레이다 강우 자료와 강우계 자료를 이용한 강우 추정
Indicator Co-Kriging에 기초한 우량 자료의 합성을 중심으로

서 동 준

1. 서론

기존의 관악산 레이다외에 부산, 제주, 동해 및 군산의 레이다가 가동 되기 시작하면서 이들로부터 얻어지는 자료의 수문 연구 및 현업에의 이용에 대한 관심이 높아 지고 있다. 레이다 우량 자료는 공간적으로 연속적이라는 장점이 있지만 우량계 자료와 비교하면 오차가 상당히 커서, 레이다 우량 자료만을 이용하는 것보다는 우량계 자료와 함께 이용하여 우량을 추정하는 것이 바람직하다. 본 연구에서는 indicator co-kriging(Journel 1983)에 기초한 레이다 우량 자료와 우량계 자료와의 합성을 위한 최적 추정 기법을 제시하였다.

2. Indicator Co-Kriging을 이용한 강우 추정

아래와 같이 indicator 확률 변수들, $I_g(u;z_c)$, $I_r(u;z_c)$, $J_r(u)$를 정의하자:

$$i_g(u;z_c) = \begin{cases} 1 \text{ if } z_g(u) \leqq z_c \\ 0 \text{ if } z_g(u) > z_c \end{cases} \quad (1)$$

$$i_r(u;z_c) = \begin{cases} 1 \text{ if } z_r(u) \leqq z_c \\ 0 \text{ if } z_r(u) > z_c \end{cases} \quad (2)$$

$$j_r(u) = \begin{cases} 1 \text{ if } z_r(u) > 0 \\ 0 \text{ if } z_r(u) = 0 \end{cases} \quad (3)$$

여기서 $z_g(u)$는 지점 u에서의 우량계 우량, $z_r(u)$는 지점 u에서의 레이다 우량, z_c는 cutoff 우량이다. 이상적으로는 $E[I_g(u_0;z_c) | I_g(u_i;z_c)=i_g(u_i;z_c), i=1,...,N_g; I_r(u_j;z_c)=i_r(u_j;z_c), j=1,...,N_r]$ (여기서 N_g는 우량계 자료의 수이고 N_r은 레이다 우량 자료의 수이다)를 추정하는 것

한국 과학 기술 연구원 시스템 공학 연구소

이 바람직하나, 본 연구에서는 상기의 조건 기대치에 $T=\{j(u_i)=1, i=0,\ldots,Ng;\ j(u_j)=1, j=1,\ldots,N_r\}$의 조건 사상을 더해 $E[I_g(u_0;z_c)|I_g(u_i;z_c)=i_g(u_i;z_c), i=1,\ldots,N_g;\ I_r(u_j;z_c)=i_r(u_j;z_c), j=1,\ldots,N_r;T]$를 추정한다. 이렇게 하면 우량 추정이 가능한 지역이 레이다가 나타내는 강우 지역에 한정된다는 단점이 있으나 강우 현상의 intermittency를 직접 고려하지 않아도 되고 신뢰도 높은 통계치를 추정할 수 있다는 장점이 있다 (Seo 1992). Estimator는 최적 선형 추정에서와 마찬가지로 아래와 같은 형태를 취한다:

$$
\begin{aligned}
i_g^*(u_0;z_c) &= F_{g|T}(z_c) \\
&+ \sum_{i=1}^{N_g} w_{gi}(z_c)\ \{\, i_g(u_i;z_c) - F_{g|T}(z_c) \,\} \\
&+ \sum_{j=1}^{N_r} w_{rj}(z_c)\ \{\, i_r(u_j;z_c) - F_{r|T}(z_c) \,\} \qquad (4)
\end{aligned}
$$

