The Journal of GIS Association of Korea, Vol. 11, No. 4, pp. 327-340, December 2003

공간데이터베이스의 품질평가 방법에 관한 연구
- 토지데이터베이스를 중심으로 -

김미정* · 안종천** · 조우현**

A Study on the Quality Evaluation Method of Spatial Database
- Focusing on Land Database -

Mi-Jeong Kim* · Jong-Cheon An** · Woo-Hyun Cho**

요 약

공간데이터베이스의 품질평가는 공간데이터베이스의 특징에 따라 적절한 품질요소와 평가방법이 적용되어야 한다. 본 연구에서는 공간데이터베이스의 하나인 토지데이터베이스에 대한 구체적인 품질평가 방안을 제시하는데 목적을 두고 있다. 이를 위하여 토지데이터베이스를 구성하는 지형데이터베이스, 지적데이터베이스, 용도지역지구데이터베이스의 구축 단계별 품질요소를 선정하고, 선정된 품질요소에 대하여 평가방법을 제시하였다. 토지데이터베이스의 품질요소는 위치정확성, 속성정확성, 일관성, 완결성, 시간정확성, 신뢰성, 이력 등이며, 다양한 통계적 및 수학적 방법들을 검토하여 각 품질요소의 객관적인 측정방법을 제안하였다.

주요어 : 공간데이터베이스, 토지데이터베이스, 품질요소, 품질평가

ABSTRACT : Quality elements and evaluation methods should be considered according to the characteristics of spatial database. The purpose of this study is to propose specific methods for quality evaluation focusing on land database which are an important parts of spatial database. Through the study, objective methods of quality evaluation for selected quality elements are specified, which are based on the construction processes of the topographical database, cadastral database, and zoning database. Position accuracy, attribute accuracy, consistency,

* 국토연구원 GIS연구센터 책임연구원
** 국토연구원 GIS연구센터 연구원

completeness, temporal accuracy, believability, and lineage are selected as the quality elements of land database. A various statistical and mathematical skills are proposed for measurement and assesment methods of quality elements.

Keywords : Spatial Database, Land Database, Quality Element, Quality Evaluation

1. 서 론

다양한 분야에서 공간데이터베이스를 사용하기 위해서는 품질수준이 사용목적에 적절한지를 판단하는 것이 중요하다. 공간데이터베이스의 품질수준에 따라 사용목적의 달성정도가 달라질 수 있고 사용할 수 없을 수도 있다. 또한 데이터베이스의 유형에 따라 어떤 품질요소를 평가하는가도 중요한 사항이다. 예를 들어 지질데이터베이스의 경우 시간의 흐름에 따라 거의 변하지 않기 때문에 현재성이 얼마나 반영되었는지를 판단하는 시간정확성은 상대적으로 중요성이 떨어진다. 반면에 지적데이터베이스의 경우 토지이동 및 합병이 빈번하게 일어나므로 이를 반영하는 시간정확성은 매우 중요한 요소라고 할 수 있다.

공간자료에 대한 품질평가와 관련하여 대부분의 연구들은 공간데이터베이스보다는 수치지도를 대상으로 이루어졌다. 이와 같은 연구 중에 많은 연구가 수치지도에 대한 최종 결과물(김갑진 외, 1999; 김병국 외, 1999)을 대상으로 품질평가 방법을 제시하고 있다. 최종 결과물에 대한 품질평가에서는 이력, 위치정확성, 속성정확성, 논리적 일관성, 완결성, 시간정확성 등의 품질요소를 고려하였지만, 수치지도의 세부적인 제작단계별 품질에 대한 접근이 미흡하였다. 한편 수치지도 제작단계별 품질평가를 다룬 연구(이현직 외, 1998; 조우석 외, 1997; 조윤숙 외, 2003)에서는 위치정확성을 중심으로 품질을 평가하였기 때문에 제작 단계별로 나타날 수 있는 다양한 품질요소에 대한 평가가 미흡하였다. 이들 연구에서는 수치지도작성과 관련된 규칙 및 내규 등(건설교통부, 1995; 국립지리원, 1995) 제도적인 규정에 의한 성과검수차원에서 연구가 이루어졌기 때문에 품질수준에 대한 수요자의 다양한 요구수준을 반영한 객관적인 품질평가 방법론이라고 하기에는 한계가 있다. 또한 이들 연구의 대상이 공간데이터베이스라기 보다는 수치지도를 중심으로 이루어졌기 때문에, 공간데이터베이스의 특성을 반영한 품질요소를 고려하지 못하였다.

토지데이터베이스는 다양한 종류의 공간자료와 속성자료가 일련의 복잡한 과정을 통하여 구축된 것으로, 양질의 품질을 확보하기 위해서는 최종 성과물에 대한 일률적인 평가보다는 각 단계별로 품질을 평가하는 것이 바람직하다. 최종 성과물에 대하여 품질을 평가할 경우 품질의 저하가 어느 단계에서 어떤 행위로 발생했는지 정확히 파악할 수 없기 때문이다.

