The Archaeological Informatics and Spatiotemporal Digital Archive System (AISDAS)

– 1. GIS-based investigations in Kaman-Kalehöyük and surrounding areas using AISDAS –

Hiro'omi TSUMURA and Shinya SUZUKI
Kyoto *Chiba*

INTRODUCTION

Since the 1990s, the integration and standardization of spatiotemporal information for archaeological sites has been the most important and urgent theme in the field of digital heritage concepts. Many Western archaeologists have led these studies via the CAA (Computer Application and quantitative methods in Archaeology) and ECAI (Electronic Cultural Atras Initiative) and other projects, having also achieved excellent results in Turkey. For example, Peter Jablonka and his "Troia-Projeckt" team constructed a forefront system for constructing archaeological VR (Virtual Reality) that can construct a historical intrasite scene from archaeological excavation data (Jablonka *et al*. 2003). However, because data relating to excavation at an archaeological site and data from general surveys of surrounding areas are archived in a variety of different formats (commonly in analogue format), integration and standardization of different data sources are becoming more difficult as the amount of information increases. Such difficulties especially arise in projects that extend for periods of 20 years or more, as with the Kaman-Kalehöyük project. To analyse several historical and archaeological phenomena with intra- and intersite spatiotemporal informatics, the problem of data format must be managed in advance. Accordingly, since 2003 we have been exploring the introduction of GIS/GPS/CAD/Digital Photogrammetry that can operate from standardized digital spatiotemporal coordinates. The current report summarizes our preliminary approach to dealing with the enormous amount of spatiotemporal information generated from the Kaman-Kalehöyük project, as undertaken during the 2003 and 2004 seasons.

BASIC CONCEPT OF AISDAS

The most important aspect of our Archaeological Informatics Project (Archaeological Informatics System: AIS) is the standardization of various large data sets related to the Kaman-Kalehöyük project to enable the reconstruction of historical scenes at the site. For Kaman-Kalehöyük, it is considered that a combination of information technologies is effective. Figure 1 shows the differences between the existing information storage system and the next generation of data/knowledge/intelli gence storage structures that make use of GIS.

In the current information storage system, our information occurrence conduct is connected with inconsistent media formats. For example, if we wish to restore a historical scene from artefact data, first we individually access physical and/or chemical data about the archaeological artefacts in digital text data format when permitted by the information makers (conservators). However, in reconstructing and understanding the accumulation process of the deposit, we must check the distribution maps and layer matrix data on-site. Therefore, in the next step, we search for the thematic analogue image maps with the artefact number key. To determine the comprehensive three-dimensional layer constitution of archaeological components, we may also have to verify the analogue distribution/plan maps of structures and the photo-archives by the key of the grid number or from dates. The most difficult task involved in restoring the historical scene is constructing a three-dimensional image of the site structure in the mind and making a corresponding image from the scientific reproduction process. The upper structures of buildings and their

