한국습지학회지 제13권 제2호, 2011년 8월 (189 - 197)

시공간적 변동성을 고려한 강우의 결측치 추정 방법의 비교

김 병 식* / 노 희 성** / 김 형 수***+

The Comparison of Estimation Methods for the Missing Rainfall Data with spatio-temporal Variability

Kim, Byung Sik* / Noh, Hui seong** / Kim, Hung Soo***+

요약 : 본 논문에서는 지상강우의 결측치를 추정하는 방법들 중 역거리 가중치법(IDWM), 역지수 가중치법(IEWM), 상관계수가중치법(CCWM), 인공신경망(ANN)기법, 레이더 자료를 이용한 결측치 추정 방법을 비교하여 각각의 적용성을 검토하였다. 임진강 유역을 대상지역으로 하여 각 방법을 적용한 결과, 강우의 결측치 추정에 있어서 기존의 방법 중 상관계수 가중치법(CCWM)과 인공신경망(ANN)기법에 의한 RMSE가 0.46~1.79의 범위를 보였고, 레이더자료를 이용하여 강우의 결측치를 추정한 경우 RMSE가 0.05~2.26의 범위를 보였다. 레이더 강우자료가 지점 강우자료와 달리 강우의 공간상관성을 반영하고 있음을 볼 때, 지점강우 자료를 이용한 결측치 추정 기법보다 레이더자료를 이용한 결측치의 추정기법이 그 적용성에서 우수하다고 판단되어진다.

핵심용어 : 지점강우, 레이더강우, 역거리가중치법, 역지수가중치법, 상관계수가중치법, 인공신경망

Abstract : This paper reviewed application of data-driven method, distance-weighted method(IDWM, IEWM, CCWM, ANN), and radar data method estimated of missing raifall data. To evaluate these methods, statistics was compared using radar and station rainfall data from Imjin-river basin. The range of RMSE values calculated for CCWM, ANN was 1.4 to 1.79mm, and the range of RMSE values estimated data used for radar rainfall data was 0.05 to 2.26mm. Spatial characteristics is considered to Radar rainfall data rather than station rainfall data. Result suggest that estimated data used for radar data can impove estimation of missing raifall data.

Keywords : *ground rainfall, Radar rainfall, IDWM, IEWM, CCWM, ANN*

1. 서 론

강우량 자료는 수문 분석에서 있어 가장 기본이 되는 자료이기 때문에 그 이용에 있어서 매우 중요한 부분이다. 그러나 현실적으로는 완벽한 강우자료를 구축하기에는 많은 어려움이 있으며 이에 따라 수공 관련 전문가들은 강우자료의 결측치 추정에 대한 문제를 겪게 된다.

기존에 지점강우의 결측치에 대한 추정에는 일반적으로 전통적인 방법인 산술평균법, 역거리 가중치법, 역지수 가중치법, 상관계수 가중치법이 사용되었고, 최근에는 크리깅 기법이나 인공신경망(ANN)기법이 사용되고 있다. 국내의 연구 성과를 살펴보면, 김응석 등(1999)이 평창강 유역의 11개 지점 시자료를 바탕으로 1개의 호우사상을 산술평균법, 연정상 강우량법, 역거리 가중치

+ Corresponding author : sookim@inha.ac.kr
* 정회원・강원대학교 방재전문대학원 조교수·공학박사・E-mail : hydrokbs@kangwon.ac.kr
** 정회원・인하대학교 사회기반시스템공학부 박사과정・E-mail : heesung80@hanmail.net
*** 정회원・인하대학교 사회기반시스템공학부 교수・E-mail : sookim@inha.ac.kr

법, 거리고도비율법, 선형계획법, 수정 연정상 강우량법, 크리깅 기법을 이용 및 분석하여 적용성을 평가하였고, 황세운 등(2006)은 충주댐과 춘천댐 유역의 일자료를 대상으로 역거리 가중치법, 크리깅과 코크리깅(Co-Kriging)기법의 적용성을 평가하였다. 안상진 등(2000)은 보청천 유역 12개 지점의 시자료를 대상으로 고전적인 방법이외에 신경망기법을 이용한 방법을 추가하여 비교평가를 수행하였다.

