

포인트와 라인 세그먼트의 해석적 에러 검증을 위한 확률기반 시뮬레이션 모델에 관한 연구

A Study on Stochastic Simulation Models to Internally Validate Analytical Error of a Point and a Line Segment

홍 성 철 * 주 용 진**
Sung Chul Hong Yong Jin Joo

요약 해석적 또는 시뮬레이션 오차 모델은 공간 데이터가 가지는 위치오차의 분포를 설명 하는데 유용하다. 그러나 두 오차 모델은 위치오차를 모델링을 하기위하여 다른 접근 방법을 이용하므로 정의된 조건 내에서 올바른 위치오차를 예측 하는지 확인하는 내적 검증을 필요로 한다. 이에 본 논문은 오차 타원과 에러밴드 모델을 이용하여 제시한 포인트와 라인 세그먼트 시뮬레이션 오차 모델을 내부적으로 검증하는 방법을 제안하였다. 시뮬레이션 오차 모델은 분산-공분산 행렬(variance-covariance matrix)의 변수에 의해 규정된 확률분포에 따라 몬테카를로 시뮬레이션을 이용하여 위치오차들을 생성한다. 검증 절차에서는 시뮬레이션 모델에 의한 위치오차의 집합을 해석적 오차 모델에 의한 이론적 위치오차와 비교하였다. 결과적으로 제안된 시뮬레이션 오차 모델은 정의된 위치오차에 따라 동일한 공간 데이터의 위치적 불확실성을 실현함을 확인할 수 있었다.

키워드 : 해석적 오차 모델, 시뮬레이션 오차 모델, 내적검증기법

Abstract Analytical and simulation error models have the ability to describe (or realize) error-corrupted versions of spatial data. But the different approaches for modeling positional errors require an internal validation that ascertains whether the analytical and simulation error models predict correct positional errors in a defined set of conditions. This paper presents stochastic simulation models of a point and a line segment to be validated with analytical error models, which are an error ellipse and an error band model, respectively. The simulation error models populate positional errors by the Monte Carlo simulation, according to an assumed error distribution prescribed by given parameters of a variance-covariance matrix. In the validation process, a set of positional errors by the simulation models is compared to a theoretical description by the analytical error models. Results show that the proposed simulation models realize positional uncertainties of the same spatial data according to a defined level of positional quality.

Keywords : Analytical error models, Simulation error models, Internal validation

1. 서론

자원 관리, 도시 계획, 교통량 분석 등 여러 분야에서 지리정보시스템(GIS)이 활용되면서 공간 데이터를 신속·정확하게 취득하기 위한 다양한 연구가 이루어지고 있다[12, 16]. 하지만 GIS는 활용목적에

따라 다양한 형태와 축척을 가지는 공간데이터를 이용하며, 정확도가 의심스러운 공간데이터를 이용할 경우 공간 문제를 해결하기 보다는 신뢰성이 낮은 결과를 생산함으로써 의사결정 과정에 중대한 오류를 초래한다. 따라서 공간데이터의 품질은 GIS를 활용하기 전에 반드시 고려해야 할 요소 중의 하

[†]This research was supported by a grant(11 High-tech Urban G11) from High-tech Urban Development Program funded by Ministry of Land, Transport and Maritime Affairs of Korean government.

* Sungchul Hong, Ph.D. Civil & Environmental Engineering, Yonsei University. sungchul@hotmail.com (Corresponding Author)

** Yong Jin Joo, Assistant Professor, Aerial GeoInformatics, Inha Technical College. jjj@inhac.ac.kr

나이며, 공간 데이터의 품질을 결정하는 위치오차와 위치오차의 전파과정을 이해하기 위해서는 에러 모델에 대한 연구가 필요하다.

공간 데이터 품질은 활용 목적에 따른 『사용 적합성』에 의해 정의되며, 공간, 시간, 주제 등 지리적 차원의 종합적인 이해를 요구한다[10]. NCDCDS (National Committee for Digital Cartographic Standards)는 수치지도제작의 표준을 정립하기 위해 공간데이터의 이력, 위치 정확성, 속성 정확성, 논리적 일관성과 완전성 등의 기본 구성 요소를 정의하였다[22].

이 중에서 벡터기반 공간데이터(Vector-based spatial data)의 위치정확도는 지도상의 객체와 현실 세계에서의 객체의 위치 근접도로 정의된다[13]. 위치오차의 측정에는 참조 데이터(Reference data)가 참값(True value)으로 가정된다는 점에서 정확도와 관련이 있지만, 참값은 보통 알 수 없기 때문에 정확도와 위치오차는 종종 다른 의미로써 사용되어왔다. 넓은 의미에서, 오차는 측정값과 참값의 차이로 정의되지만 [18], 객체위치에 대한 참값을 모를 경우 위치 불확실성(Uncertainty)으로 표현된다.

오차모델(Error model)은 공간 데이터의 위치오차의 분포를 수치적 또는 도식적으로 표현하는데 이용되며, GIS에서는 공간데이터가 가지는 위치오차를 공간연산을 통해 결과물로 전파시키는데 사용된다. 그 동안 포인트와 라인 세그먼트의 위치오차를 다루는 오차모델에 관한 연구가 선행되어왔으며, 이는 일반적으로 해석적(Analytical error model) 또는 시뮬레이션(Simulation error model) 오차모델로 구분된다.

분산-공분산 행렬(Variance and covariance matrix) 및 오차전파의 법칙(Law of error propagation)에 의해서 유도된 해석모델은 포인트와 라인 세그먼트의 위치 오차의 분포를 에러 타원(Error ellipse) 또는 에러 밴드모델(Error band model)을 이용하여 표현한다. 반면 시뮬레이션 오차모델은 생성된 위치오차를 동일한 공간데이터에 중첩시킴으로써 위치 오차의 분포를 나타낸다. 두 오차모델은 위치오차의 분포를 표현함에 있어서 상이한 접근 방식을 사용한다. 따라서 해석적 또는 시뮬레이션 오차모델이 동일한 조건에서 올바른 위치 오차를 예측하는지 확인하는 것은 필수적이다.

