

회귀기법을 이용한 도시홍수위 예측모형의 개발

Development of Urban Flood Water Level Forecasting Model Using Regression Method

정 동 국* / 이 범 희**
Jeong, Dong Kug / Lee, Beum Hee

Abstract

A regression water level forecasting model using data from stage and rainfall monitoring stations is developed to solve the difficulties which real-time forecasting models could not get the reliabilities by assuming future rainfall duration and intensity. The model could forecast future water levels of maximum 2 hours after using data from monitoring stations in Daejeon area. It shows stable forecasts by its maximum standard deviation is 5 cm, average standard deviations are 1~4 cm and most of coefficients of determination are larger than 0.95. It shows also more researches about the stationary of watershed which assumed in this regression method are necessary.

Keywords : flood warning model, urban area, regression model, real-time model, Daejeon

요 지

실시간 홍수예측모형의 구성에서 장래 강우 양상(지속기간, 강우강도 등)에 대한 가정으로 인하여 홍수예측의 신뢰성을 높이기 어려웠다는 점을 해결하기 위하여 현재까지의 강우, 수위 및 상류지역의 수위를 기반으로 장래 수위를 예측할 수 있는 회귀모형을 구성하였다. 대상유역인 대전광역시 도심하천 구간에서 각 수위 및 강우관측소들 간의 자료들을 활용하여 최대 2시간 후의 수위변화를 예측할 수 있는 모형을 구성하였다. 각각의 선행시간에 대하여 예측한 결과 실제 실측치를 예측하는 과정에서 표준편차가 최대 5 cm, 평균 표준편차가 1~4 cm에 머무르고 있는 점 및 결정계수의 값이 대부분 0.95 이상을 나타내는 점 등을 살펴보면 전체적으로 예보모형이 안정적으로 운영이 되고 있음을 알 수 있었다. 다만 본 회귀모형의 특성이 유역반응의 정상성을 가정함을 감안한다면 어느 정도 기간까지 정상성을 유지할 수 있는가의 문제는 추후 연구가 더욱 필요할 것으로 보인다.

핵심용어 : 홍수예보모형, 도시유역, 회귀모형, 실시간모형, 대전

1. 서 론

최근 들어 우리나라를 둘러싼 수문특성의 변화 및 도시의 발달은 돌발성 호우의 발생과 급격한 수위상승

효과로 인하여 도시유역에서의 홍수로 인한 피해가 급증하고 있다. 현재 홍수예측 업무에 활용하고 있는 수문학적 홍수예측모형인 저류함수 모형의 적용을 통해서 유역면적 300 km² 내외의 작은 대상 유역에 대한

* 한남대학교 공과대학 건설시스템공학과 교수
Prof., Dept. of Civil & Environmental Engrg., Hannam Univ., Daejeon 306-791, Korea
** 교신저자, 배재대학교 공과대학 건설환경철도공학과 부교수
Corresponding Author, Associate Prof., Dept. of Civil, Env. & Railroad Engrg., Paichai Univ., Daejeon 302-735, Korea
(e-mail: bhlee@pcu.ac.kr)

홍수예측 시 주요 예보지점까지의 도달시간이 매우 짧아 홍수예측 업무를 수행하는데 선행시간을 충분히 확보할 수 없는 단점이 있다. 더구나 과거에서처럼 제외지로부터 하천의 제방을 월류하여 침수가 발생하는 피해만이 아니라 하천공간 내의 하상도로 및 하상 주차장 등을 활용한 도시 활동이 늘어나는 추세에 따라 하상공간간의 침수여부에 대한 관심이 더욱 증진되고 있는 현실에서 단순히 범람여부의 예측만이 아니고 하상도로 운행 차량의 차단 및 주차된 차량의 안전지대로의 이동 등 피해감소를 위한 선행시간의 확보도 중요하게 되어 얼마나 빠른 시간 안에 충분한 선행시간을 두고 정확한 범람여부를 예측할 수 있는가가 더욱 중요한 관점이 되고 있다.

이러한 목적에 따라 간단한 입력자료 만으로도 홍수예측을 수행할 수 있도록 통계학적인 기법을 활용하여 홍수예보 업무에 활용할 수 있도록 하였으며, 연구 대상지역으로는 도시화된 하천유역으로서 대전광역시의 갑천 유역에 대하여 홍수 예보 모형을 구성하고 적용성을 검토해 보았다.

최근 국외에서의 연구 추세를 살펴보면 도시홍수 예보의 활용성을 높이기 위해 복잡한 도시유역의 홍수특성을 반영하도록 수리 및 수문모형의 연계를 통하여 유출모형을 개선하고 있는데 Liang *et al.* (1995)은 싱가포르에 대하여 SWMM을 적용하였고, Knebl *et al.* (2005)은 HEC-HMS와 HEC-RAS를 이용하여 도시유역에 대한 홍수 모의 체계를 개발하였다. 국내의 경우에 있어서는 신현석 등(2007)이 부산 온천천 유역을 대상으로 SWMM과 HEC-RAS를 이용하여 지속시간별 한계유출량과 홍수발생 강우량을 산정하고 이를 통하여 해안도시지역에서의 홍수예보체계 구축방안을 검토한 바 있다. 이와 함께 1990년대 이후에 수문수자원 및 환경공학 분야에서 인공신경망(Artificial Neural Networks: ANNs) 이론의 활발한 적용은 강우-유출과정과 같은 복잡한 비선형 과정들의 모형화에 대한 적용타당성을 나타내고 있으며(박성천과 하수정, 2003),

예측을 위한 도구로서 그 활용방법이 다양하게 이용되고 있다. 지금까지의 수문수자원 분야에서 인공신경망 이론의 적용은 그 과정 및 결과에 있어 그 타당성이 검증되었으며, 이의 실질적인 적용을 위한 연구들 역시 홍수기의 실시간 댐 유입량 및 하도홍수 예측모형의 구축(윤강훈 등, 2004a; 2004b)을 통해 이루어지고 있다(박성천 등, 2006). 그러나 이러한 이론적 검증과 실질적인 적용을 위한 연구에도 불구하고 인공신경망에 대한 이론적인 체계가 정립되어 있지 않은 현장실무자가 쉽게 접근하기 어렵다는 점과 상용모형의 적용을 위한 다양한 매개변수 추정과정의 어려움, 미래의 수위변동을 예측하기 위한 장래의 강우량 및 강우 형태 등에 대한 가정이 필요하다는 문제가 발생할 수 있으므로 본 연구에서는 현재까지의 실시간 강우 및 수위 관측 자료로부터 수위의 변동을 직접 예측할 수 있는 간단한 방법을 제시하도록 하였다.