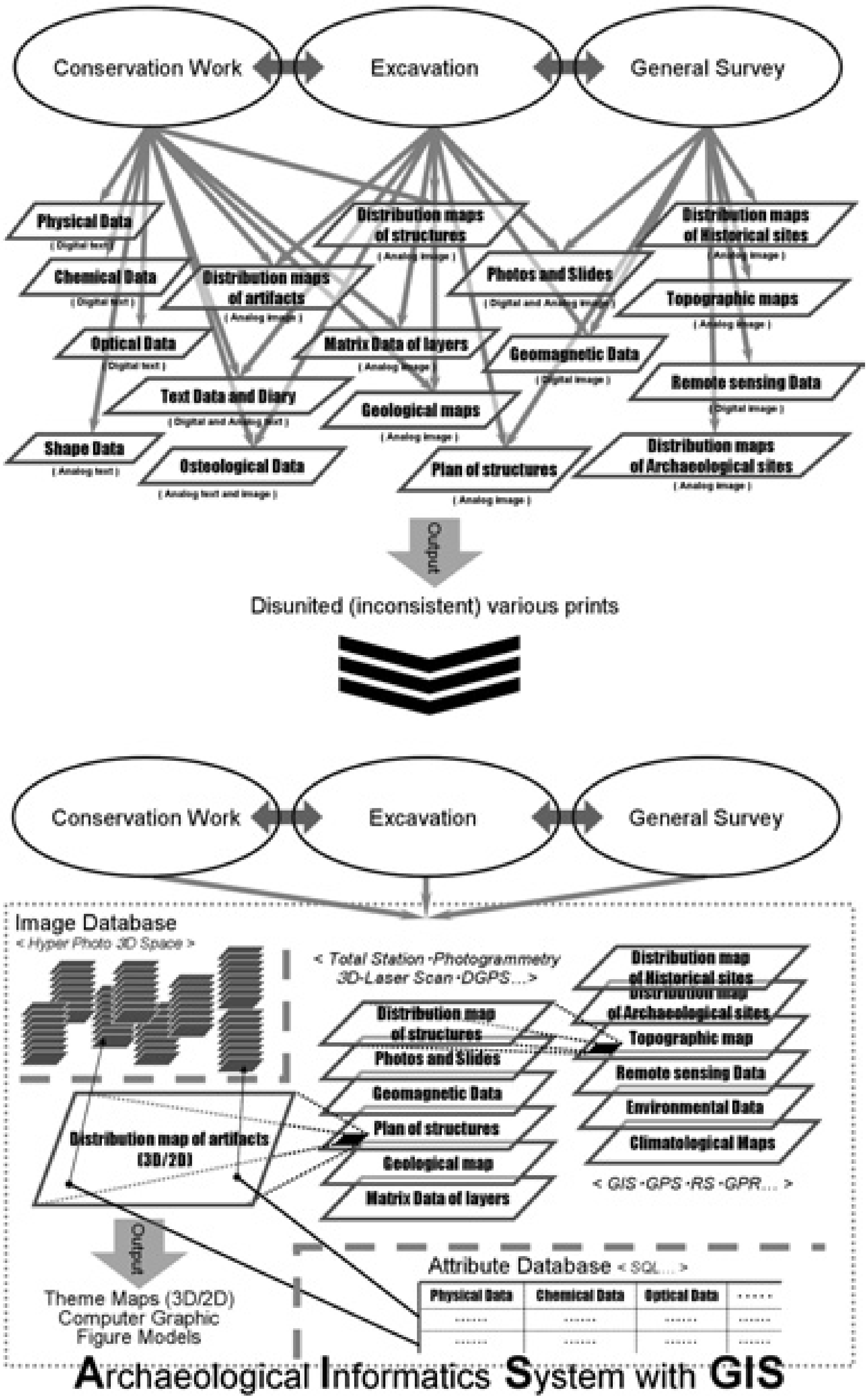


Fig.1 Comparison of different information storage structures

four-dimensional remodelling sequence involves almost no proof of the validity of information except for personal on-site observations and photographs taken from various angles.

We face various problems in the above process. For example, we may be unable to gain permission from conservators to access the data, or maps may be inconsistent with neighbouring maps. While a series of photographs reveals the two-dimensional nature of the area of interest, it does not reveal the points from where each photograph was taken (therefore, we are unable to calculate the nadir distance), and we are unable to determine the precise three-dimensional relationship of objects within the scene. The most troublesome problem with data in various formats is that the data are scattered throughout the archive without order, although they are now to be arranged in terms of their relationships to other data. To know the details of only a single historical scene, it is necessary to search through the entire data set. Tsumura and Teramura (2002) and Tsumura (2005) proposed the AIS concept as a preliminary solution to the above situation. Until now, AIS and GIS have been employed only in part. The bottom half of Fig. 1 describes the situation in the case of applying the AIS concept to the Kaman-Kalehöyük project.

In this AIS concept, the most important technological concern is that all the information can be integrated and standardized in terms of four-dimensional coordinates: the three-dimensional space coordinates and the linear one-dimensional time coordinate. For instance, an unearthed artefact has three-dimensional XYZ coordinates at the site, and we can measure these coordinates using a surveying instrument such as a total station. In this situation, if we use universal coordinates such as UTM (Universal Transverse Mercator projection) rather than site-peculiar coordinates, we can integrate the data with general survey data via GPS. If the excavated scene is captured by stereo photographs, we can construct various two-dimensional and three-dimensional thematic maps via photogrammetric digitizing. Physical/chemi cal/optical data that are analysed by conservators are related to virtual spatiotemporal zones by universal coordinates that are also calculated by photogrammetric digitizing. Likewise, as photogrammetric photographs have geometric coordinates, if we rectify these to geo-coordinates, Hyper Photo 3D Space (Tanaka *et al*. 2005) can be constructed as an image database from the geo-coordinate data. Thus, all the information from Kaman-Kalehöyük and the surrounding area can be integrated with spatiotemporal coordinates. Every attribute of the information is saved in a text database with SQL (Structured Query Language); this then has a unique relationship with the map layers via a set of ID numbers.