이에 본 논문의 목적은 포인트와 라인 세그먼트의

위치 오차를 확률적으로 시뮬레이션하는 방법, 즉 시뮬레이션 위치오차 모델을 제시하는 것이다. 이를 위해 해당 해석적 오차모델인 오차타원과 에러 밴드 모델과 비교 분석하는 내적 검증 기법¹⁾(Internal validation)을 제안하였다. 시뮬레이션 기반의 위치오차 모델과 해석적 위치 오차모델은 정의된 분산-공분산 행렬에 따라 위치오차를 생성 또는 표현한다. 따라서 내적 검증을 위한 실험에서는 다양한 형태의 분산-공분산 행렬을 정의하였고 이를 기반으로 생성된 포인트와 라인 세그먼트의 위치 오차들의 분포를 해당 해석적 오차모델인 오차타원과 에러 밴드모델과 비교하였다. 이때 시뮬레이션 오차 모델의 유효성은 생성된 위치 오차의 분포를 해석적 에러 모델과 비교하였을 때, 차이점이 허용 한계 내에 존재할 때 적정하다고 판단하였다. 전체적인 연구 절차와 이와 관련된 내용은 다음 Figure 1과 같다.

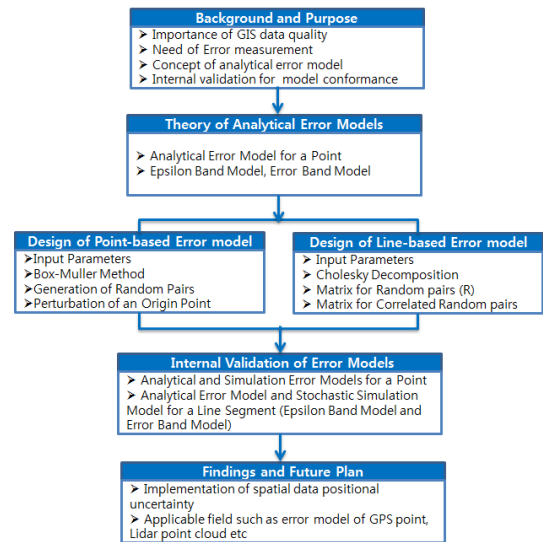


Figure 1. Overview of Study Procedure

2. 해석적 오차모델의 이론적 검토

해석적 오차모델은 오차전파 법칙을 공분산 행렬

- 1) 모델 검증은 크게 내적 검증(Internal validation)와 외적 검증(External validation)로 구분됨. 내적 검증은 이상적인 혹은 이론적인 환경에서 모델의 적정성과 모델 예측 값의 유효성을 평가하는 절차를 말하며, 외적 검증은 내적으로 검증된 모델을 실 데이터 또는 외부데이터에 적용함으로써 모델의 적정성과 예측 값의 유효성을 평가 하는 절차를 의미함

에 적용함으로써 유도된다. 분산 및 공분산 행렬은 공간 데이터를 구성하는 기본 요소인 포인트들의 위치오차와 위치 오차 간의 관계를 정량적으로 설명하는데 이용되고 있다. 포인트에 대한 해석적 에러모델로는 오차타원이, 라인 세그먼트에 대한 에러모델로는 엡실론 밴드 모델과 에러밴드 모델이 있다.

2.1 포인트 기반 오차모델

측정 기술로부터 취득된 포인트 객체는 일반적으로 2차원 혹은 3차원 좌표 시스템을 이용하여 표현된다. 2차원 좌표 상에 표시될 경우, 포인트의 위치는 x와 y 방향에서 오차의 영향을 받는다. 포인트 에러 모델은 측량, 측지학, 그리고 사진 측량의 분야에서 오랜 기간 동안 연구되었으며, 이변량 정규 분포(Bivariate normal distribution)에 기초한 오차 타원이 포인트 측정에 대한 전형적인 오차모델로 사용되고 있다[28].

오차 타원은 측정된 지점을 중심으로 X와 Y 축을 따르는 위치 오차 분포를 설명하며, 동일 객체를 같은 측정기술과 같은 조건하에서 측정하였을 때 동일한 위치에서 취득할 수 있는 확률을 표현한다. 오차 타원은 분산-공분산 행렬의 매개 변수에 따라 다양한 모양, 크기 및 방향을 가진다. 두 임의 벡터 사이에 강한 상관관계가 있을 때 오차 타원은 좁고 긴 모양이 된다. 반대로 두 분산이 상관관계가 없는 경우 오차 타원은 원형이 되며, 이는 포인트 관측 오차의 x와 y 방향 성분이 동일한 정확도를 가지고 독립적이라는 의미를 가진다.

2.2 라인 세그먼트 기반 오차모델

2.2.1 엡실론 밴드 모델(Epsilon Band Model)

라인에 대한 가장 일반적인 오차모델은 라인길이 계산 및 일반화(Generalization)를 위해서 만들어진 엡실론 밴드 모델이며[21, 22, 23], GIS에서는 불확실성 및 정확도 평가 그리고 일반화에 활용되고 있다 [9, 11, 4, 7, 15, 26]. 엡실론 밴드 모델은 시작점과 종점을 포함하는 라인으로부터 일정한 거리("ε")를 가지는 버퍼를 생성함으로써 위치 오차를 표현한다. 하지만 엡실론 밴드 모델을 이용한 위치오차 표현은 직관적이며 통계적으로 결정되지 않는다. Caspary[8]과 Dutton [14]은 라인 세그먼트의 양 끝점이 독립적이고 같은 분산을 가지고 있다는 가정하에서 시뮬레이션을 하였을 때, 엡실론 밴드 모델

의 위치오차가 시뮬레이션 모델과 일치하지 않음을 확인하였다.

