大 韓 土 木 學 會 論 文 集
第30卷 第6B號 · 2010年 11月
pp. 525 ~ 532

일강우량의 적정 베리오그램

On Proper Variograms of Daily Rainfall Data

박민규* · 박창열** · 신기일*** · 유철상****
Park, Minkyu · Park, Changyeol · Shin, Key-Il · Yoo, Chulsang

Abstract

Kriging is widely applied to dealing with the spatial distribution of rainfall, however its prediction results are different according to the selection of variogram type. This study investigated adequate variogram for daily rainfall. The comparative results show that kriging prediction with covariates is better than that without covariates. The Matérn correlation function, which is the most general type variogram, is recommended if adequate variogram is difficult to determine.

Keywords : *variogram, kriging, geostatistical statistics, Matérn correlation function*

요 지

크리깅은 강우현상의 공간적 분포문제를 다루는데 가장 널리 사용되는 방법이나 선택한 베리오그램에 따라 예측결과가 상당히 큰 차이를 보여준다. 본 연구에서는 이러한 이유로 일강우량을 대상으로 적정 베리오그램을 고찰하였다. 그 결과 일강우량자료만을 이용하는 것보다 강우발생과 상관성이 높은 이차변수를 함께 고려하여 베리오그램을 유도하는 것이 보다 적정한 결과를 주는 것으로 나타났다. 한편, 한정적인 적용사례이나 크리깅 적용에 필요한 적정 베리오그램의 결정이 어려운 경우 가장 일반적인 형태의 베리오그램인 Matérn 상관함수를 사용하면 상대적으로 양호한 결과를 얻는 것으로 나타났다.

핵심용어 : 베리오그램, 크리깅, 지구통계학, Matérn 상관함수

1. 서 론

강우현상의 공간적 변동성에 대한 해석은 수자원 계획 및 관리의 중요한 관심사이며 이러한 강우량의 공간변동 해석 결과는 한정된 범위의 관측자료로부터 유역평균강수량을 추정하거나 결측치를 보완하는 방법으로 사용되어 왔다(Tabios and Salas, 1985). 국내 수자원 분야에서는 이러한 경우에 주로 티센법이나 역거리가중법의 방법이 이용되어 왔다(윤용남, 2007). 그러나 이러한 방법들은 단순히 면적비나 거리를 인자로 적용하기 때문에 강우량자료가 보이는 공간변동의 통계적인 특성과 무관한 결과를 얻을 수 있다. 이에 대한 대안으로 사용되는 것이 지구통계학적 분석방법(Geostatistical Analysis)이다. 이는 크리깅(Kringing) 이론으로 주로 불리워지고 있는 방법을 의미하며 현재 많은 분야에서 폭넓게 활용되고 있다. 국내 수자원 분야에서 이 기법은 강우 관측망 최적설계(이재형과 윤양규, 2002), 강우의 공간분포 해석(윤강훈 등, 2004), 면적평균강우량의 추정(유철상과 정광식, 2001), 강우의 공간 분포에 따른 유출 분석(윤용남 등, 2002) 등에 활용되었다. 이 외에도 기상예보 등에서 매일 접하게 되는 등우선도의 작성 등이 이 방법에 기초하여 이루어지고 있다.

크리깅 이론은 Krige(1951)에 의해 처음 소개되었고 Matheron(1971)에 의해 발전한 기법이다. 이 방법은 기지점의 값들에 일정한 가중함수를 적용하여 얻은 선형결합의 형태로 임의 지점에서의 값을 추정하는 방법이라고 할 수 있다. 이때 가중함수는 오차분산이 최소화되도록 한다는 전제하에 베리오그램과 자료간의 거리값으로 일종의 공식처럼 얻을 수 있다. 이 경우 가장 중요한 역할을 하는 것이 베리오그램(variogram)으로 공간데이터의 상관관계를 나타내는 척도로서 알려져 있다. 일반적으로 관측자료로에서 얻은 경험적 베리오그램에 이론적 베리오그램이라고 불리워지는 일정한 모형을 적합시켜 사용하고 있으며, 이론적 베리오그램에는 지수모형(exponential model), 구형모형(spherical model) 그리고 가우시안 모형(Gaussian model) 등이 주로 사용되고 있다. 어떤 이론적 베리오그램 모형을 사용하느냐에 따라 동일한 관측값에 대해서도 크리깅의 적용결과가 달라지기 때

*정회원 · 고려대학교 공과대학 건축/사회환경공학부 연구교수 (E-mail : mkhoin@korea.ac.kr)
**고려대학교 공과대학 건축/사회환경공학부 박/사과정 (E-mail : changyeol@korea.ac.kr)
***한국외국어대학교 자연과학대학 정보통계학과 교수 (E-mail : keyshin@hufs.ac.kr)
****정회원 · 교신저자 · 고려대학교 공과대학 건축/사회환경공학부 교수 (E-mail : envchul@korea.ac.kr)

문에 적정 베리오그램을 판단하는 문제는 가장 중요한 부분이라고 할 수 있다(신기일, 2008, 최종근, 2007). 따라서 공간분포에 관한 연구는 주어진 자료에 대한 베리오그램이 어떤 특징을 보이는지에 대하여 이루어지고 있다. 그러나 강우량에 대해서는 적정한 베리오그램이 어떤 형태인지에 관한 연구가 드물게 진행되었고 어떤 베리오그램 모형을 사용하는 것이 좋은지에 관한 연구 역시 수행되지 않았다. 한편, Goovaerts(2000)는 강우량만을 이용하여 유도된 베리오그램을 이용하는 것보다 이차변수로 표고자료를 함께 고려해 유도한 베리오그램을 이용하는 경우에 크리깅의 적용 결과가 향상될 수 있음을 보여주었다.

본 연구에서는 등우선도 작성이나 공간적인 보간을 위한 목적으로 일강우량 자료에 크리깅을 적용할 때 필요한 적정 베리오그램에 대해 검토하였다. 기상청 기상대 자료를 이용해 베리오그램을 유도하여 기상청 AWS 지점의 강우량을 예측하였다. 다양한 이론적 베리오그램에서의 강우량 예측값과 AWS 지점에서의 실제 관측값과 비교하여 베리오그램의 적정성을 판단하였다. 이때 강수량 자료만을 이용하는 경우와 강우량과 관련된 기상인자를 함께 고려하여 추정하는 경우를 비교하여 강우의 공간분포를 예측하는 방법의 정확도를 개선하는 방법에 대해 고찰하였다.

