We are witnessing nowadays that the last decade of the past century, as well as the first years of the present one, have brought technology expansion with respect to spatial data gathering and processing which makes a physical basis for management of spatial development. This has resulted in enlargement of the spatial data market. New technologies, presented in computer applications, have greatly expanded the number of users of these products.

The philosophy of spatial data collecting has changed; analogue maps and plans printed on paper have been replaced by digital data bases, which enable their presentation in a way that is the best for a particular user. Further, digital spatial data bases provide the possibility of their further upgrading by users.

The two aspects, with respect to circumstances mentioned above, are very important in the process of data bases production and distribution. Firstly, the users of these data bases should be the ones who decide which of the available bases could satisfy their requirements, or in other words, what is the data quality level necessary for a certain application. On the other hand, the visualization of digital data bases could often mislead, since review of data bases could present data with better accuracy then the actual one. Thus, certain methods that would point to a quality of the selected data in the process of their analysis should be available to users.

Specific, already adopted international standards, or specially developed procedures and methodologies, so called de facto standards, could be used in this data processing, enabling the estimation of these data quality.

The development of Open GIS concept requires the adoption of widely accepted standards for spatial data quality. It is recommended that ISO standards should be accepted, firstly TC211 standards, which are related to geographic information and geomatics. The realization of projects on ISO standards should be finished by 2006, so, all participants of these data bases should be both familiar with this project and ready to adapt to the given solutions.

The basic components defining quality of data bases are explained by this work, and the results of the standardization regarding the procedures and methodology of their quality assessment, obtained so far, are also presented.

INTRODUCTION

The fundamentals for decision making in projecting and spatial planning are made of spatial data, first of all cartographic data, such as cadastre, topographic and thematic maps, then statistical data and all available associated documentations. In the last decade of the twentieth century, these documentations have been changed by geospatial data (in digital format) organized in GIS encirclement, suitable for data processing and visualization for all data users. Progress of information technology, as well as revolution in data collecting technologies (GPS, digital photogrammetry, laser scanning, InSAR etc.), play a leading role in spatial data base management. The final result was the change in technology of accessing, maintaining and processing of spatial data. Term spatial data refers to data about positions, attributes, and relationships of features in space (9).

Since then the traditional cartography products have no longer an irreplaceable role in spatial and sustainable development planning. Software which use spatial data in digital form (GIS, COGO and CAD utilities) are unavoidable tools for spatial planners, architects and other experts.

The geographic information systems (GIS) have largely been used as technology for complex management which, as a mean for coping with enormous volume of data associated with geographic information and the extensive calculations, need to rectify and analyze these data in decision-making context (2).
Different sources of spatial data lead to different levels of data quality. Great amount of data are obtained by encoding analog maps of different scales or satellite images of different resolutions. Those parameters have great influence on some data quality elements.

Spatial databases of completely unknown qualities of accuracy are greatly enhanced by inter-mixture of digital maps and associated attribute data with different errors and uncertain characteristics with specific error reducing or error amplifying properties of particular sequence of GIS operations (12).

SPATIAL DATA AND SUSTAINABLE DEVELOPMENT MANAGEMENT

The use of spatial information produces a direct or indirect possibility of increased efficiency in all sectors of public administration, in political decision-making process as well as in the private sector (13). There are innumerable applications where spatial information are used, like: cadastral projects, land management, transportation infrastructure, health monitoring programmes, monitoring of environmental issues, environmental impact assessment, statistical analyses, conservation projects, natural resource management, etc. All those activities could be classified as a part of the broader work called sustainable development planning.

Spatial Data is data with a direct or an indirect geographic reference to the surface of the earth. Combining data from one or more different sources creates information. More than 80% of all information in society has a spatial reference. Spatial Data and information are strategically important for decision-makers at all levels of Sustainable Development Management (13).

