DESIGN OF HAZTURK DATA CATALOG AND REPOSITORY

Selim Serhan Yildiz[1], **Muhammed Şahin**[1], Himmet Karaman [1], Elif Demir[1], Serpil Ates[1]

Istanbul Technical University

Istanbul-TURKEY

Keywords: HAZTURK, Repository, Data Catalogue, Emergency Management, Earthquake Hazards.

Abstract

Study of "Design of HAZTURK Data Catalog and Repository" is carried out to store and serve the data which is necessary for HAZTURK (Hazards Turkey) earthquake loss assessment software. Data are classified by relevant object areas and object types. For this purpose, TABIS Object Catalog, which has completed before, is used as a reference frame. Five object areas and 29 object types are formed in this study. Names, descriptions, planimetric object types, data types and cardinalities of object types are determined and attributes are added to each object type. Abbreviations, codes, names, cardinalities and data types are determined. Object types are designed in shape, raster, CSV and XML file formats; attributes are designed in integer, string and double data formats. Firstly, HAZTURK Data Catalog that obtains all of the standardizations which are mentioned above was created. Then, HAZTURK Repository was created by designing shape files and raster files in conformity of standardizations. Thus, classification and standardization processes were completed for HAZTURK software, which was designed to handle the loss estimation system for Turkey with standardized data format managed by a master unit, as in other countries. WGS 84 geographic coordinate system, which coordinates of any object on Earth, can be used directly without any zonal information was selected as the reference coordinate system in the study for the design of the repository. This study will be an effective tool for the users, when HAZTURK begins to be used around the country in the near future.

INTRODUCTION

The Design of HAZTURK Data Catalog and Repository study was carried out in order to collect and classify datasets which are integrated and used in HAZTURK (Hazards Turkey) earthquake loss assessment software. In this context, Turkey Disaster Information System (TABIS) Data Catalog is used as reference which was completed earlier. Standards of data which will be collected and classified in order to use in earthquake loss estimation studies are determined. Turkey has a huge earthquake risk because of its geographic location. Kocaeli and Düzce earthquakes, which were occurred in 1999 with on the moment magnitude of 7.2 and 7.4 respectively, have shown that our country needs to be analysed at before, during and after earthquake disaster stages and these analyses should be performed under a single roof of an integrated system. Such systems have been in use abroad since 1990 and provide benefits for earthquake disaster. Hazards United States (HAZUS) and Mid-America Earthquake Visualization (MAEVIZ) systems, which were developed by various earthquake research institutes and supported by Federal Emergency Management Agency (FEMA) throughout the United States of America, can be given as examples. These systems were initiated to estimate earthquake damage based on various earthquake scenarios. Then different disaster types are also included to the system in scope. Use of those systems is fairly become widespread through their benefits that are supplied in preparedness, damage reduction, intervention and rehabilitation stages of disaster management. Standardization and design studies of datasets, that are predicted to be used in these systems, must be completed in order to handle the loss estimation system for Turkey with standardized data format managed by a master unit, as in other countries. For this aim standardization and design work of datasets that is predicted to be used in HAZTURK system was completed. These designed data standards were turned into a data catalog. Thus, the studies performed with HAZTURK will comply with data standards. Detailed explanations and examples of object types and their attributes are prepared according to this catalog. Repository that is used in the system was created with

the help of prepared guide. This sample repository is complies with designed data catalog and its reference TABIS data catalog.

TABIS OBJECT CATALOG

"Turkey Disaster Information System (TABIS) Reference Model predicts two basic model components in vector structure." These are;

- Digital Space Model (DSM)
- Digital Disaster Model (DDM)

Both digital models create space by individually allocating elements on an object oriented fundamental that means atomizing space in database. Atomized data related to both digital models are produced in the form of object catalog. Hereafter, this catalog will be referred as Turkey Disaster Information System Object Catalog (TABIS-OC).

TABIS-OC consists of two basic components that are prepared in parallel to approaches of DSM and DDM.

- TABIS-Basic Topographic Spatial Object Area Catalog (TABIS-TOC)
- TABIS-Disaster Management Object Area Catalog (TABIS-DOC) (Sahin et al, 2002).