2. 대상수계 현황

대전시는 갑천, 유등천, 대전천 등 3대 하천이 도심을 관통하고 있고, 대청호가 주변에 위치하고 있어 타 도시에 비해 풍부한 수자원을 보유하고 있다. 유등천과 대전천이 모두 갑천의 지류이므로 공식적으로 갑천 유역으로 표시되고 있는 이 지역은 전국적으로 보았을 때 상대적으로 큰 홍수재해로 알려진 바는 없으나 140만 명의 거주인구가 이들 하천을 중심으로 생활하고 있으며, 1971년 이래로 1985년 1,224명, 1987년 2,337명의 이재민을 내는 등 매년 크고 작은 홍수피해를 입고 있어 이 지역에 대한 홍수예보 시스템의 개발이 필요하고 보았다.

대전시의 하천 현황은 Table 1에 나타나 있는 바와 같이 금강, 대전천 등 국가하천과 구계천 등 지방하천으로 구성되어 있으며, 이 중 대전천은 도심을 통과하며 유역면적 87.8 km², 지방하천 구간 연장 7.7 km의 하천으로 유등천에 합류된다.

Table 1. The Present State of Channels in Daejeon

classification	Channels	Length (km)	Repair record			Remark
			Total Length	Repair	%	
Total	31	203.6	312.9	232.9	74.4	Repair means both banks length
National stream	4	74.7	67.4	58.1	86.2	Geum R., Gap cheon, Yudng cheon, Juwon cheon
	1	7.7	14.9	14.4	96.6	Daejeon cheon
Regional	26	121.2	230.6	160.4	69.6	Gugye cheon <i>et al.</i>

Table 2. Informations of Monitoring Stations (WAMIS, 2009)

Group	Name	No.	Type	Agency	Open	Address
Rainfall stations	Daejeon	⑧	S/R	KMA	1969-01-01	Daedug-Gu Gooseong 2-Dong Daejeon Meteorological Administration
	Bangdong	①	T/M	MLTM	2001-05-01	Yuseong-Gu Doogye Br.
	Yuseong	⑥	T/M	MLTM	2001-05-01	Yuseong-Gu Mannyun Br.
	Indong	④	T/M	MLTM	2001-05-01	Dong-Gu In-Dong Inchang Br.
	Hoidug	⑦	T/M	MLTM	1993-06-01	Daedug-Gu Wonchon Br.
Stage Stations	Gasowon	⑤	T/M	MLTM	2001-05-01	Seo-Gu Gasowon-Dong Gasowon Br.
	Bangdong	①	T/M	MLTM	2001-05-01	Yuseong-Gu Doogye Br.
	Bogssoo	③	T/M	MLTM	2001-05-01	Seo-Gu Bogssoo-Dong Bogssoo Br.
	Yongchon	②	T/M	MLTM	2001-05-01	Seo-Gu Yongchon Br.
	Yuseong	⑥	T/M	MLTM	1985-03-01	Yuseong-Gu Mannyun Br.
	Indong	④	T/M	MLTM	2001-05-01	Dong-Gu In-Dong Inchang Br.
	Hoidug	⑦	T/M	MLTM	1917-06-01	Daedug-Gu Wonchon Br.

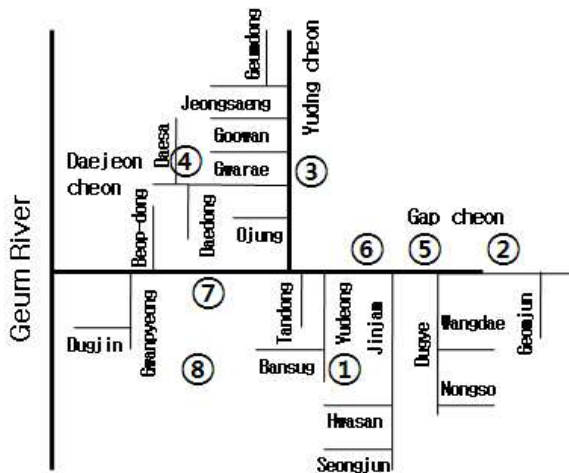


Fig. 1. Schematic Diagram of Channels and Stage/Rainfall Stations in Daejeon (Daejeon, 2003)

대전광역시 3대 하천 및 주변 수위/강우 관측소의 개략적 위치도는 Fig. 1과 같다. 그림의 원번호는 각각 강우 및 수위관측소의 위치를 설명하고 있으며 각종 관측소의 정보는 Table 2와 같다. 강우 관측소중 대전관측소를 제외한 나머지 4곳의 강우관측소는 수위관측소와 같이 위치하고 있다.

3. 회귀분석 이론 및 예측모형을 위한 회귀분석 기법의 적용

회귀분석(Regression Analysis)은 어떤 하나의 변수 값을, 다른 변수들의 값들을 이용하여 예측/제어하고

싶은 경우에 이용되는 주요한 수법이다. 예측하고 싶은 변수를 목적변수(종속변수)라 부르고, 예측에 사용되는 변수를 설명변수(독립변수)라 부른다.

회귀분석의 첫 단계로는 산포도(혹은 산점도)의 작성을 통하여 회귀모형을 직선(선형회귀)으로 나타낼 것인지 곡선(비선형회귀)으로 나타낼 것인지를 결정한다.

회귀식의 유효성(실제로 설명변수가 목적변수를 설명하는 정도)을 평가하려면, 기여율(R^2 ; 결정계수, coefficient of determination)을 계산해 보면 되며, 각 설명변수의 목적변수 y 에 대한 영향력의 대소를 판단하기 위한 편회귀계수의 유의성 검토에는 t 값과 p 값(유의 확률)을 사용할 수 있다. t 값이 높은 변수일수록 목적변수 y 를 예측(설명)하는 데에 공헌도가 높다고 하며, t 값을 대신하여 F 값을 사용하는 경우도 있다.

$$t\text{값} = \frac{\text{편회귀계수의 값}}{\text{계수의 표준오차}} \quad (1)$$

$$F\text{값} = (t\text{값})^2 \quad (2)$$

이외에 p 값에 의하여 편회귀계수의 유의성을 판정할 수 있으며, 이러한 과정을 통하여 걸러진 유의하지 않은 변수는 “목적변수 y 를 예측하는데 불필요한 변수이다”라고 결론짓는다. 경험적으로는 F 값이 2 이상이면 유효한 변수, 2 미만이면 불필요한 변수로 선택할 것을 권하고 있으나(노형진과 정한열, 2001) 이 기준은 어디까지나 경험적인 것으로 유효한 변수를 간과할 위험성을 적게 하고 싶다면 F 값이 1 이상의 것을 선택하더라

도 무방할 것이다.