여기서 $F_{g|T}(z_c) \equiv E[I_g(u)|T]$, $F_{r|T}(z_c) \equiv E[I_r(u)|T]$ 그리고 $w_{gi}(z_c)$들과 $w_{rj}(z_c)$들은 산정되어야 할 가중치들이다. $w_{gi}(z_c)$들과 $w_{rj}(z_c)$들은 조건 오차 분산, $E[\{i_g(u_0;z_c) - i_g^*(u_0;z_c)\}^2|T]$,를 최소화함으로써 구해지며 다음과 같이 주어진다:

$$
[W_g(z_c)\ \ W_r(z_c)] = [Q_{0g}(z_c)\ \ Q_{0r}(z_c)] \begin{bmatrix} Q_{gg}(z_c) & Q_{gr}(z_c) \\ Q_{rg}(z_c) & Q_{rr}(z_c) \end{bmatrix}^{-1} \qquad (5)
$$

여기서 $W_g(z_c)$와 $W_r(z_c)$는 가중치 vector들이며 각각의 i번째와 j번째 요소는 $w_{gi}(z_c)$ and $w_{rj}(z_c)$이고, $Q_{0g}(z_c)$와 $Q_{0r}(z_c)$는 공분산 vector들이며 각각의 i번째와 j번째의 요소는 $Cov[I_g(u_0;z_c), I_g(u_i;z_c)|T]$와 $Cov[I_r(u_0;z_c), I_r(u_j;z_c)|T]$이고, $Q_{gg}(z_c)$, $Q_{gr}(z_c)$, $Q_{rg}(z_c)$ 그리고 $Q_{rr}(z_c)$는 공분산 행렬들로서 공분산 vector들과 유사하게 정의된다. 구하고자 하는 각 cutoff level에서의 indicator 변수의 추정치, $i_g^*(u_0;z_c)$, 들로부터 지점 u_0에서의 추정 우량, $z_g^*(u_0)$, 는 아래로부터 구해진다:

$$
z_g^*(u_0) = \int_0^{\infty} u\ di_g^*(u_0, u) \qquad (6a)
$$

$$
\approx \sum_{k=1}^{K} z_{c,k} \cdot [i_g^*(u_0;z_{c,k+1}) - i_g^*(u_0;z_{c,k})] \qquad (6b)
$$

여기서 $z_{c,k'}$는 구역, $[z_{c,k+1}, z_{c,k}]$,의 중간값을 나타낸다. 식 (5)로부터 최적 가중치를 구하기 위해서는 레이다 우량의 indicator 공분산, 우량계 우량의 indicator 공분산 그리고 레이다 우량과 우량계 우량간의 indicator 공분산이 추정되어야 한다. 통상적으로 레이다 우량은 그 자료의 수가 많아 indicator 공분산의 추정에 큰 어려움은 없다. 그러나 우량계 자료는 그 수가 충분히 크지 않아 신뢰도 높은 우량계 우량의 indicator 공분산과 레이다 우량과 우량계 우량간의 indicator 공분산의 추정이 매우 어렵다. 따라서 본 연구에서는 우량계 우량의 indicator 공분산과 레이다 우량과 우량계 우량간의 indicator 공분산을 실측 자료들로부터 직접 추정하는 대신에 레이다 우량과 우량과 우량간에 가장 단순한 형태의 오차 구조(식 (10))을 가정하여 레이다 우량 자료로부터 산정하는 방법을 택하였다.

3. 레이다 우량의 indicator 공분산

레이다 우량의 indicator 공분산은 아래와 같이 정의된다:

$$\begin{aligned}
&\mathrm{Cov}[I_r(v;z_1), I_r(u;z_c)] \\
&= \Pr[Z_r(v) \leqq z_1, Z_r(u) \leqq z_c] - \Pr[Z_r(v) \leqq z_1] \Pr[Z_r(u) \leqq z_c] && (7a) \\
&= \Pr[Z_r(u) \leqq z_c] \{ \Pr[Z_r(v) \leqq z_1 | Z_r(u) \leqq z_c] - \Pr[Z_r(v) \leqq z_1] \} && (7b)
\end{aligned}$$