따라서 본 연구의 목적은 토지데이터베이스의 특성에 적합한 품질요소를 단계별로 선정하고 이를 측정할 수 있는 객관적인 평가 방안을 제시하는데 있다. 이를 위하여 토지데이터베이스의 특징에 따라 품질요소를 선정하고, 토지데이터베이스를 구성하고 있는 지형데이터베이스, 지적데이터베이스, 용도지역·지구데이터베이스의 구축단계별 품질평가를 위하여 품질요소별 품질평가방안을 제시하였다.

2. 공간데이터베이스 품질요소와 측정방법

2.1 공간데이터베이스 품질요소

1) 공간데이터베이스 품질 개념

일반적으로 전산환경에서 데이터베이스는 자료의 중복 등록을 피하기 위하여 특정 업무나 프로그램에 예속시키지 않고 다양한 업무에 여러 가지 형태로 이용할 수 있도록 설계된 자료 파일의 집합체라고 할 수 있다. 데이터베이스 품질은 데이터베이스의 바람직한 정도 또는 우수성이라고 정의되며, 데이터베이스 자료에 관한 품질과 데이터베이스 서비스에 관한 품질로 구분할 수 있다(이국희 외, 1998). 데이터베이스에서 자료의 품질이란 데이터베이스가 담고 있는 데이터 자체의 바람직한 정도를 의미하며, 데이터 구조, 데이터 값, 데이터 표현 등 세 가지 관점에서 바람직한 정도를 분석할 수 있다. 반면, 데이터베이스의 서비스 품질이란 데이터 자체가 아니라 데이터가 사용자에게 서비스되는 과정에서의 매체, 수단, 방법, 그리고 공급처 지원 등의 바람직한 상태를 의미하는 것으로 서비스의 완전함, 정확함, 신속함, 편리함, 이해 가능함, 친절함, 다양함 등을 망라하는 개념이다.

데이터베이스의 품질을 사용목적과 관계를 고려하여 용도에 대한 적합성이라고 정의하기도 하는데, 이는 소비자가 제품을 사용함으로써 그 목적이 성공적으로 달성된 정도를 말하는 것으로 품질의 결정 주체를 소비자로 보고 소비자의 관점을 반영한 것이다. 또한 품질을 요구조건에 대한 일치성이라고 정의하는 관점도 있는데, 이는 품질을 적절히 관리하기 위해서 품질을 측정할 수 있어야 한다는 것으로 생산자의 관점을 반영한 것이다 (Wang R. Y. et. al., 2001).

한편 공간데이터베이스는 데이터베이스를 구성하는 내용물이 공간과 관련되어 있으며, 크게는 실제세계의 지형지물을 점, 선, 면 형태로 나타낸 공간자료와 이들 지형지물에 대하여 문자, 숫자 등으로 나타낸 속성자료로 구성되어 있다. 따라서 데이터베이스 품질에 대한 일반적인 정의를 고려할 때 공간데이터베이스 품질도 데이터베이스에 담겨져 있는 자료 자체의 바람직한 정도를 의미하는 것으로 자료 구조, 자료 값, 자료 표현이라는 측면에서 접근할 수 있다. 그러나 데이터베이스에 담겨져 있는 자료 자체의 바람직한 정도는 사용목적에 따라 품질요소들이 조금씩 다르게 나타날 수 있기 때문에 이 연구와 관련된 토지데이터베이스의 품질정의는 토지관련 행정업무의 사용목적에 적합하게 활용

할 수 있는 정도라고 할 수 있다.

하지만, 품질은 늘 고정된 것이 아니라 시대적인 변화에 따라 변화하는 것이기 때문에 동적인 상태에서의 품질을 고려해야 한다. 특히, 정보화시대에서 공간데이터베이스의 품질은 전산환경에서의 단순한 구축만이 아니라 다양한 활용성을 통한 고부가가치의 창출이 중요하게 작용하기 때문에 품질에 대한 정의도 기준, 결정 주체, 사용목적 등을 고려하여 동적인 측면에서 접근되어야 한다.

2) 공간데이터베이스 품질요소

공간데이터베이스 품질을 구성요소 관점에서 접근하는 연구(Guptill, S. C. and Morrison, J. C.(eds), 1999; International Cartographic Association, 1995; ISO TC211, 2001; SDTS, 1992; Veregin, H., 1998)에서는 품질평가 요소로 완결성, 일관성, 위치정확성, 시간정확성, 주제정확성, 해상도, 이력, 보안성 등 다양하게 제시하고 있다. 하지만, 이들 연구에서는 품질평가를 위한 품질요소의 선정에 초점을 두었기 때문에 품질평가를 위한 측정방법, 평가기준 등과 같은 구체적인 방법론의 제시가 미흡하다. 그러나 공간데이터베이스의 특성에 따라 품질요소를 선정하고 선정된 품질요소를 측정할 수 있는 측정방법과 평가기준을 적용하여야 하기 때문에 품질요소와 측정방법 및 평가기준은 동시에 고려되어야 한다(표 1).

공간데이터베이스의 품질요소 고찰과 관련하여 공간데이터베이스가 가지는 자료의 특성, 사용 목적, 평가의 대상범위, 자료 사용자, 자료의 표현 등이 검토되어야 한다.