This AIS system is useful for archaeological informatics and as a data archiving system. When we want data concerning "Sector XVIII", we can designate this in a PC window and the information that is spatiotemporally related to "Sector XVIII" is displayed in three-dimensional layer structures. Data from adjoining sectors are automatically combined with the attribute database, and Hyper Photo 3D Space is loaded simultaneously. If we wish to view the data from a macro perspective, it is easy to change the scale of the map. The AIS system responds to a request by loading several topographic maps, remote sensing data, or aerial photos. If a micro view is desired, the AIS system displays several distribution maps of specific artefacts or figures and photos of the artefact. As this AIS system is based on GIS, it is possible to undertake structure distribution pattern analysis, site location analysis, and various other spatiotemporal analyses. We term this comprehensive digital archive system and the concept of archaeological informatics "AISDAS".

THE BEGINNINGS OF AISDAS

The completion of AISDAS will probably take such a long time that earlier data collected will not be standardized..In a practical approach to AISDAS, we began by examining the base analogue maps of Turkey because these provide the most important data for constructing the virtual space. We can generally use the government-authorized maps (1/200000 scale) in Turkey that are controlled by Harita Genel Komutanliği and were made by the USSR. As the scale of these maps is inadequate for enabling the construction of DEM (digital elevation model) maps, the topographic 2.5-dimensional

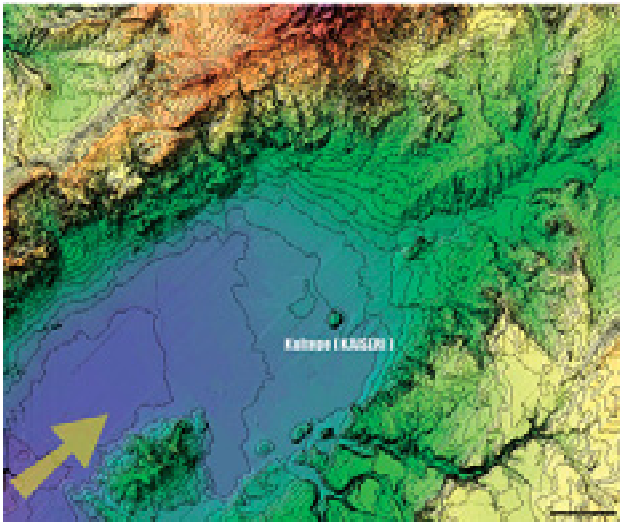


Fig.2 DEM (base) and counter (draped) maps of the Kaiseri area

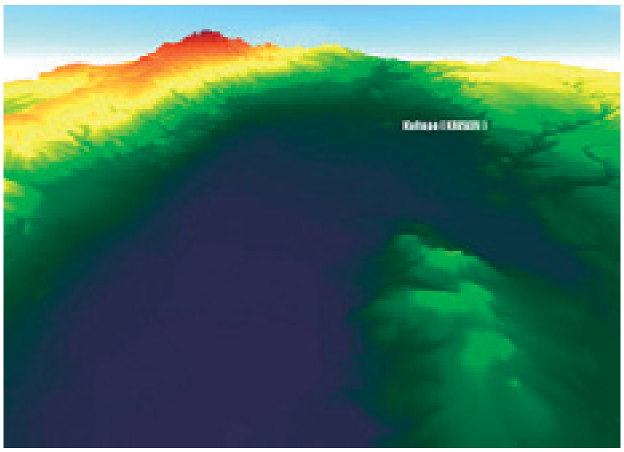


Fig.3 Bird's-eye view of the Kaiseri area

digital base data makes it possible to analyse the environment in GIS;

we had to search for additional data. Although we did find government-manufactured maps (1/1000000, 1/500000, 1/250000, 1/100000, 1/50000, and 1/25000 scale), their compilation is still in progress. Figure 2 shows the archived DEM and counter map of the Kaiseri area, and Fig. 3 shows a bird's-eye view of the area. These data were digitized during the 2003 season.