회귀분석 기법을 적용하여 도시하천에서의 수위 상승을 예측하되 가용한 한정된 정보 (현재까지의 관측강우, 현재수위 및 상류지역의 관측수위)를 기반으로 장래 수위를 예측하여 홍수 예보를 제시할 수 있도록 하였다.

이를 위하여 대전광역시의 갑천, 유등천, 대전천 등 3대 도심하천 구간에서 각 수위 및 강우관측소들 간의 자료들을 활용하고, 현재시점까지의 관측 자료들을 이용하여 10분부터 20분, 30분, 60분, 90분, 120분 후의 수위변화를 예측할 수 있는 모형을 구성하였다. 연구 과정에서는 3시간 및 4시간 후의 수위변화에 대한 예측 모형까지를 구성해 보았으나 2시간을 넘어가는 경우의 예측은 그 결과의 정확성이 큰 의미를 갖지 못하여 최종적으로 제외하였다. 여기에서 기상청의 강우예보 자료를 활용하지 않은 이유로는 강우에 대한 예측이 총량 예보에 의한 방식으로 이루어지고 있어 본 예측 강우의 지속시간이 얼마인지를 정확히 추정하기 어려우며 기존 모의모형에서 요구하는 전체 강우자료의 입력이 어려웠다는 점(즉 실시간 예측에서 장래 강우 형태의 예측이 불가능함)을 들 수 있으며, 기존의 강우-유출 모형이 아닌 통계적 모형을 이용하여 홍수예경보 모형의 구성을 시도한 것은 강우의 형태, 강우 강도 등에 대한 정보가 없는 경우 기존의 강우-유출모형을 적용하기 어렵고, 이러한 모형을 적용에 따른 각종 매개변수의 추정 과정이 복잡하며, 이러한 모형의 수행 및 분석과정을 통하여 예보를 시행할 경우 돌발 홍수의 특성을 지니고 있는 도시홍수에 대한 충분한 예비시간(대피를 위한 선행시간)의 확보가 어렵다는 점을 들어 측정된 현재까지의 관측 자료로 부터 원하는 선행시간 이후의 수위변화를 직접 제시할 수 있는 모형을 구성하고자하는 목표를 이유로 들 수 있다.

이에 따라 대전광역시 유역에서의 수위, 강우자료들은 앞에서 제시하였던 강우 및 수위관측소들의 자료를 활용하도록 하였으며, 이 경우 Fig. 1 에 제시된 대전광역시 수위관측소들의 개략적인 배치상황으로부터 예측하고자 하는 대상지점의 수위예측은 그 지점의 현재까지의 수위자료와 상류에 위치한 수위관측소들의 자료, 그리고 주위의 강우자료들을 활용하여 예측하도록 하였다. 이에 따라 회귀분석에 의하여 수위의 변화를 예측하려는 식의 형태로는

$$Y_{i,n} = b_{i,n} + \sum_{t=0}^{p_i} \sum_{j=0}^{q_i} \Phi_{i,m,j,t} X_{j,t} + \sum_{t=0}^{p_i} \sum_{j=0}^{r_i} \Psi_{i,n,j,t} R_{j,t} \quad (3)$$

을 사용하도록 하였다.

- 여기에서,
- $Y_{i,n}$: 수위를 예측하고자 하는 대상지점 i 에서 n 시간단계 후의 수위
- $X_{j,t}$: j 지점 수위관측소에서 t 시간단계 전에 관측된 수위자료
- $R_{j,t}$: j 지점 강우관측소에서 t 시간단계 전에 관측된 강우자료
- $b_{i,n}$: 대상지점 i 에서 n 시간단계 후의 수위를 예측하기 위한 선형회귀모형에서의 상수
- $\Phi_{i,n,j,t}$: i 지점의 상류에 위치한 j 지점의 t 시간단계 전 수위자료와 대상지점 i 의 n 시간 후 예측 수위간의 회귀계수
- $\Psi_{i,n,j,t}$: i 지점의 주변에 위치한 j 지점의 t 시간 전 강우자료와 대상지점 i 의 n 시간단계 후 예측 수위간의 회귀계수
- t : 기존에 관측된 자료의 관측시간단계. 현재시간을 0으로 두었을 때 회귀식 계산을 위하여 사용된 j 지점에서의 p_i 시간까지의 선행자료 시간단계 ($t = 1, 2, \dots, p_i$), 본 연구에서는 현재, 10분전, 20분전, 30분전, ... 60분전까지 10분 간격으로 측정.
- q_i : i 지점 상류에 위치하고 있는 수위관측소의 수이다.
- r_i : i 지점 주변에 위치하고 있는 강우 관측소의 수이다. 다만 주변의 모든 수위 및 강우관측소를 포함하는 것이 아닌 선택된 강우관측소의 수로서 추후 변수 선택과정을 통하여 결정된다.

수위 및 강우관측소 지점은 앞의 Table 2 에서와 같이 1번 방동, 2번 용촌, 3번 복수, 4번 인동, 5번 가수원, 6번 유성, 7번 회덕지점으로 정의하였으며, 대상지점 인근 강우관측소를 중심으로 해석하는 과정에서 대전강우 관측소는 제외되었다. 대상 지점에 대한 수위의 예측은 10분부터 2시간(120분) 후까지의 수위변화를 10, 20, 30, 60, 90, 120분의 6단계로 예측하는 식들로 구성하였으므로 예측시간 단계 n 의 경우 1은 10분, 2는 20분, 4는 60분, 6은 120분 후의 수위를 예측하는 것을 의미한다. 수위 및 강우 관측값의 경우는 각각의 대상지점에 대하여 현재시각으로부터 1시간 이전까지 10분 간격의 관측 자료를 사용하는 것으로 하여 $t=0$ 은 현재, 1은 10분, 2는 20분 및 6은 60분전의 관측 값을 의미하도록 제시하여 이후 $R_{1,2}$, $X_{3,4}$, $Y_{5,6}$ 은 각각 방동강우관측소의 20분전 관측값, 복수 수위관측소의 40분전 관측값, 가수원 수위관

측소의 120분 후의 수위 예측값 등을 의미하는 것으로 제시하도록 하였다.