<표 1> 공간데이터베이스 특성에 따른 품질요소

공간데이터베이스 특성		평 가 항 목
사용목적에 따라		해상도, 보안성, 타당성
평가대상 범위	미시적 대상	위치정확성, 속성정확성, 논리적일관성, 해상도
	거시적 대상	완결성, 시간정확성, 이력
자료사용자	상업사용자	완결성, 무결성, 유동성, 이해의 용이성
	자료분석가	정확성, 단순성
	행정업무자	통합성
	응용시스템 개발자	자료의 구현성
자료 표현방법	문자자료	구현성, 이해의 용이성, 보안성
	공간자료	위치정확성, 속성정확성, 시간정확성, 해상도, 이력

자료의 특성은 시간, 공간, 주제적인 측면과 자료를 이용하여 분석을 충분히 할 수 있는지, 또는 자료 표현이 일관성이 있게 표현되었는지에 대한 측면에서 적합성을 고려할 수 있다(Guptill S. C. and Morrison J. L., 1999). 따라서 자료의 특성을 반영하는 품질요소는 위치정확성, 속성정확성, 시간정확성, 완결성, 일관성 등으로 분류할 수 있다.

사용 목적에 따른 품질요소는 해상도, 보안성, 타당성 등을 품질요소로 분류할 수 있다(Aronoff S., 1985). 즉, 해상도의 경우 지리정보에 대한 선명도를 목적으로 할 경우에는 해상도가 좋아야 하며, 자료 공개에 보안이 필요한 경우 보안성이 중요하게 작용할 수 있기 때문이다.

평가의 대상범위에 따른 품질요소는 크게 거시적인 측면과 미시적인 측면으로 나누어 살펴볼 수 있다. 먼저, 거시적인 측면에서는 자료 전체를 하나의 평가대상으로 선정하여 종합적으로 평가하는 것을 의미하기 때문에 완결성, 시간정확성, 이력 등이 해당된다. 반면 미시적인 측면에서는 객체 하나하나에 대하여 개별적으로 품질을 평가하는 것으로 위치정확성, 속성정확성, 논리적 일관성, 해상도 등의 품질요소가 해당된다.

자료 사용자에 따른 품질요소는 자료를 누가 사용하는지에 따라서 품질요소가 선정될 수 있다. 일반적으로 자료의 사용자는 상업적 사용자, 자료분석가, 행정업무자, 응용시스템개발자 등으로 나누어지며 이들 사용자별로 품질요소가 분류된다(Moody D. L., 1998). 즉, 상업적 사용자의 경우에는 완결성, 무결성, 유동성, 이해의 용이성이 해당되며, 자료분석가에게는 정확성과 단순성, 행정 업무자에게는 통합성, 응용시스템개발자에게는 자료의 구현성이 중요한 품질요소로 고려된다.

자료의 표현 방법에 따른 품질요소는 크게 문자자료와 공간자료로 구분할 수 있다. 문자자료는 일반적인 정보시스템에 활용되거나 생산되는 자료로서 공간적인 위치개념이 포함되어 있지 않으므로 시스템에서 구현이 용이한지, 이해는 쉽게 할 수 있는지, 보안성은 있는지 등에 대한 품질요소에 대하여 품질평가가 주로 이루어지고 있다. 반면, 공간자료에서는 위치정확성, 속성정확성, 해상도, 시간정확성, 이력 등이 품질요소로 고려되고 있다.

2.2 품질요소의 측정방법

품질요소의 평가는 요소의 특성에 따라 다양한 통계적 또는 수학적인 측정방법을 사용하게 된다. 여기서는 공간데이터베이스에서 사용되는 다양한 품질요소를 검토하여 토지데이터베이스의 품질평가에 적합한 품질요소를 선정하고, 각 품질요소를 평가하기 위한 다양한 측정방법들을 고찰하였다.

1) 위치정확성

위치정확성은 수치화된 객체의 위치가 실제세계의 참위치에 얼마나 근접한지를 나타내는 정도로 절대 정확성과 상대 정확성으로 구분된다(Montogomery G. E. and Schuch H. C., 1993).

절대정확성은 수치지도에서의 어떤 대상물의 위치가 실제 지상에서의 참위치(절대좌표계에서의 위치)와 얼마나 일치하는가를 나타내는 것이다. 이는 다시 X, Y좌표나 경위도 좌표를 이용하여 측정하는 수평적 정확성과 등고선, DEM 등을 이용하여 측정하는 수직적 정확성으로 나누어진다. 상대정확성은 자료의 위치를 다른 자료의 위치를 기준으로 나타내는 것으로 두 객체의 거리와 실제 지상의 거리를 비교하여 정확성을 측정할 수 있다. 이는 다시 자료내부의 상대적 위치일치성을 판단하는 위치정확성, 자료내부의 상대적 길이일치성을 판단하는 길이정확성, 자료내부의 상대적 면적일치성을 표현하는 면적정확성으로 나누어진다. 상대정확성은 면적비교, 객체의 둘레비교, 객체 중심점에서 버텍스까지의 거리비교 등을 통하여 이루어진다.