2. 배경이론

2.1 베리오그램의 개요

베리오그램은 공간분포된 자료의 거리에 따른 상관관계를 해석하기 위한 통계적 분석 방법이다. 신기일(2008)이 정리한 베리오그램에 관한 결과를 살펴보면 다음과 같다. 일반적으로 통계학에서 사용되는 공분산함수(covariance function, autocovariance) 또는 자기상관함수(autocorrelation function)를 공간자료에도 적용할 수 있으나 공간분포의 해석에서는 상관함수보다는 아래의 식과 같은 분산의 정도를 나타내는 베리오그램이 더욱 유용하게 사용되는 경우 많다.

$$2\gamma(\hat{h}) = \frac{1}{N}\sum_{i=1}^{N_h}(z(U_i)-z(U_i+h))^2 \qquad (1)$$

여기서 $\gamma(\hat{h})$는 거리와 각도를 나타내는 벡터 h에 따른 세미베리오그램(semivariogram)으로 베리오그램 $2\gamma(\hat{h})$의 절반에 해당한다. 일반적으로 세미베리오그램과 베리오그램은 서로 혼용하여 사용된다. $z(U_i)$는 특정위치 U_i에서의 자료값이며 $z(U_i+h)$는 $z(U_i)$에서 일정한 거리와 각도를 나타내는 벡터 h에 위치한 자료값이다. N_h는 분석하고자 하는 공간영역내에서 벡터 h만큼 떨어져서 위치한 두 자료쌍의 총 수이다. 이때, 벡터 h의 적용에 따라 방향에 상관없이 거리만을 고려하여 해석하는 등방성(isotrophy) 베리오그램과 몇 개의 방향으로 나누어 방향별로 거리를 고려하는 이방성(anisotrophy) 베리오그램의 경우로 구분할 수 있으며 본 연구에서는 강우장(rainfall field)이 등방성이라고 가정하고 거리에 대한 함수로서만 고려하였다.

베리오그램은 앞서 소개한 자기상관함수와는 다르게 두 값의 차이가 크게 나면 값이 커지고 두 값의 차이가 작게 나면 값이 작아지므로 자기상관함수와 선대칭의 관계라고 할

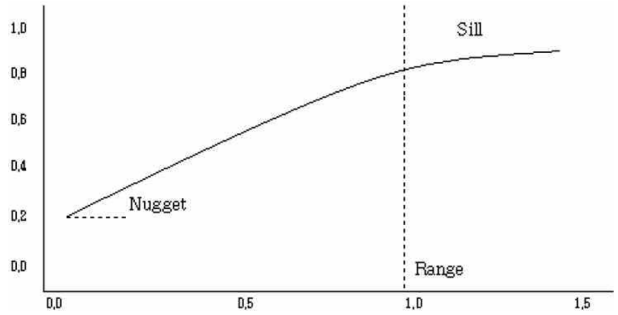


Fig. 1 Typical shape of variogram

수 있다. 따라서 일반적으로 h가 증가할 때 "0"으로 수렴하는 자기상관함수와 달리 베리오그램은 증가하다가 어떤 값에 수렴하는 형태를 가진다. 즉, 자기상관함수가 유사성의 척도(measure of similarity)라면 베리오그램은 비유사성의 척도(measure of dissimilarity)라고 볼 수 있다.

베리오그램의 일반적인 형태는 Fig. 1과 같이 종축의 거리 h가 증가할 때 횡축의 분산의 정도인 세미베리오그램 $\gamma(\hat{h})$가 함께 증가하는 형상을 보인다. 일반적으로 베리오그램은 몇 개의 특징적인 값인 너깃(nugget), 문턱값(sill), 그리고 상관거리(range, correlation length)에 의하여 그 특성과 모양이 결정된다. 너깃은 덩어리라고도 불리워지며 일반적으로 거리 h=0 이면 두 변수의 차의 분산으로 $\gamma(0)=0$ 되어야 한다. 그러나 자료에 따라서는 h가 "0"에 가까워질 때 "0"으로 수렴하지 않는 경우가 있다. 즉 $\gamma(h\to 0)\neq 0$인 경우가 있으며 이러한 경우를 너깃 효과(nugget effect)가 있다고 말하고 이러한 불연속의 크기를 너깃이라 부른다. 너깃의 물리적 의미는 극한이론에서 두 점간의 위치가 완전히 일치하지 않으면서 무한히 가까워지더라도 일정한 값의 차이가 나타나는 것으로 볼 수 있다. 일반적으로는 너깃은 기지점들 사이의 가장 짧은 거리보다 작은 거리에서의 상관관계에 대한 정보가 부족하기 때문에 발생하는 효과라고 볼 수 있으며 크리깅의 적용시 비교적 작은 거리상에서 급격한 값의 변화가 나타날 때 이를 적절하게 설명하지 못하는 원인변수로 알려져 있다. 최근에는 이러한 너깃효과를 모형화하는 방법을 기존의 이론적 베이오그램과 결합하여 크리깅 방법의 정확성을 향상시키는 혼합모형(mixed model)에 관한 연구도 상당한 진척을 이루고 있다(Journel and Huijbregts, 1978, Isaaks and Srivastava, 1989, Cressie, 1991). 문턱값은 공간 자료가 특정한 조건(약한 의미의 정상성 조건)을 만족하게 되면 베리오그램에서 h가 증가해도 분산의 정도가 무한히 증가하지 않고 특정한 값에 수렴하게 되며 이 경우의 가장 큰 분산의 정도의 값을 부르는 용어이다. 만약 약한 의미의 정상성이 만족되지 않으면 문턱값은 존재하지 않고 계속해서 증가하는 경향을 보이게 되며 크리깅을 적용하는데 적정한 조건의 베리오그램이 없는 것으로 간주된다. 상관거리는 구간 또는 상관구간으로도 불리우며 두 지점간의 상관관계가 거의 "0" 이 될 때에 해당되는 거리 h를 의미한다. 베리오그램을 살펴보면 문턱값이 얻어질 때의 거리가 상관거리로 정의될 수 있다. 일반적으로 상관거리보다 큰 거리에 있는 두 자료간에는 상관관계가 존재하지 않는 것으로 본다. 마지막으로 부분적 문턱값(partial sill)은 문턱값과 너깃의 차이이다. 만약 너깃이 "0"이면 부분적 문턱값은 문턱

값과 같은 값을 갖게 된다.

2.2 이론적 베리오그램

자료로부터 추출한 경험적 베리오그램은 그대로 사용하기 어렵기 때문에 일정한 형상을 표현하는 이론적인 모형식에 적합시켜 사용하게 된다. 관측자료로부터 식 (1)을 적용하여 얻어진 것을 경험적 베리오그램(empirical variogram), 경험적 베리오그램의 값을 이론적인 식에 적합시킨 것을 이론적 베리오그램(theoretical variogram)이라 부른다. 이론적 베리오그램은 너깃 θ_1, 문턱값 θ_2 그리고 상관거리 θ_3의 세 모수를 이용하여 정의할 수 있는 다음의 세 가지 형태가 가장 널리 사용된다(신기일, 2008).