모형의 구성은 SPSS(Ver. 10.0)를 통하여 Fig. 2 와 같은 순서로 회귀분석 방법을 적용하여 보았다.

3.1. 자료의 특성분석 및 상관분석

모형의 구성 및 적용을 위하여 2005년 10월부터 2007년 12월까지 3년간의 강우 및 수위 관측자료(금강홍수 통제소 홈페이지, 2009; 국가수자원관리 종합정보홈페이지, 2009)를 모형의 구성에 사용하고, 이를 다시 2007년(모형 구성년도)과 2008년 자료에 대하여 적용을 하여보았다. 자료의 구분은 회귀분석 기법의 경우 대상특

성의 시간적 변동성이 없다고 보는 정상모형이므로 오래전의 자료에 대한 검토는 해석기법의 특성에 따라 의미가 없으며, 2005년 말부터 10분단위의 수위관측 정보가 제공되기 시작한 점을 고려하여 최근의 자료들만을 활용하여 모형을 구성하고, 이를 검증해보기 위한 개념으로 임의로 구분하였다.

2005년 말부터 2007년까지의 강우 및 수위관측 자료들 중 무강우 기간의 자료를 제외하고 실제 강우의 발생기간에 대한 강우 및 수위 관측 자료들만을 선정한 결과 Table 3 과 같이 각 관측소 별 10분 간격 관측 자료들 중 모두 17,550개(2,925시간 분량)의 자료를 해석하였다.

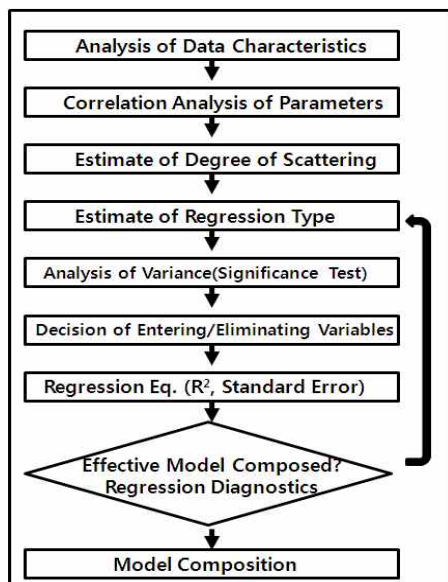
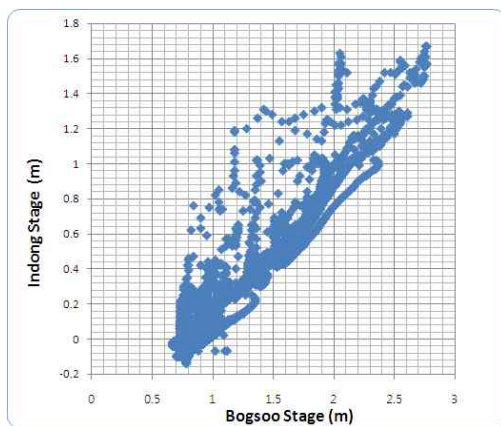


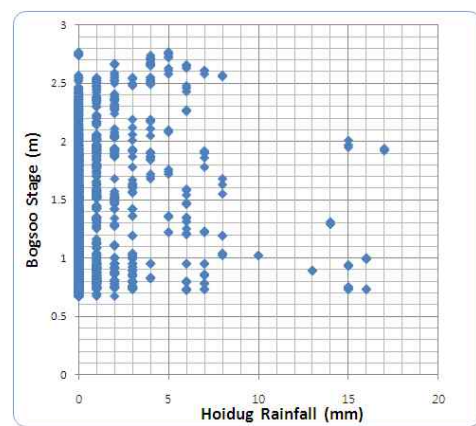
Fig. 2. Procedure of Regression Analysis

Table 3. Stage/Rainfall Data Characteristics

	Mean (stage: m, Rain: mm)	S.D.	Data (N)
X ₁ (Bangdong stage)	0.8373	.1770	17,550
X ₂ (Yongchon stage)	0.9016	.1474	
X ₃ (Bogsoo stage)	0.9239	.3088	
X ₄ (Indong stage)	0.0724	.2468	
X ₅ (Gasowon stage)	1.8870	.1030	
X ₆ (Yuseong stage)	1.2624	.2439	
X ₇ (Hoidug stage)	0.8864	.2840	
R ₁ (Bangdong Rain)	0.1350	.9062	
R ₄ (Indong Rain)	0.1364	.9565	
R ₆ (Yuseong Rain)	0.1304	.8080	
R ₇ (Hoidug Rain)	0.1308	.8774	



(a) Indong Stage vs Bogsoo Stage



(b) Bogsoo Stage vs Hoidug Rainfall

Fig. 3. Scatter Diagrams

회귀분석에서는 설명변수끼리 서로 독립일 것을 요구하고 있다. 그러나 계획적으로 수집한 데이터가 아니고 과거의 기록 등을 데이터로 사용하는 경우 설명변수끼리 완전히 무관한 상태를 확보하기는 매우 어려우며, 이 경우 설명변수간의 상관계수가 ±1이 되거나 이에 가까운 경우 신용도가 높은 회귀계수의 산출이 어려운 경우가 발생하게 된다. 이를 검토하기 위하여 각각의 지점별 예측시점 수위에 대한 설명변수인 관측시점 수위 및 강우자료(11개 강우 및 수위관측소별 7시점 자료로 모두 77개 설명변수) 간의 상관도를 구해보았다. 상관도가 -0.149 ~ 0.153으로 상대적으로 높게 나온 경우인 인동수위($X_{4,0}, \dots, X_{4,6}$)와 복수수위($X_{3,0}, \dots, X_{3,6}$) 관계 및 상관도가 -0.031 ~ 0.089로 낮게 나온 복수수위와 회덕강우($R_{7,0}, \dots, R_{7,6}$)의 사례를 Fig. 3에 제시하였다. 상관도가 높게 나온 일부 설명변수들의 경우에는 뒤의 진입 및 제거변수의 선정과정을 통하여 설명변수들을 제거해 줌으로써 다중공선성(multi-colinearity)문제를 대처할 수 있다.

3.2. 회귀식 형태의 선정 및 분산분석

잔차와 설명변수의 산포도로부터 회귀식의 형태를 검토할 수 있다. 본래 잔차와 설명변수의 산포도에서는

어떠한 경향도 보이지 않는 것이 바람직하다.(노형진과 정한열, 2001) 만일 설명변수와 잔차의 산포도에서 2차적인 경향 등이 나타난다면 회귀식의 구성에 해당 설명변수에 대한 자승항이 추가되어야 함을 의미한다고 볼 수 있다.