위치정확성의 측정은 정량적인 방법으로 측정할 수 있으며 기준이 되는 자료와의 오차정도를 표현하는 것이므로 평균제곱근오차를 구하여 측정할 수 있다. 절대정확성의 경우 좌표간의 거리를 산정하여 평균제곱근오차를 구하고, 오차가 신뢰도 구간에 들어가는지의 여부를 통하여 평가할 수 있다. 상대정확성의 경우에는 면적의 변화, 객체를 형성하는 길이의 변화, 객체간의 거리 변화를 절대정확성과 마찬가지로 평균제곱근오차를 구하고 신뢰도 구간에 포함되는지의 여부를 통하여 평가할 수 있다. 평가기준은 신뢰도 구간에 따라 정해지며, 신뢰도 구간은 데이터의 사용목적에 따라 각각 다르게 정해진다.

2) 속성정확성

속성정확성은 자료에 있어서 임의의 위치나 위치군 및 지표면의 형상에 대한 사실(표고, 온도, 지질, 지목, 자연 및 인공물의 종류 및 이름 등)을 나타내는 것으로 원래의 속성값이 정확하게 유지되었는지에 대한 평가가 이루어져야 한다(Guptill S. C. and Morrison J. L.,(eds) 1999).

속성정확성을 평가하기 위해서는 분류정확성, 정량적 속성정확성, 정성적 속성정확성 등으로 구분하여 분석할 수 있다. 먼저, 분류정확성은 개개의 자료에 대한 분류코드와 속성값의 정확성을 분석하는 것이고, 정량적 속성정확성은 지형지물속성 중에서 정량적 값으로 정의된 것의 정확도를 분석하는 것이다. 마지막으로 정성적 속성정확성은 지형지물속성 중에서 토지이용현황, 식생형태 등과 같이 정성적 값으로 정의된 것의 정확도를 분석하는 것이다.

분류정확성에 대한 평가는 속성을 분류한 계급의 수, 각 지역의 형태와 크기, 표본지점에 대한 선정방법이 고려되어야 하기 때문에 실제로 적용하기는 어렵다. 정량적 속성정확성은 위치정확성 측정방법과 마찬가지로 평균제곱근오차를 구하여 평가할 수 있으며, 정성적 속성정확성은 실제 지형지물에 대한 속성값과 자료로 구축된 속성값을 비교하여 오차매트릭스를 만들어서 Kappa계수를 이용하여 일치도를 산정할 수 있다.

속성정확성은 속성값 내용의 불일치, 속성값 내용의 누락, 속성값의 범위 불일치 등의 항목을 평가함으로써 품질평가를

수행할 수 있다. 속성정확성의 평가기준은 평균제곱근오차를 산정하여 측정하는 정량적 속성정확성은 신뢰구간을 선정하여 평가할 수 있으며, 정성적 속성정확성은 Kappa계수의 수치에 1에 가까울수록 일치한다고 볼 수 있다.

3) 시간정확성

시간정확성은 사용하고자 하는 업무에 적절하게 사용할 수 있는 최신의 자료가 작성되어 있는지 여부를 의미한다. 시간정확성은 최신성, 시간적 유효성을 통하여 평가하게 된다. 최신성은 객체나 자료셋의 갱신일자로 판단하며, 시간적 유효성은 자료가 유효한 시간이 얼마인지를 제공하는 것이다.

시간정확성에 대한 평가는 하나의 데이터가 얼마나 최신성을 가지고 있느냐가 측정되는 것으로서 최신성을 나타내기 위하여 자료사용시점과 입력시점을 비교하여 측정할 수 있다.

$$C = (D\text{-}I) + (P\text{-}D) \quad \cdots\cdots\cdots\cdots \quad (식\ 1)$$

C = 자료의 최신성, D = 자료의 공개시기,
I = 자료의 구축시기, P = 현재

시간정확성은 일정한 범위를 가지고 있는 것이 아니라 데이터의 특성에 따라 다르게 나타나는데 예를 들어 지적도의 경우 최신성이 자료 사용시점과 자료구축시기가 동일하게 나타나야 하므로 최신성은 0으로 나타나게 된다. 따라서, 최신성은 0에 가까울수록 좋다고 볼 수 있으며 0보다 크면 클수록 시간정확성은 떨어진다고 볼 수 있다.

4) 완결성

완결성은 추상화된 대상과 자료 집합으로 표현된 객체와의 관계에 대한 특성으로 자료의 누락이나 계측 가능한 자료의 질에 대한 오차가 없어야 함을 의미한다. 완결성은 다시 자료완결성과 모델완결성으로 나눌 수 있고, 자료완결성은 다시 객체완결성과 속성완결성으로 나눌 수 있다(Veregin, H., 1998).

자료완결성은 데이터베이스와 사양서 사이에 나타나는 누락오차로 사양서에 설명된 모든 객체가 다 있다면 자료완결성이 충족되었다고 할 수 있다. 객체완결성은 실세계의 모든 객체들이 누락 혹은 초과 없이 표현되었는가를 통하여 분석할 수 있으며, 속성완결성은 모든 속성들이 누락 혹은 초과 없이 표현되었는가와 관련된다. 모델완결성은 사용목적에 맞는 모델이 구축되었다면 모델완결성이 충족되었다고 할 수 있다.