The sustainable exploitation of environmental resources requires:

- Data to be available, to be up-to-date, reliable and usable indicating the quality, quantity and spatial location of the various resources and the size and spatial distribution of the population which depend on these resources;
- The availability of tools to support the transformation of data into understandable information for decision-makers, from the national and international levels to the grassroots levels;
- Rethinking of both inter-organizational and intra-organizational relations in order to improve the use of common data and the reuse of data.

In view of the fact that spatial data plays very important role in decision making, it is necessary to bear in mind the quality of the above data as well as its applications.

Since a dataset is produced for different applications rather than for a specific application, the quality of dataset can be assessed only by knowing the data quality elements, as well as by the data quality overview element (6).

QUALITY ASSESSMENT OF SPATIAL DATABASES

New technologies of the computer systems applications have changed traditional role of cartographic products. They have enabled inclusion of more participants in making spatial data products, and what is more important, they have significantly expanded the number of users of such products. The introduction of GIS in the mapping process has also produced a completely new kind of users different from the traditional map users.

The philosophy of map production has changed, maps sheets printed on paper have been finally replaced by databases, with data structures which enable visualization of different manners, that depend on the purpose and necessity. The dominant opinion among such data users is that digital data are of a higher quality than conventional map data.

The fact is that in distribution, digital data bases offer opportunity to be updated by users. It is very common case that spatial data bases are compiled by different firms which are engaged in collection and processing of a data.

Two aspects are very important in the cycle of production and distribution of spatial data bases. Firstly, users have to decide which of responsible spatial data bases may satisfy their needs, in other words, which level of data quality is sufficient for particular application; on the other side computer visualization of digital data often mislead users because the views obtained may present more accurate data than they really are. That is the main reason why users must have certain methods which would point out the quality in the stage of data analysis and visualization.

The interest for data quality assessment is emphasized for the following reasons (17):

- Increased data production by private sector. Historically, mass production of geospatial data was the domain of governmental agencies. Unlike these agencies, private companies are not required to conform to already known quality standards.
- Increased use of GIS as a decision-support tool. This trend has led to realization of the potential deleterious effects of using poor quality data, including the possibility of litigation if the minimal quality standards are not attained.
- Increased reliance on secondary data sources. This has been fuelled by a reduction in accessibility and constraints resulting from network accessibility and the development of standards for data exchange.

Data quality is itself a difficult term to categorize. The term data quality is broader than the accuracy of the data. Accuracy plays a large part in evaluating quality, but there are related issues which must also be considered. There are many varying classification schemes developed by research organizations in order to describe data quality. The objective of the different categorizations is to separate data suitability into distinct components.

Nowadays, the term ‘data quality’ has been replaced with the concept of uncertainty. According to dictionary, word ‘uncertain’ means not known, unreliable, changeable, or erratic. A variety of terms have been used, almost interchangeably, to communicate spatial uncertainty, including: error, accuracy, precision, vagueness, ambiguity, and reliability. Uncertainty can be defined in terms of either an affirmative or negative character. Uncertain data possess attributes of either accuracy (an affirmative attribute, means closeness of agreement between a test result and the accepted reference value) or error (a negative attribute, measured in terms of discrepancy). Accuracy is more readily quantified, in comparison with a model (geodetic, statistical, cartometric,
thematic, etc.) constructed for a specific purpose. Uncertainty is not simply a flow to be avoided or ignored; it is rather an inherent attribute of data manipulation processes (2).

**COMPONENTS OF THE DATA BASES QUALITY**

Work on evaluation of data bases quality started in the early eighties of the last century. At that time the only well-known standards which could be used as starting point were already existing cartographic standards. International Cartographic Association (ICA) was the first organization involved in research with the aim of standardizing various aspects of digital data bases.

At the ICA meeting held in September 1991 in Bournemouth U.K. the Commission on Spatial Data Quality was established, and its main goals were to:
- develop and document a comprehensive set of data quality criteria,
- develop and document a standardized rating scheme against those criteria,
- develop a methodology for data quality testing,
- publish an ICA manual for assessing digital spatial data quality.