여러 설명변수들과 잔차의 산포도를 살펴보고 이들 중 가수원 수위와 잔차의 산포도를 Fig. 4에 사례로 제시하였다. 이와 같이 설명변수들과 잔차 사이에 어떠한 경향도 나타내고 있지 않으므로 선형회귀분석식을 사용하여 모형을 구성하는 것으로 하였다.

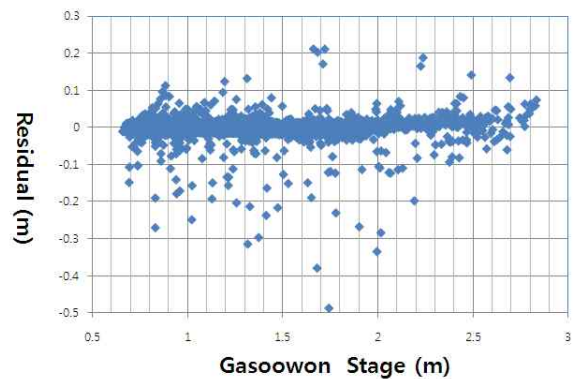


Fig. 4. Scatter Diagrams (Residual vs. stage)

Table 4. Regression Analysis Results of Bangdong Model

Model		Sum of square	Degree of Freedom	Mean square	F	significance probability
10 Minute	Regression	549.272	30	18.309	1485637	0
	Residual	0.216	17487	1.23E-05		
	Sum	549.488	17517			
20 Minute	Regression	548.7643	31	17.70208	430207.4	0
	Residual	0.71951	17486	4.11E-05		
	Sum	549.4838	17517			
30 Minute	Regression	547.975	32	17.124	198984.5	0
	Residual	1.505	17485	8.61E-05		
	Sum	549.48	17517			
60 Minute	Regression	544.089	28	19.432	63180.23	0
	Residual	5.379	17489	3.08E-04		
	Sum	549.468	17517			
90 Minute	Regression	536.792	26	20.646	28513.88	0
	Residual	12.665	17491	7.24E-04		
	Sum	549.457	17517			
120 Minute	Regression	527.835	23	22.949	18578.2	0
	Residual	21.61	17494	1.24E-03		
	Sum	549.445	17517			

분산분석의 결과는 회귀식의 유의성(통계적으로 의미가 있는지의 여부)을 검토할 수 있는 자료가 될 수 있다. 본 연구를 통하여 산정된 모형들의 분산분석 결과에서 F의 유의확률이 유의수준(0.05) 이내로 들어오는지 확인하여 유의성을 검토할 수 있다. 이곳에서는 Table 4의 방동지점에 대한 분산분석 결과에서 F의 유의확률이 유의수준 이내로 들어옴을 통하여 본 회귀식이 의미가 있음을 제시하였다.

3.3. 진입 및 제거변수의 선정

설명변수가 목적변수의 예측(설명)에 도움이 되는지를 Eqs. (1) and (2)를 적용하여 유의성을 판정하되, 유의하면($F \geq 2$) 유효한 변수로 판단하고, 유의하지 않으면 불필요한 변수로 판단하였다. SPSS에 의하여 축차변수선택법을 적용하면 변수의 단계적 선택에 따른

변수의 진입 및 제거가 이루어지게 된다. 방동 30분 예측모형의 구성과정에서도 방동수위 및 방동, 유성, 인동, 회덕 강우관측 자료들만이 유효한 변수들로 모형을 구성하는 것을 Table 5에 제시하였다. 그러나 이들 자료 군에서도 $X_{1,3}$, $R_{4,3}$, $R_{4,2}$ 등의 변수들은 t값의 비교를 통하여 제거되었다.

3.4. 회귀식의 제시 및 회귀진단

회귀분석 결과를 바탕으로 대전광역시외의 중심부를 관통하는 감천유역에 적용 가능한 도시 지류하천 홍수 예측시스템을 구성하였다. 회귀분석에 의하여 각각의 예측 대상 지점에 대한 회귀계수 산정 결과 중에서 방동 수위 예측을 위한 회귀모형은 Table 6과 같으며, 다른 지역에 대해서도 같은 방식으로 식을 구성하였다.

Table 5. Entering/Eliminating Results of Bangdong Model

selection order	var.	t	selection order	var.	t	selection order	var.	t
1	(Const.)	40.341	2	X1,6	103.919	3	X1,5	-27.258
4	R1,5	54.737	5	R1,6	44.428	6	R1,4	55.751
7	R1,3	19.626	8	R1,0	-10.958	9	R6,3	4.717
10	R1,2	7.231	11	R4,0	-10.061	12	R6,0	9.135
13	R4,5	-9.177	14	R6,5	5.719	15	R6,2	4.45
16	R7,4	5.101	17	X1,0	4.629	18	X1,4	-5.114
19	R7,1	-6.131	20	R6,1	8.954	21	R4,1	-5.705
22	R6,6	6.134	23	R4,6	-3.641	24	R7,3	5.007
25	R7,6	-3.898	26	R7,0	-3.527	27	R1,1	-1.932
28	R7,4	3.127	29	X1,3	-1.908	30	X1,5	-1.767
31	X7,5	-1.45	32	R6,4	3.756	33	R4,2	-3.683
Eliminating variables								
var.		t	var.		t	var.		t
X1,2		0.554	R4,3		0.699	R4,4		-1.147

Table 6. Regression Coefficients of Bangdong Model

		constant	Bangdong stage						
			X1,0	X1,1	X1,2	X1,3	X1,4	X1,5	X1,6
Bangdong	Y1,1	0.00392	0.02271	-0.04880	0.05984	0.02243	-0.15395	-0.44061	1.53328
	Y1,2	0.00949	0.03235	-0.03762		0.11997	-0.21670	-0.80923	1.89886
	Y1,3	0.01597	0.08174	-0.09876		0.11168	-0.19101	-0.99091	2.06646
	Y1,6	0.03528	0.22884	-0.19388			-0.27701	-0.93283	2.12910
	Y1,9	0.05432	0.48192	-0.37882			-0.31800	-0.86975	2.01453
	Y1,12	0.07243	0.56173	-0.34917			-0.25086	-0.83305	1.77825

Table 6. Regression Coefficients of Bangdong Model (Continued)