완결성을 측정하기 위해서는 Boolean방식을 이용하여 논리적인 사실이 참인지 거짓인지를 판단하고 평가해야 한다. 객체완결성에 대해서는 추상화영역의 모든 객체에 대한 누락 여부를 통하여 평가할 수 있으며, 속성완결성은 다음의 사항에 대하여 논리적인 측정을 수행하는 것으로 모두 참인 값으로 측정되어야 한다. 즉, 메타자료에는 정의되어 있지 않지만 사양서에서는 입력하도록 되어 있는 객체에 대한 속성값의 누락여부, 메타자료에는 정의되어 있지만 입력되어 있지 않은 속성값의 존재 여부(DBMS함수로 검사), 객체의 일부에만 속성값이 누락되어 있는

경우(자동적으로 체크 가능), 복잡한 형태의 속성값이 부분적으로 누락되어 있는 경우 등을 분석하여야 한다.

완결성에 대한 측정내용은 하나라도 오류가 생기면 공간분석 및 활용에 이용할 수 있는 데이터가 제공되지 않으므로 평가기준은 100% 만족되어야 한다.

5) 일관성

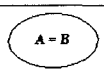
일관성은 자료들간의 논리적 관계가 얼마나 잘 구성되었는지를 나타내는 것으로 단순한 속성의 범위, 기하학적 조건, 위상적 조건으로부터 공간관계에 대한 특별한 일관성 규칙에 이르기까지 여러 단계로 구분할 수 있으며 이 규칙을 어길 때 일관성은 깨어지게 된다. 일관성은 다시 자료 일관성, 위상 일관성, 표현 일관성으로 나누어질 수 있다 (Guptill S. C. and Morrison J. L.,(eds) 1999).

자료일관성은 자료 무결성을 의미하는 것으로 트랜젝션, 도메인 일관성 등에 대하여 평가가 이루어진다. 여기서 트랜젝션은 데이터베이스 오퍼레이션에 해당하는 것으로 성공적으로 오퍼레이션이 수행되었는지의 여부를 나타내는 것이고, 도메인 일관성은 공간객체의 속성이 논리적으로 올바른지에 대한 것으로, 예를 들어 속성값이 날짜라면 그 범위가 1에서 31에 해당해야 하고, 음수는 해당하지 않는다는 법칙에 위배되지 않아야 한다.

위상일관성의 경우에는 공간참조자료의 위상표현이 가장 일반적인 방법으로 cell complex(0-cell : 점, 1-cell : 선, 2-cell : 면)를 보면 위상일관성 법칙이 5가지로 나타난다.

표현일관성은 객체간의 수학적 구성법칙으로 연결성 매트릭스를 통해서 정의할 수 있는데, 각 객체의 연결성은 <표 2>과 같은 특성을 가지게 되며, 연결성 매트릭스에 위배되지 않아야 표현일관성이 확보되었다고 할 수 있다.

<표 2> 표현일관성의 연결성 메트릭스(Guptill S. C and Morrison J. L. 1999)

	A	B	C	D	E
A	 (equal)	(meet)	(meet)	(disjoint)	(disjoint)
B	(meet)	(equal)	(covers)	(contains)	(contains)
C	(meet)	(coverd by)	(equal)	(disjoint)	(disjoint)
D	(disjoint)	(inside)	(disjoint)	(equal)	(contains)
E	(disjoint)	(inside)	(disjoint)	(inside)	(equal)

6) 이 력

자료의 이력은 맨 처음 원시자료에 대한 수집으로부터 편집, 수정, 전환 그리고 새로운 분석처리에 따른 변환까지 이르는 일련의 과정에 대한 자료의 내역을 의미한다. 자료의 이력이 가지는 궁극적인 목적은 과거로부터 현재까지 제작된 자료로부터 보다 유용한 자료를 만들기 위한 정보를 축적하는데 있다고 볼 수 있다.

자료 이력에 대한 정확도를 측정하기 위하여 표본을 추출하지 않고 전체 자료를 대상으로 자료원, 제작절차, 투영 및 좌표변환의 파라미터, 결정인자 및 판단기준 등이 기준에 맞게 제작되었는지를 평가할 수 있다(한선희, 김영표, 2002). 그 외에 내용의 원본과 추가 및 삭제사항, 전환방법 등의 처리 과정에 대하여 기술되어 있는지도 평가에 포함될 수 있다(최병남 외, 1997).

이력은 원시자료에 대한 정보와 구축된 자료에 대한 정보를 구축단계별로 상세하게 제공하는 것이 가장 이상적이다. 이 연구에서는 이력에 포함되어야 할 항목의 선정을 위해 국가표준(한국정보통신기술협회, 2002)을 준거로 사용할 수 있다. 이력의 측정 방법은 이력내용이 모두 포함되어 있는지의 여부를 통하여 수행하게 되며 평가기준은 100% 충족되는 것이 이상적이지만 사용목적에 따라 달리 적용될 수 있다.