HAZTURK DATA REQUIREMENT

A great deal of various data is required for use of HAZTURK software. These data must be classified so as to be used within the software in a healthy and proper manner. The data that is acquired in accordance with operation methodology and data format of HAZTURK software are classified and processed. Data classification concept describes the data set construction in relation with format, extension and type that can be used by program. As a result of data classification, data are classified as Disaster, Building, Geology, Topography, Boundary, Fault line, Attenuation and Mapping As a result of classification, data formats of data sets are outlined in Table 1. Data set matching is a classification procedure for building fragility curves and processes of this data set will be included in next phase of project.

Table 1. Data classification system (Karaman, 2008)

Data set	*Data Format*	*File Extension*	*Data Type*
Disaster	ASCII Raster	*.asc, *.txt	ASCII
Building	ArcGIS Shape file	*.shp	Point
Geology	ArcGIS Shape file	*.shp	Polygon
Topography	ASCII Raster	*.asc, *.txt	ASCII
Boundary	ArcGIS Shape file	*.shp	Polygon
Attenuation	Table	*.csv	Table
Other	ArcGIS Shape file	*.shp	Polyline, Point, Polygon
Matching	XML	*.xml	XML

Data classification provides opportunity for obtaining a developed program that will be able to recognize the data and analyze in shorter time. Classification enables the data used and its outputs to be in a certain format and data compatibility in every environment where this software used. The other purpose of data classification is to maintain harmonized operation of data and data set used in software and analysis with each other and also the software. The generation of correct results by process of data, in terms of reference system or data format, is directly related with data classification. If data

classification is not developed correctly, performed damage analysis will have the possibility of generating results with low accuracy as well as analysis system may not work. Considering this possibility, the performance of these classified data is controlled upon evaluation. This performance assessment is performed in accordance with whether acceptance of data by analysis or not, whether run of analysis by data or not and whether generation of logical results based on data at the end of analysis or not. Under the scope of project, data are classified so as to provide the most appropriate performance with respect to its format, data types, and coordinate systems and processes. Data processing is performed with the purpose of data compliance that will be included in data classes and data sets engendered dependent on data classification carried out to defined standards and to cover defined operation region. Under this scope, the reference system, type and format of data acquired is reached to defined standards. Sample attribute models that are developed for data set are given in the Table 1. It is a prerequisite that the compliance of data set that will be composed in order to be used in program this models for generation of results from analyzes to be conducted. One of the most time consuming step of data processing is the preparation of disaster data set. Disaster data set is constituted with the acquisition of peak ground acceleration values of Istanbul with the attenuation model prepared in ArcGIS software and by use of existing attenuation relations (Sahin et al, 2007).

NEED FOR HAZTURK DATA CATALOG

Earthquake damage estimation software HAZTURK uses several types of data with different features during operation. A data catalog that clearly contains the data standards and data associations is needed to classify, collect, create and store data. This catalog has an important role in designing data collecting forms when unavailable data is needed to use. Standards of object attributes are in a certain order. This condition will provide a great convenience in creating the repository. Thus, HAZTURK needs a data catalog in all applications such as data collecting, data entry and data storing in a secure and reliable way. On this point of study, standardization and design work of datasets that are predicted to be used in HAZTURK was completed. Thus, data standards used in studies carried out with HAZTURK were incorporated in a catalog.

FORMATION OF HAZTURK DATA CATALOGUE

For storage and using of data necessitated by HAZTURK software, firstly data fields and object type must be defined in a hierarchal structure. In this direction, mentioned data is divided under five object fields such as data related with buildings, geological data, topographical data, data related to vulnerability and data related to economy. The data type, definition, planimetric object type and cardinality of every object type are determined and attributions are added to object type in terms of its feature (Figure 1).