여기서 z_1와 z_c는 서로 다른 cutoff level들이다. 본 연구에서는 nugget 효과를 포함하는 레이다 우량의 조건 공분산을 아래와 같이 나타낼 수 있다고 가정하였다 (Seo 1992):

$$\begin{aligned}
&\mathrm{Cov}[I_r(u+h;z_1), I_r(u;z_c) | T] \\
&= \Pr[Z_r(u) \leqq z_c | T] \{ \Pr[Z_r(u) \leqq z_1 | Z_r(u) \leqq z_c, T] \\
&- \Pr[Z_r(u) \leqq z_1 | T] \} e^{-h/L} && (8a)
\end{aligned}$$

$$= \begin{cases} \Pr[Z_r(u) \leqq z_c | T] \{1 - \Pr[Z_r(u) \leqq z_1 | T] - n(z_1, z_c)\} e^{-h/L} & \text{if } z_1 \leqq z_c \\ \Pr[Z_r(u) \leqq z_1 | T] \{1 - \Pr[Z_r(u) \leqq z_c | T] - n(z_1, z_c)\} e^{-h/L} & \text{if } z_1 < z_c \end{cases} \quad (8B)$$

여기서 $n(z_1, z_c)$가 nugget 효과이고 L은 상관 거리이다. $\mathrm{Cov}[I_r(u+h;z_1), I_r(u;z_c) | T]$는 $\Pr[Z_r(u) \leqq z_c | T]$로 정규화되어 아래와 같이 쓰여질 수 있다

$$\Pr[Z_r(u+h) \leqq z_1 | Z_r(u) \leqq z_c, T]$$

$$P \begin{cases} \Pr[Z_r(u) \leqq z_1 | T] \; (1 - e^{-h/L}) + \mu(z_1, z_c) \; e^{-h/L} & \text{if } z_1 \leqq z_c \\ \Pr[Z_r(u) \leqq z_1 | T] \; \{ 1 - e^{-h/L} + \mu(z_1, z_c) \; e^{-h/L} / \Pr[Z_r(u) \leqq z_c | T] \} & \text{if } z_1 < z_c \end{cases} \quad (9)$$

여기서 $\mu(z_1, z_c)=1-n(z_1, z_c)$. 본 연구에서 우량계 우량의 indicator 공분산과 레이다 우량과 우량계 우량간의 indicator 공분산을 구하기 위해 이용된 $Pr[Z_g(v) \leqq z_g | Z_r(u) \leqq z_r, T]$과 $Pr[Z_g(u+h) \leqq z_{gc} | Z_g(u) \leqq z_g, T]$의 산정 방법은 아래와 같다.

4. Pr [Zg(v) ≦ zg | Zr(u) ≦ zr, T]의 산정

우량계 우량과 레이다 우량이 다음과 같은 관계를 갖는다고 가정하자:

$$Z_g(u) = \beta Z_r(u) + \varepsilon(u) \tag{10}$$

여기서 β는 레이다 우량이 갖는 bias의 역수이고(이미 알고 있다고 가정), $\varepsilon(\ )$는 무작위 오차 항이다. 확률 변수들, $Z_g(u)$, $Z_r(u)$, $\varepsilon(u)$가 취하는 값들을 각각 z_{gu}, z_{ru}, e_u라 표기한다면 이들 사이에는 $e_u \geq -\beta z_{ru}$의 부등식이 성립하며, 조건 확률 $Pr[Z_g(v) \leqq z_g | Z_r(u) \leqq z_r, T]$는 아래와 같이 주어진다:

$$Pr[Z_g(v) \leqq z_g | Z_r(u) \leqq z_r, T]$$

$$= \int_0^{\infty} \int_{-\beta z_{rv}}^{z_g - \beta z_{rv}} f_{(v), Zr(v)|}(e_v, z_{rv} | Z_r(u) \leqq z_r, T)\ de_v\ dz_{rv} \tag{11a}$$

$$= \int_0^{\infty} \left[\int_{-\beta z_{rv}}^{z_g - \beta z_{rv}} f_{(v)|Zr(v),}(e_v | z_{rv}, Z_r(u) \leqq z_r, T) de_v \right] f_{Zr(v)|}(z_{rv} | Z_r(u) \leqq z_r, T)\ dz_{rv} \tag{11b}$$