7) 신뢰성

신뢰성은 자료에 대한 믿음 혹은 경험적인 신뢰도로서 객관적인 판단과 주관적인 판단이 함께 이루어져야 한다. 신뢰성을 나타내는 요소는 원시자료에 대한 신뢰도, 데이터베이스가 표준을 준수하는가와 관련된 표준에 대한 비교평가, 그리고 이전의 사용경험에 기반을 둔 평가로 세분하여 볼 수 있다. 신뢰성의 측정 방법은 사용자의 경험에 의존한 경우이므로 데이터베이스의 신뢰도를 평가하기 위해서는 사용자를 대상으로 하여 설문조사를 실시하여 평가할 수 있다.

3. 토지데이터베이스 품질평가 방안

3.1 토지데이터베이스의 특징과 품질평가방법

토지데이터베이스는 일반적인 데이터베이스와는 달리 행정업무 및 민원서비스에 사용할 목적으로 구축된 데이터베이스로 토지 관련 공간자료와 속성자료가 결합된 공간데이터베이스이다. 따라서 토지데이터베이스의 품질을 평가하기 위해서는 토지데이터베이스가 어떠한 자료들을 대상으로 구축되었는지, 또한 무슨 목적에 사용할 것인지 등 다양한 특성이 고려되어야 한다. 이와 같이 토지데이터베이스의 특성을 검토하여 각 특성이 반영될 수 있는 품질요소가 평가되어야 한다.

먼저, 토지데이터베이스는 공간자료를 저장 및 관리하는 공간데이터베이스이다. 공간을 포함하고 있는 공간데이터베이스는 좌표값을 가지고 있어서 객체마다 공간적 위치를 가지고 있으며 이들 공간자료는 속성자료와 연계되어 있다. 이러한 토지데이터베이스를 이용하여 토지 관련 공간분석

및 정책결정 등에 이용할 수 있다. 따라서 공간자료의 위치정확성을 측정함으로써 공간데이터베이스의 품질이 평가되어야 한다.

둘째, 토지데이터베이스는 주요 사용 목적이 행정업무 및 민원서비스에 사용하는 것이다. 따라서 이러한 구축목적에 적합하기 위해서는 필지 하나하나에 해당하는 면적, 위치 등의 절대값 뿐만 아니라 필지에 대한 가장 최근의 변동사항을 반영하는 시간정확성도 중요하다. 또한 용도지역지구간의 관계 등이 논리적으로 어긋나지 않도록 일관성 등이 평가되어야 한다.

셋째, 토지데이터베이스는 다양한 자료를 통합하여 사용하고 있다. 즉, 지형, 지적, 용도지역지구데이터베이스 등 다양한 자료가 통합되어 이루어져 있다. 또한 토지데이터베이스를 기본으로 하여 타 시스템에서 연계활용하고 있으므로, 데이터베이스에 대해 변화가 나타나면 이 변화가 다른 데이터베이스에도 영향을 미친다. 따라서 데이터가 누락되었거나 중복되는 경우가 없어야 하며 이는 데이터베이스의 완결성이 확보되어야 한다는 것을 의미한다.

넷째, 토지데이터베이스는 다양한 축척의 원시자료들을 사용목적에 맞게 변형하여 사용한다. 토지데이터베이스의 원시자료는

<표 3> 토지데이터베이스의 품질요소와 품질평가방법

품질요소		평가항목	평가방법	평가기준	비고
위치정확성	절대정확성	x,y 좌표값	RMSE 산정	정규분포의 신뢰도 선정	산정
	상대정확성	상대적 위치정확성, 길이정확성, 면적일치성			
속성정확성	정성적정확성	속성내용 불일치, 범위불일치	오차메트릭스 Kappa계수 산정	절대값	산정
시간정확성		최신성	최신성 수식	절대값	최신성 수식 산정
완결성	데이터완결성	객체완결성	Boolean값 (객체누락 및 중복)	100% 참값	도면/자동검사
		속성완결성	속성값누락여부(자료사양서리스트와 메타자료 리스트 비교, 자동체크, 수동검사)	100% 존재	자동/화면검사
	모델완결성	사용목적에 적절한 사양서 작성여부	사양서 비교	상대적	문서검사
일관성	데이터일관성	자료무결성	오차율 산정	100% 참값	자동검사
	위상일관성	단위 및 입력, 교차, 오일러 공식, 노드연결, 폴리곤연결	Boolean값	100% 참값	자동검사
	표현일관성	위상적 표현일관성	연결성 매트릭스 작성	100% 참값	자동검사
이력		원시자료 구축절차	이력의 세부내역 작성여부	상대적	문서/자동검사
신뢰성		원시자료신뢰도	설문조사를 통한 평가	절대값	설문조사
		표준과 비교			수동검사
		경험 기반 신뢰도			설문조사

수치지형도, 수치지적데이터, 용도지역지구도 등으로 이들은 또한 각각의 데이터가 다양한 축척으로 구축되어 있다. 이와 같은 다양한 축척을 이용하여 연속 및 편집데이터베이스를 만들기 때문에 이들 구축 과정에 대한 이력이 평가되어야 한다.