<table>
<tr><td colspan="3">HAZTURK Data Catalogue</td><td colspan="3">Version 1</td></tr>
<tr><td colspan="3">A. Data Related with Buildings</td><td colspan="3">Object Group:</td></tr>
<tr><td>Code:</td><td colspan="2">Name</td><td colspan="3"></td></tr>
<tr><td>A.1</td><td colspan="2">Data Related with Buildings</td><td colspan="3"></td></tr>
<tr><td colspan="3">Object Type:</td><td colspan="3"></td></tr>
<tr><td>Code</td><td colspan="5">Name:</td></tr>
<tr><td>A.1.</td><td colspan="5">Data Related with Buildings</td></tr>
<tr><td colspan="6">Description: Descriptive data related with buildings contained in study zone. Geographical name will be named with the name of official region which data in format of shape file belongs to.

Planimetric object type: Point

Name:</td></tr>
<tr><td colspan="3">Data Type:</td><td>Shape File</td><td colspan="2">Cardinality: 1:1</td></tr>
<tr><td colspan="2">NAME:</td><td colspan="4">Geographical Name (Official Name)</td></tr>
</table>

Figure 1. Sample Object Type Features (Yildiz, 2008)

Data related with structure are divided under six object types such as data related with buildings, data related with bridges and viaducts, data related with road networks, data related with embedded pipelines, data related with water depots and data related with power stations and transformers.

Geological data are divided under seven object types such as geology and engineering geology map, ground water depth map, fault line map (tectonic map), liquefaction susceptibility map, ground class map according to Turkey earthquake regulation, ground classification map according to NEHRP and map of Vs (0-30m.) shear wave velocity.

Topographical data is divided under three object type such as slope map, digital elevation model and administrative boundary map.

Data related with fragility is divided under six data type such as building fragility curves, fragility curve for transportation structures, fragility curve for embedded pipelines, building fragility curve mapping data, fragility curves for transportation structures and fragility curves for embedded pipelines mapping data.

Monetary data are divided under seven object types such as building reinforcement cost data, building repair cost data, bridge reinforcement cost data, and bridge repair cost data, building value data, building content value data and real estate tax data (Figure 2 and Figure 3).

HAZTURK Data Catalogue		Version 1
Object Field:	Classification of Object Fields	
A.	Data related with Building	
Description:	Covers the structures such as building, bridge, road, service networks located in the study zone that can be damaged after earthquake. The objects to be included in catalogue are selected from entities in the HAZTURK software that can be utilized in earthquake damage analysis.	
Code:	Name:	Page Number
A.1	Data Related with Building	3
A.2	Data Related with Bridge and Viaducts	6
A.3	Data Related with Road Networks	8
A.4	Data Related with Embedded Pipe Lines	10
A.5	Data Related with Water Depots	11
A.6	Data Related with Power Stations and Transformers	12
B	Geological Data	
Description:	Covers the data such as ground class, ground geology etc. geological and seismological in generation of Earthquake Risk Maps. These data are used in analysis like liquefaction susceptibility and earthquake scenarios.	
Code:	Name:	Page Number
B.1.	Geology and Engineering Geology map	13
B.2.	Ground Water Depth Map	16
B.3.	Fault Line map(Tectonic Map)	17
B.4.	Liquefaction Susceptibility Map	18
B.5.	Ground Class Map according to Turkey Earthquake Regulation	19
B.6.	Ground Classification Map according to NEHRP	23
B.7.	Vs (0-30m.) Shear Wave Velocity Map	24
C.	Topographical Data	
Description:	Covers the data types that define significantly the topographical character of region due to their presence over the physical earth surface “fundamental” and will be basis for multi purpose geographical information systems and other data groups that are deemed as required related with these object with respect to HAZTURK software.	

Code:	Name:	Page Number
C.1.	Slope Map	25
C.2.	Digital Elevation Model	26
C.3.	Administrative Boundary Map	27

Figure 2. HAZTURK Object Types-1 (Yildiz, 2008)

