

자료기반 수문 예측모형의 성능평가 기법 개발

황석환*, 김종훈**, 정성원***

요 지

자료기반 수문예측 모형은 서로 자기상관이 다른 자료계열에 대해 예측결과만으로 모형의 성능에 대한 상대비교가 어렵다. 그러나 관측치와 예측치간의 평균 오차만을 기준으로 판단하는 기존의 모형 성능평가 기법은 대부분 이러한 자료기반 예측모형의 특성을 고려하지 못하고 있다. 따라서 본 논문에서는 자료기반 수문 예측모형의 성능을 보다 객관적으로 평가할 수 있는 새로운 모형 성능평가 기법인 상대상관계수(Relative Correlation Coefficient; *RCC*) 제시하였다. *RCC*는 자기상관계수에 대한 관측치와 예측치간의 상관계수의 비로 산정되며, 자기상관정도에 따라 예측성능의 결과가 달라진다. 본 논문에서는 다양한 자기상관을 가지는 선형, 비선형 자료계열에 대해 자료기반 수문모형을 적용하여 기존 모형평가 기법의 한계를 제시하였다. 그리고 기존의 성능평가 기법과 *RCC*를 비교분석하여 자료기반 수문예측모형의 성능평가에 있어 *RCC*가 보다 객관적이고 일관성 있는 성능평가가 가능함을 보였다.

핵심용어 : 예측, 수문모형, 성능평가

1. 서론

수문학에서, 일반적으로, forecasting은 실시간 예측을 의미하고, prediction은 불특정 시점의 상태를 추정하는 것을 의미한다(Klemes, 1986). 본 논문에서는 전자의 실시간 예측(forecasting) 모형의 성능 평가 기법에 대한 개선방법을 제시하고자 한다. 특히, 자료기반의 수문 예측모형의 성능을 평가하기 위한 기존의 방법들의 장단점을 평가하고, 이를 보완할 수 있는 방법을 제시하고자 한다. 수문학적 모형을 크게 두가지로 구분하면 물리적 기반의 수문모형과 자료기반의 수문모형으로 구분할 수 있다. 물리적 기반의 수문모형은 모형의 입력과 출력간의 관계가 수학적으로 명확하게 설명 가능하다는 장점이 있으나, 입력자료의 수가 많거나 입력자료와 출력자료간에 비선형성이 강한 경우 모형화가 쉽지 않다는 단점이 있다. 따라서, 근래 컴퓨터의 발달과 함께 부각된 모형이 자료기반 모형이다. 자료기반 모형의 대표적인 예는 신경망 모형이다. 이러한 자료기반 모형은 실시간 모형의 근간을 이루고 있으며 입력자료와 출력자료간의 관계에 강하게 의존한다는 점에서 자기회귀 모형과 밀접한 관계를 가지고 있다. 이러한 자료기반 모형은 주로 입출력간의 상관성을 규정하기 위해 학습 알고리즘을 사용하기 때문에 입력자료의 종류가 많고 비선형성이 강한 경우 물리적 기반의 모형에 비해 상대적으로 좋은 성능을 발휘한다. 그러나 입출력간의 관계가 블랙박스로 처리된다는 점과 대부분 자기학습 알고리즘의 한계인 지역해 문제가 존재한다는 단점을 가지고 있다. 모형의 성능을 평가하기 위한 가장 직관적인 방법은 자료를 도시하여 비교하는 방법이다. 그러나 판단 기준이 주관적이라는 단점이 있다. 모형의 성능을 보다 객관적으로 평가하기 위해서 일반적으로 사용하는 방법은 예측과 실측간의 오차를 측정하는 수학적 방법이다. 이러한 수학적 모형평가 방법은 목적에 따라 다양하게 개발되어 왔다. 그러나, 일반적으로 쓰이는 방법은 몇가지 정도에 불과하다. 최근, 기술의 발전과 관측자료의 축적으로 자료기반 수문 예측모형이 많이 사용되고 있다. 따라서 자료기반 수문 예측모형의 성능

* 한국건설기술연구원 수자원연구실 연구원

Researcher, Korea Institute of Construction Technology, GyeongGi-Do 411-712, Korea
(e-mail: sukany@kict.re.kr)

** 고려대학교 공과대학 건축·사회환경공학부 교수

Prof. Dept. of Architectural, Civil & Environmental Eng., Korea Univ., Seoul 136-713, Korea
(e-mail: jaykim@korea.ac.kr)

*** 한국건설기술연구원 수자원연구실 책임연구원

Chief Researcher, Korea Institute of Construction Technology, GyeongGi-Do 411-712, Korea
(e-mail: swjung@kict.re.kr)