2.2.2 에러밴드모델(Error Band Model)

에러밴드 모델은 정규분포를 따르는 라인세그먼트의 두 끝점의(end points) 불확실성에 의해서 결정된다. 엡실론 밴드 모델과는 달리, 에러 밴드 모델은 두 끝점에서 최대 폭을 가지며 라인 세그먼트의 중간 지점에서 최소 폭을 가진다. Shi [26]는 공분산 행렬(Covariance matrix)에 오차전과 법칙을 적용하여 에러 밴드 모델을 유도하였으며, Alesheikh[1, 2] 등은 두 끝점의 상관관계(Correlation)를 설명하는 에러밴드모델을 유도하였다. 에러밴드모델은 라인 세그먼트를 선형함수를 따르는 무한한 점들의 집합으로 간주하며 공분산 행렬을 이용하여 각 끝점이 가지는 위치오차와 위치오차의 상관관계를 설명한다. 따라서 에러밴드는 오차전과법칙을 공분산 행렬과 선형함수에 적용하여 생성된 라인세그먼트 점들의 오차타원을 통합함으로써 라인 세그먼트가 가지는 위치 불확실성을 표현한다.

특히, Leung[17] 등은 공분산 행렬기반의 분석식에 의해 유도되는 에러 밴드를 제안하였으며 엡실론 밴드 모델(epsilon band model)과 여러 종류의 에러 밴드 모델을 표현하였다. 또한 최대 허용 한도(Maximal allowable limit)개념을 이용하여 위상위반(Topology violation)을 피하였고 오차 분석의 일관성을 유지하였다. 이 모델은 끝점이 강한 상관관계를 가질 때 엡실론 밴드 모델이 되고, 그들 사이에 상관관계가 약할 때 에러 밴드 모델이 됨을 보여준다.

3. 확률 시뮬레이션 오차모델

시뮬레이션 오차모델은 다수의 위치 오차를 생성함으로써 공간데이터가 가지는 공간적 변화량, 즉 위치 불확실성을 표현한다. 특히, 몬테카를로 시뮬레이션을 이용한 위치오차는 확률모델을 기반으로 생성되므로 통계적으로 평가되고 추론될 수 있고 해석적 오차모델과 비교 분석될 수 있다[3].

3.1 포인트 기반 오차 모델 설계

포인트 객체의 위치 오차는 x, y 좌표의 관측 값과 관측 값의 표준 편차 및 상관 계수를 따르는 우연 오차를 생성함으로써 정의될 수 있다[6]. 수식 1에서 U_1 과 U_2 는(0, 1) 구간을 가지는 균등분포(Uniform

Distribution)에서 발생한 독립적인 난수(Random variable)이며, x' 와 y' 는 평균 0과 단위 분산을 가진 정규 분포로부터 발생한 독립적인 임의 변수의 쌍이다[5].

$$\begin{aligned} x'' &= \sqrt{-2 \ln U_1} \cos 2\pi U_2 \\ y'' &= \sqrt{-2 \ln U_1} \sin 2\pi U_2 \end{aligned} \quad (1)$$

정규분포의 표준 편차($\sigma x'$, $\sigma y'$)를 가지는 난수(x' , y')는 수식(2)에 의해 생성된다.

$$\begin{aligned} x' &= \sigma_{x'} x'' \\ y' &= \sigma_{y'} y'' \end{aligned} \quad (2)$$

정의된 표준 편차와 상관 계수를 가진 난수들의 분포는 기본적으로 오차 타원의 장축(X'), 단축(Y')과 일치하기 때문에, X, Y 좌표 상에서 정의된 σx 와 σy 는 좌표 변환 방정식(3)과 (4)를 사용하여 X'-Y'좌표 공간의 $\sigma x'$ 와 $\sigma y'$ 로 좌표변환 돼야 한다.

$$\begin{aligned} \sigma_{x'}^2 &= \sigma_x^2 \cos^2 \theta + 2\gamma \sigma_x \sigma_y \cos \theta \sin \theta + \sigma_y^2 \sin^2 \theta \\ \sigma_{y'}^2 &= \sigma_y^2 \sin^2 \theta - 2\gamma \sigma_x \sigma_y \cos \theta \sin \theta + \sigma_x^2 \cos^2 \theta \end{aligned} \quad (3)$$

$$\theta = \frac{1}{2} \arctan \left[\frac{2\gamma \sigma_x \sigma_y}{\sigma_x^2 - \sigma_y^2} \right] \quad (4)$$

변환 된 표준편차 $\sigma_{x'}$ 와 $\sigma_{y'}$ 는 식(2)에서 Box-Muller 변환법에 의해서 생성된 난수(x'' , y'')의 스케일 인자(Scale factor)로 사용된다. x' 와 y' 는 수식(5)를 이용하여 다시 x 와 y 로 좌표 상으로 좌표변환 된 됴므로써 관측 값(x , y)의 위치오차가 된다.

$$\begin{aligned} x &= x' \cos \theta - y' \sin \theta \\ y &= x' \sin \theta + y' \cos \theta \end{aligned} \quad (5)$$

포인트 객체의 위치오차를 시뮬레이션하는 절차는 Figure 2와 같다. 독립적이고 정규분포를 따르는 임의의 난수(x' , y')는 평균값이 0이라는 전제하에 생성되며, 이러한 난수(x' , y')들의 집합은 오차 타원의 장·단반경과 일치되는 주축(x' - y')을 따른다. 따라서 표준 편차($\sigma_{x'}$, $\sigma_{y'}$)와 상관계수(γ)를 따르는 난수(x , y)를 얻기 위해서는 좌표변환이 필요하다. 위치 오차는 좌표 변환 된 난수들을 원 좌표 값에 추가함으로써 생성된다.