1) 구형 베리오그램(Spherical variogram)

$$\gamma(h,\theta) = \theta_1 + \theta_2\left(\frac{3}{2}\frac{h}{\theta_3} - \frac{1}{2}\left(\frac{h}{\theta_3}\right)^3\right), \quad 0 < h < \theta_3 \tag{2}$$

$$\gamma(h,\theta) = \theta_1 + \theta_2 \quad h > \theta_3$$

2) 지수형 베리오그램(Exponential variogram)

$$\gamma(h,\theta) = \theta_1 + \theta_2\left(1 - \exp\left(-\frac{h}{\theta_3}\right)\right), \quad 0 < h \tag{3}$$

3) 가우시안 베리오그램(Gaussian variogram)

$$\gamma(h,\theta) = \theta_1 + \theta_2\left(1 - \exp\left(-\frac{h^2}{\theta_3}\right)\right), \quad 0 < h \tag{4}$$

구형과 지수형 베리오그램은 정확한 상관거리를 가지고 있지 않으며 근사적인 상관거리를 구하여 사용하는 것으로 알려져 있다. 이때 사용하는 상관거리를 의사상관거리(pseudo-range)라고 부르며 계산상으로 얻어진 문턱값의 95%에 해당하는 값을 가지는 거리로서 정의된다. 상기에서 소개된 여러 개의 이론적인 베리오그램이 존재한다는 것은 선택의 어려움을 의미한다. 즉 분석자료에 잘 맞는 이론적인 베리오그램을 선택하는 것은 쉬운 일이 아님을 의미한다. 최근에는 위에서 언급한 이론적인 베리오그램을 일반화시킨 베리오그램으로 부르고 있지만 엄밀한 의미에서 베리오그램이 아닌 상관함수에 형태를 가지는 Matérn 상관함수를 이용하는 경우가 많다. Matérn 상관함수는 Matérn 베리오그램이라는 용어와 혼용되어 사용되는 경우가 많다(신기일, 2008).

4) Matérn 상관함수(Matérn correlation function)

$$C_\theta(h) = \frac{\sigma^2}{2\Gamma(\kappa)}\left(\frac{2\sqrt{\kappa}\,\|h\|}{\rho}\right)K_\kappa\left(\frac{2\sqrt{\kappa}\,\|h\|}{\rho}\right) \tag{5}$$

여기서 모수는 분산 σ^2, 베리오그램의 상관거리를 결정하는 ρ 그리고 상관함수의 형상을 결정짓는 κ의 세 가지이다. $\Gamma(\ )$는 감마함수를 의미하고 $K_\kappa(\ )$는 차수가 κ인 제2종 수정 Bessel 함수이다. 모수 중에서 κ가 커지면 함수가 부드러워지는 것으로 알려져 있다. κ가 0.5이면 지수형, 무한대이면 가우시안과 같아지는 것으로 알려져 있다.

경험적 베리오그램에 이론적 베리오그램 모형을 적합시키는 방법은 크게 두 가지로 구분할 수 있다. 첫 번째는 모형 속에 포함된 오차의 분포와 상관구조를 가정하여 적용하는 최대우도추정법이고 두 번째는 앞서 소개된 이론적 베리오그램에 관한 비선형 함수의 매개변수를 회귀분석에서 처럼 최소제곱법(least square method)을 이용하여 추정하는 방법이다. 최대우도추정법은 자료가 정규분포를 따라야 하고 계산시 일정하게 수렴하는 베리오그램을 구하지 못하는 경우가 종종 발생한다고 알려져 있다. 따라서 이론적 베리오그램의 추정방법으로는 일반적으로 아래와 같은 가중 최소제곱 추정법(weighted least squares estimation)을 많이 사용한다(신기일, 2008). 가중 최소제곱 추정법은 경험적 베리오그램과 이론적 베리오그램의 값의 차의 제곱이 최소가 되도록 모수를 찾는 방법의 일종으로 경험적 베리오그램을 구성하는 자료가 등분산성을 만족하지 못할 경우에도 적정한 모수를 구해주는 것으로 알려져 있다. 가중 최소제곱 추정법을 수식으로 나타내면 다음과 같다.

$$\min \quad \theta \quad \sum_{h_i=1}^{k} \frac{N_h}{[\gamma(h_i,\theta)]}\left(\gamma(\hat{h}_i) - \gamma(h_i,\theta)\right)^2 \tag{6}$$

여기서 θ는 이론적 베리오그램의 모수, $\gamma(\hat{h}_i)$는 h_i번째 이론적 베리오그램의 추정치, $\gamma(h_i,\ \theta)$는 h_i번째 경험적 베리오그램의 값, 그리고 N_h는 최소제곱 추정법에 이용된 경험적 베리오그램을 구성하는 자료의 총개수이다.

본 연구에서는 상기 식 (6)에 의한 가중 가중 최소제곱 추정법으로 이론적 베리오그램의 모수를 추정하였다.

2.3 크리깅 기법

크리깅의 의미는 통계학적으로 최적의 공간 보간법(optimal spatial interpolation)을 의미한다(신기일, 2008). 크리깅은 관심있는 지점의 값을 알기위해 기지값들의 선형조합을 이용하는 기법이라고 앞서 소개한 바 있으며 이를 간단히 수식으로 정리하면 다음과 같다(신기일, 2008, 최종근, 2007).

$$z^* = \sum_{i=1}^{n} \lambda_i z_i \tag{7}$$

여기서 z^*는 위치가 알려진 지점에서의 크리깅을 이용한 예측치, i는 이미 그 위치와 값을 알고 있는 i번째 강우량 자료, λ_i는 i번째 강우량 자료에 대한 가중치, 그리고 n은 크리깅 예측을 위해 사용한 강우량 자료의 총개수이다. 가중치 결정을 위해 예측값과 참값 사이의 오차가 최소가 되도록 하며(minimum variance) 많은 경우에 추정값이 편향되지 않아야(unbiased) 한다는 조건을 추가로 사용한다. 선형회귀나 티센법, 역거리가중법과 같은 내삽법의 경우 사용하는 변수들 사이의 명확한 관계를 이용하는 반면 크리깅은 베리오그램이라는 공간적으로 분포된 자료들의 공분산 특성을 나타내는 별도의 내재하는 관계를 이용하게 된다. 따라서 주어진 동일한 이론적 베리오그램에 대하여 크리깅은 오직 하나의 추정값이 결정론적으로 정해지므로 추계학적인(stochastic) 기법과도 명백한 차이가 있다(최종근, 2007).