In terms of DEM, data such as GTOPO or SRTM issued by the USGS can be downloaded from the USGS web site. GTOPO30 is modelled by latticed elevation data at 1 km resolution The widths of lattices in SRTM-1 and SRTM-3 are 30 m and 90 m, respectively. As these data have already been rectified by the USGS, we are able to input them only with geocoding. We have completed this process for maps covering 60% of Turkey.

The above topographical digital data enable us to analyse the topographical environmental at Kaman-Kalehöyük and surrounding areas. Figure 4 shows the results of analysis of these data. Figures 4a –d show the distribution of elevation, slope grade, slant aspect, and hill-shading value, respectively. The slope grade and slant aspect were calculated using the following function for each pixel value in the DEM:

$$S = \sqrt{(\delta Z/\delta X)^2 + (\delta Z/\delta Y)^2} \quad \cdots \text{ Slope grade}$$

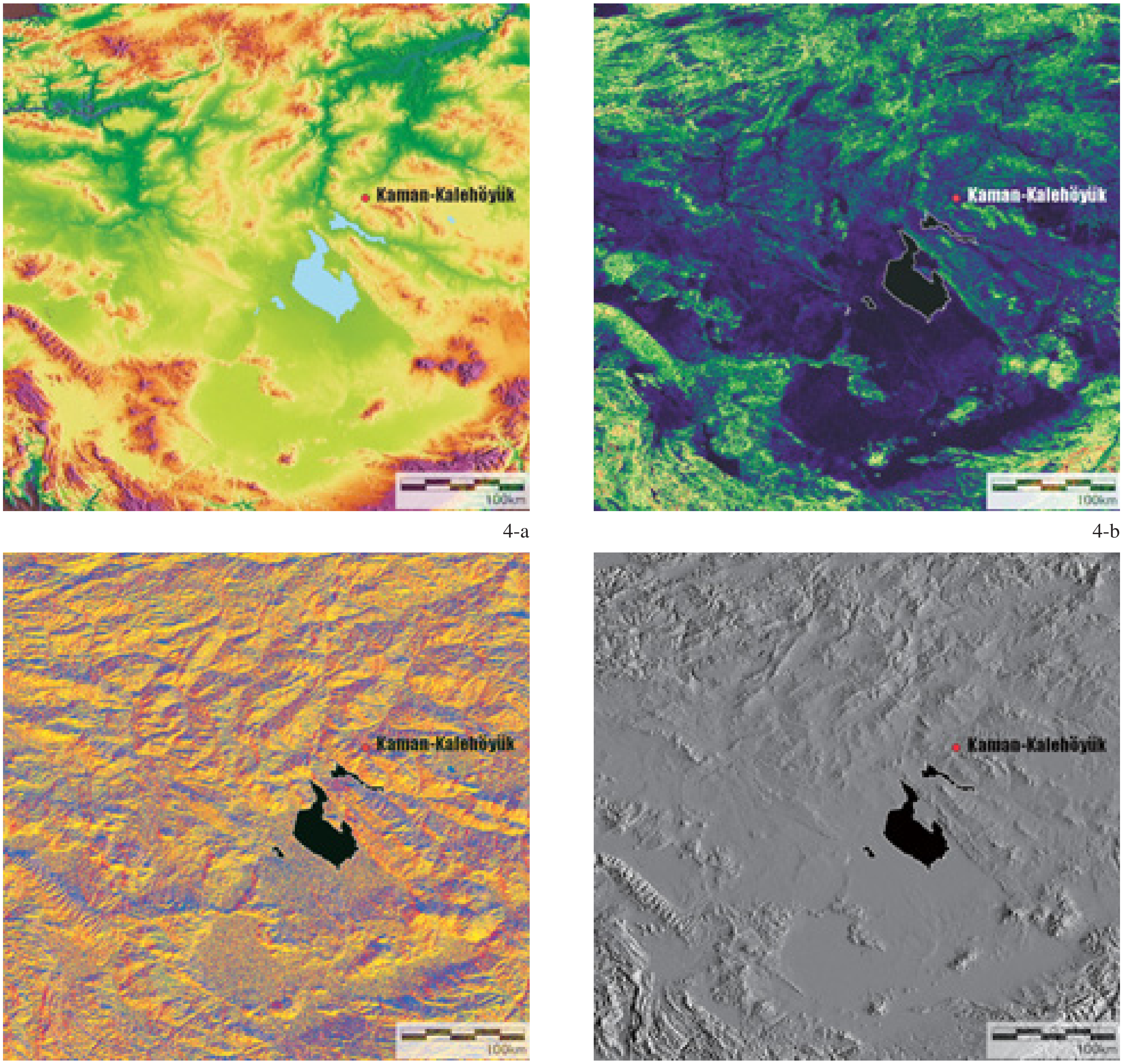
$$A = \frac{-\delta Z/\delta Y}{\delta Z/\delta X} \ (-\pi < A < \pi) \quad \cdots \text{ Slant aspect}$$