HAZTURK Data Catalogue		Version 1
Object Field:	Classification of Object Fields	
D.	Data related with vulnerability	
Description:	Includes the parameters related with vulnerability functions of all structures that may be damaged due to earthquake. In this object field, vulnerability curves that are useful for estimation of rate of damage level of structures due to earthquake and matching values that determines this vulnerability curves is for which type of building are located.	
Code:	Name:	Page Number:
D.1.	Building Vulnerability Curves	29
D.2.	Vulnerability Curve for Transportation Structures	34
D.3.	Vulnerability Curve for Embedded Pipe Lines	37
D.4.	Mapping Data of Building Vulnerability Curve	38
D.5.	Mapping Data of Vulnerability Curves for Transportation Structures	42
D.6.	Mapping Data of Vulnerability Curves for Embedded Pipe Lines	44
E.	Data Related with Economy	
Description:	In this object filed, unit values of buildings that will be used in determination of economic losses that will occur after earthquake, reinforcement of structures, repair costs are located.	
E.1.	Building Reinforcement Cost Data	45
E.2.	Building Repair Cost Data	47
E.3.	Bridge Reinforcement Cost Data	49
E.4.	Bridge Repair Cost Data	50
E.5.	Building Unit Value Data	51
E.6.	Building Content Value Data	52
E.7.	Real Estate Tax Data	53

Figure 3. HAZTURK Object Types-2 (Yildiz, 2008)

The abbreviation, name, cardinality, description and data type of attributes included in object type are defined. If attribute has values for each value, its code is determined. The name of every attribute is composed of text and its abbreviation is composed of three block letters. Because all data in the data catalogue must be used in HAZTURK software, its cardinality is determined as 1:1(each attribute can have only one value). Data type of object types and attributes are determined as integer, string, double, shape file, CSV file or XML file (Figure 4). Shape file is a vector based GIS data. All vector GIS data included in HAZTURK is in this format. Shape file format is composed of three data types: Point, line and polygon. CSV file is a table data format. All table data included in HAZTURK is in this table format. In the first line, there are column names (field names), second line included data type related with column. Mapping files are in XML format.(Karaman et al., 2007). Mapping file matches a certain data with a certain vulnerability curve or data related with economy as per user criteria. For example, definitions such as “for a wooden building which is higher than 2 floors, vulnerability curve of Joe is employed” are utilized. (Sahin et al., 2007).

Abbreviation	Code	Name	Cardinality	Data Type
YPT		Structure Type (Continuation)	1:1	String
	RM 1	Reinforced Masonry Bearing Walls with Wood or Metal Deck Diaphragms		
	RM2	Reinforced Masonry Bearing Walls with Precast Concrete Diaphragms		
	URM	Unreinforced Masonry Bearing Walls		
	MH	Mobile House		
YAY		Construction Year (Building Age)	1:1	String
		The year of building will be written		
KAT		Number of Floors	1:1	Integer
		Total number of floors will be written		
NK		Occupancy Class	1:1	String
		Occupancy Class will be written		
KRT		Essential Facilities	1:1	String
		Essential Facilities within building will be written		
BDE		Building Value	1:1	Double
		The Market Price of Building will be written		
BID		Building Content Value	1:1	Double
		All total value including but not limited to all machinery, computer, valuable item, will be written.		
BBK		Independent Unit Number	1:1	Integer
		The number of independent unit in compliance with Property Law		
TAL		Area	1:1	Double
		The base area of building will be written as m^2 unit.		

Figure 4. Sample Attribute Features

DESIGN OF HAZTURK REPOSITORY

For utilization of data necessitated by HAZTURK software, object types and attribute data must be in format and data type defined in catalogue. Accordingly, shape files are created in defined format for object types such as; data related with buildings, data related with bridges and viaducts, data related with road networks, data related with embedded pipelines, data related with water container and data related with power stations and transformers, geology and engineering geology map, ground water depth map, fault line map, liquefaction susceptibility map, ground classification map according to Turkey earthquake regulation, ground classification map according to NEHRP and map of Vs (0-30m.) shear wave velocity and boundary map (Figure 5).