다섯째, 토지데이터베이스는 공간분석을 수행하는 기능을 가지고 있다. 토지데이터베이스는 공간자료와 속성자료의 연계와 위상관계 등이 구축되어 있기 때문에 토지관련 공간분석, 공간계획, 도시계획 등에 사용할 수 있다. 공간분석을 수행하기 위해서는 객체간의 위상관계가 설정되어야 한다. 따라서 공간분석을 수행하기 위해서는 데이터베이스의 일관성과 완결성에 대한 품질이 평가되어야 한다.

이상과 같은 토지데이터베이스의 특징에 따라 선정한 품질요소는 <표 3>과 같다.

3.2 토지데이터베이스 구축과정 및 평가품질요소

1) 지형데이터베이스

지형데이터베이스 구축과정은 원시자료인 수치지형도로부터 필요한 지형지물 레이어 추출, 도면접합을 통한 연속지형도 구축, 기하학적 오류 수정, 위상구조 형성, 공간자료와 속성자료의 연계, 토지지형데이터베이스 등록으로 구분된다. 이들 단계 중 수치지형도(원시자료), 기본지형레이어, 연속지형레이어, 무결점연속지형레이어, 지형데이터베이스 등 총 5 단계를 품질평가를 위한 단계로 선정하였으며, 각 단계별로 평가되어야 하는 품질요소는 <표 4>와 같고, 품질요소별 평가방법은 앞에서 제시한 바와 같다.

2) 지적데이터베이스

지적데이터베이스[1]는 토지이용계획확인서의 발급, 공시지가 업무 참조뿐만 아니라 용도지역지구의 결정과 변경이 발생할 경우 기본도 등의 용도로 사용하기 위하여 구축되는 것으로, 다양한 축척의 원시자료를 사용하여 복잡한 단계를 거쳐 최종 지적데이터베이스가 구축된다. 이 연구에서는 낱장지적레이어 및 경계점좌표등록, 정도곽낱장지적레이어, 좌표변환낱장지적레이어, 좌표변환연속지적레이어, 통일원점연

<표 4> 지형데이터베이스 구축단계별 품질요소

구 축 단 계	품 질 요 소
수치지형도	절대위치정확성, 범주속성정확성, 시간정확성, 완결성, 일관성, 이력
기본지형레이어	완결성, 이력
연속지형레이어	절대위치정확성, 범주속성정확성, 이력
무결점연속지형레이어	절대위치정확성, 일관성, 완결성, 이력
지형데이터베이스	범주속성정확성, 시간정확성, 완결성, 일관성, 신뢰성, 이력

1) 낱장 지적도가 아니라 연속지적데이터베이스와 편집지적데이터베이스를 지칭함

<표 5> 지적데이터베이스 구축단계별 품질요소

구 축 단 계	품 질 요 소
낱장지적레이어	절대위치정확성, 상대위치정확성, 범주속성정확성, 절대값속성정확성, 시간정확성, 완결성, 일관성, 이력
경계점좌표등록	절대위치정확성, 시간정확성, 이력
정도곽낱장지적레이어	절대위치정확성, 이력
좌표변환낱장지적레이어	절대위치정확성, 이력
좌표변환연속지적레이어	상대위치정확성, 범주속성정확성, 완결성, 일관성, 이력
통일원점연속지적레이어	상대위치정확성, 범주속성정확성, 완결성, 일관성, 이력
경계점좌표등록레이어	속성정확성, 완결성, 일관성, 이력
연속지적레이어	상대위치정확성, 속성정확성, 완결성, 일관성, 이력
연속지적데이터베이스	일관성, 이력, 신뢰성
편집지적데이터베이스	상대위치정확성, 이력

속지적레이어, 경계점좌표등록레이어, 연속지적레이어, 연속지적데이터베이스, 편집지적데이터베이스 등 총 9단계를 품질평가를 위한 단계로 선정하였으며, 각 단계별 품질요소를 고려하여 품질평가 방안을 제시하였다<표 5>.

3) 용도지역지구데이터베이스

이 연구에서는 용도지역지구도 및 대장/조서, 낱장용도지역지구래스터 파일, 낱장용도지역지구레이어, 지형기반용도지역지구레이어, 연속지적기반용도지역지구레이어, 연속지적기반무결점용도지역지구레이어, 연속지적기반용도지역지구데이터베이스, 지형지적기반용도지역지구데이터베이스 등 총 10 단계를 품질평가를 위한 단계로 선정하였다. 또한 <표 6>과 같이 각 단계별 품질요소를 고려하였고, 단 품질요소별 평가방법은 앞에서 제시한 바와 같다.