그리고 최근에 호우사상 및 산지 집중호우를 예측하기 위해 위성 및 레이더 영상분석이 많이 이용되면서 이들 자료를 이용한 강우보정 방법에 대한 연구가 활발히 진행되고 있다. 안상진 등(2003)은 관악산 기상레이더의 반사강도와 강우강도를 분석하여 국내 조건에 맞는 관계식을 제시하였고, 김병식 등(2007)은 임진강 강우레이더를 이용하여 안성천 유역의 지점 시자료를 대상으로 조건부 합성방법을 이용하여 강우를 모의한 후, 관측 우량과 비교평가를 수행하였다.

국외의 연구성과를 살펴보면, Hodgson 등(1989)은 역거리 가중치법(IDWM)에서 거리 계산의 수를 줄이기 위해 learned search approach를 채택하여 역거리 가중치법(IDWM)을 수정하였다. Shepard 등(1968)은 지형학적인 측면을 통합적으로 고려하기 위하여 수정 역거리 가중치법(IDWM)을 제안하였다. 이러한 역거리 가중치법(IDWM)의 성공여부는 양의 공간자기상관성의 존재 여부에 따라 결정된다(Griffith, 1987; Vasiliev, 1996). 그러므로 대상지역과 가까운 곳에서 계측한 자료는 멀리 떨어진 곳에서 관측한 자료보다 비슷한 값을 나타낼 것이라 추측할 수 있다(Sullivan and Unwin, 2003). 그러나 거리만으로는 공간자기상관성을 측정할 수 없고, 음의 자기상관성의 존재 여부는 결측치의 추정을 위한 역거리 가중치법(IDWM)의 적용을 제한하는데 중요한 역할을 한다. 역거리 가중치법(IDWM)을 사용하기 위한 또 하나의 중요한 논점은 결측치 추정을 위한 특정 지점의 주변 관측 지점을 선택하는 것이다. 거리에 따른 관측값을 사용하기 위한 지점 선택은 자의적 판단에 편중되어 있다. Navone과 Ceccatto (1994)는 인도 전역의 여름철 몬순 강우 예측을 위하여 인공신경망 기법을 사용하였다. 본 연구에서는 기존까지 지점강우의 결측치를 추정하는 방법인 역거리 가중치법(IDWM), 역지수 가중치법(IEWM), 상관계수가중치법(CCWM), 인공신경망(ANN)기법의 결과와 더불어 최근 그 이용성이 커지고 있는 레이더 강우를 이용한 방법을 비교하여 강우의 결측치를 추정함에 있어서 최선의 방안을 제시하고자 한다.

2. 결측강우 자료의 추정기법

2.1 역거리 가중치법(IDWM, Inverse Distance Weighting method)

역거리 가중치법은(IDWM)은 인근 관측지점의 우량을 거리에 반비례로 가중하여 결측지점(보정지점)의 우량을 결정하는 방법이다.

$$\theta_m = \frac{\sum_{i=1}^{n} \theta_i d_{mi}^{-k}}{\sum_{i=1}^{n} d_{mi}^{-k}} \quad (1)$$

여기서, θ_m : m지점의 관측값, θ_i : I지점의 관측값, n : 관측지점의 수

d_{mi} : I지점으로부터 m지점까지의 거리(km),

k : 거리 가중 계수(1.0~6.0), 일반적으로 2.0을 사용

2.2 역지수 가중치법(IEWM, Inverse Exponential Weighting method)

역지수 가중치법은 일반적으로 surface generation를 위한 정량적 지리학에서 사용되는 방법이다(Goodchid, 1986; Sullivan and Unwin, 2003). 역거리 가중치법(IDWM)의 요소 중 d_{mi}^{-k}

가 e_{mi}^{-kd}로 대체되어 있는 형태로 일반적으로 k값은 2를 사용한다.

$$\theta_m = \frac{\sum_{i=1}^{n} \theta_i e_{mi}^{-kd}}{\sum_{i=1}^{n} e_{mi}^{-kd}} \quad (2)$$

2.3 상관계수 가중치법(CCWM, Coefficient of Correlation Weighting method)

상관계수 가중치법(CCWM)은 각 관측소의 강우자료의 시간적 상관성을 이용하는 방법으로 우선 가까운 관측소를 이용하여 두 관측소의 상관분석을 수행한 후, 계산된 상관계수를 가중치로 하여 결측지점의 우량을 보정하게 된다. 이 방법은 결측치를 추정할 때, 역거리 가중치법(IDWM)이 거리를 사용하는 반면에 상관계수를 사용하기 때문에 공간적 상관성보다는 자료의 상관성을 더 고려하는 방법이라고 할 수 있다.