일반적으로 사용되는 크리깅은 주로 정규 크리깅(ordinary kriging) 그리고 일반 크리깅(universal kriging)의 두 가지이다. Krige(1951)에 의해 최초로 유도된 단순 크리깅(simple krige)은 크리깅 예측결과를 포함한 전체자료의 평균을 산정하면 기지값들의 평균과 서로 다르게 나타난다고 알려져 있어 사용되지 않는다. 이러한 문제로 평균이 편향되지 않으면서 오차분산을 최소로 하는 방법으로 제안된 것이 정규 크

리깅이다. 정규 크리깅을 적용할 때 식 (7)과 같은 선형결합식에서 가중치 λ는 베리오그램과 기지점에서 미지점까지의 거리를 이용하여 라그랑쥐 승수법을 적용해 다음과 같은 행렬식으로 유도될 수 있다.

$$\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \Gamma^{-1}\left(\gamma - \frac{\bar{1}^T\Gamma^{-1}\gamma - 1}{\bar{1}^T\Gamma^{-1}\bar{1}}\right)\bar{1} \tag{8}$$

여기서 λ는 위치와 값을 알고 있는 각 자료에 곱하게 되는 가중치들의 행렬 표현, Γ는 기지점과 미지점의 거리에 따른 베리오그램의 값들의 행렬 표현으로 $\Gamma = \{\gamma(X_i - X_j)\}$ 로 표현할 수 있다. X_i는 기지점의 위치, X_j는 미지점의 위치로 $X_i - X_j$는 두 지점간의 거리를 나타낸다. $\bar{1}^T = (1, 1, \ldots, 1)$ 로 단위행렬을 의미한다.

위에서 소개된 정규 크리깅은 자료들이 분포한 위치에 상관없이 그 평균이 같은 값을 나타낸다고 가정한다. 그러나 실제 강우량 자료들은 공간적으로 특정 경향을 나타내거나 평균이 지역에 따라 변화하는 경우가 많다. 크리깅 적용시 이러한 공간적 분포특성을 고려하지 못하면 결국 잘못된 답을 얻을 수 있다. 이 경우에 주어진 자료가 보여주는 일정한 분포경향을 찾아내어 원자료에서 이를 제거한 잔차를 얻을 수 있고, 얻어진 잔차에 대해 크리깅 기법을 적용한 후 분포경향을 합산하여 예측하는 방법을 일반크리깅이라고 한다. 따라서 일반 크리깅은 외적 분포경향을 가지는 크리깅(kriging with external drift)이라고도 불리워지며 이를 수식으로 나타내면 다음과 같다.

$$z(x)^* = D(x)^* + R(x)^* \tag{9}$$

식 (9)에서 $D(x)^*$는 자료가 나타내는 공간적 분포경향(drift or trend)이며 $R(x)^*$는 각 자료값이 가지는 잔차이다. 단일변량을 대상으로 크리깅을 적용할 때 공간적 분포경향의 모형식으로는 좌표값에 대한 1차 다항식 또는 2차 다항식을 주로 사용한다. 한편, 분석대상 변수와 상관성이 높은 자료를 2차변수로 이용하여 공간적 분포경향을 추정하기도 한다. 본 연구에서는 정규크리깅과 2차변수를 이용한 일반크리깅을 적용하여 비교하였다.

3. 대상자료의 공간적 분포특성

기상청의 기상대 자료와 AWS 자료를 적정 베리오그램을 유도하여 비교검토하기 위한 목적으로 수집하였다. 기상대 자료는 이론적 베리오그램을 유도하는데 이용하였고, AWS 자료는 베리오그램별 크리깅 예측값의 적정성을 판단하기 위한 비교목적으로 이용하였다. 2002년 8월 31일 태풍 루사(RUSA) 내습시를 대상으로 강원도 북부와, 제주도 및 섬지역을 제외한 우리나라 중남부 지역의 50개소의 기상대와 312개소의 AWS 자료를 수집하였다. 이때 강우량 이외에 수집가능한 기상인자 중 일강우량과 상관성 높을 것으로 보여지는 일평균기온, 일평균풍속, 상대습도, 해면기압, 이슬점온도의 5개 종류의 자료를 수집하였다. Fig. 2는 기상청 측후소(KMA)와 AWS의 관측위치를 보여주는 그림이며 Fig. 3은 수집된 일강우량의 등우선도를 보여주고 있다. Fig. 3은 기상대와 AWS 자료를 모두 이용하여 작성한 것으로 강우량의 분포가 전국에 걸쳐 원활한 형상의 일정한 경향성을 가지고 있다기 보다는 특정 지역에서 집중적으로 나타나는 군집특성(clustering feature)을 보여주고 있다.

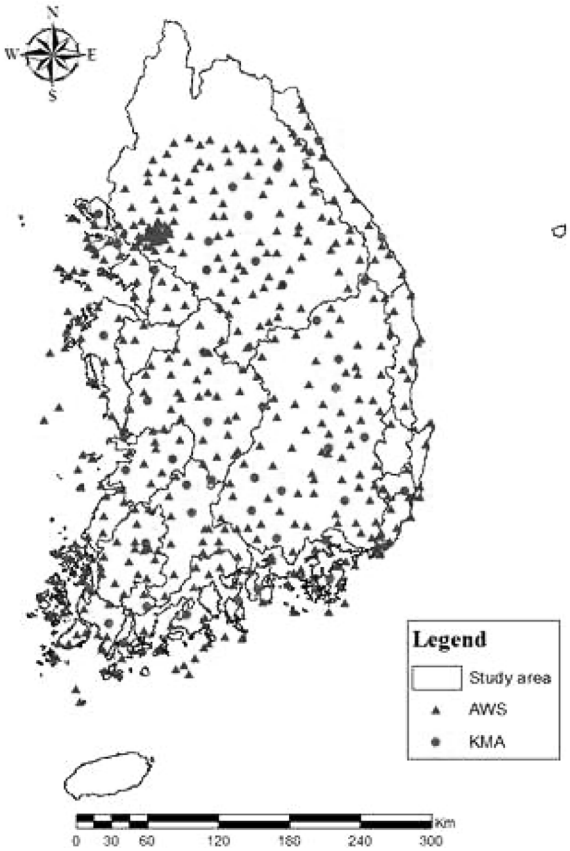


Fig. 2 location map for weather station

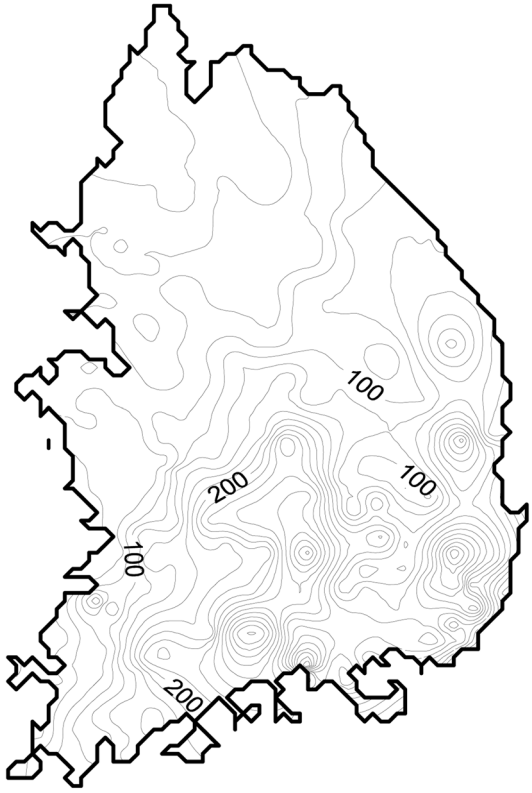


Fig. 3 Daily rainfall map (measured data)