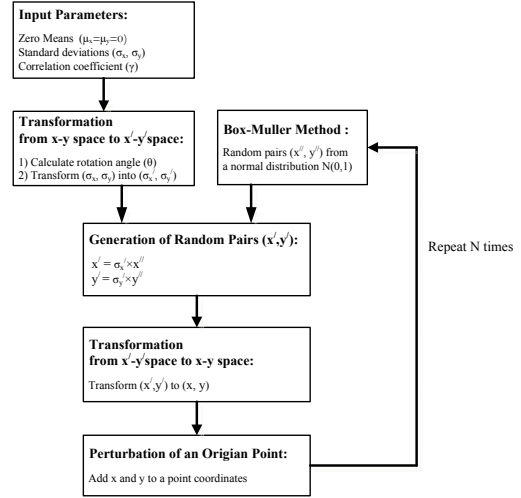


Figure 2. General Steps to Simulate Positional Errors in a Point

3.2 라인기반 오차 모델 설계

라인 세그먼트의 위치오차 시뮬레이션 모델은 Cholesky 분해를 이용하여 끝점들의 위치오차를 생성한다. 몬테카를로 시뮬레이션에서 Cholesky 분해는 상관관계 행렬(Correlation matrix) 'C'에서 정의된 포인트들의 상관관계를 만족하는 난수를 생성하는데 이용된다.

$$C = \begin{bmatrix} 1 & \rho_{x_1 y_1} & \rho_{x_1 x_2} & \rho_{x_1 y_2} \\ \rho_{x_1 y_1} & 1 & \rho_{y_1 x_2} & \rho_{y_1 y_2} \\ \rho_{x_1 x_2} & \rho_{y_1 x_2} & 1 & \rho_{x_2 y_2} \\ \rho_{x_1 y_2} & \rho_{y_1 y_2} & \rho_{x_2 y_2} & 1 \end{bmatrix} \quad (6)$$

상관관계 행렬 'C'에서 첨자 1과 2는 각각 라인 세그먼트의 첫 번째와 두 번째 포인트를 의미한다. 예를 들어, $\rho_{x_1 y_2}$ 는 첫 번째 포인트의 x 좌표와 두 번째 포인트의 y좌표의 상관관계를 나타낸다. 상관관계가 있는 난수들은 공간적 상관관계를 가지는 위치오차로 간주되며 라인 세그먼트 끝점들의 좌표에 추가함으로써 시뮬레이션 될 수 있다.

수식(7)에서 Cholesky 분해는 상관 행렬 "C"를 하삼각행렬(Lower triangular matrix) "L"과 하삼각행렬의 전치행렬(Conjugate transpose) "L^T"로 분해한다. 상관관계 있는 난수 들은 난수 행렬(Random number matrix) "R"과 하삼각행렬 "L"을 곱함으로써 생성 될 수 있다.

$$C=LL^T$$

$$R_{cor}=LR$$

C : Correlation Matrix (7)

L : Lower Triangular Matrix

R : Random Number Matrix

R_{cor} : Correlated Random Number Matrix

그러나 Cholesky 분해에서 상관관계 행렬은 대칭적 정치행렬(Symmetric positive-definite matrix)이어야 하며 $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ 과 $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ 과 같은 상관관계 행렬을 만족시키는 난수를 생성하지 못한다. 이러한 한계를 극복하기 위해, 고유 분해 기법(Eigen decomposition)이 다음과 같이 사용된다[27].

$$A=Q\Lambda Q^T$$

A : Real Symmetric Matrix (8)

Q : Eigenvector Orthogonal Matrix

Λ : Eigenvalue Diagonal Matrix

수식(8)은 수식(9)로 대체될 수 있다.

$$A=BB^T \quad \text{where} \quad B=Q\Lambda^{1/2} \quad (9)$$

“B”행렬은 수식(7)의 하삼각행렬 “L”에 해당하므로, 상관관계가 있는 난수는 수식(10)의 Cholesky 분해의 대체 형식을 이용하여 생성 될 수 있다.

$$R_{cor}=BR \quad (10)$$

수식(9)에서, 고유치(Eigen value) 대각 행렬 내의 요소는 양의 준정부호(Semi-definite) 행렬이어야 한다. 시뮬레이션을 위한 절차는 Figure 3과 같다. 공분산 행렬에서 유도된 표준 편차 행렬 “S”와 상관행렬 “C”는 라인 세그먼트의 위치오차 분포의 특성을 정의하는 입력변수이다. 상관 행렬 “C”는 수식 9에 의해서 분해된다. 행렬 “R”의 독립적인 난수는 Box-Muller 변환법으로 생성된다. 이 난수 행렬은 수식 10의 행렬 B와 함께 상관관계가 있는 난수행렬 “Rc”을 생성하며 표준편차 행렬은 “S”는 행렬 “Rc”의 난수들의 척도인자(Scale factor)로써 사용된다. 시뮬레이션에서 위치오차를 가지는 라인 세그먼트는 생성된 난수들을 라인 세그먼트 포인트들의 해당 X와 Y 좌표에 상관 난수들을 추가함으로써 실현된다.