Bangdong Rain							
$R_{1,0}$	$R_{1,1}$	$R_{1,2}$	$R_{1,3}$	$R_{1,4}$	$R_{1,5}$	$R_{1,6}$	
0.00027	0.00026	0.00055	0.00124	0.00254	0.00180	0.00104	
-0.00060	0.00058	0.00106	0.00267	0.00569	0.00531	0.00346	
-0.00223	-0.00039	0.00148	0.00411	0.00950	0.00913	0.00738	
-0.00429	-0.00397	-0.00238	0.00302	0.01862	0.02054	0.01950	
-0.00344	-0.00286	-0.00208	0.00167	0.02139	0.02428	0.02559	
-0.00148			0.00267	0.02377	0.02656	0.02764	
Yuseong Rain							
$R_{6,0}$	$R_{6,1}$	$R_{6,2}$	$R_{6,3}$	$R_{6,4}$	$R_{6,5}$	$R_{6,6}$	
0.00069	0.00075	-0.00027	0.00022	-0.00069	0.00112		
0.00138	0.00172	0.00036		-0.00063	0.00105	0.00141	
0.00238	0.00231	0.00116	0.00118	-0.00042	0.00146	0.00157	
0.00378	0.00372	0.00328	0.00366	0.00436	0.00493	0.00403	
0.00368	0.00297	0.00295	0.00413	0.00866	0.00993	0.00916	
0.00236	0.00204	0.00258	0.00370	0.00966	0.01234	0.01221	
Indong Rain							
$R_{4,0}$	$R_{4,1}$	$R_{4,2}$	$R_{4,3}$	$R_{4,4}$	$R_{4,5}$	$R_{4,6}$	
-0.00038	-0.00016	0.00018	-0.00026		-0.00071		
-0.00114	-0.00053			-0.00039	-0.00118	-0.00049	
-0.00185	-0.00105	-0.00027			-0.00178	-0.00072	
-0.00310	-0.00124	-0.00108	-0.00079	-0.00058	-0.00173	-0.00081	
-0.00292				-0.00084	-0.00241	-0.00166	
-0.00213					-0.00150	-0.00126	
Hoidug Rain							
$R_{7,0}$	$R_{7,1}$	$R_{7,2}$	$R_{7,3}$	$R_{7,4}$	$R_{7,5}$	$R_{7,6}$	R ^{*2}
-0.00054	-0.00039	0.00090	0.00022		-0.00034		
-0.00055	-0.00112	0.00102	0.00117	0.00026	-0.00052	-0.00048	0.99900
-0.00071	-0.00124	0.00063	0.00108	0.00105	-0.00039	-0.00081	0.99700
	-0.00154			0.00107			0.99000
	-0.00087	0.00103		0.00112			0.97700
0.00147		0.00185		0.00213			0.96100

이에 따르면 방동지점의 30분 후 예측 수위는

$$\begin{aligned}
 Y_{1,3} = & 0.01597 + 0.08174X_{1,0} - 0.09876X_{1,1} + \\
 & 0.11168X_{1,3} - 0.19101X_{1,4} - 0.99091X_{1,5} + 2.06646X_{1,6} - \\
 & 0.00223R_{1,0} - 0.00039R_{1,1} + 0.00148R_{1,2} + 0.00411R_{1,3} + \\
 & 0.00950R_{1,4} + 0.00913R_{1,5} + 0.00738R_{1,6} + 0.00238R_{6,0} + \\
 & 0.00231R_{6,1} + 0.00116R_{6,2} + 0.00118R_{6,3} - 0.00042R_{6,4} + \\
 & 0.00146R_{6,5} + 0.00157R_{6,6} - 0.00185R_{4,0} - 0.00105R_{4,1} -
 \end{aligned}$$

$$\begin{aligned}
 & 0.00027R_{4,2} - 0.00178R_{4,5} - 0.00072R_{4,6} - 0.00071R_{7,0} - \\
 & 0.00124R_{7,1} + 0.00063R_{7,2} + 0.00108R_{7,3} + 0.00105R_{7,4} - \\
 & 0.00039R_{7,5} - 0.00081R_{7,6}
 \end{aligned} \tag{4}$$

와 같이 표시할 수 있음을 의미한다. 본 식의 계수들은 SPSS의 계산결과로 산정된 계수들을 그대로 제시하였으며, 7개 지점에 대하여 6가지 시점의 예측치를 구할 수 있는 총 42개의 식을 구성하였다.

4. 모형의 구성 및 적용

2005년에서 2007년까지의 강우-유출 자료에 대하여 선형회귀모형을 구성하고 이를 2007 ~ 2008년도 강우-유출 자료에 적용하여 본 모형의 예측모형으로서의 적용성을 검증하였다. 정밀도에 대한 판단은 표준편차(Standard Deviations)를 사용하여 나타내었다. 표준편차는 모의치인 z_t 와 참값 v 와의 관계를 이용하여 나타낸 값으로 본 연구에서는 예측치와 실측치를 비교하였다.

이외에도 최대절대오차, 최대상대오차 등을 제시하였다.

$$\text{절대최대오차} = \text{Max} | \text{예측치} - \text{실측치} | \quad (5)$$

$$\text{최대상대오차} = \text{Max} (\text{절대오차}/\text{실측치}) \quad (6)$$

앞에서 각각의 예보지점에 대하여 설정된 선형회귀분석식에 의하여 모의한 결과는 다음 Table 7 과 같다.