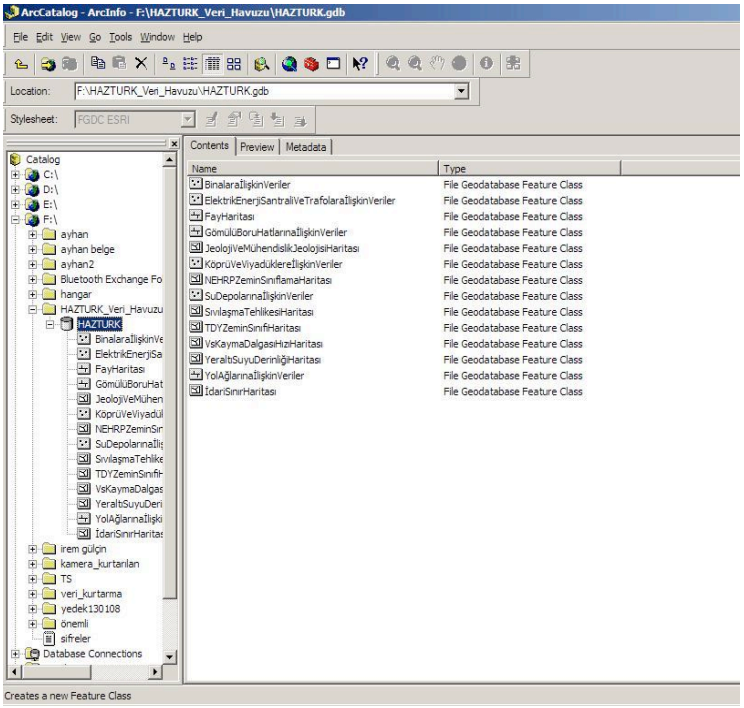

Figure 5. Shape Files of Object Types (Yildiz, 2008)

During formation of these files, attributes of every object type are saved as defined data type within file (Figure 6). While creating the shape files, WGS 84 geographic coordinate system is used (Figure 7). The reason behind using this coordinate system is that utilization of any object coordinates on the earth surface directly in the software without necessitating any zoning. In the created files, values of attributes are not contained, so files are ready for data entry. Processes are carried out through the ArcGIS software.

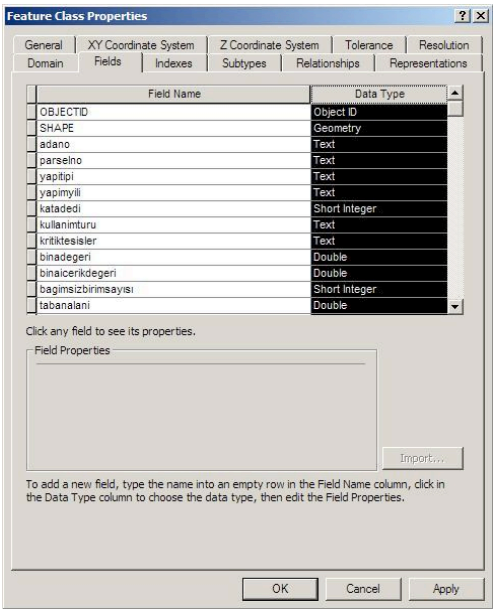

Figure 6. Sample Attributes in the Shape File (Yildiz, 2008)

HAZTURK software uses data in raster, CSV, XML as well as shape file format. Slope map, digital elevation model object types are in raster format. Fragility curves for buildings, fragility curves for transportation structures and fragility curves for embedded pipelines object types are in CSV format. Building fragility curve mapping data, fragility curves for transportation buildings mapping data,

fragility curves for embedded pipelines mapping data, building reinforcement cost data, building repair cost data, bridge reinforcement cost data, bridge repair cost data, building unit value data, building content value data and real estate tax value data objects types are in XML format.

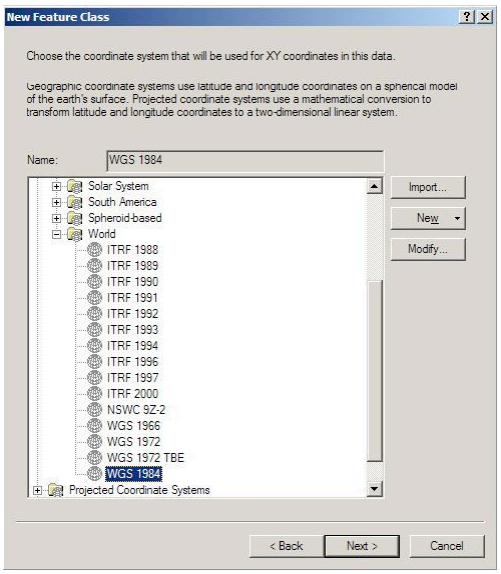

Figure 7. Sample Shape File Coordinate System Determination (Yildiz, 2008)

According to planimetric object type of object types, shape files are composed as point, linear and areal (Figure 8). Attributes contained in object type are added shape file according to data type determined in catalogue.