본 연구에서는 식 (11b)의 내부 적분의 계산을 위해 아래와 같이 가정하였다:

$$f_{(v)|Zr(v),}(e_v | z_{rv}, Z_r(u) \leqq z_r, T)$$

$$\approx f_{(v)|Zr(v),}(e_v | z_{rv}, T) \tag{12}$$

또한 $f_{(v)|Zr(v),}(e_v | z_{rv}, T)$는 아래와 같이 truncated 정규 확률 밀도 함수라 가정하였다 (서 동준 1992):

$$f_{(v)|Zr(v),}(e_v | z_{rv}, T)$$

$$= \frac{1}{(2\pi)^{1/2}\ \sigma \{1-\phi(-\beta z_{rv})\}} \exp\{-(e_v - m)^2 / 2\sigma^2\} \tag{13}$$

여기서 m과 σ^2은 untruncated 정규 확률 밀도 함수의 평균과 분산이며

$$\phi(x) = \int_{-\infty}^{x} \frac{1}{(2\pi)^{1/2}\sigma} \exp\{-(x-m)^2/2\sigma^2\}\, dx \tag{14}$$

조건 확률 $Pr[Z_g(v) \leqq z_g | Z_r(u) \leqq z_r, T]$는 이제 아래와 같이 주어진다:

$$\begin{aligned} &Pr[Z_g(v) \leqq z_g | Z_r(u) \leqq z_r, T] \\ &= \int_0^{\infty} \phi(-\beta z_{rv}, z_g-\beta z_{rv})\ f_{z_r(v)|}\ (z_{rv}|Z_r(u) \leqq z_r, T)\ dz_{rv} \end{aligned} \tag{15}$$

여기서

$$\phi(-\beta z_{rv}, z_g-\beta z_{rv}) = \int_{-\beta z_{rv}}^{z_g-\beta z_{rv}} f_{(v)|z_r(v),}\ (e_v|z_{rv}, T)\ de_v \tag{16}$$

식 (15)의 적분을 행하는 데에는 그대로 수치 적분을 하는 것과 부분 적분을 한 후 수치 적분을 하는 것의 두 가지 방법이 있다. 만약 nugget 효과가 존재하지 않는다면 그대로 수치 적분하는 것이 더 효율적이나, 본 연구에서와 같이 nugget 효과를 고려한다면 먼저 부분 적분을 행하는 것이 편리하며 이 경우 $Pr[Z_g(v) \leqq z_g | Z_r(u) \leqq z_r, T]$는 아래와 같이 주어진다:

$$\begin{aligned} &Pr[Z_g(u+h) \leqq z_g | Z_r(u) \leqq z_r, T] \\ &= (1-e^{-h/L})\ Pr[Z_g(u) \leqq z_g | T] \\ &+ e^{-h/L}\ Pr[Z_g(u+) \leqq z_g | Z_r(u) \leqq z_r, T] \end{aligned} \tag{17}$$