$$\theta_m = \frac{\sum_{i=1}^{n} \theta_i R_{mi}}{\sum_{i=1}^{n} R_{mi}} \quad (3)$$

여기서, R_{mi} : 지점 m과 지점 i 자료 간의 상관계수

2.4 인공신경망기법(ANN, Artificial Neural Networks)

인공신경망(ANN)기법은 인간 두뇌의 학습기능을 모형화한 정보처리 시스템이므로 외부로부터 지식이나 정보가 입력되면 동적상태반응을 일으켜 필요한 출력을 생성시키는 네트워크라 할 수 있다. 다음 Fig. 1은 신경망이 정보 또는 지식을 습득하는 방법인 학습을 통하여 계산된 출력간에 오차가 있는 경우 그 오차에 대하여 계속적으로 반복하면서 각 결선에 대한 매개변수를 최적화하기 위해 연결강도를 조정하여 최종적인 결과를 얻는 것을 나타내는 모식도이다. 각 연결강도의 계산은 신경망의 목적함수를 학습시간 t에 대한 오차함수라고 할 경우, 출력값(O)과 실제값(Y)과의 제곱오차 합 $E(t)$으로 표현할 수 있으며, 여기서 학습이란 $E(t)$가 최소가 되도록 연결내의 가중치를 적절히 조정하는 것을 의미한다.

$$E(t) = \frac{1}{2}\sum (Y(t) - O(t))^2 \quad (4)$$

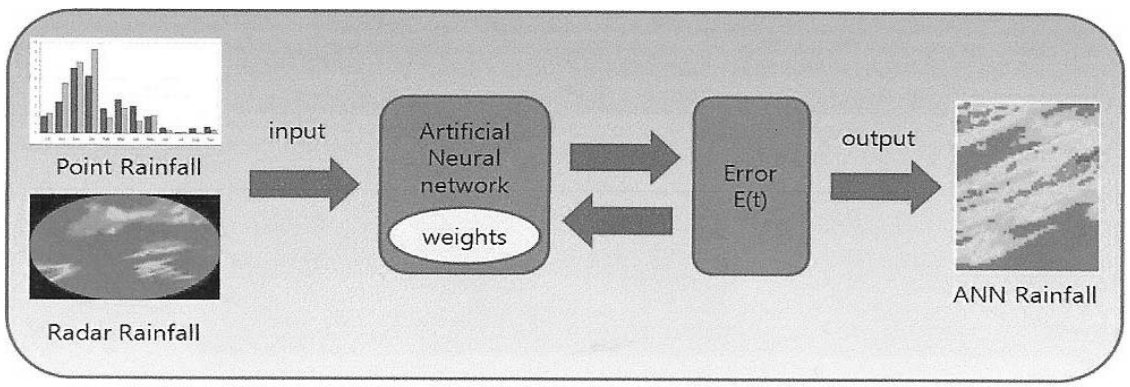


Figure 1. Concept of ANN algorithm

3. 임진강레이더 자료의 현황

임진강 강우레이더는 국내 최초로 수문관측을 위해 설치된 레이더로 기존의 기상 예보를 위해 설치된 기상레이더와는 그 운영의 성격에 차이를 두고 있으며, 임진강 하류지역의 홍수예보 및 수자원 관리를 위해 홍수예보시스템에 연계되어 레이더 관측자료를 생산하고 있다. 임진강 하류지역 수자원의 효율적 관리를 위해서는 북한상류지역의 강우량 관측이 필요하나 우량관측자료를 얻기는 어려운 실정이다. 임진강 강우레이더의 자세한 제원은 Tabel. 1과 같다.

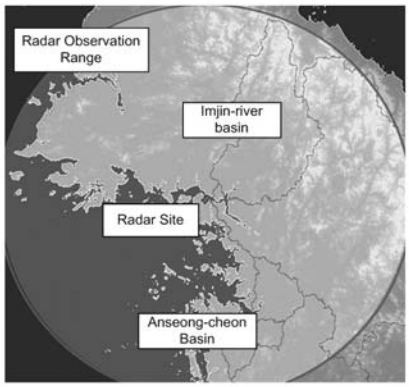


Figure 2. Imjin River rainfall radar

Table 1. Summary of Imjin River rainfall radar characteristics and observation modes