(a) Total rainfall (b) 2002.08.31.12:20 (c) 2002.08.31.12:40

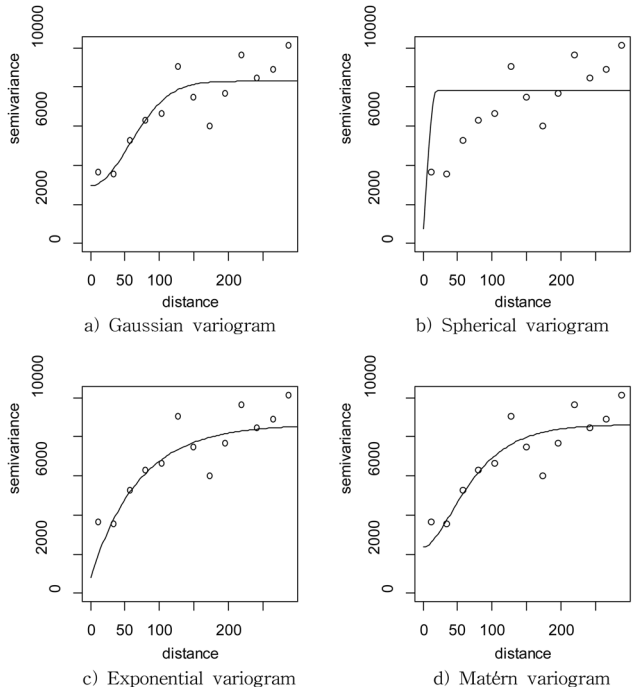
주) ① 태백산맥 : 추가령곡에서 동해안을 따라 낙동강 하구의 대대포 부근까지 이르는 산맥, 500 km 이상에 걸쳐서 1,000 m 이상의 산들로 구성.
② 광주산맥 : 태백산맥의 철령 부근에서 분기하여 서울 부근에 이르는 산맥. 400~1,200 m의 산들로 구성.
③ 차령산맥 : 태백산맥의 오대산에서 갈라져서 충북의 북부, 충남의 중앙을 남서방향으로 뻗은 산맥, 250~1,500 m의 산들로 구성
④ 노령산맥 : 소백산맥의 중부 추풍령 부근에서 남서방향으로 전주시와 순창군의 중간을 지나는 산맥, 300~1,100 m의 산들로 구성
⑤ 소백산맥 : 태백산맥에서 갈라져 남서방향으로 뻗어 내려오다 영남과 호남지방의 경계를 이루는 산맥,600~1,900 m의 산들로 구성

Fig. 4 Orographic effect of rainfall in typhoon RUSA

본격적인 분석에 앞서 분석대상이 되는 태풍 루사에 의한 강우특성을 살펴보면 다음과 같다. 제15호 태풍 '루사'는 8월 23일 약한 소형태풍으로 괌섬 북동쪽 약 1,800 km 부근 해상에서 생성되어, 이후 강한 대형 태풍으로 발달하였으며, 8월 26일 15시에는 일본 오키나와 섬 동남쪽 약 1,860 km 부근 해상으로 진출하였다. 8월 31일 12시에는 제주도 서귀포 동쪽 약 58 km 부근 해상까지 진출하였으며, 18시에는 전남 고흥반도 부근에 상륙하여 2002년 9월 1일 03시에 충북보은을 경유하여, 12시에는 강원도 인제 부근을 지나 15시에는 속초 북동쪽 130 km 부근해상에서 열대성 저기압으로 약화되어 소멸되었다(건설교통부, 2000). Fig. 4는 남서쪽으로부터 유입된 호우가 5번 소백산맥과 1번 태백산맥에 가로막혀 이동이 제한되는 산지효과가 발생하고 있음을 보여주고 있다.

4. 공간자료의 탐색적 분석을 통한 베리오그램 추정

본 연구에서 이론적 베리오그램은 기상청의 기상대 자료만을 이용하여 유도하였다. 경험적 베리오그램을 점으로 도시하여 이론적 베리오그램을 선의 형태로 함께 도시한 결과는 Fig. 5와 같다. 이때 추정된 베리오그램의 모수는 Table 1에 정리하였다. 특정한 베리오그램의 우월성을 판단하는데 많이 사용되는 로그우도함수는 본 연구의 경우 큰 차이가 없어 활용하기 어려운 것으로 파악되었다. Fig. 5를 살펴보면 구형 베리오그램은 적합결과가 상당히 낮고, 가우시안, 지수형 그리고 Matérn 베리오그램의 경우 구형 베리오그램의 경우보다는 적합도가 높다. 그림으로 판단하기에 적정하지 못한 구형 베리오그램을 제외한 나머지 베리오그램으로 추정된 모수를 평가해 보면 가우시안의 경우에는 상관거리가가 약 80 km, 지수형의 경우에는 약 70 km으로 얻어진 반면 Matérn의 경우에는 약 40 km로 상대적으로 작은 상관거리를 가지는 것으로 분석되었다.



a) Gaussian variogram b) Spherical variogram
c) Exponential variogram d) Matérn variogram

Fig. 5 Theoretical variogram for rainfall without covariates

Fig. 5에서 경험적 베리오그램의 형태를 살펴보면 자료들 간의 거리가 멀어짐에 따라 분산의 정도가 커졌다가 작아지는 진동형태를 유지하면서 점점 커지고 있다. 따라서 분석자에 따라서는 문턱값이 일정한 값으로 수렴하지 않고 계속 증가하고 있다고 판단할 수 있다. 이 경우 이론적 베리오그램의 모수가 계산상으로 얻어지기는 했으나 경험적 베리오그램을 적정하게 반영해주는 이론적 베리오그램이 존재하지 않는다고 볼 수 있고, 이러한 베리오그램을 이용한 크리깅 적용결과 역시 적정하지 못한 것으로 판단될 수 있다.

경험적 베리오그램이 진동하는 원인을 고찰해 보면 다음과 같다. 경험적 베리오그램이 진동하는 것은 강우의 거리에 따른 상관관계가 강우량 값 뿐만 아니라 강우의 발생특성과

Table 1. Parameters of theoretical variogram for rainfall without covariates

Variogram	Nugget	Partial sill	Range	Log-likelihood
Gaussian	2,956.8	5,336.3	80.1	-281.2
Spherical	773.9	7,041.4	19.2	-289.1
Exponential	805.5	7,829.8	70.9	-281.3
Matérn	2,355.0	6,271.1	38.7	-281.1

상관성 높은 어떤 특징적인 변수의 영향이 함께 받고 있기 때문이다(Goovaerts, 2000). 예를 들면 평지와 산지가 교대로 나타날 경우 상대적으로 평지의 강우량 보다 산지의 강우량이 큰 경우가 많다. 따라서 평지와 평지 또는 산지와 산지의 강우량 크기는 상관성이 높은 반면에, 평지와 산지는 강우량 크기의 상관성이 떨어진다고 볼 수 있다. 이 경우 강우가 발생하는 공간적인 특성을 설명해 줄 수 있는 변수를 도입하여 해석에 이용할 경우 베리오그램을 안정화시키고 크리깅의 적용성을 높일 수 있다. Goovaerts(2000)는 이러한 배경에 따라 지형표고를 이차변수로 도입하여 적용한 연구결과를 제시한 바 있다.