In 1982 USA established National Committee on Digital Cartographic Data Standards (NCDCDS) under the auspices of the American Congress of Surveying Mapping (ACSM). “A Draft Proposed Standard for Digital Cartographic Data” deliberated by this Committee was the first comprehensive report related to digital cartographic data quality. This report clearly points out that “The purpose of the Quality Report is to provide detailed information for a user to evaluate the fitness of data for a particular use. This style of standard can be characterized as ‘truth in labeling’, rather than fixing arbitrary numerical thresholds of quality.” (8).

A modified version of this document has been accepted by the National Institute of Standards and Technology (NIST) as the Federal Information Processing Standard-FIPS 173.

For the first time this report clearly designates five quality components of spatial data bases:  
- lineage,
- positional accuracy,
- attribute accuracy,
- completeness,
- logical consistency.

ICA commission on Spatial Data Quality has accepted these five components as initial elements of spatial data quality. Two additional components are added on latter:  
- semantic accuracy,
- temporal information.

**Lineage or genealogy** of data includes description of data sources, methods which are used for data bases creation, including all data transformations and transactions used in this process. This component must contain all data important for both, data sources and for the data upgrading process.

Lineage is usually the first component given in data quality reports, because all other quality components are subordinate on data genealogy and vice versa. The final purpose of lineage report is to keep precious information of data history for future users.

**Positional Accuracy** – Spatial and geometric accuracy of data. For point features, spatial accuracy is represented by discrepancy between encoded location and the location defined in the specification. It might be expressed in measures accuracy along coordinate axis or as a sum of them (Figure 1).

Metrics of spatial accuracy depends on dimensional entity which is considered. In the case of the point features it is represented with usual statistical means, like root mean square error (RMSE), standard deviation or confidence interval etc. For line and area features situation is more complex, because positional accuracy is a result of positional accuracy of points which define lines or segments used as generalized shape of reality. The accuracy of the linear features like roads, contours etc. are very often presented by ε-bands (Figure 2.)

In the case of the positional accuracy three distinct levels are presented:
- absolute or external accuracy – closeness of reported coordinate values to values accepted as/or being true,
- relative or internal accuracy – closeness of the relative positions of features in a dataset to their respective relative positions accepted as/or being true,
- gridded data position accuracy – closeness of gridded data position values to values accepted as/or being true.

Two methods for assessing positional accuracy are available:
- direct comparison with referent values with higher hierarchy level,
- assessing of accuracy with indirect methods.

**Attribute accuracy** is defined as the closeness of attribute values to their true value. In some documents it is mentioned as thematic accuracy. In contrast to positional accuracy, which considers spatial components of the features, attribute accuracy, its measures and methods of assessment, depend on domain of attributes. It can be defined as an accuracy of
quantitative attributes and the correctness of non-quantitative attributes and of the classifications of features and their relationships.

The term attribute accuracy is often replaced with attribute uncertainty. Attribute accuracy can be nominal, ordinal or interval. In case of the nominal or ordinal domain of attributes, frequently used accuracy indicators are percent correctness of their attributes. The completeness of a data illustrates the degree to which the information describes adequately (in terms of temporal precision, frequency and process history) spatial phenomena.

Report of completeness describes the exhaustiveness of set of features and their attributes in spatial data bases. It reports on how much detailed features and entities are presented in data base meaning to their spatial attribute characteristics. Completeness is divided on:

- completeness of a data,
- completeness of a model.

The report on completeness describes the relationship between objects represented and the abstract universe of all such objects. The completeness of a data illustrates the commission of an entities in spatial data base related to their number in the real world. It encompasses the completeness of entities and completeness of their attributes. The completeness of a model indicates the level of credibility of how accurate is a data base. Model completeness is in a correlation with semantic accuracy. It is application-dependent and therefore it is an aspect of fitness-for-a use.