Characteristics of radar(TDR-43250C)		Observation mode	
Radar	TDR-43250C	Scan mode	Multi Elevation Volume Scan
Frequency	5.645 GHz	Elevation number	12(0.4 ~2.2)
PW/PRF	$2\mu s/500\,Hz$	Z-R equation	$Z= 31R^{1.17}$
Transmitter	Klystron	CAPPI Elevation	$1km$
Peak power	$250KW$	Antenna velocity	15deg/sec(2.5rpm)
Signal processor	RVP-8	Observation range	$170km$(max. $400km$)
Antenna	Offset style		
Veam width	0.95		

4. 적 용

4.1 대상유역 및 강우자료 현황

본 논문의 대상지역은 임진강 유역 Fig. 3으로써 총 유역 면적 약 8,117.5㎢로 이 중 약 63%인 5,108.7㎢가 북한 지역에 속해 있으며, 지상강우관측소는 TM관측소 15개를 선정하였다. 강우관측소 및 경·위도는 Table 2와 같다. 자료기간은 2006년 7월 14일 00:00부터 7월 15일 00:00시까지 30분 자료를 이용하였다.

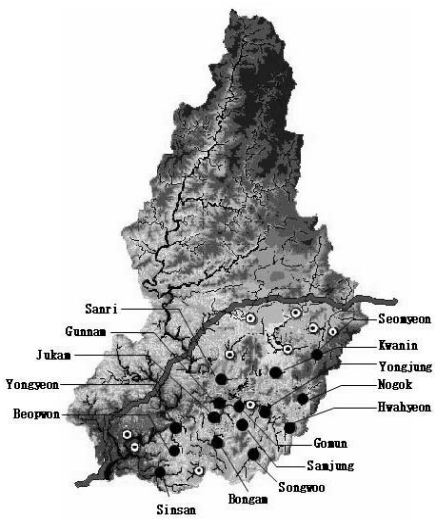


Figure 4. Imjin-river basin

Table 2. Rainfall Stations

	Station	Longitude	Latitude
St.1	Sinsan	126-51-14	37-46-40
St.2	Yongyeon	126-51-36	37-55-33
St.3	Beopwon	126-52-55	37-51-14
St.4	Bongam	127-00-01	37-54-21
St.5	Jukam	127-00-23	38-05-05
St.6	Gunnam	127-01-10	38-03-13
St.7	Sangri	127-04-39	38-08-59
St.8	Gomun	127-06-55	38-03-46
St.9	Samjung	127-07-59	37-57-40
St.10	Songwoo	127-08-44	37-49-37
St.11	Yongjung	127-13-06	38-00-23
St.12	Kwanin	127-16-08	38-07-20
St.13	Hwahyeon	127-17-13	37-53-56
St.14	Nogok	127-20-18	38-18-18
St.15	Seomyeon	127-25-05	38-09-56

4.3 분석 및 결과

본 논문에서는 역거리 가중치법(IDWM), 역지수 가중치법(IEWM), 상관계수 가중치법(CCWM), 인공신경망기법(ANN)을 이용하여 미계측 지점의 결측치를 산출할 수 있는 체계를 구축하였고, 본 연구자가 개발한 F2MAP 모형을 이용하여 보정 레이더 자료를 생성하고 그 자료를 결측치 추정 방법에 이용하였다. 또한, 성능을 평가하기 위한 방법으로 교차검정(Cross-Validation)을 적용하였다. 모의한 결과의 제곱평균제곱근오차(RMSE)를 Table. 3에 정리하였고, 각 지점별 RMSE 그래프는 Fig. 5과 같다. 각 방법별 RMSE와 Box Plot은 Fig. 6에 나타내었으며, Fig. 4는 본 연구의 흐름도이다.