(Tan S = Slope degree, Tan A = Direction degree)

In this function, δZ, δX, and δY indicate the remainder of elevation pixel values, distance in an E–W direction, and distance in a N–S direction, respectively. The hill-shading values were calculated from a given slope grade and slant aspect value. Using this DEM, various other topographic features that are basic physical features of the natural environment can be analysed. This is very important information that should not be missed when seeking the location of an archaeological site. For example, using the above analysis, Tsumura (2005) found that the location of an archaeological site is influenced more by the slope grade than by the distance from a fountain or spring. This result shocked some traditional archaeologists and geographers who maintain a qualitative understanding of site location and are unfamiliar with quantitative approaches. In the present AISDAS project, topographic analyses that use DEM have been finished for 80% of the area of Turkey.

For other environmental information, remote sensing data are also being compiled. These data can be roughly classed into two groups: data from satellites and data from airplane surveys. At Kaman-Kalehöyük, aerial photographs are taken annually from balloon. Unfortunately, we have not yet been able to digitize and compile these aerial photographs, as they are enormous and their geocoding is not yet finished. They will soon be digitized and compiled into AISDAS.

A further type of data is satellite remote sensing



4-a 4-b 4-c 4-d

Fig.4 Thematic analyses figures about topographic features in the middle of Anatolia

data from multispectrum observations. Satellite remote sensing can use various bands of light such as infrared or ultraviolet rays, while aerial photography only considers the visible spectrum. Several satellites use synthetic aperture radar (SAR) for observing the face of the earth. If these data are suitably analysed, the environment of the face of the earth can be assessed with indexing. For example, although the monitoring of environmental change during a season or data concerning such change represents indispensable knowledge about a site location, it is difficult to compile this information on-site, as we are unable to assess environmental elements that cannot be observed under visible rays, such as the moisture content of soil or heat balance. In addition, it is not always possible to collect data over a prolonged period. The satellites are able to