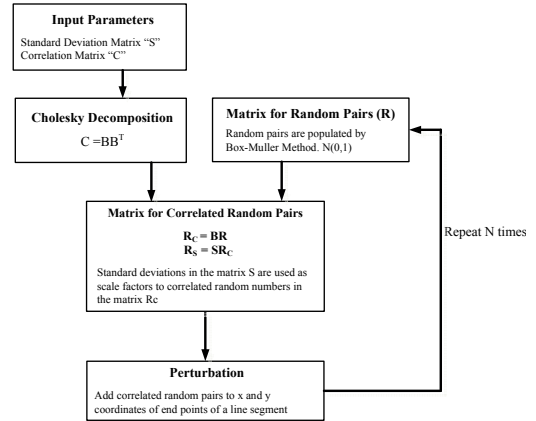


Figure 3. General Steps to Simulate Positional Errors in a Line Segment

4. 오차모델의 내적 검증

내부 검증은 해석적 또는 시뮬레이션 기반의 오차 모델이 주어진 조건에 동일한 위치오차를 생성하는지 확인하는 것을 목적으로 한다. 해석적 모델은 위치오차의 분포를 오타타원 또는 에러밴드를 이용하여 도식적으로 설명하지만, 시뮬레이션 모델은 가정된 오차 분포에 따라 다량의 위치오차를 생성하고 공간데이터 위에 중첩시킴으로써 표현한다. 따라서 해석적 오차 모델의 이론적 위치오차 분포는 시뮬레이션 모델에 의한 위치 오차의 집합과 비교될 수 있다.

시뮬레이션 오차모델은 공분산 행렬과 몬테카를로 시뮬레이션을 이용하여 만족할만한 위치오차의 분포를 얻을 때까지 위치오차를 반복적으로 생성한다(Figure 4). 위치오차의 분포는 원 공간데이터(Original spatial data)와 위치오차가 존재하는 공간데이터(Perturbed spatial data)를 비교함으로써 시각적으로 확인할 수 있다. 시뮬레이션 오차모델은

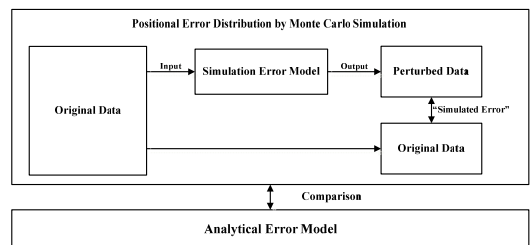


Figure 4. Internal Validation of Analytical and Simulation Error Models

해석적 위치오차 모델과 비교하여 차이점이 허용 한계 내에 존재할 때 내부적으로 검증된다.

4.1 포인트 위치 오차모델의 검증

포인트 위치 오차모델의 내부 검증은 시뮬레이션된 위치오차들의 분포를 오차타원과 비교함으로써 이루어진다(Figure 5).

시뮬레이션 모델은 X와 Y 좌표상의 이변량 정규

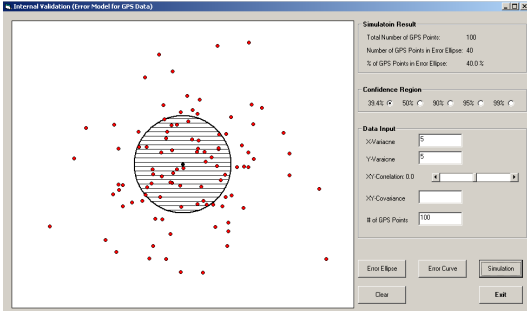


Figure 5. Evaluation of Analytical and Simulation Error Models for a Point

분포를 따르는 위치오차를 생성한다. 따라서 위치오차가 존재하는 포인트 객체들이 오차타원 내에 속하는 확률은 오차타원의 신뢰면적(Confidential Region)과 일치해야 한다. 예를 들어, Figure 5의 오차타원은 표준편차 1, 상관관계 0을 가지며 39.4% 신뢰구역을 가진다. 이는 동일한 특성을 가진 시뮬레이션 모델이 위치 오차를 가진 포인트들을 생성할 때, 포인트들의 39.4%가 오차타원 내에 존재해야 함을 의미한다.

Figure 6은 시뮬레이션 위치오차모델의 내부검증 결과를 보여준다. 좌측 결과는 동일한 분산($\sigma_x^2 = \sigma_y^2 = 1$)을 가지는 이변량 정규 분포에 따라 위치오차가 시뮬레이션되었을 때 결과를 보여주며, 우측 결과는 위치 오차가 다른 분산($\sigma_x^2 = 2, \sigma_y^2 = 3$)을 가진 이변량 정규 분포에 의해서 시뮬레이션되었을 때 결과이다. 이때, 각각의 시뮬레이션 결과에서는 주어진 분산을 유지하면서 상관계수를 -1에서 1까지 변화시켰다. 시뮬레이션 모델과 해석모델과의 내부 검증을 위해 동일한 조건하에서 10회 시뮬레이션한 결과들을 39.4%, 50%, 95% 신뢰영역을 가지는 오차타