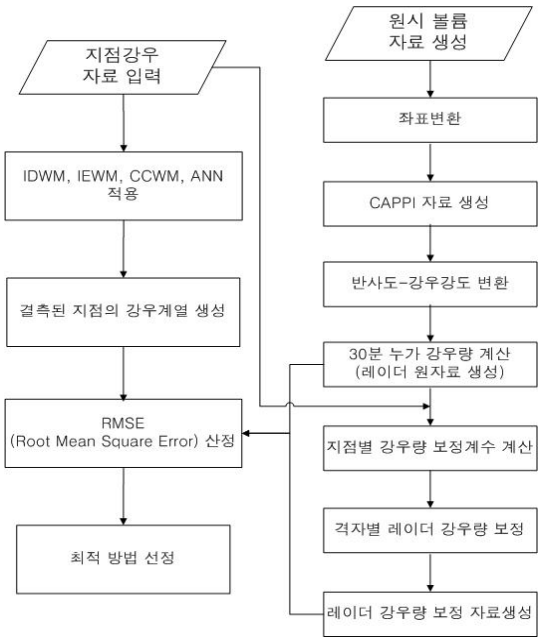
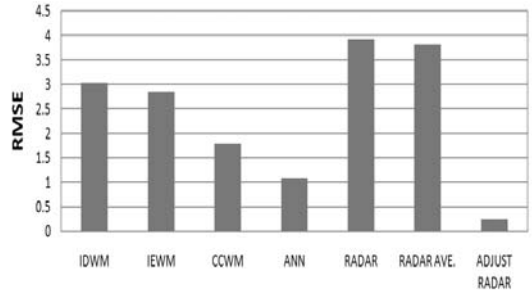


Figure 4. Study flow chart

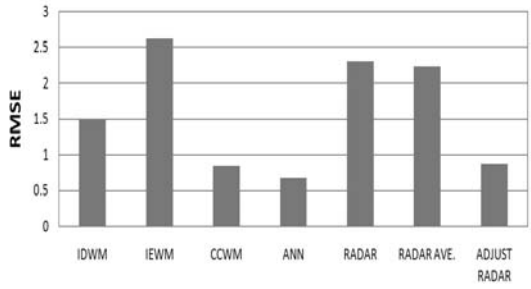
Table 4. Comparison of estimation data(RMSE)

(RMSE : mm)

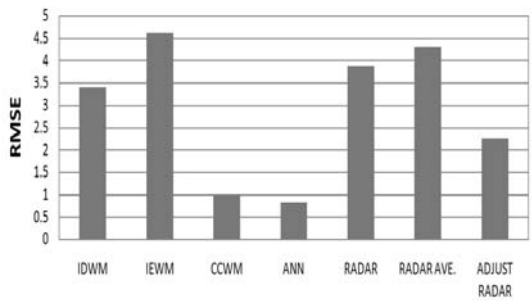
Estimation \ Station		st1	st2	st3	st4	st5	st6	st7	st8	st9	st10	st11	st12	st13	st14	st15
station rainfall	IDWM	2.67	4.27	2.12	1.37	2.36	4.03	2.62	2.81	2.55	3.41	1.49	1.41	3.03	1.73	1.95
	IEWM	2.72	4.89	3.87	2.41	3.4	5.29	3.36	4.63	3.79	4.63	2.62	2.06	2.85	1.78	5
	CCWM	0.78	0.75	0.77	0.57	0.79	0.97	1.1	0.81	0.77	0.99	0.84	0.81	1.79	1.21	0.91
	ANN	0.68	0.9	0.79	0.46	0.53	0.84	0.96	0.79	0.81	0.82	0.68	0.75	1.09	1.04	0.84
Radar rainfall	RADAR	5.56	5.28	5.18	4.41	5.41	5.6	5.34	3.95	3.05	3.89	2.3	2.44	3.92	2.88	3.4
		5.36	5.23	4.91	4.37	4.93	5.2	5.35	3.52	2.98	4.31	2.24	2.62	3.81	2.75	3.29
	ADJUST RADAR	1.09	0.87	1.4	0.4	0.86	0.3	1.8	1.99	1.39	2.26	0.87	0.05	0.25	0.27	0.68