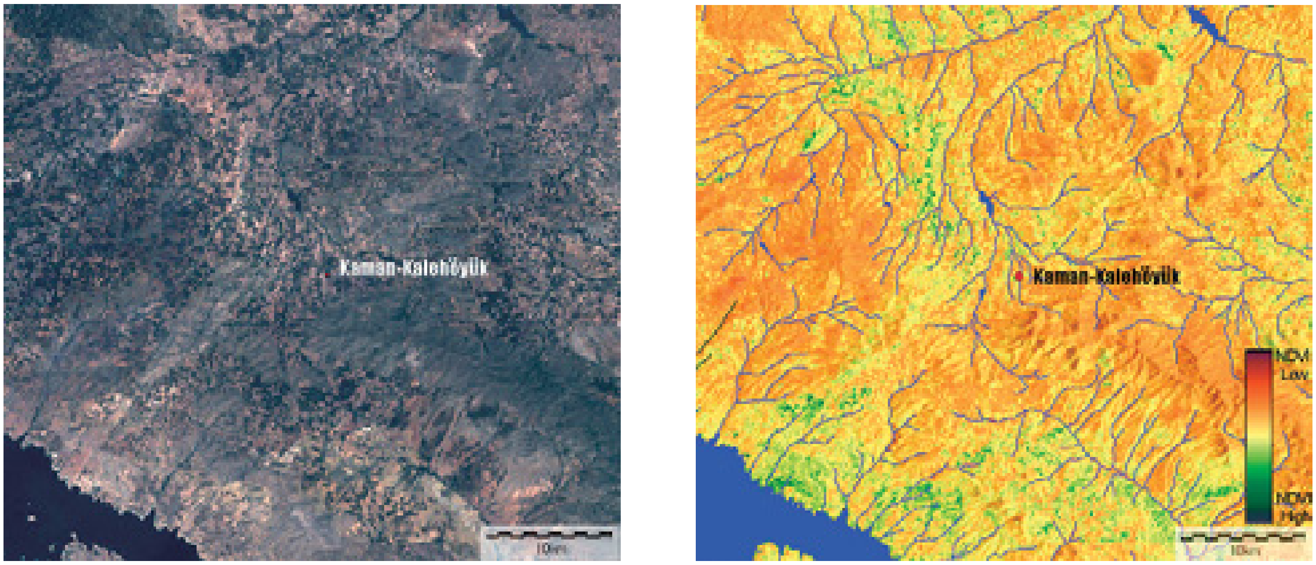


Fig.5 True-colour image (left) and NDVI map (right) of the area surrounding Kaman-Kalehöyük

monitor environmental change over a season without any physical effort on our behalf because they are orbiting the earth periodically.

There are various kinds of satellites and related data, e.g., Landsat, SPOT, IKONOS, and ASTER. These kinds of satellite remote sensing data, except for high-resolution data such as IKONOS, can be downloaded from the GLCF web site and used on a standalone PC. As these satellites each have a different revolution period and cloudy skies prevent them from observing the surface of the earth, the required data are not always available. For areas where there are no data and images from the USGS, NASA, and GLCF web servers, a new observation mission may need to be requested. The entire Turkey area was finally captured by satellite remote sensing data in 2003. These images have not yet been finished because of their very large data capacity. About 75% of the area has been finished to date.

An environmental assessment of the area surrounding Kaman-Kalehöyük is currently being processed at Doshisha University, Japan. Our target areas are land cover and land use, the moisture content of soil, heat balance, vegetation, the salt content of soil and water, changes in temperature, and the normalized difference vegetation index (NDVI (Normalized Difference Vegetation Index)), among other factors. These thematic data will provide useful information for considering the characteristics of site location. The left-hand figure in Fig. 5, for example, is a true-colour image of the area surrounding Kaman-Kalehöyük composed of three Landsat band data sets with completed rectification and geocoding. The true-colour image, as a composite of data from different bands, shows the colour of land cover under visible light. By observing these data microscopically, we can determine the land use patterns of the target area. As it is difficult to distinguish vegetation from soil solely based on colour in this area, a thematic map of NDVI values was created to measure vegetation activity in this area. The right-hand figure in Fig. 5 shows the NDVI values. Brown to yellow colours in the figure indicate areas of low vegetation activity, while yellow to green colours represent areas of high activity. Analysis of these data enables us not only to distinguish vegetation from soil, but also to indirectly determine the fertility of soils in the region. The most important aspect of this approach is to assess the data probabilistically or quantitatively rather than visually or qualitatively. When we consider the site location in terms of symbiotic relationships within the natural environment, we must consider that local people have required a large amount of firewood since the Iron Age; such an intelligent or analytical approach is therefore indispensable. In conducting our