The term ‘fitness for a use’ is referred to decision making for accessing whether a database meets the needs of a particular application. Logical consistency represents degree of achieved reliability of logical rules and connections in data structures. It deals with structural integrity of a given data set based on a formal framework for the modeling of spatial data and the relationships among objects.

Subelements of a logical consistency are:

- conceptual consistency – adherence to rules of the conceptual scheme,
- domain consistency – adherence of values to the value domains,
- format consistency – degree to which data is stored in accordance with the physical structure of the dataset,
- topological consistency – is usually assumed to refer to the lack of topological errors (e.g., unclosed polygons, dangling nodes, etc).

Semantic accuracy describes the number of features, relationships and attributes that have been correctly encoded in accordance with a set of feature representation rules, or in other words it means the quality that geographics are described in accordance with selected model.

Temporal information gives date of data observation, type of update (creation, modification, deletion, unchanges), and validity periods for spatial data records. Aspects of temporal elements are already presented in the first primary elements of data quality. The quality of temporal information can be appraised by the degree to which the information describes adequately (in terms of temporal precision, frequency and process history) spatial phenomena.

Traditionally temporal characteristics of spatial objects are handled as special or thematic attributes of the appropriate object types (classes). For example, cadastral data of a certain property includes date of property transaction that is crucial data.

ISO STANDARDS FOR GEOGRAPHIC INFORMATION

Many international associations interested in using such data bases were working on their own standards for assessing data quality. A study conducted by technical committee TC211 for geographical information and geomatics of International Organization for Standardization (ISO) became very important at the beginning of the nineties of the last century. The main goal of TC211 was to harmonize all responsible standards related to spatial data bases. Study of TC211 was very closely connected with the work of other international associations engaged in GIS and geomatics, like Fédération Internationale des Geometres (FIG), International Association of Geodesy (IAG), International Society of Photogrammetry and Remote Sensing (ISPRS), Open GIS Consortium (OGC), United Nations Geographic Information Working Group (UNGIWG), International Civil Aviation Organization (ICAO), Global Spatial Data Infrastructure (GSDI), Digital Geographic Information Working Group (DIGIWG), World Meteorological Organization (WMO), Committee on Earth Observation Satellites (CEOS), Global Spatial Data Infrastructure (GSDI) etc. All these enumerated user communities are the external liaison organizations of ISO/TC 211 Geographical information / Geomatics.

Implementation details are left to other organizations such as the Open GIS Consortium. The OGC is an international membership organization composed of many private companies, government agencies, and academic institutions, committed to development of geospatial data and geoprocessing standards. The OGC is working to develop the ‘interoperable geoprocessing’, which refers to ability of digital systems to:

- freely exchange all kinds of spatial informa-
tion about Earth and about objects and phenomena on, above, and below the Earth's surfaces.

- cooperatively, over networks, run software capable of manipulating such information

Since its establishment in 1994, ISO/TC211 (Secretariat NTS, Norway) Committee has been steadily increasing. There are now 28 P (participating) members and 30 O (observer) members.

Coordination with regional organization, like technical committee TC287 for geographical information of a Comité Européen de Normalisation (CEN), was very important for this study. The aim of TC287 committee was to develop "...a structured set of standards which specifies a methodology to define, describe, structure, interrogate, update, codify, transform and transfer data and metadata that represent geographic information'.

At the moment ISO TC211 works on 40 standards related to acquiring, processing, analyzing, and accessing of Geographical information. Work on those standards is subdivided into 9 working groups.

Working group (WG3) of TC211 responsible for Data Administration was concerned with standards important for the field of spatial data quality, and the most significant among them are:

- ISO 19113 Quality principles
- ISO 19114 Quality evaluation procedures
- ISO 19115 Metadata

Interesting project concerning the spatial data quality has been led under working group 9 (WG 9) which is in charge of Information management, with the following interesting standards:

- ISO 19138 Data quality measures
- ISO 19139 Metadata – Implementation specification

The ISO standards define principles for reporting quality of spatial data bases and designate components for submitting the report of their quality. These standards also take care of the access, organization and information of quality report. They are intended for both the companies which are involved in data bases production, and for spatial data users. The first ones use them in assessment and reporting of how much those data describe reality, formally or implicit, and others use them while making decision if some particular base consist of quality and useful data for some application. Realization of ISO standard projects must be completed until the end of 2006.