을 객관적으로 검증하기 위한 성능평가 기법에 대한 고려도 필요하다. 그러나 현재, 자료기반 수문 예측모형의 성능을 평가하기 위해 사용하는 기법들은 일반적인 수문모형의 성능평가 기법과 구분 없이 적용하고 있는 것이 현실이다. 자료기반 수문 예측모형은 과거자료를 입력자료로 사용한다는 점에서 일반적인 개념적 수문모형과는 다른 특성을 가지고 있다. 자료기반 수문예측모형은 과거자료를 입력자료로 사용하기 때문에 자료의 자기상관 정도에 따라 그 예측 결과가 현저히 달라질 수 있다. 즉, 대부분의 실시간 자료기반 모형의 경우 입력계열과 출력계열 간의 명확한 물리적 역학관계가 아닌 상관성에 대한 의존도가 크기 때문에, 자료계열의 상관정도가 예측의 정확도에 기여하는 바가 매우 지배적인 경우가 많다. 따라서 입력계열과 출력계열의 상관성, 특히 선형상관, 이 강한 경우 예측의 정확도는 모형의 성능에 좌우되지 않는다는 점이다. 그러나 관측치와 예측치간의 평균 오차만을 기준으로 판단하는 기존의 모형 성능평가 기법은 대부분 이러한 자료기반 예측모형의 특성을 고려하지 못하고 있다. 따라서 본 논문에서는 자료기반 수문 예측모형의 성능을 보다 객관적으로 평가할 수 있는 새로운 모형 성능평가 기법인 상대상관계수(Relative Correlation Coefficient; *RCC*) 제시하고자 한다.

2. 이론적 배경

일반적으로 수문모형의 성능을 평가하기 위한 성능평가방법은 크게 세가지로 분류해 볼 수 있는데, 첫번째가 오차량 측정방법이고 두번째가 상대오차 측정방법 그리고 세 번째가 대조모델오차 측정방법이다.

2.1 오차량 측정방법(Scale-dependent error measures; *SDM*)

이 방법은 예측치와 실측치간의 오차 $e_i = y_i - \hat{y}_i$ 를 산정하여 예측의 정확도를 판정하는 방법으로, 평균제곱오차(mean square error; *MSE*)가 대표적인 예라 하겠다. 그러나 이 방법의 경우, 실측치와 예측치간의 평균오차(mean error; *ME*)나 평균제곱오차를 사용하기 때문에, 특히, 평균제곱오차를 사용하는 경우, 성능평가 결과가 이상치(outlier)에 크게 영향을 받으며 통계적 특성이 서로 다른 자료계열 간에는 상대비교가 어렵다는 단점이 있다.

2.2 상대오차 측정방법(Measures based on relative errors; *MBR*)

이 방법은 예측치와 실측치간의 오차를 기준치에 대한 오차(기준오차)의 비로 나누어 오차의 크기를 표준화한 방법이다. 상대오차는 $r_i = e_i/e_{\langle b \rangle i}$ 로 표현되며, 여기서 기준오차는 $e_{\langle b \rangle i} = y_i - y_b$ 로 산정된다. 일반적으로 기준치 y_b 는 기준모델의 실측치 평균, 이전 실측값 또는 0 등이 사용되고, 어떤 값을 기준치로 설정하는가에 따라 상대오차의 가중치를 조절하게 된다. 가장 대표적인 방법으로는 *MAPE*(mean absolute percentage error; p_i)가 있는데 식($p_i = 100 \times e_i/y_i$) 자체로만 보면 상대오차에서 기준치 $y_b = 0$ 을 사용한 경우로, 실측치의 크기로 표준화한 오차를 나타내며 일반적으로 백분율(%)로 표시한다. *MAPE*의 경우는 *RMSE*와 달리 제곱오차를 사용하지 않기 때문에 이상치에 대해 상대적으로 덜 민감하다는 장점이 있다. 그리고 기준치의 비로 환산된 값을 사용하기 때문에 서로 다른 자료계열간에 비교도 가능하다. 그러나, 기준오차가 0이 되는 경우 상대오차가 발산하게 되는 단점이 있다.

2.3 대조모델오차 측정방법(Relative Measures; *RM*)

이 방법은 상대오차 측정방법과 유사하나, 다른 점은 기준치에 대한 상대오차를 직접 측정하는 것이 아니고 기준 시나리오(benchmark model; 대조 모델)에 대한 모의 시나리오의 오차량 비를 측정하는 방법이다. 즉, 평균제곱오차 간의 비로 나타내며 이를 수식으로 표현하면 다음과 같다.

$$F^2 = MSE = \text{mean}(e_i^2) = \text{mean}\{(y_i - \hat{y}_i)^2\}, F_b^2 = MSE_b = \text{mean}(e_{\langle b \rangle i}^2) = \text{mean}\{(y_i - y_b)^2\} \quad (1)$$

$$\text{Relative variance} = \frac{\text{coefficient of agreement}}{\text{coefficient of efficiency}} = \frac{MSE}{MSE_b} = \frac{F^2}{F_b^2} \quad (2)$$

이 경우의 대표적인 예는 $RRMSE$ 와 NSC (Nash-Sutcliffe coefficient of efficiency, 1970)로 각각, 대조모델의 기준치 $y_b = 0$ 와 $y_b = \bar{y}$ 인 경우이다. 자기회귀 모형의 경우, 현재와 변동이 없는 상태(Naive)에 대한 예측의 정확도의 개선정도를 나타내기 위해 기준치로 이전 실측치를 사용($y_b = y_{i-1}$)하는 PI (Persistence Index)가 쓰이기도 한다. 즉, PI 는 예측하고자 하는 시점에 대하여 최소정보의 상태(현재의 상태)와의 비를 나타낸다. 그림 1은 대조모델오차 측정방법인 $RRMSE$, NSC 와 PI 간의 관계를 도식적으로 나타낸 것이다.