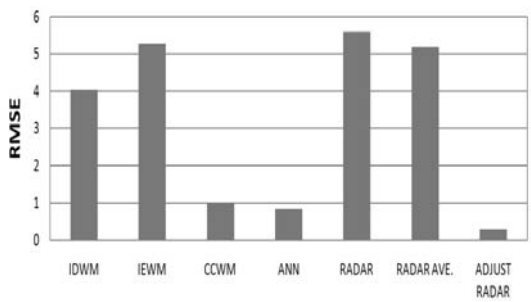
(a) Hwahyun



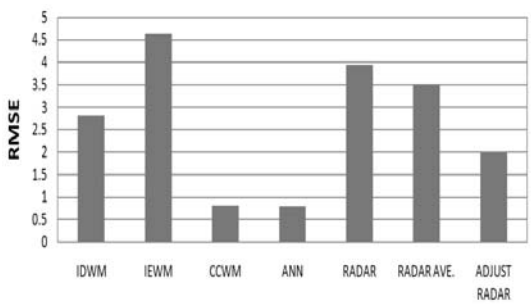
(b) Yeongjung



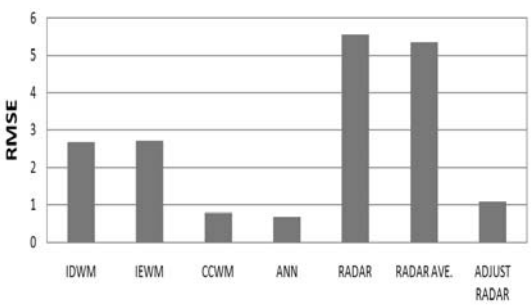
(c) Songwoo



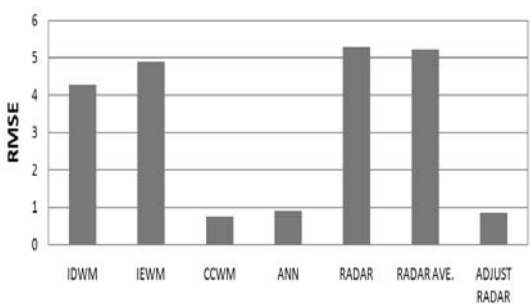
(d) Gunnam



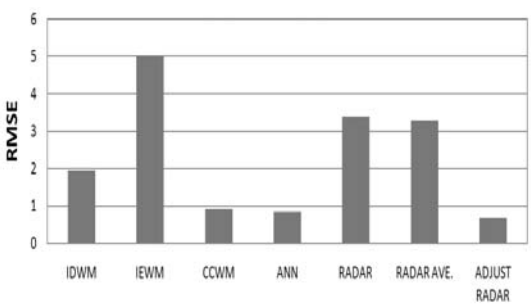
(e) Gomun



(f) Sinsan

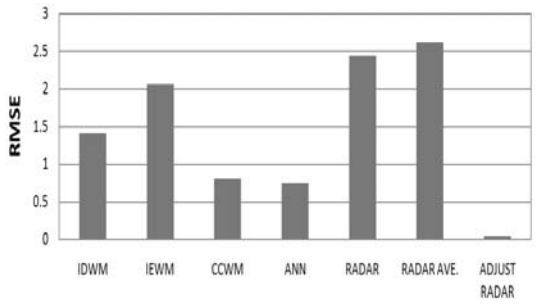


(g) Yongyeon

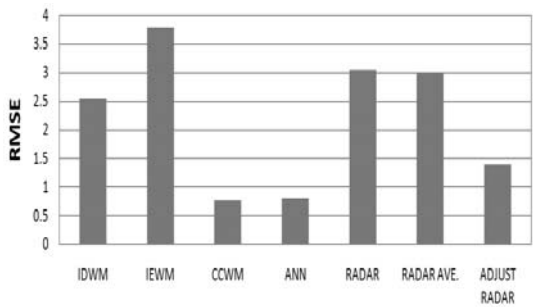


(h) Seomyeon(d) Gunnam

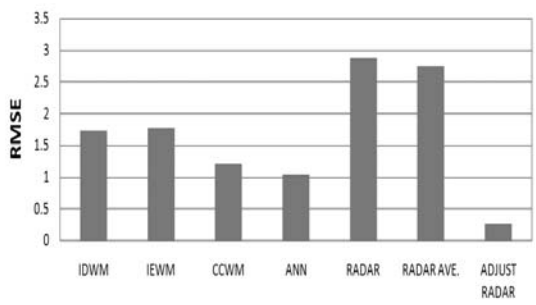
Figure 5. Comparison of estimation data(station)