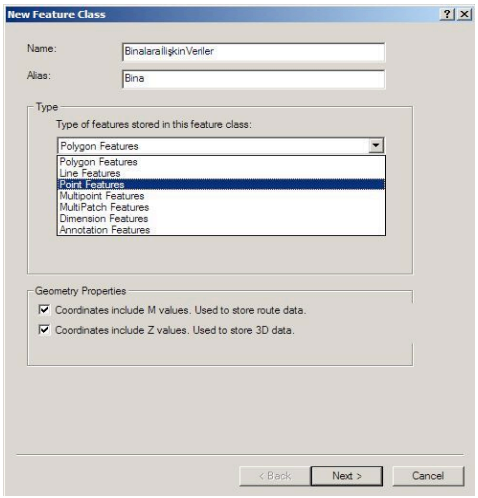

Figure 8. Sample Shape File Planimetric Object Type Determination (Yildiz, 2008)

For slope map and digital elevation model object types, raster formatted data sets are formed as stated in data catalogue (Figure 9).

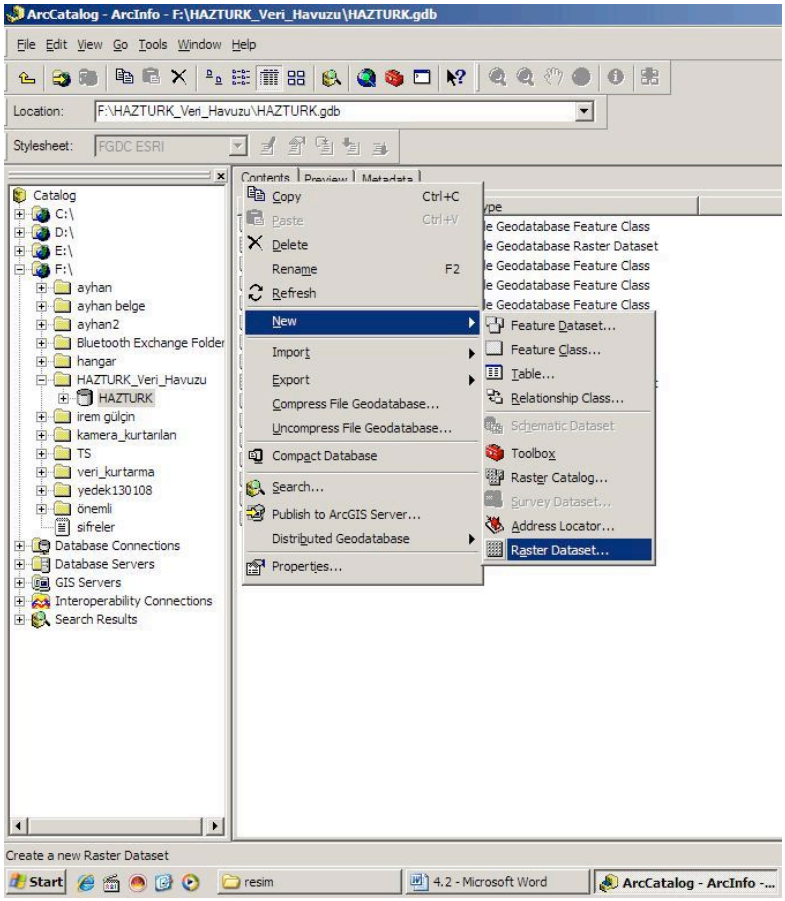

Figure 9. Sample Raster File Formation (Yildiz, 2008)

While engendering raster files, WGS 84 geographical coordinate system is utilized (Figure 10).