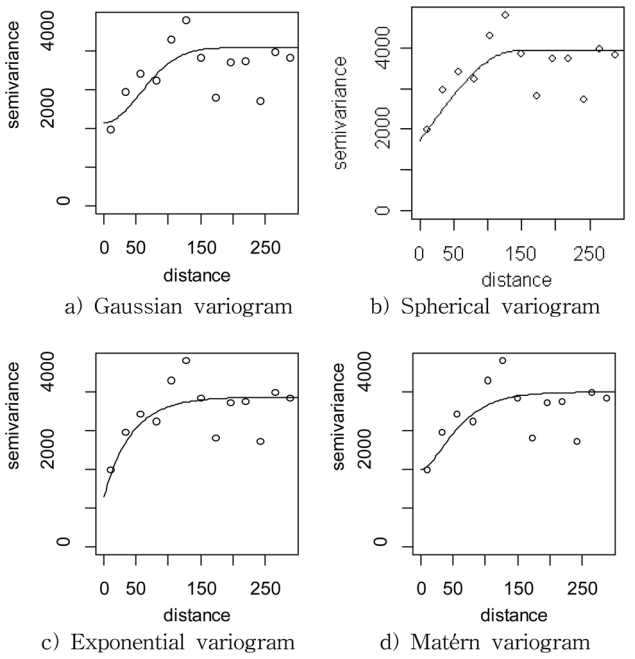
본 연구에서는 표고자료를 대체하여 강우량의 변동특성을 설명해 줄 수 있는 다양한 기상변수를 시행착오적으로 검토하였다. 그 결과 일평균풍속, 상대습도, 이슬점 온도의 세 가지 변수를 이용하는 것이 가장 개선된 베리오그램을 얻을 수 있는 것으로 검토되었다. 선택된 이차변수와 강수량 사이의 관계를 살펴보면 다음과 같다. 풍속의 경우 단시간동안 관측된 강우량만을 고려한다면 강우의 공간분포와의 상관성이 크지 않을 것이다. 그렇지만, 일강우량처럼 상당한 기간동안 관측되는 강우량의 경우 구름의 이동 등이 일평균풍속의 영향을 받을 것이므로 상당한 거리에 걸쳐서 지속적인 영향을 미치는 주요인자가 될 것이다. 상대습도와 이슬점 온도의 경우는 현재 강우효율을 평가할 때 쓰이는 주요한 개념인 가강수량(precipitable water)을 계산하는데 있어 사용되는 주요한 두 가지 변수가 된다는 점에서 강우의 효율성 즉 강수량의 크기와 상대적으로 큰 관계를 가지는 것으로 판단할 수 있다(윤용남, 2007).

선택된 세 가지 이차변수들은 일반적으로 그 측정값이 표고값의 변화에 영향을 받는 것으로 알려진 대표적인 기상인자이다. 그렇지만 이들 기상인자들의 공간분포와 강우량의 공간분포 사이에 직관적으로 파악할 수 있는 일련의 정량적인 관계를 보이기는 어려운 문제다. 이보다는 강우량 자료만을 이용한 베리오그램이 일반적으로 잘 수렴하지 않는 문제와 관련해서 대안으로서 기상인자들을 활용할 수도 있다는 것에 의의가 있다. 따라서, 본 연구에서 2차변수로 선택한 기상인자들을 이용한다 하더라도 항상 개선된 베리오그램을 추정할 수 있다고 보기는 어렵다. 그렇지만 강우량만을 이용한 분석이 적절하지 못할 경우 대안으로서 고려할 수 있는 방안으로서 의미를 가진다.

앞서 소개된 세 가지 기상인자를 이차변수로서 도입하여 이론적 베리오그램을 다시 추정해 본 결과는 Fig. 6 및 Table 2와 같다. 이차변수를 고려하여 얻어진 일강우량의 경험적 베리오그램을 살펴보면 단순히 강수량만을 고려할 때 지속적으로 증가하는 양상을 보인 것과 달리 뚜렷하게 일정

Table 2. Parameters of theoretical variogram for rainfall without covariates

Variogram	Nugget	Partial sill	Range	Log-likelihood
Gaussian	2,146.0	1,938.0	79.7	-262.7
Spherical	1,723.3	2,199.1	141.9	-262.6
Exponential	1,283.5	2,582.6	42.9	-263.6
Matérn	1,998.2	1,993.5	31.8	-263.3



a) Gaussian variogram
b) Spherical variogram
c) Exponential variogram
d) Matérn variogram

Fig. 6 Theoretical variogram for rainfall with covariates

한 값으로 수렴하고 있음을 확인할 수 있다. 따라서 강우량 변수를 단독으로 사용하여 베리오그램을 유도한 경우와 달리 이론적 베리오그램이 경험적 베리오그램을 잘 따르는 형태로 유도되었다. 이론적 베리오그램의 모수를 살펴보면 구형의 경우에는 상관거리가 약 140 km로 크게 얻어졌으나 가우시안의 경우에는 약 80 km로 강우량만을 고려하는 경우와 비슷하게 나타났다. 그러나 지수형과 Matérn의 경우에는 강우량만을 사용하는 경우보다 작은 약 40 km와 30 km로 얻어졌다.

5. 최적 베리오그램의 검토

앞서 얻어진 이론적 베리오그램의 추정결과 만으로는 어떤 베리오그램이 적절한지를 판단하기 어렵다. 따라서 본 연구에서는 다양한 베리오그램을 이용하여 크리깅 기법으로 AWS 관측지점의 강우량을 예측해 하고 실제 관측된 강우량과 비교하여 그 결과를 고찰하기로 한다. 우선 비교대상이 되는 참값으로는 우리나라 중남부 지역의 50개소의 기상청 관측소와 이들 관측소 사이의 312개의 AWS를 모두 이용하였으며 Fig. 3과 같다. 이에 비해 Fig. 7과 Fig. 8은 50개의 기상청 관측소 자료만을 이용하여 베리오그램을 유도하고 AWS의 강우량을 추정해서 등우선을 작도한 결과이다. 따라서 Table 3에 제시된 검정통계량은 실제 관측된 AWS 강우량과 기상청 자료로 추정한 강우량을 서로 비교한 결과이다.