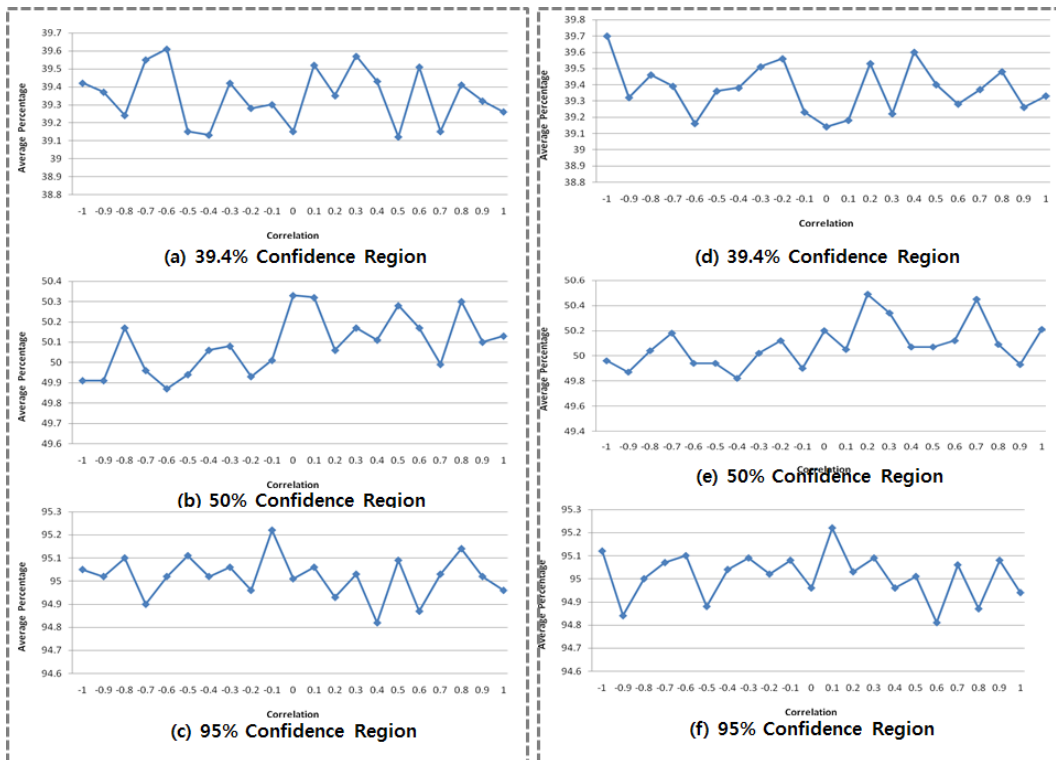


Figure 6. Average Percentage of Points to Fall Within Error Ellipse (The Left-hand : $\sigma_x^2 = \sigma_y^2 = 1$, The Right-hand : $\sigma_x^2 = 2, \sigma_y^2 = 3$)

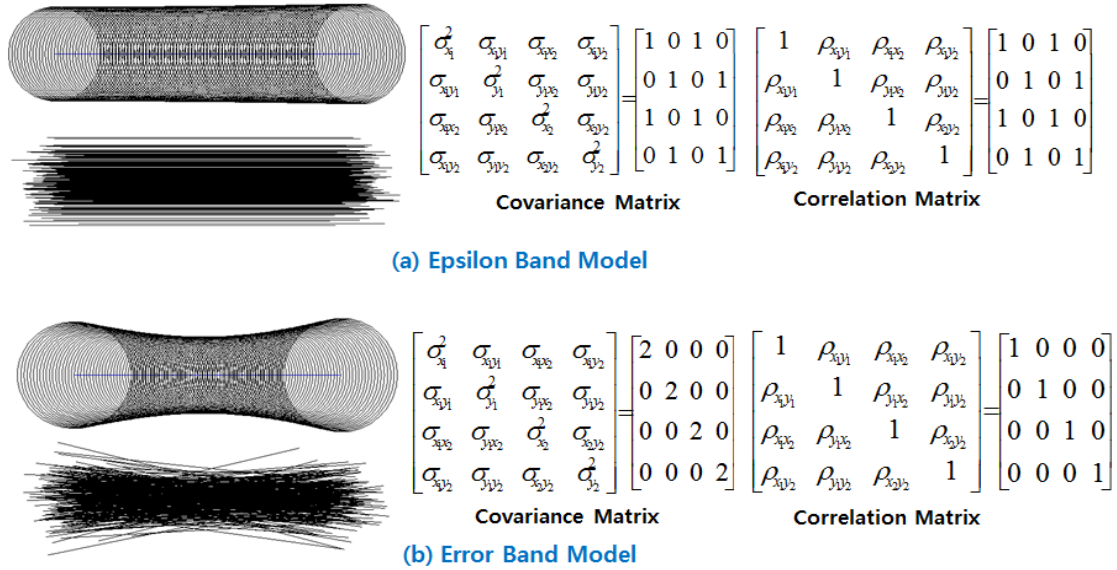


Figure 7. Analytical Error Model and Stochastic Simulation Model for a Line Segment :
Epsilon Band Model and Error Band Model

원과 비교하였으며 각각의 시뮬레이션 마다 10,000의 위치 오차들을 발생시켰다. 오차 타원은 이변량 정규 분포의 단면이므로, 발생된 위치오차들은 동일 오차 타원의 여러 신뢰 구역과 검증하는 것이 필요하다.

시뮬레이션 결과는 평균 비율이 해당 신뢰영역을 기준으로 0.5 % 이내에 근접함을 보여준다. 차트 내에 커브들은 임의적으로 변동되지만, 평균 비율과 신뢰구간 사이의 차이는 0.28 %에서 0.49 %에 불과하다. 이러한 작은 차이는 발생된 위치오차의 분포가 이론적인 오차분포를 반영함을 나타내며 시뮬레이션된 위치 오차가 포인트의 x, y축의 설정된 상관 계수 적 특성을 구현하는 것을 보여준다.

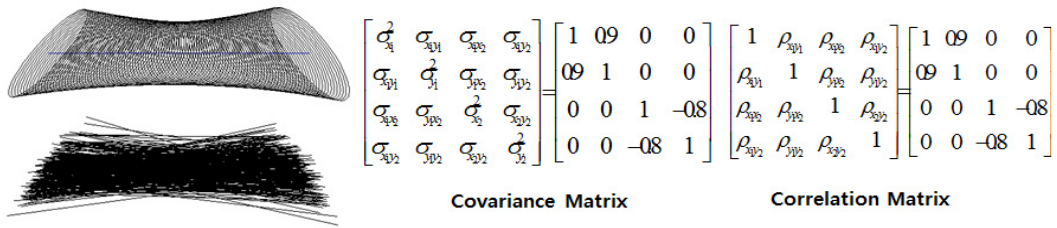
4.2 라인 세그먼트 오차 모델의 검증

라인 세그먼트의 시뮬레이션 위치오차 모델의 유효성 검사를 위해 Leung 등 [17]이 제안한 에러밴드 모델을 채택하였다. 시뮬레이션 오차 모델과 해석적 오차 모델은 동일한 분산-공분산 행렬을 기반으로 생성되며 300회 시뮬레이션 된 위치 오차의 분포를 99%의 신뢰영역을 가진 해당 에러밴드 모델과 비교하였다. 다음 Figure 7은 시뮬레이션 된 위치오차의 분포를 엡실론 밴드 모델과 에러밴드 모델과 비교한 결과를 보여준다.