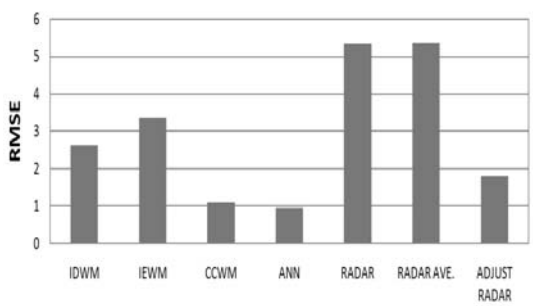
(i) Kwanin



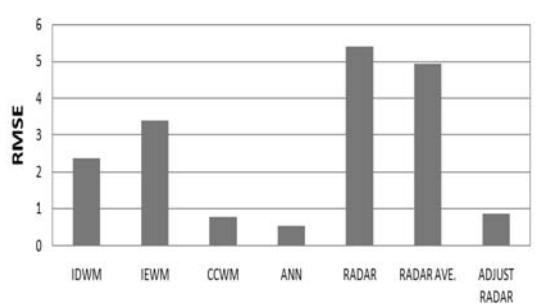
(j) Samjung



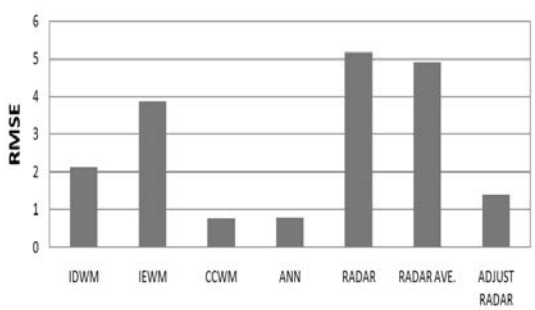
(k) Nogok



(l) Sanri

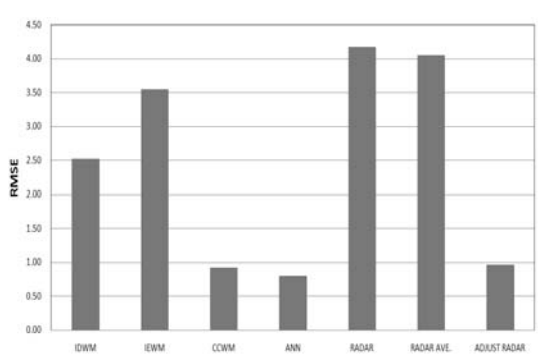


(m) Jukam

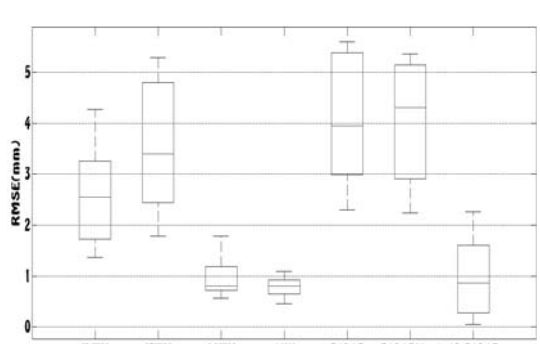


(n) Beopwon

Figure 5. Comparison of estimation data(station) (계속)



(a) RMSE Average of Methods



(b) BoxPlot of Methods

Figure 6. Comparison of estimation data(Method)

Table 8. Comparision of descriptive statistics(Method)

statistics \ Methods	IDWM	IEWM	CCWM	ANN	RADAR	RADAR Ave.	ADJUST RADAR
Mean	2.52	3.55	0.92	0.80	4.17	4.06	0.97
Standard error	0.23	0.29	0.07	0.04	0.30	0.28	0.18
Standard deviation	0.90	1.14	0.29	0.17	1.18	1.10	0.68
Variance	0.80	1.30	0.08	0.03	1.38	1.21	0.47
Range	2.90	3.51	1.22	0.63	3.30	3.12	2.21
Minimum	1.37	1.78	0.57	0.46	2.30	2.24	0.05
First-quantile	1.84	2.67	0.775	0.715	3.225	3.135	0.35
Second-quantile	2.55	3.4	0.81	0.81	3.95	4.31	0.87
Third-quantile	2.92	4.63	0.98	0.87	5.31	5.065	1.395
Maximum	4.27	5.29	1.79	1.09	5.60	5.36	2.26

5. 결 론

본 논문에서는 강우의 결측치를 추정하는 방법들 중 역거리 가중치법(IDWM), 역지수 가중치법(IEWM), 상관계수가중치법(CCWM), 인공신경망(ANN)기법, 레이더 자료를 이용한 결측치 추정방법을 비교하여 각각의 적용성을 검토하였다.

본 논문의 결과를 요약하면 다음과 같다.