Table 7. Forecasting Results of 2007~2008 Cases

	Preceding time (Min.)	Max. abs. Error (m)	Max. rel. Error (%)	S.D. (m)	R ²
Gasowon	10	0.048	2.07	0.00181	1.000
	20	0.103	4.45	0.00296	0.999
	30	0.155	6.71	0.00418	0.998
	60	0.230	9.64	0.00730	0.995
	90	0.249	10.43	0.00892	0.992
	120	0.249	10.51	0.01067	0.988
Bangdong	10	0.109	11.95	0.00347	1.000
	20	0.136	13.19	0.00628	0.999
	30	0.201	17.46	0.00908	0.997
	60	0.433	28.48	0.01742	0.990
	90	0.693	45.60	0.02688	0.977
	120	0.733	46.98	0.03512	0.961
Yongchon	10	0.333	17.73	0.00415	0.996
	20	0.628	33.41	0.00807	0.993
	30	0.752	37.41	0.01146	0.989
	60	0.808	40.19	0.02003	0.976
	90	0.897	44.63	0.02729	0.959
	120	0.896	47.98	0.03331	0.941
Indong	10	0.388	747.76	0.01011	0.998
	20	0.675	1158.62	0.01814	0.995
	30	0.773	1502.51	0.02392	0.991
	60	0.766	2667.41	0.03443	0.980
	90	0.873	2703.09	0.04323	0.969
	120	1.133	2721.36	0.05226	0.955
Bogssoo	10	0.134	10.58	0.00483	1.000
	20	0.324	18.65	0.00932	0.999
	30	0.488	24.58	0.01345	0.998
	60	0.675	28.98	0.02355	0.994
	90	0.714	30.14	0.03187	0.989
	120	0.710	37.18	0.03971	0.983
Yuseong	10	0.910	34.61	0.00787	0.999
	20	0.877	33.07	0.01125	0.998
	30	0.844	31.60	0.01391	0.997
	60	0.732	33.98	0.01883	0.994
	90	0.673	35.76	0.02208	0.992
	120	0.593	30.95	0.02401	0.990
Hoidug	10	0.492	35.14	0.00723	0.999
	20	0.556	38.47	0.01281	0.998
	30	0.560	41.36	0.01821	0.995
	60	0.584	48.87	0.03110	0.986
	90	0.999	57.43	0.04181	0.974
	120	1.017	58.02	0.05100	0.961

전체적으로 2007년 1월부터 2008년 8월까지의 자료에 대하여 각각의 선행시간을 가지고 예측한 결과 실제 실측치를 예측하는 과정에서 표준편차가 인동 및 회덕지점에서 2시간 예측을 하였을 때 5 cm로 최대가 발생하였을 뿐, 나머지 지점 및 선행시간 예측의 경우에 있어서 표준편차가 1~4 cm에 머무르고 있는 점 및 결정계수의 값이 용chon지점 2시간 예측에서 0.941로 최저를 나타낼 뿐 대부분 0.95 이상을 나타내는 점 등을 살펴보면 전체적으로 예측시스템이 안정적으로 운영이 되고 있음을 알 수 있다.

여기에서 인동지점의 경우에서는 최대 상대오차가 1,150 %에서 2,720 %에 이르는 등 매우 높은 상대오차를 보여주고 있음을 알 수 있는데, 이는 인동지점의 경우 수위의 기준점이 실제 수위보다 높게 잡혀있어 관측치가 (-)값에서 (+)값을 오르내리는 과정에서 실측치가 0 m에 아주 근접한 값을 나타내는 경우 상대오차가 과

장되게 나타나는 현상이므로 큰 의미를 갖는 값은 아닌 것으로 보인다. 다만, 인동 및 회덕지점의 장기예측 및 유성지점의 단기예측에서 최대 절대오차 값(0.8 ~ 1.1 m)이 크게 나타나는 현상에 대해서는 그 원인을 살펴볼 필요가 있다. 선행시간 이전에 예측한 수위 값과 실제 시간에서의 유출수위를 비교하는 과정에서 이렇게 큰 오차가 발생하는 경우는 비가 오지 않는 상태에서 10분에서 120분까지의 선행시간 유출예측이 이루어지고 난 이후에 강우가 시작됨으로 인해서 나타나는 오류인 것으로 보인다. 즉, 유출의 예측시점에서는 현재까지의 강우 조건으로 예측하였으나 예측시점과 실제 유출이 발생하는 시점사이에 강우가 발생됨으로 이의 영향을 받은 실제 유출로 인하여 큰 오차가 발생하는 것으로 나타났다. 그러나 매 10분 간격으로 자료의 입력과 동시에 미래 수위를 재 예측하는 본 시스템의 특성상 강우의 발생과 동시에 기존의 예측치를 신속하게 수정해주

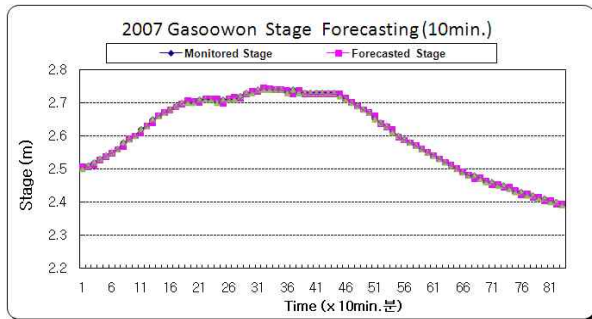


Fig. 5. Forecasting Result of Gasoowon (10 min.)

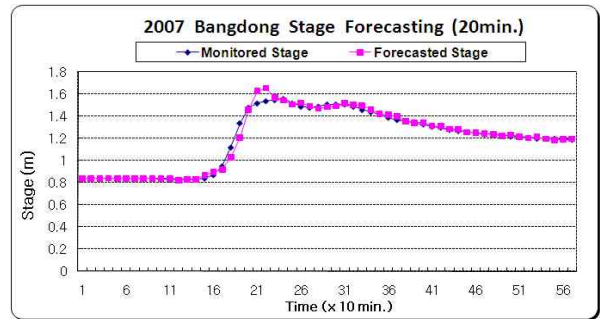


Fig. 6. Forecasting Result of Bangdong (20 min.)

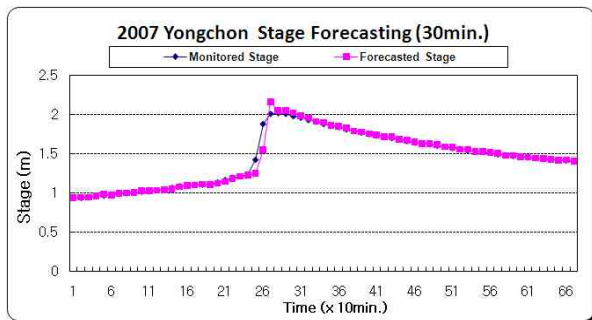


Fig. 7. Forecasting Result of Yongchon (30 min.)

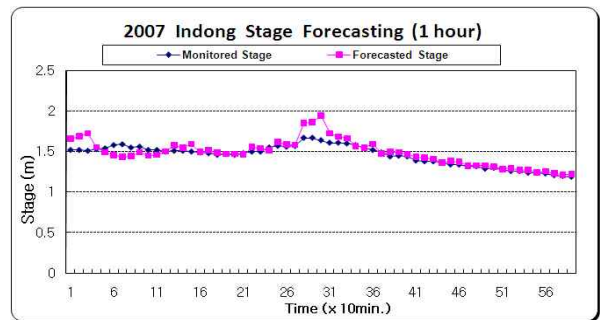


Fig. 8. Forecasting Result of Indong (60 min.)