It is worth mentioning that standards and associations engaged in this field of research at the national level, like Spatial Data Transfer Standard (SDTS) developed by the United States Geological Survey (USGS) are based on the principles that users should be able to characterize fitness-for-use for a given application based on data quality documentation; likewise Canadian Spatial Archiving and Interchange Format (SAIF), or Digital Geographic Information Exchange Standards (DIGEST) which represents a NATO effort to develop standard geospatial data exchange formats for military applications.

METHODS FOR EVALUATION AND REPORT OF DATA QUALITY

The quality of a dataset might be described by using two components:
- data quality elements;
- data quality overview elements.

Data quality elements, together with data quality subelements and the descriptors of data quality subelements, describe how well a dataset meets the criteria, set forth in its product specification and how well it provides quantitative quality information. Quantitative quality information shall be reported as metadata in conformity with the requirements of ISO 19115, or by using a quality evaluation report in compliance with the requirements of ISO 19114 (6). According ISO 19113 recognized data quality elements are:

- completeness,
- logical consistency,
- positional accuracy,
- temporal accuracy,
- thematic accuracy,
- quantitative attribute accuracy.

ISO 19113 establishes the principles for describing the quality of geographic data and specifies components for reporting quality information. It also provides an approach for organizing information about data quality.

Data quality overview elements provide general, non-quantitative information, and they shall be reported as metadata in conformity with requirements of ISO 19115. Data quality overview elements are:

- purpose - which gives information on the reasons for creating the dataset and on the intended use of the dataset,
- usage - provision of information on the kind of application for which the dataset has been used,
- lineage - describes the history of the dataset.

ISO 19114 standard provides a framework of procedures for the evaluation of quality that is applicable to digital geographic datasets and it is consistent with principles defined in ISO 19113. The data quality evaluation procedures are used for determining and reporting data quality information, either as a part of data quality metadata only, or as a quality evaluation report as well. This standard is intended for persons and firms which are involved in spatial data bases production, as framework for data quality reports, as well as, for users who, on the basis of these reports, evaluate how much these data are useful for particular application.

The objective of ISO 19115 Standard is to provide a structure for describing digital geographic data. This International Standard is intended to be used by information system analysts, program planners, and developers of geographic information systems, as well as others, in order to understand the basic principles and the overall requirements for standardization of geographic information. This International Standard defines metadata elements, and establishes a common set of metadata terminology, definitions, and extension procedures which provide comfortably and effective use of such data. All these standards are applicable to all types of digital geographic data. Their principles can be extended to many other forms of geographic data, such as maps, charts and textual documents.

Data quality evaluation methods are divided into two main classes, direct and indirect. Direct methods determine data quality through comparison of the data with internal and/or external reference information. Indirect methods infer or estimate data quality by using information on the data, such as lineage. The direct evaluation methods are further subcategorised by the source of the information needed to perform the evaluation. Figure 3. depicts this classification structure.
The direct evaluation method is further subdivided into internal and external, depending on whether the reference data are parts of the existing data base, or they belong to external data.

Meta-information or metadata are vitally important for informing and warning the data users of the data limitations. The purpose of a metadata standard (ISO 19115) is to provide a common set of terminology and definitions for documentations related to metadata. Metadata (an attached / associated description of the data; data about data) statements should provide users with an indication of these data quality issues and fitness for use in (common applications). They are vitally important to informing and warning data users about limitation of the data. The main reason to document data is to maintain an organization’s investment in its geospatial data. Organizations that do not document their data often find that, over time or because of personal changes, they no longer know the content or quality of their data (4).