(1) 결측치 추정방법 중 상관계수 가중치법(CCWM) (RMSE:0.57~1.79), 인공신경망(ANN) (RMSE:0.46~1.09), 보정레이더(RMSE:0.05~2.26)기법이 어느 정도 신뢰성이 있는 값을 보여주었다. 반면에 역거리 가중치법(IDWM), 역지수 가중치법(IEWM), RADAR 격자값 및 평균값은 RMSE의 값이 1.37~5.60의 범위를 보이며 상당히 큰 오차를 보였다.

(2) 일반적으로 레이더의 강우자료가 지점강우자료에 비해 정량적인 부분에 있어서 정확성이 떨어지기는 하지만, 이를 위한 해결책으로 기존에 설치되어 있는 지상강우 관측망과 강우레이더 관측망을 조합한 보정레이더를 사용하고 있다. 본 연구에서도 레이더 자료를 바로 적용하는 것 보다 보정레이더를 사용할 경우 결측치를 추정함에 있어서 더 타당한 RMSE값을 얻을 수 있었다.

(3) 레이더를 이용한 지속적이고 장기적인 강우자료의 구축으로 이러한 문제를 해결하게 된다면, 신뢰성 있는 결측치의 추정과 더불어 이를 이용한 강우-유출모형의 적용에 있어서도 타당한 결과를 유추할 수 있을 것으로 판단된다.

감사의 글

본 연구는 국토해양부가 출연하고 한국건설교통기술평가원에서 위탁시행한 건설기술혁신사업(08기술혁신F01)에 의한 차세대홍수방어기술개발연구단의 연구비 지원에 의해 수행되었습니다.

참 고 문 헌

김병식, 홍준범, 김형수, 최규현, 조건부 합성방법을 이용한 레이더 강우와 지상 강우자료의 조합, **한국수자원학회논문집**, 제27권, 제3B호, pp. 255-265, 2007.

김응석, 김형수, 김중훈, 점 강우량 결측시 보정방법에 관한 비교 연구, **한국수자원학회 학술발표대회논문집**, 한국수자원학회, pp. 374-381, 1999.

안상진, 김진극, 레이더 반사강도와 강우강도의 Z-R 관계식 산정, **한국수자원학회논문집**, 제36권, 제1호, pp. 13-21, 2003.

안상진, 연인성, 김광일, 신경망을 이용한 강우예측, **한국수자원학회논문집**, 제20권, 제5-B호, pp. 711-722, 2000.

윤강훈, 서봉철, 신현석, 크리깅 기법을 이용한 낙동강 유역 홍수강우의 공간해석 연구, **한국수자원학회논문집**, 제37권, 제2호, pp. 233-240, 2004.

황세운, 박승우, 장민원, 조영경, Co-Kriging 기법을 이용한 일강우량 공간분포 모델링, **한국수자원학회논문집**, 제39권, 제8호, pp. 669-676, 2006.

건설교통부, **전국 강우레이더 설치 및 홍수예보시스템 개선 기본계획 수립(3차년도)**, 2003.

Ling Liu, Ozsu, M. Tamer, *Encyclopedia of Database Systems*, Springer, 2009.

Griffith, D.A., Spatial Autocorrelation: A Primer. **Association of American Geographers**, Washington, DC, 1987.

Hodgson, M.E., Searching methods for rapid grid interpolation, **Professional Geographer** 411, 51-61, 1989.

Navone, H.D., Ceccatto, H.A., Predicting indian monsoon rainfall: a neural network approach, Climate Dynamics 10, 305-312, 1994.

Ramesh S.V. Teegavarapu and V. Chandramouli, Improved weighting method- deterministic and stochastic data-driven models for estimation of missing precipitation records, **Journal of Hydrology.**, 312, pp. 191-206, 2005.

Ramesh Teegavarapu, Mohammad Tufail, Lindell Ormsbee., Optimal function forms for spatial interpolation of precipitation data, **Enviromental Informatics Archives**, Vol.4, pp. 343-353, 2006.

Shepard, D., A two-dimensional interpolation function for irregularly spaced data, Proceedings of the Twenty-Third National Conference of the Association for Computing Machinery, pp. 517-524, 1968.

Sullivan, D.O., Unwin, David J., *Geographical Information Analysis*, Wiley, New York, 2003.

Vasiliev, I.R., *Visualization of spatial dependence: an elementary view of spatial autocorrelation*, In: Practical Handbook of Spatial Statistics, CRC Press, Boca Raton, 1996.

○ 논문접수일 : 2011년 03월 28일
○ 심사의뢰일 : 2011년 03월 31일
○ 심사완료일 : 2011년 06월 28일