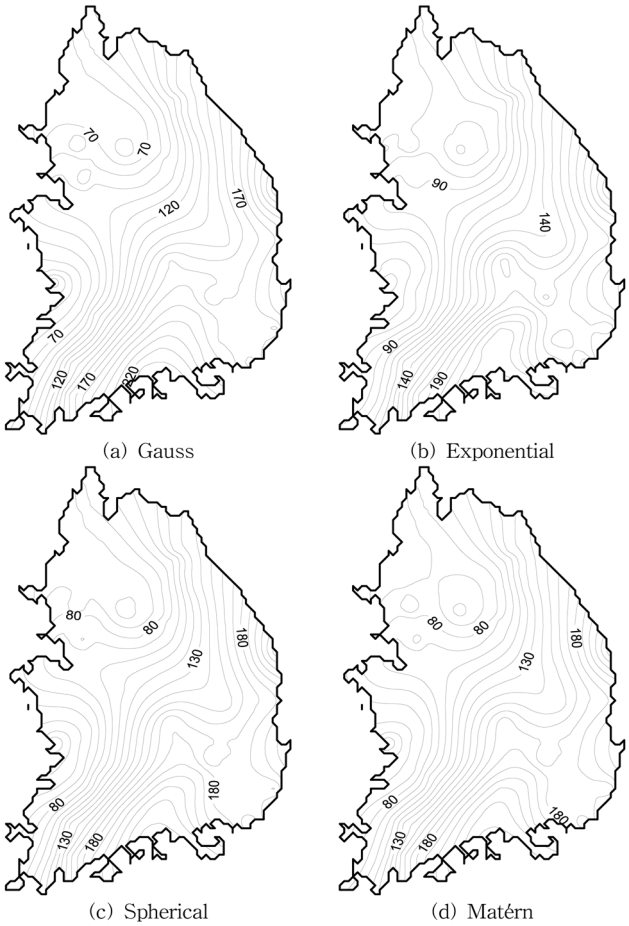


Fig. 7 Rainfall maps using ordinary kriging(variogram without covariates)

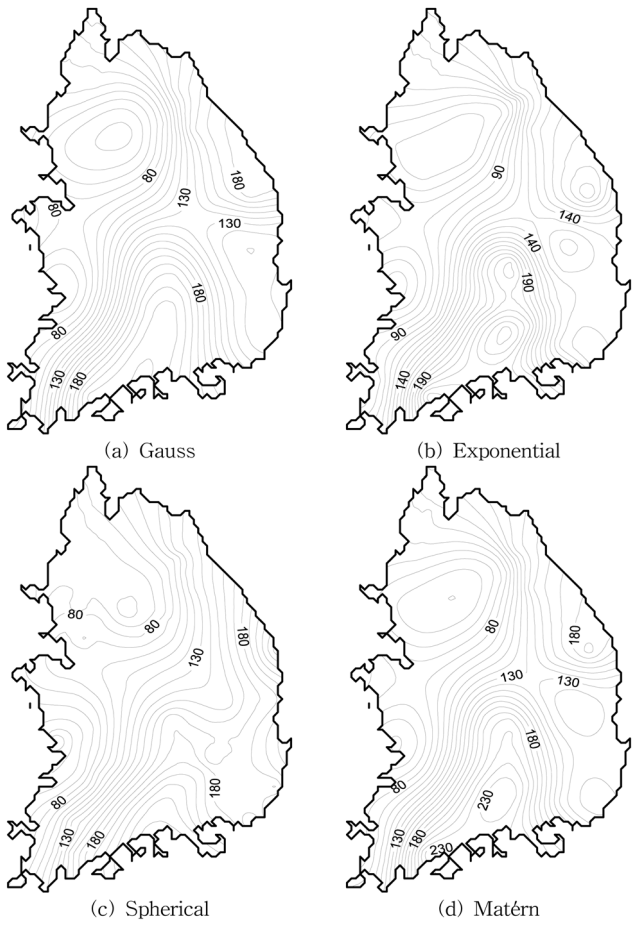


Fig. 8 Rainfall maps using universal kriging(variogram with covariates)

우선 예측된 강수량을 이용하여 Fig. 7 및 Fig. 8의 등우선도를 작성하고 Fig. 3의 관측 강우량의 등우선도와 비교하였다. 또한 AWS 지점의 관측강우량에 대한 예측결과를 정량적으로 확인하기 위해 비교통계량을 계산하였다. 비교통계량으로는 RE(Relative Error), MAPE(Mean Absolute Percent Error) 그리고 PBIAS(Percent of Bias)를 이용하였으며 그 형태는 식 (7)~(9)과 같다.

$$RE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{R_{true,i}-R_{comp,i}}{R_{true,i}}\right)^2 \tag{7}$$

$$MAPE(\%) = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|R_{true,i}-R_{comp,i}\right|}{R_{true,i}}\times 100 \tag{8}$$

$$PBIAS = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{R_{true,i}-R_{comp,i}}{R_{true,i}}\right) \tag{9}$$

Fig. 7은 강우량 자료만을 이용하여 얻어진 이론적 베리오그램을 기초로 정규크리깅을 적용하여 얻은 강우량의 등우선도이다. Fig. 3의 관측 강우량의 등우선도에 비해 전국적으로 분포된 상당히 원만한 곡선의 형태를 보이며, 지역별로 집중되는 강우의 형태는 거의 보여주지 못하고 있다. 한편, Fig. 8은 강우량 자료와 이차변수를 함께 고려하여 얻어진 이론적 베리오그램을 기초로 일반크리깅을 적용하여 얻은 강우량의 등우선도이다. 관측강우량의 등우선도에는 미치지 못하지만 상대적으로 지역별로 집중되는 강우의 형태가 조금씩 나타나고 있다. 특히 지수형 베리오그램과 Matérn 상관함수를 사용한 경우의 등우선도의 형태가 다른 베리오그램을 사용한 경우보다 지역별로 집중되는 등우선 형태를 보다 잘 나타내주는 것으로 보여진다.

Table 3. Comparative results of kriging application with various variograms

Class	Variogram	RE	MAPE	PBIAS
Ordinary kriging	Gaussian	130.7	139.5	-125.1
	Spherical	257.0	128.6	-114.1
	Exponential	96.2	256.2	-240.3
	Matérn	102.2	129.5	-115.0
Kriging with external drift	Gaussian	39.6	118.4	-104.9
	Spherical	42.2	121.3	-108.2
	Exponential	44.7	120.9	-107.4
	Matérn	39.7	119.8	-106.6

예측값과 실제 관측값의 차이의 정도를 나타내는 비교통계량을 계산하여 정리한 결과는 Table 3과 같다. Table 3에서 먼저 정규 크리깅의 적용결과를 살펴보면 베리오그램별로 예측결과를 나타내는 비교통계량의 값의 차이가 상당히 큰 것을 발견할 수 있다. 이는 주어진 경험적 베리오그램의 값에서 적정한 이론적 베리오그램을 유도하기가 상대적으로 어렵다는 것을 의미한다. 대체로 구형 베리오그램과 Matérn 상관함수를 이용한 경우가 상대적으로 작은 값을 가지고 있어 금회 호우사상에 대해 예측력이 높았다고 할 수 있으나 이중에서 구형 베리오그램의 경우 Fig. 5에서 단순히 경험적

베리오그램과 비교시 가장 적합하지 못한 것으로 판단된 바 있어 베리오그램의 유도와 크리깅 예측결과가 상충되게 나타났다고 볼 수 있다.