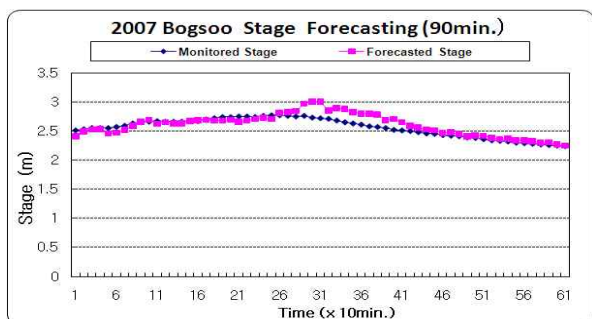


Fig. 9. Forecasting Result of Bogsoo (90 min.)

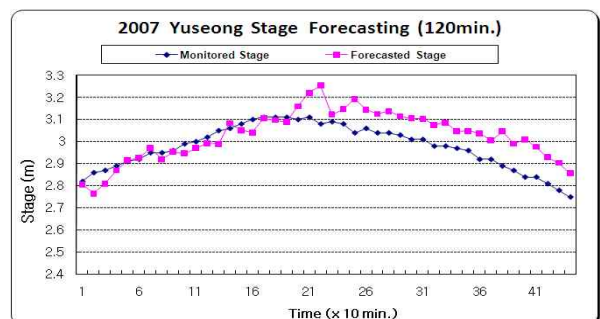


Fig. 10. Forecasting Result of Yuseong (120 min.)

므로 무강우로 인한 예측오차는 Fig. 5 ~ Fig. 10 에서와 같이 즉시 수정되어 큰 문제를 나타내지는 않는 것으로 보인다. 본 그림들은 2007년의 경우에서 각 지점 및 선행시간별 침투유출 수위의 실제 관측치와 모의양상을 나타내었다.

5. 결 론

그동안 실시간 홍수예측모형의 구성에서 가장 어려웠던 문제점으로서 실시간 홍수예측을 위하여 장래 강우 양상을 어떻게 가정할 것인가 하는 문제였다. 즉, 기존의 모의모형 기반의 홍수예측을 위해서는 강우 형태(지속기간, 강우강도 등)에 대한 정보를 가정하여야 하나 이의 가정에서 이미 많은 오차를 포함할 수 있어 이를 기반으로 하는 홍수예측의 신뢰성을 얻는 것이 어렵다는 점을 들 수 있다. 따라서 이러한 문제점을 해결할 수 있는 방안으로서 현재까지의 강우, 현재수위 및 상류지역의 수위를 기반으로 장래 수위를 예측하고, 이를 통하여 홍수 예정보를 제시할 수 있도록 하였다.

대상유역인 대전광역시의 갑천, 유등천, 대전천 등 3대 도심하천 구간에서 각 수위 및 강우관측소들 간의 자료들을 활용하고, 현재까지의 관측 자료들을 이용하여 최대 2시간 후의 수위변화를 예측할 수 있는 모형을 구성하였다. 대전지역 도시하천의 경우 강우 발생 후 도시하천 전반에서의 침투 유출의 발생이 2시간 이내에 이루어짐을 통하여 2시간 이상의 수위예측은 큰 의미가 없는 것으로 보였다.

각각의 선행시간을 가지고 예측한 결과 실제 실측치를 예측하는 과정에서 표준편차가 최대 5 cm, 평균 표준편차가 1~ 4 cm에 머무르고 있는 점 및 결정계수의 값이 대부분 0.95 이상을 나타내는 점 등을 살펴보면 전체적으로 예측시스템이 안정적으로 운영이 되고 있음을 알 수 있었다. 다만 본 연구에서 적용결과가 매개변수의 구성 후 3년까지는 큰 오류없이 적용되고 있음을 보여주고는 있으나 회귀모형의 특성이 유역반응의 정상성을 가정하여 구성된 것을 감안한다면 어느 정도 기간까지 정상성을 유지할 수 있는가의 문제는 추후 연구가 더욱 필요할 것으로 보인다.

감사의 글

이 논문은 2009년도 한남대학교 교비학술연구조성비 지원에 의하여 연구되었음.

참 고 문 헌

- 국가수자원관리 종합정보홈페이지(WAMIS) (2009). <http://www.wamis.go.kr/>, 한강홍수통제소
- 금강홍수통제소 홈페이지 (2009). <http://www.geumriver.go.kr/>, 건설교통부 금강홍수통제소
- 노형진, 정한열(2001). 한글 SPSS 10.0 기초에서 응용까지, 형설출판사
- 대전광역시(2003). 대전광역시 물수요 관리 종합계획 수립
- 박성천, 오창열, 김동렬, 진영훈(2006). “인공신경망 이론을 이용한 홍수유출 예측시스템 개발, -GULFFS 개발 및 적용-.” **대한토목학회논문집**, 대한토목학회, 제26권 제2-B호, pp. 145-152.
- 박성천, 하수정 (2003). “GANN에 의한 하천의 수질예측.” **대한토목학회논문집**, 대한토목학회, 제23권 제6-B호, pp. 507-514.
- 신현석, 박용운, 홍일표 (2007). “미계측 해안 도시유역의 홍수예정보 시스템 구축 방법 검토 -부산시 온천천 유역 대상-.” **한국수자원학회논문집**, 한국수자원학회, 제40권, 제6호, pp.447-458.
- 윤강훈, 서봉철, 신현석 (2004a). “신경망을 이용한 낙동강 유역 홍수기 댐유입량 예측.” **수자원학회논문집**, 한국수자원학회, 제37권 제1호, pp. 67-75.
- 윤강훈, 서봉철, 신현석 (2004b). “신경망을 이용한 낙동강 유역 하도유출 예측 및 홍수예정보 이용.” **한국수자원학회논문집**, 한국수자원학회, 제37권 제2호, pp. 145-154.
- Knebl M.R., Yang, Z.L., Hutchison, K., and Maidment, D.R. (2005). “Regional scale flood modeling using NEXARD rainfall, GIS, and HEC-HMS/RAS: a case study for the San Antonio River Basin Summer 2002 storm event.” *Journal of Environmental Management*, Vol. 75, pp. 325-336.
- Liong, S.Y., Chan, W.T., and Ram J.S. (1995). “Peak-flow forecasting with genetic and SWMM.” *Journal of Hydraulic Engineering*, Vol. 121, Issue 8, pp. 613-617.

논문번호: 09-104	접수: 2009.09.06
수정일자: 2010.01.13/01.29	심사완료: 2010.01.29