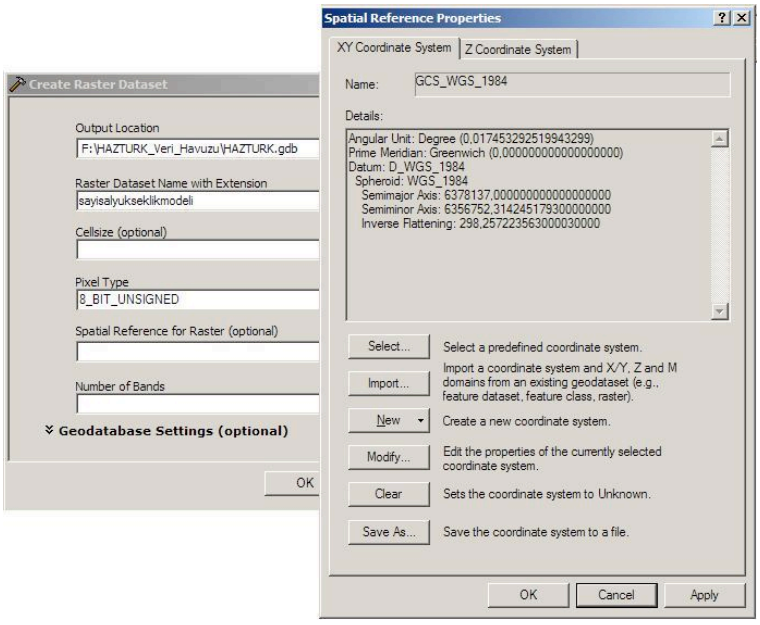

Figure 10. Sample Raster File Coordinate System Determination (Yildiz, 2008)

The band number of raster files are adjusted as three (RGB) and its pixel depth as eight bit. (Figure 11).

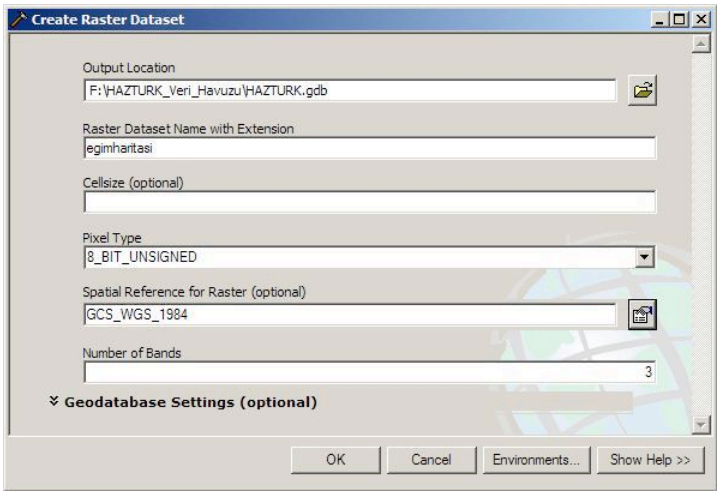

Figure 11. Sample Raster File Band Number and Pixel Depth Determination (Yildiz, 2008)

CONCLUSION AND SUGGESTIONS

With this study, the necessary standardization regarding collection, classification, storage and utilization of data, needed for HAZTURK earthquake damage estimation software, is obtained. With this regard, HAZTURK data catalogue is constituted and data pool design is constructed. Any kind of process to be conducted in relation with data can be performed easily in compliance with standards by use of this data catalogue. Due to designed data pool, the entry, storage and use of collected data can be maintained practically.

This study is carried out for the data management that is proposed to be used in HAZTURK in the first place. With the help of HAZTURK software, damage estimation of many objects in the earth can be sustained. To do this, the features and attributes of object must be determined firstly. Software can be developed according to these determined attributes and features. Earthquake damage analysis, which has only been carried out in one pilot zone in Istanbul so far, should be applied in all regions of Istanbul and subsequently in across Turkey due to its vital importance in terms of Turkey.