한편, 일반 크리깅의 적용결과를 살펴보면 베리오그램의 종류와 상관없이 적용오차가 거의 비슷하게 얻어졌다. 또한, 정규 크리깅을 적용한 경우에 비해 모든 경우에서 훨씬 작은 오차를 보이고 있다. 이는 경험적 베리오그램의 값이 안정적인 형태로 얻어질 경우 서로 다른 이론적 베리오그램을 선택하더라도 크리깅 적용결과의 차이는 그리 크지 않다는 것을 의미한다. 또한 강우발생과 상관성 높은 이차변수를 함께 고려하여 일강우량의 경험적 베리오그램을 유도할 수 있다면 상대적으로 강우량만을 이용하는 경우보다 안정적인 형태의 경험적 베리오그램을 얻을 수 있다는 것을 보여준다. Fig. 8의 결과와 함께 고찰할 때 유역평균강우량의 추정시에는 어떤 베리오그램을 쓰더라도 차이가 크지 않은 반면 특정 지점의 값을 보간하는 목적으로 이용하는 경우에는 상대적으로 국지적으로 집중되는 호우형태를 잘 모의해주는 지수형과 Matérn 상관함수의 사용이 적절할 것으로 보여진다.

공간적인 분포를 해석하거나 검토해야 할 경우 무분별하게 Surfer(Golden software, Inc., 2002) 등 패키지 프로그램에서 기본적인 옵션에만 기초하여 크리깅을 적용하는 사례가 발견된다. 그러나 본 연구의 검토결과처럼 대상자료에 대한 베리오그램의 적정성을 확인하지 못한 상태에서 적용할 경우 그 해석결과는 상당히 큰 오차를 유발할 수 있다고 판단된다. 한편, 본 연구에서는 기존에 널리 사용되어온 이론적 베리오그램인 구형, 지수형, 가우시안 이외에 이들 베리오그램 모형의 일반적인 형태로 알려진 Matérn 상관함수를 함께 적용하여 비교하였다. 적용 호우사상의 수가 한정적이기는 하나 다른 이론적 베리오그램 모형에 비해 정규 크리깅과 일반 크리깅의 경우 모두 비교통계량이 우수한 것으로 나타났다. 따라서 일강우량에 대한 적정한 베리오그램의 판단이 어려운 경우에 Matérn 상관함수를 이용한 크리깅 적용을 고려할 수도 있다고 보여진다.

6. 결 론

크리깅 방법은 공간적으로 분포된 자료를 다루는 문제에서 가장 널리 사용되고 있는 방법이나 적정 베리오그램이 유도되지 못할 경우 상당히 큰 오차를 수반한다고 알려져 있다. 본 연구에서는 이러한 이유로 일강우량의 적정 베리오그램을 고찰하였다. 그 결과 일강우량 자료만을 이용하는 경우보다 강우발생과 상관성이 높은 이차변수를 함께 고려하여 베리오그램을 유도하는 것이 실제 관측자료와 비교할 때 훨씬 현실성 있는 예측을 가능하게 해 주는 것으로 나타났다. 이때, 일강우량의 공간분포와 상대적으로 관련성이 크다고 보여지는 이차변수로는 일평균풍속, 상대습도, 이슬점 온도을 선정할 수 있었다. 한편, 한정적인 적용사례이기는 하나 크리깅의 적용에 앞서 베리오그램의 적정성을 판단하기 어려운 경우 가장 일반적인 형태의 베리오그램인 Matérn 상관함수를 이용하는 방안을 우선 고려할 수 있을 것으로 보여진다.

국내에서는 등우선도 등의 작성 등을 위하여 크리깅을 적용하게 될 경우 적정 베리오그램에 대한 검토 없이 상용프로그램에서 제공하는 기본적인 옵션을 이용하는 경우가 많다. 그러나, 본 연구내용으로 미루어 볼 때 크리깅은 정규크리깅, 일반크리깅 등 상당히 다양한 적용방법을 가지고 있으며 동일한 적용방법이라도 베리오그램에 따라 결과의 차이가 크게 발생할 수 있다. 즉, 대상자료의 성격에 부합하지 않는 베리오그램을 사용한 크리깅의 적용은 상당힌 큰 오차를 수발할 수 있다. 따라서, 확률강우량의 산정시 분포형에 대한 적합도 검정이 선행되어야 하는 것처럼, 크리깅 적용결과의 신뢰도를 제고하기 위해서는 적정 베리오그램의 결정과정이 선행되어야 할 것으로 판단된다.

참고문헌

신기일(2008) **공간통계학 강의자료**

유철상, 정광식(2001) 면적평균강우량의 추정 및 추정오차, **한국수자원학회논문집**, 한국수자원학회, 제34권, 제4호, pp. 317-326.

윤광훈, 서봉철, 신현석(2001) 크리깅 기법을 이용한 낙동강 유역 홍수강우의 공간해석 연구, **한국수자원학회논문집**, 한국수자원학회, 제37권, 제2호, pp. 233-240.

윤용남, 김중훈, 유철상, 김상단(2002) 공간분포된 강우를 사용한 유출 매개변수 추정 및 강우오차가 유출계산에 미치는 영향 분석, **한국수자원학회논문집**, 한국수자원학회, 제35권, 제1호, pp. 1-12.

윤용남(2007) **수문학**. 청문각.

이재형, 유양규(2002) 면적강우량 산정을 위한 관측망 최적설계 연구, **한국수자원학회논문집**, 한국수자원학회, 제35권, 제2호, pp. 187-194.

최종근(2007) **지구통계학**. 시그마프레스.

Cressie, N.A.C (1991) *Statistics for Spatial Data,* John Wiley & Sons, Inc.

Diggle, P.J. and Ribeiro, P.J. (2007) *Model-based Geostatistics.* Springer.

Golden software, Inc. (2002) *Surfer user's guide version 8.*

Goovaerts, P. (2000) Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall, *Journal of Hydrology*, Vol. 228, No. 1-2, pp. 113-129.

Isaaks, E.H. and Srivastava, R.M. (1989) *An Introduction to Applied Geostatistics,* Oxford University Press.

Journel, A.G. and Huijbregts, Ch. J. (1978) *Mining Geostatistics,* Academic Press.

Krige, D.G. (1951) A statistical approach to some basic mine valuation problems on the Witwatersrand", *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, Vol. 52, pp. 119-139.

Matérn, B. (1960) Spatial variation, *Technical report*, Statens Skogsforsningsinstitut.

Matheron, G. (1971) The theory of regionalized variables and its applications, *Technical Report 5,* Cahiers du Centre de Morphologie Mathematique.

Tabios, G.Q. and Salas J.D. (1985) A comparative analysis of techniques for spatial interpolation of precipitation, *Water Resources Bulletin,*, Vol. 21, No. 3, pp. 365-380.

(접수일: 2009.2.6/심사일: 2009.4.7/심사완료일: 2010.9.10)