여기서

$$\begin{aligned} &Pr[Z_g(u) \leqq z_g | T] \\ &= \int_0^{\infty} -\phi'(-\beta z_{rv}, z_g-\beta z_{rv}) Pr[Z_r(v) \leqq z_{rv} | T] dz_{rv} \end{aligned} \tag{18}$$

$$\begin{aligned} &Pr[Z_g(u+) \leqq z_g | Z_r(u) \leqq z_r, T] \\ &= \int_{z_r}^{\infty} -\phi'(-\beta z_{rv}, z_g-\beta z_{rv})\ \mu(z_{rv}, z_r) dz_{rv} \\ &+ \frac{\int_0^{z_r} -\phi'(-\beta z_{rv}, z_g-\beta z_{rv})\ \mu(z_{rv}, z_r)\ Pr[Z_r(v) \leqq z_{rv} | T]\ dz_{rv}}{Pr[Z_r(u) \leqq z_r | T]} \end{aligned} \tag{19}$$

$$\begin{aligned} &Pr[Z_g(u+) \leqq z_g | Z_r(u) \leqq z_r, T] \\ &\equiv \lim Pr[Z_g(u+\delta) \leqq z_g | Z_r(u) \leqq z_r, T] \end{aligned} \tag{20}$$

$\delta \to 0+$

식 (18), (19)에서 $\phi'(-\beta z_{rv}, z_g-\beta z_{rv})=\partial\phi(-\beta z_{rv}, z_g-\beta z_{rv})/\partial z_{rv}$. h=0에서의 조건 확률, $Pr[Z_g(u) \leqq z_g | Z_r(u) \leqq z_r, T]$, 는 식 (19)에서 nugget 효과를 생략함으로써 구해진다.

5. Pr [Zg (v) ≦ zgc | Zg (u) ≦ zg, T] 의 산정

$Pr[Z_g(v) \leqq z_{gc} | Z_g(u) \leqq z_g, T]$는 $Pr[Z_g(v) \leqq z_g | Z_r(u) \leqq z_r, T]$의 산정과 유사하게 아래와 같이 구해진다:

$$Pr[Z_g(v) \leqq z_{gc} | Z_g(u) \leqq z_g, T]$$

$$= \int_0^{\infty} \int_{-\beta z_{rv}}^{z_{gc}-\beta z_{rv}} f_{\varepsilon(v), zr(v)|}(e_v, z_{rv} | Z_g(u) \leqq z_g, T) \; de_v \; dz_{rv} \tag{21a}$$

$$= \int_0^{\infty} \left[\int_{-\beta z_{rv}}^{z_{gc}-\beta z_{rv}} f_{\varepsilon(v)|zr(v),}(e_v | z_{rv}, Z_g(u) \leqq z_g, T) de_v \right] f_{zr(v)|}(z_{rv} | Z_g(u) \leqq z_g, T) \; dz_{rv} \tag{21b}$$

$Pr[Z_g(v) \leqq z_g | Z_r(u) \leqq z_r, T]$의 산정 때와 유사하게 아래와 같이 가정한다:

$$f_{\varepsilon(v)|}(e_v | z_{rv}, Z_g(u) \leqq z_g, T) \approx f_{\varepsilon(v)|}(e_v | z_{rv}, T) \tag{22}$$

위의 가정으로 $Pr[Z_g(v) \leqq z_{gc} | Z_g(u) \leqq z_g, T]$는 아래와 같이 주어진다:

$$Pr[Z_g(v) \leqq z_{gc} | Z_g(u) \leqq z_g, T] = \int_0^{\infty} \phi(-\beta z_{rv}, z_{gc}-\beta z_{rv}) \; f_{zr(v)|}(z_{rv} | Z_g(u) \leqq z_g, T) \; dz_{rv} \tag{23}$$

부분 적분을 행하면 $Pr[Z_g(u+h) \leqq z_{gc} | Z_g(u) \leqq z_g, T]$는 다음과 같이 주어진다:

$$Pr[Z_g(u+h) \leqq z_{gc} | Z_g(u) \leqq z_g, T] = (1-e^{-h/L}) \; Pr[Z_g \leqq z_{gc} | T] + e^{-h/L} \; Pr[Z_g(u+) \leqq z_{gc} | Z_g(u) \leqq z_g, T] \tag{24}$$