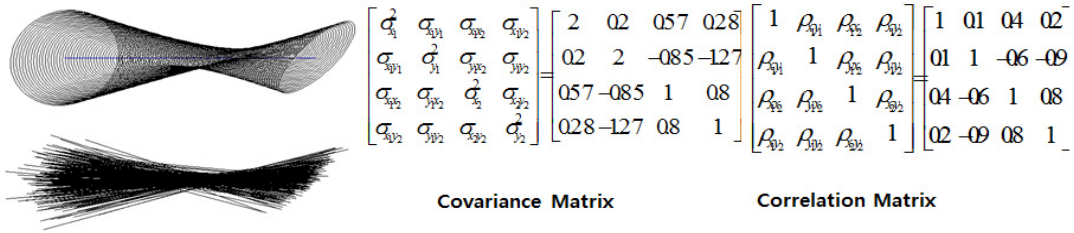
Figure 7(a)에서, 공분산 및 상관관계 행렬은 라인 세그먼트의 두 끝점이 동일한 분산을 가지고 강한 상관관계가 있음을 나타낸다. 한 끝점의 X와 Y 좌표는 서로 독립적이지만 다른 끝점의 동일 좌표와는 높은 상관관계가 있다. 따라서 시뮬레이션 오차모델은 동일한 방향과 거리를 가지는 위치오차들을 생성하므로 시뮬레이션 된 라인 세그먼트들은 원 라인세그먼트와 평행하다. 또한 시뮬레이션 오차모델은 정규분포를 기반으로 위치오차를 생성하므로 시뮬레이션 된 세그먼트들은 원 세그먼트에 가까울수록 높은 빈도수를 가지며 멀어질수록 빈도수는 낮아진다.

Figure 7(b)에서, 상관관계 행렬은 각 끝점의 X, Y 좌표들이 서로 독립적이고 상관관계가 없음을 나타낸다. 따라서 해석 위치 오차모델인 에러밴드 모델과 같이 시뮬레이션 결과는 중간에 최소 폭과 끝 지점에서 최대 너비를 가진다.

Figure 8에서는 시뮬레이션 위치오차 모델을 두 가지 유형의 에러 밴드 모델과 비교하였다. Figure 8(a)에서, 상관관계 행렬은 두 끝점이 각각 독립적이라는 점에서 Figure 7(b)의 에러 밴드 모델과 같지만 끝점의 X, Y 좌표는 각각 0.9와 -0.6의 정부 상관 관계를 가진다. 따라서 시뮬레이션 된 라인 세그먼트 집합은 Figure 7(b)의 결과처럼 중간에 최소 폭을 가지지만, 각 끝점의 위치오차 분포는 반대 방향



(a) Advanced Error Band Model with Independent End Points of a Line Segment



(b) Advanced Error Band Model with Correlated End Points of a Line Segment

Figure 8. Analytical Error Model and Stochastic Simulation Model for a Line Segment :
Advanced Error Band Model for Different Cases

으로 편향되어 있다.

Figure 8(b)에서, 분산 및 상관행렬은 라인세그먼트 끝점의 위치오차가 다른 분산을 가지지만 완전한 상관관계를 가지는 것을 보여준다. 특히, 왼쪽 끝점의 분산이 오른쪽 끝점의 분산보다 크며, 왼쪽 끝점의 y 좌표는 오른쪽 끝점의 x와 y 좌표들과 부의 상관관계가 있다. 따라서 시뮬레이션에서는 부의 상관관계로 인해 끝 점의 위치오차는 서로 반대 방향으로 생성되며, 시뮬레이션된 라인 세그먼트의 왼쪽이 위치 불확실성이 오른쪽보다 크다. 결과적으로 시뮬레이션 위치오차 모델은 공분산 행렬에 규정된 조건에 부합하는 라인 세그먼트에 위치 불확실성을 실현함을 알 수 있고, 원 라인 세그먼트에 부과된 변위의 집합은 에러 밴드 모델에 의해 예측된 오차 분포를 따름을 확인할 수 있다.

5. 결론 및 시사점 도출

본 논문에서는 정의된 확률 분포 및 특성에 따라 위치 오차를 생성하는 시뮬레이션 기반의 위치오차 모델 들을 소개하고 오차전파 법칙에 의해서 유도되는 해석적 위치오차 모델과 내부 검증을 위한 방법론을 제시하였다. 즉, 내부 검증을 위해 포인트와 라

인 세그먼트의 위치 오차들의 분포는 해당 해석적 위치오차 모델인 오차타원과 에러밴드 모델과 비교하였다. 결과적으로 각각의 시뮬레이션 모델이 분산-공분산 행렬에서 규정된 조건에 부합하는 동일한 공간 데이터의 위치 불확실성을 구현함을 보여주었다. 특히 라인세그먼트의 위치오차 모델은 구현이 간단하므로 여러 개의 포인트를 가지는 라인에 쉽게 적용 가능하다.

