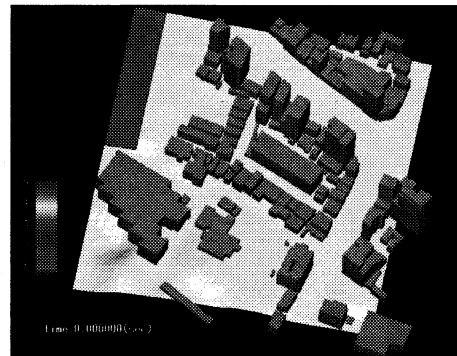
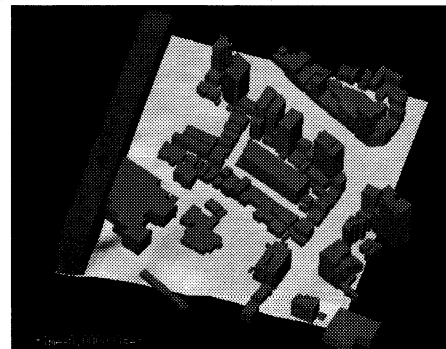


Fig. 11 Grid for area near Tokyo Dome.



a) flood simulation



b) volcano lava simulation

Fig. 12 Snap shot of flood and volcano eruption simulation.

- place, *Journal of Mechanical Design*, ASME, **125**, 2-13, 2003.
- 7) D.S. Ram and D.L. Robert, *Distributed and Integrated Collaborative Engineering Design*, Sarven Publishers, Glenwood, 2002.
- 8) W. Michael, *Introduction to Multiagent Systems*, Wiley, New York, 2002.

(Received: April 14, 2008)