여기서

$$\Pr[Z_g \le z_{gc}] = \int_0^\infty -\phi'(-\beta z_{ru}, z_{gc}-\beta z_{ru})\Pr[Z_r(u) \le z_{ru} | T] dz_{ru} \quad (25)$$

$$\Pr[Z_g(u+) \le z_{gc} | Z_g(u) \le z_g, T] = \frac{\int_0^\infty -\phi'(-\beta z_{ru}, z_{gc}-\beta z_{ru})\Pr[Z_g(u+) \le z_g, Z_r(u) \le z_{ru} | T] dz_{ru}}{\Pr[Z_g(u) \le z_g | T]} \quad (26)$$

$$\Pr[Z_g(u+) \le z_{gc} | Z_g(u) \le z_g, T] \equiv \lim_{\delta \to 0+} \Pr[Z_g(u+\delta) \le z_{gc} | Z_g(u) \le z_g, T] \quad (27)$$

h=0에서의 조건 확률은 정의에 의해 다음과 같이 주어 진다:

$$\Pr[Z_g(u) \le z_{gc} | Z_g(u) \le z_g, T] = \begin{cases} \dfrac{\Pr[Z_g(u) \le z_{gc} | T]}{\Pr[Z_g(u) \le z_g | T]} & \text{if } z_{gc} \le z_g \\ 1 & \text{if } z_{gc} > z_g \end{cases} \quad (28)$$

6. 결어

본 연구에서는 indicator co-kriging에 기초한 레이다 우량 자료와 우량계 자료의 합성 기법을 제시하였다.

강우 지역과 비강우 지역을 모두 고려하는 경우, 강우 현상의 intermittency로 인해 신뢰도 높은 통계치의 추정이 매우 어려워, 본 연구에서는 레이다가 나타내는 강우 지역내에 포함되는 자료만을 이용함으로써 과거 자료를 이용한 통계치의 추정을 용이하게 하였다.

우량계 우량의 indicator 공분산 및 레이다 우량과 우량계 우량간의 indicator 공분산은 비교적 많은 수의 자료를 가지고도 신뢰도 높은 추정을 하기 어려워 본 연구에서는 레이다 우량과 우량계 우량간에 선형 오차를 갖는 구조를 가정하여 레이다 우량의 통계치들로부터 레이다 우량과 우량계 우량간의 indicator 공분산 및 우량계 우량의 indicator 공분산을 구하도록 하였다.

본 연구에서 제시된 indicator co-kriging에 기초한 합성 기법은 비교적 개념이 단순하고

algorithm이 간단하며 통계치들을 자료로부터의 추정에만 의존하는 방법보다 공분산 행렬의 positive definiteness 조건을 만족시키기 용이하다는 장점이 있어 레이다 우량 자료와 우량계 자료의 실시간 합성에 유용할 것으로 보인다. 그러나 본 연구에서 제시된 기법은 simple co-kriging에 의한 우량 자료 합성에 비해 계산량이 10배 정도 많아 계산량 감소를 위한 단순화에 대한 연구가 요청된다.

또한 본 연구에서 제시된 합성 방법은 레이다 우량 자료의 bias를 미리 알아야 한다는 제약 조건이 있어 레이다 원시 자료를 레이다 우량 자료로 변환하는데 종래의 Z-R 변환보다 보다 효과적인 방법의 이용이 요청된다.

참고 문헌

서 동준 (1992), 레이다와 우량계를 이용한 우량 추정(I), 한국 과학 기술 연구원 시스템 공학 연구소 연구 보고서.

Journel, A. (1983), Nonparametric estimation of spatial distributions, Math. Geology, 15(3), 445-468.

Seo, D.-J. (1992), Rainfall Estimation Using Radar Data and Rain Gage Measurements - An Indicator Approach, Proceedings of the 6th IAHR International Symposium on Stochastic Hydraulics, Taipei, 1992.