The metadata content elements included in the standard were determined on the basis that metadata serve the following roles (1):
- availability: Information to determine if information exist for a geographic location,
- fitness for use: Information to determine if data meet a specific need,
- access: Information needed to acquire a set of data,
- transfer: Information needed to process and use a set of data.

THE SERBIAN EXPERIENCE

Our country belongs to the group of participating countries and is involved in the work of ISO standards right from the beginning through the Institution for Standardization of Serbia&Montenegro (ISSM). At the national level there is Commission for Standards for Geographic Information/Geomatics (KS Z 211), established in 1995 under condensation of ISSM. The goal of this Commission is work on national standards concerning geomatics and also participation in related bodies of ISO TC211. Unfortunately, during the last few years the work of this Commission has become extinct, and thus our participation in ISO TC211 research could be considered as de facto standards. As an example one can mention regulations for which Republic Geodetic Authority of Serbia (RGA) is responsible. In such regulations like act standard for Digital Geodetic Map recognizes some of the mentioned elements of spatial data quality like topological, geometric and attribute consistency, positional accuracy etc. As the subsequent step that is supposed to be undertaken one can mention an immediate implementation and harmonization existing technical regulations with ISO standards, particularly with respect to those already mentioned above. At the institutions such as Republic Geodetic Authority of Serbia there is a widespread opinion that the work on these standards is indispensable and that during the next period something must be done along these lines.

CONCLUSION

The availability of spatial datasets at the local, regional and global level is indispensable for sustainable development planning. Exchange of information through different levels of Spatial Data Infrastructures (SDI) could be made attainable through vertical line of information, from the local, through the national, regional and up to the global level.

Great and rapid development of technologies and methods of surveying and mapping by using contemporary means, such as integrated geographic systems, satellite positioning, remote sensing digital networks, for sharing and disseminating of data, is of great importance with respect to sustainable development.

The rapid development of technologies and methods in surveying and mapping, such as integrated geographic information systems, remote sensing, satellite positioning systems and digital networks for sharing and disseminating of data, provides a strong and important tool for decision making for Sustainable Development. Accessible and relevant geographic information will play an important role in planning, executing and monitoring development.

The ability to access the information and to interpret this information by decision makers at all levels requires global planning and implementation of sustainable development. This is really the best way to evaluate risks and consequences of all relevant information. Recognizing quality elements, such data bases and their level determination play important role in decision making and planning.

The International Organization for Standardization (ISO) Technical Committee 211 for Geographic Information / Geomatics is developing an integrated suite of standards to promote global interoperability.

In the near future one can expect that data quality assessment is going to be increased. Recent efforts of national agencies which are trying to develop workable data quality standards for spatial data is the best indication for such statement. Increased standardization could enhance the opportunity of communicating data quality characteristic of transferred data.

One aspect of data quality assessment that is likely to increase in importance over the near-term is standardization of data quality information. This is recognized in recent efforts of the national agencies to develop workable data quality standards for geographical databases. Increased standardization would serve to enhance the ability to communicate the data quality characteristic of transferred data. Flexible standards are required allowing different levels of quality available to the intended use of the database.

The International Federation of Surveyors (FIG) decided at Congress in Brighton in 1998 to form a Task Force in order to prepare an FIG statement on how the Federation will implement the concept of sustainable development in the twenty-first century, exposed at The United Nations Rio Conference in 1992 in document known as Agenda 21. Conclusions of FIG Agenda 21 could be underlined as working guideline for all organizations participating in life cycles of spatial data bases, especially in collecting those data.

*Considerable data exist, but access to data is
often hampered by lack of standardization, coherence and adequate services for data retrieval, including information about what data exist and where data are kept.” **FIG Agenda 21** emphasizes: “Work with international bodies such as the International Standards Organization (ISO) to develop and implement suitable standards for the exchange of geographic information.” (11).

**References**