<표 6> 용도지역지구데이터베이스 구축단계별 품질요소

구 축 단 계	품 질 요 소
용도지역지구도	속성정확성, 시간정확성, 이력
낱장용도지역지구래스터파일	속성정확성, 완결성, 이력
낱장용도지역지구레이어	완결성, 일관성, 이력
지형기반용도지역지구레이어	절대위치정확성, 일관성, 이력
연속지적기반용도지역지구레이어	범주속성정확성, 이력
연속지적기반무결점용도지역지구레이어	속성정확성, 시간정확성, 완결성, 일관성, 이력
연속지적기반용도지역지구데이터베이스	완결성, 일관성, 이력
지형지적기반용도지역지구데이터베이스	이력

4. 결 론

공간데이터베이스의 객관적인 품질평가를 위해서는 종이도면 혹은 전산 파일에 대하여 수행하던 기존의 평가방법만으로는 부족하다. 기존에 수행되던 평가방법은 측량 관점에서 이루어졌기 때문에 GIS기반의 공간데이터베이스의 특성을 반영하지 못하고 있다. 따라서 공간데이터베이스의 특성을 반영한 품질평가가 이루어질 수 있도록 평가방법이 제시되어야 한다.

본 연구에서는 GIS기반의 공간데이터베이스의 하나인 토지데이터베이스의 품질을 평가하는 방안을 모색하였다. 토지데이터베이스는 공간데이터베이스라는 특징과 함께 행정업무나 민원서비스 등의 용도로 사용되는 목적을 가지고 있기 때문에 이에 적합한 품질평가가 이루어져야 한다. 특히, 현재 시군구에서 구축중인 토지데이터베이스는 지자체뿐만 아니라 국가의 공간정보 인프라 역할을 담당하고 있으므로, 이에 대한 품질 수준이 객관적으로 평가되고 확보되어야 한다. 또한 토지데이터베이스의 품질을 평가함으로서 현재 구축되고 있는 토지데이터베이스의 품질을 향상시킬 수 있는 구체적인 방안이 모색될 수 있다.

토지데이터베이스의 품질평가를 위해 토지데이터베이스의 특징을 먼저 파악하고 이러한 특징에 따라 평가해야 할 품질요소를 선정하여 품질요소별 평가방안을 제시하였다. 특히, 토지데이터베이스를 구성하고 있는 지형데이터베이스, 지적데이터베이스, 용도지역지구데이터베이스의 구축과정에 따라 단계별로 필요한 품질요소를 선정하고, 각각의 품질을 평가할 수 있는 방안을 제시하였다는 점에서 선행 연구와 차별성을 가지고 있다.

토지데이터베이스의 품질평가에 적합한 평가방안의 제시를 통하여 현재의 품질수준에 대한 파악은 물론 품질수준의 제고가 필요한 단계를 파악하여, 토지데이터베이스의 전체적인 품질을 향상시키는데 기여할 수 있을 것이다. 또한, 이 연구를 통하여 전산 환경에서 자동적으로 품질평가가 가능한 품질요소가 검토되었으며, 이에 해당하는 품질요소에 대해서는 우선적인 품질평가가 이루어져 보다 양질의 품질확보에 기여할 수 있을 것이다.

참고문헌

건설교통부, 1995, 「수치지도작성작업규칙」, 경기 : 건설교통부.

국립지리원, 1995, 「수치지도작성작업내규」, 경기 : 국립지리원

김갑진 외, 1999, 「수치지도 정확도 제고를 위한 수정/갱신방안에 관한 연구」, 경기 : 국립지리원.

김병국 외, 1999, 「수치지도 품질관리 연구」, 경기 : 국립지리원.

이국희 외, 1998, 「데이터베이스 품질평가에 관한 연구」, 서울 : 한국데이터베이스.

이현직 외, 1998, 「수치지도 위치정확도에 관한 연구」, 경기 : 국립지리원.

조우석 외, 1997, 「수치지도의 정확도 향상방안 연구」, 경기 : 국토개발연구원.

조윤숙 외. 2003. 「지리정보 품질지수의 도입방안 연구」, 경기 : 건설교통부.

최병남 외, 1997, 「공간정보데이터베이스 구축을 위한 실험연구」, 경기 : 국토개발연구원.

한국정보통신기술협회, 2002, 「지리정보 유통 목록(메타데이터) 표준」, 경기 : 한국정보통신기술협회.

한선희, 김영표, 2002. 「지리정보의 정확도 측정 방법에 관한 연구」, 경기 : 국토개발연구원.

Guptill S. C. and Morrison J. L.(eds), 1999, *Elements of Spatial Data Quuality*. Oxford : Elsevier.

International Cartographic Association, 1995, *Elements of Spatial Data Quality*, Pergamon.

ISO TC211, 2001, ISO/DIS 19114 : *Geographic Information Quality Evaluation Procedures*. International Standard Committee Draft.

Montogomery G. E. and Schuch H. C., 1993, *GIS Data Conversion Handbook*. GIS World.

Moody, D. L., 1998, "Metrics for Evaluating the Quality of Entity Relationship Models", *Paper presented at a Proceedings of the Seventeenth International Conference on Conceptual Modelling(ER '98)*, Singapore.

Veregin H., 1998, "Data Quality Measurement and Assessment", *NCGIA Core Curriculum in GIScience*.(http://www.ncgia.ucsb.edu/giscc/units/u100/u100_f.html)

SDTS, 1992, *The Spatial Data Transfer Standard(FIPS-173)*.

Wang R. Y., Ziad M. and Lee Y. W., 2001. *Data Quality*. London : Kluwer Academic Publishers.