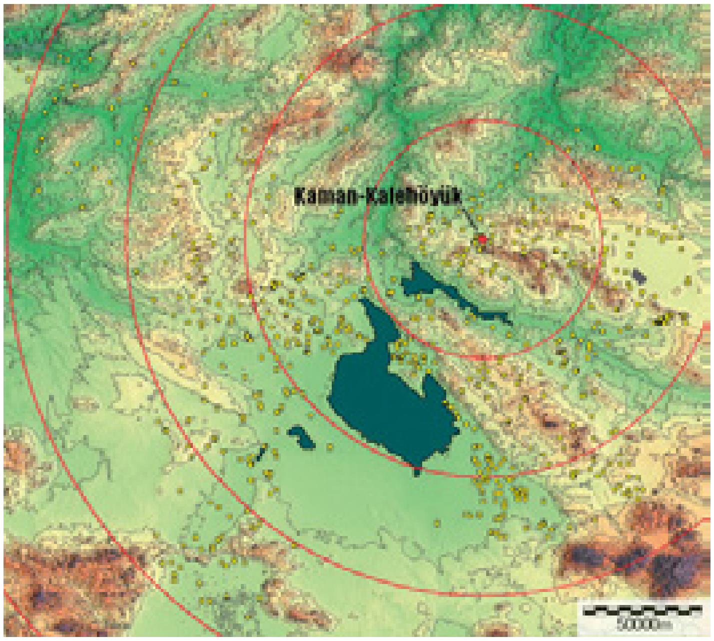


Fig.6 Site distribution map identified by general survey

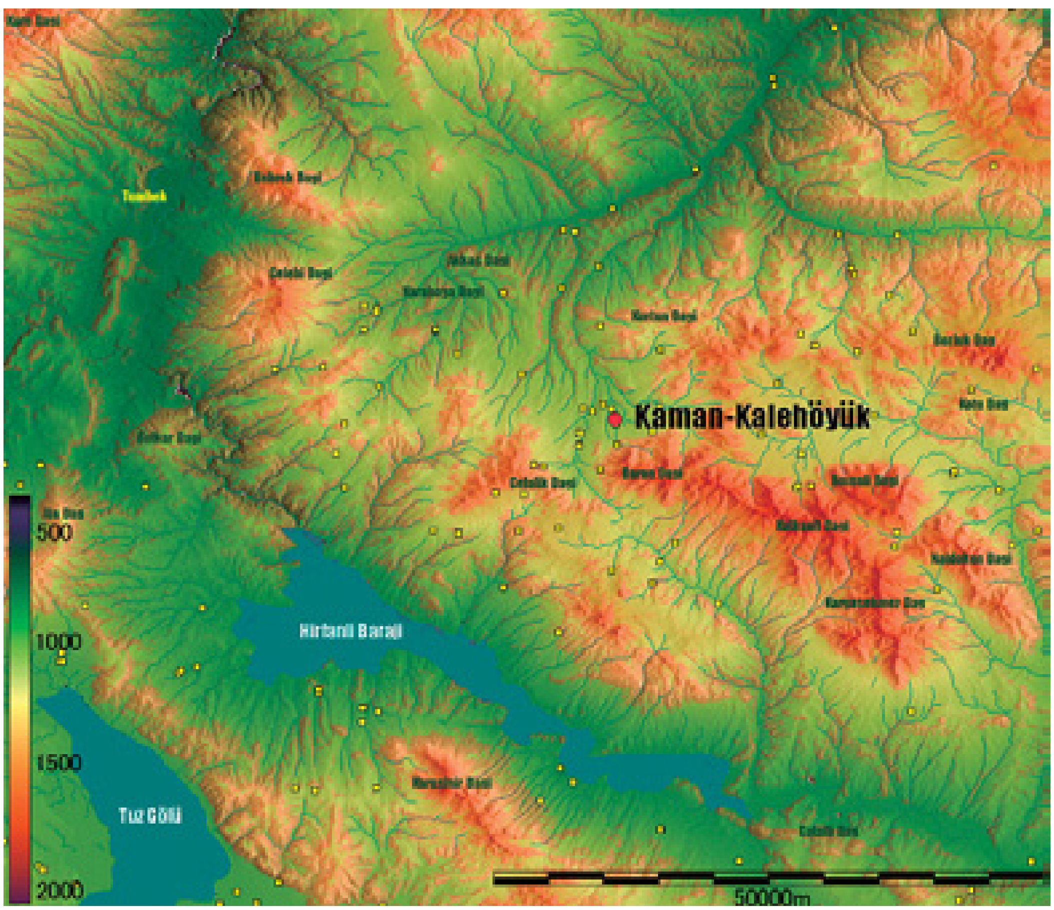


Fig.7 Site distribution map draped the classed slope grade map layer

general survey, the remote sensing data in AISDAS have provided much unpredictable information on the area. We have currently completed 15% of the total analysis.