본 논문에서는 제시한 시뮬레이션 기반의 위치 오차모델은, 포인트 모델의 경우 GPS 포인트 또는 라이다의 측량결과인 포인트 클라우드가 가지는 위치 오차를 구현할 때 응용되어 쓰일 수 있다. 또한, 라인세그먼트의 위치오차 모델은 선형을 가지는 도로 망도 등의 위치오차를 구현할 때 응용될 수 있다. 하지만 GIS에서는 공간 데이터를 생성하기위해서 다양한 측량 방법에 쓰이고 있으며, 공간 데이터의 위치 오차는 측량 방법에 의해서 다른 형태와 특성을 가진다. 또한 이러한 위치 오차들은 공간 연산에 의해서 변형되어 GIS의 결과물까지 전파된다. 따라서 향후 연구에서는 측량 기법을 고려한 다양한 위치오차 모델은 물론 공간 연산을 고려한 위치오차전파 모델의 개발이 필요할 것이다.

References

- [1] Alesheikh, A. A; Blais, J. A. R; Chapman, M. A; Karimi, H. 1999, Rigorous Geospatial Data Uncertainty Models for GISs., Spatial accuracy assessment: Land information uncertainty in natural resources., K. Lowell and A. Jaton, eds., Ann Arbor Press, Chelsea, Michigan, 195-202.
- [2] Alesheikh, A. A; Li, R. 1996, Rigorous Uncertainty Models of Line and Polygon Objects in GIS, Proceedings of GIS/LIS'96, Denver, CO, 906-920.
- [3] Ang, A. H; Tang, W. H. 1975, Probability Concepts in Engineering Planning and Design, John Wiley & Sons, New York, Vol. 2.
- [4] Blakemore, M. 1984. Generalization and error in spatial databases. *Cartographica*, 21(2):131-139.
- [5] Box, G. E. P; Muller, M. E. 1958, A Note on the Generation of Random Normal Deviates, *The Annals of Mathematical Statistics*, 29(2): 610-611.
- [6] Campbell, C. W. 1983, Efficient algorithm for generating random number pairs drawn from a bivariate normal distribution, NTIS, SPRINGFIELD, VA(USA), 1983, 17.
- [7] Carver, S. 1991, Adding error handling functionality to the GIS toolkit, *Proceedings of EGIS*, 9(1), 1.
- [8] Caspary, W; Scheuring, R. 1992, Error-Bands as Measures of Geometrical Accuracy, *EGIS*, 92: 226-233.
- [9] Chrisman, N. 1982, Theory of cartographic error and its measurement in digital data bases, 5 th Int. Symp. Comp. Assisted Cartography & Int. Soc. Photogrammetry and Remote Sensing Commission IV.
- [10] Chrisman, N. 1984, The role of quality information in the long-term functioning of a geographic information system, *Cartographica*, 21(2/3):79-87.
- [11] Chrisman, N. R. 1983, Epsilon filtering: A technique for automated scale changing, *Proceedings of the Annual Conference of the American Congress for Surveying and Mapping*, 322-331.
- [12] Cho, H. K; Wie, G. J; Cho, Y. S; Lee, S. J. 2011, Accuracy Assessment Geoposition of Airborne Line-Scanner Image, *Journal of Korea Spatial Information Society*, 19(1):51-59.
- [13] Drummond, J. 1995, Positional accuracy, *Elements of spatial data quality*, S. C. Guptaill and J. L. Morrison, eds., Oxford, Elsevier, 31-58.
- [14] Dutton, G. 1992. Handling positional uncertainty in spatial databases, *Proceedings of 5th International Symposium on Spatial Data Handling*, 460-469.
- [15] Goodchild, M. F; Hunter, G. J. 1997, A simple positional accuracy measure for linear features, *International Journal of Geographical Information Science*, 11(3):299-306.
- [16] Lee, J. H; Choi, K. A; Lee, I. P. 2010, Development of a Portable Spatial Information Acquisition System, Paper presented at the Conference of Geo-Spatial Information, 296-297.
- [17] Leung, Y; Ma, J. H; Goodchild, M. F. 2004, A general framework for error analysis in measurement-based GIS Part 1: The basic measurement-error model and related concepts, *Journal of Geographical Systems*, 6(4):325-354.
- [18] Lo, C. P; Yeung, A. K. W. 2002, *Concepts and Techniques of Geographic Information Systems*, Prentice Hall Upper Saddle River, NJ.
- [19] Mikhail, E. M; Gracie, G. 1981, *Analysis and adjustment of survey measurements*, Van Nostrand Reinhold.
- [20] NCDCCS. 1988, The proposed standard for digital cartographic data, *The American Cartographer(Special Issue)*, 15(1):58-62.
- [21] Perkal, J. 1956. On the e-length." *Bull De i' Academie Sc. Polonaise des Sciences*, 4: 399-403.
- [22] Perkal, J. 1966a, An Attempt at Objective Generalization, *Michigan Inter-University Community of Mathematical Geographers*, Discussion Paper, 10.
- [23] Perkal, J. 1966b, On the length of empirical

- curves." Discussion Paper 10, Michigan Inter-University Community of Mathematical Geographers, Community of Mathematical Geographers.
- [24] Shi, W. 1994, Modelling Positional and Thematic Uncertainties in Integration of Remote Sensing and Geographic Information Systems, Ph.D. Thesis, ITC Publication, No. 22., The Netherlands.
- [25] Shi, W; Liu, W. 2000, A stochastic process-based model for the positional error of line segments in GIS, International Journal of Geographical Information Science, 14(1):51-66.
- [26] Tveite, H; Langaas, S. 1999, An accuracy assessment method for geographical line data sets based on buffering, International Journal of Geographical Information Science, 13(1):27-47.
- [27] Watkins, D. S. 2002, Fundamentals of matrix computations, John Wiley and Sons.
- [28] Wolf, P. R; Ghilani, C. D. 1997, Adjustment Computations: statistics and least squares in surveying and GIS, Wiley, New York.

논문접수 : 2013.03.06
수정일 : 2013.04.22
심사완료 : 2013.04.25