References

Karaman, H., Sahin, M., Uçar, D., Baykal, O., Türkoğlu, H., Tarı, E., İpbüker, C., Musaoğlu, N., Göksel, Ç., Coşkun, M.Z., Kaya, Ş., Yiğiter, R., Erden, T., Yavaşoğlu, H., Bilgi, S., Üstün, B., 2002. GIS Standards of Turkey based on Emergency Management, *International Symposium on Geographic Information Systems,* pp.82-85, Istanbul, Turkey, September 23-26, 2002.

Şahin, M., Boduroğlu, H., Eyidoğan, H., Tüysüz, O., Uçar, D., Karaca, M., Ural, D., Şeker, D., Tarı, E., Kalkan, Y., Kızılsu, G., Yanalak, M., İpbüker, C., Musaoğlu, N., Coşkun, Z., Kaya, Ş., Alkan, R., Tarı, U., Karaman, H., Yavaşoğlu, H., Erden, T., Bilgi, S. 2007. Developing A Loss Estimation Program, HAZTURK Based On HAZUS (Hazards US) To Be Used Before, During And After A Disaster: A Case Study For Istanbul, I.T.U.,Istanbul

Karaman, H., Şahin, M., Elnashai, A.S., 2007. Earthquake Loss Assessment Features of MAEviz-ISTANBUL (HAZTURK), *International Symposium on Earthquake Loss Estimation for Turkey*, Istanbul, Turkey, September 24-26, 2007.

Karaman, H., Şahin, M., Elnashai, A.S., Pineda O.A., 2007. Loss Assessment Study for Zeytinburnu District of Istanbul Using MAEviz-ISTANBUL (HAZTURK), *International Symposium on Earthquake Loss Estimation for Turkey*, Istanbul, Turkey, September 24-26, 2007.

Karaman, H., 2007. Effects of the Datum and Projection of the Data for the Earthquake Loss Assessment, *International Symposium on Earthquake Loss Estimation for Turkey*, Istanbul, Turkey, September 24-26, 2007.

Karaman, H., 2008. Consequence Based Risk Management and Earthquake Loss Estimation Analysis. Ph.D. Dissertation, Istanbul Technical University, Institute of Science and Technology, Maslak, Istanbul, Turkey, November 2008.

Yildiz, S.S., 2008. Design of HAZTURK Data Catalog and Repository, *MSc Thesis,* Istanbul Technical University, Istanbul, Turkey.

Author Biographies

Muhammed Şahin

He is a professor of Surveying Technique Division in the Faculty of Civil Engineering, Istanbul Technical University. He was the head of Surveying Technique Division between the dates of 2004-2009. He has been rector of Istanbul Technical University since August 2008. He was born in Pazar, a town of Rize where he finished his primary, secondary and high schools. He graduated from the Department of Geodesy & Photogrammetry, Istanbul Technical University in 1987. He received MSc & PhD from University College London and University of Newcastle Upon Tyne, UK, respectively. He becomes an assistant professor in 1994, an associate professor in 1996 and professor in 2002. His research interests include satellite positioning techniques, monitoring of earth crust using GPS, emergency management, disaster information systems, GIS based on emergency management.

Himmet Karaman

He is a geodesy and photogrammetry engineer. He graduated from ITU-Geodesy and Photogrammetry Engineering Department, and then received his MSc degree in 2003 and PhD degree in 2008. He has been a research assistant in Surveying Techniques Division since 2001.

Selim Serhan Yildiz

He was born in Osmaniye/Turkey in 1983. He graduated from ITU-Geodesy and Photogrammetry Engineering Department, and then received his MSc degree in 2006. He is a PhD candidate since 2008. He has been a research assistant in Surveying Techniques Division since January 2009.

Elif Demir

She was born in Istanbul/Turkey in 1984. She graduated from ITU-Geodesy and Photogrammetry Engineering Department in 2008. She is an MSc candidate since 2008. She has been a research assistant in Surveying Techniques Division since January 2009.

Serpil Ateş

She was born in Trabzon/Turkey in 1985. She graduated from ITU-Geodesy and Photogrammetry Engineering Department in 2008. She is an MSc candidate since 2008.