Finally, we would like to report that we are currently assembling archaeological attribute information in addition to the above environmental information. In the Kaman-Kalehöyük project, a general survey and exploration have been carried out over the vast project area. These results and data are being archived as an analogue context sheet "card" for every sheet. These sheets document three kinds of spatial information about the archaeological site: a copied topographic map (1/25000 scale) with an indication of the site location, the latitude and longitude coordinates, and text data of the address. It is remarkable that GPS was introduced into this general survey as a world first, and that this method is still employed even though many archaeologists with limited consideration of archaeological informatics believe that GPS is not useful in archaeological field surveys. First, we undertook geocoding of site spatial information except for those sites for which latitude and longitude coordinates are assigned. Second, we undertook rectification of these coordinates to UTM. In this process, for sheets without coordinates (about 30% of the total number of sheets), we undertook the difficult task of repositioning a dot onto the geocoded and rectified map. From an information technology perspective, we believe that re-exploration will be necessary. Accordingly, Fig. 6 shows the distribution map of archaeological sites identified in the general survey that were outputted from AISDAS. To expand the spatial focus of the survey area toward the entire Turkey area, concentric circles are drawn at 5000 m distances from the Kaman-Kalehöyük area that was analysed in AISDAS.

Other information such as site type was entered into an attribute text database with an SQL structure. This database is linked with digital map layers by site ID numbers. Archaeologists who would like to view a "map of the Iron Age" will be able to access such a map by arranging the attribute database, as the SQL structure automatically links maps with the database. This linkage is very useful in archiving data, and is now fully complete. Photo images are pasted on the requested sheet as part of the provided information. These photos are hyperlinked to the database even though they are archived separately in other AISDAS data fields. About 7% of this work has been completed to date.

Our current research includes the tasks of digitizing analogue topographic maps, creating DEMs, compiling remote sensing data, analysing topographic features, and assessing environmental data. AISDAS will be useful for such purposes in any situation. For example, Fig. 7 shows an outputted thematic map from AISDAS that displays three-dimensional terrain characteristics with DEM. This thematic map is draped over the site distribution map layer, the drainage system layer, and the classed slope grade map layer, which is tinged with red. From this figure, it is easy to understand the environmental and topographic factors that explain the site location. In future general surveys, we will explore the area that is the closest to the site of drainage and the Wadi system where the slope grade value is less than five degrees. A further important point is not only to record the site location, but also to track process data via GPS because we must ensure that the next target area has not been previously explored. Needless to say, we can take a quantitative approach to site location using AISDAS. We would like to discuss this topic further in subsequent publications.

FUTURE SCHEDULE OF AISDAS

During the 2005 season, many stereo photographs were taken of archaeological structures and artefacts at the Kaman-Kalehöyük archaeological site for photogrammetric digitizing (Tsumura, 2005). These photographs are now being compiled into AISDAS. In the near future, we will compile these intrasite data alongside the archiving of the above intersite and environmental macro spatial data. It is especially important that plan maps of structures digitized using painting software and stored in Kaman-Kalehöyük Camp are geocoded as soon as possible, because of the very large volume of data.

ACKNOWLEDGEMENTS

This work was performed within the framework of the Kaman-Kalehöyük Archaeological Project. We would like to thank Dr Sachihiro Omura, Dr Masako Omura, Dr Kimiyoshi Matsumura, and members of our research team.

BIBLIOGRAPHY

Jablonka, P., S. Kirchner and J. Serangeli
2003 "Troia VR: a virtual reality model of Troy and the Troad," *The Digital Heritage of Archaeology*, pp.13–18.

Tanaka, H., M. Arikawa, R. Shibasaki and Y. Konagaya
2005 "A Tool for Creating Pseudo-3D Spaces with Hyperphoto: An Application in Ethnographic Studies," *GIS-Based Studies in the Humanities and Social Sciences*, pp.19-33.

Tsumura, H. and H.Teramura
2002 "State and perspective of 'Digital Heritage' in archaeological studies," *Theory and Application of GIS* Vol. 10, No. 2, pp.79–86.

Tsumura, H.
2005 "Reconstruction of palaeoenvironments and analysis of shell midden location/distribution with GIS predictive modelling," *Papers and Proceedings of Japan Society for Archaeological Information* Vol. 1, pp.121–128.

Hiro'omi Tsumura
Assistant Professor
Spatiotemporal Information Science and Behaviormetrics Laboratory
Faculty of Culture and Information Science
Doshisha University
tel & fax +81-774-65-7664
htsumura@mail.doshisha.ac.jp

Shinya Suzuki
Graduate School of Humanities
Chiba University
galatia_drynemetos_reburekki@msn.com