$$RRMSE(\text{Relative Root Mean Square Error}) = \sqrt{\frac{F^2}{F_b^2}} = \sqrt{\frac{\text{mean}\{(y_i - \hat{y}_i)^2\}}{\text{mean}\{(y_i - y_b)^2\}}}, y_b = 0 \quad (3)$$

$$NSC(\text{Nash-Sutcliffe } R^2) = 1 - \frac{F^2}{F_b^2} = 1 - \frac{\text{mean}\{(y_i - \hat{y}_i)^2\}}{\text{mean}\{(y_i - \bar{y})^2\}}, y_b = \bar{y} \quad (4)$$

$$PI(\text{Persistence Index}) = 1 - \frac{F^2}{F_b^2} = 1 - \frac{\text{mean}\{(y_i - \hat{y}_i)^2\}}{\text{mean}\{(y_i - y_{i-1})^2\}}, y_b = y_{i-1} \quad (5)$$

그러나, $RRMSE$, NSC , PI 도 모두 평균제곱오차(MSE)을 사용하기 때문에 아주 크거나 아주 작은 오차량에 대해 매우 민감하게 반응하는 단점이 있다. 이러한 측면에서 자료기반 모형(자기회귀 모형 포함)의 성능을 평가하기 위한 새로운 방법을 제시하고자 한다. 자료기반 모형은 입력자료와 출력자료의 상관관계에 의존하기 때문에 자료계열간의 상관정도에 따라 예측결과가 지배된다. 특히, 선형상관이 강한 경우가 두드러지는데, 선형상관이 강한 자료계열은 모형의 성능과 상관없이 작은 오차를 보이게 되는 것이 일반적이다. 예를 들어, 입력과 출력자료 계열의 상관계수(Pearson's correlation coefficient; r)가 0.99인 자기상관 모형을 사용하는 경우 선형회귀만으로도 충분히 높은 정확도의 예측이 가능하다. 이러한 경우에는 결과만으로 모형의 성능을 평가하는 것은 무리가 있다. 따라서

본 논문에서는 입력과 출력의 상관정도에 대한 예측과 실측치의 상관정도를 산정하여 모형의 성능을 평가하는 상대상관계수(Relative correlation coefficient; RCC)를 제안하고자 한다. 즉, 이 방법은 예측의 결과가 좋은 경우에 자료자체의 강한 선형성으로 인한 결과인지 아니면 모형의 성능이 우수한 것인지를 손쉽게 판단가능하게 해준다. 특히 비선형성이 강한 자료계열에 대해 예측 성능을 평가하는데 매우 유용하다. 왜냐하면 비선형성이 강한 자료계열의 경우 입력과 출력의 선형상관계수는 작을 것이고 모형의 성능이 좋다면 예측은 정확하여 실측과 예측의 선형상관계수는 클 것이기 때문에 당연히 RCC 는 커지고, 선형성이 강한 경우는 이와 반대의 결과를 보이기 때문이다.

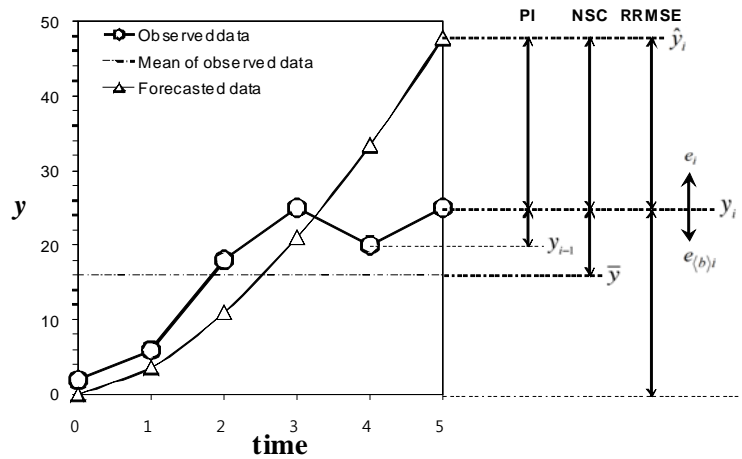


Fig. 1. the relationships among PI, NSC, and RRMSE

$$RCC = \frac{\text{degree of agreement}}{\text{degree of persistence}} = \frac{\left(\frac{\text{covariance}}{\text{multiplied standard deviations}} \right)}{\left(\frac{\text{autocovariance}}{\text{multiplied standard deviations}} \right)} = \frac{r_{\hat{y}y}}{r_{yyb}} = \frac{\frac{s_{\hat{y}y}}{\sqrt{s_{yy}} \sqrt{s_{\hat{y}y}}}}{\frac{s_{yyb}}{\sqrt{s_{yy}^*} \sqrt{s_{y_b y_b}}}} \quad (6)$$

3. RCC의 성능 평가

본 논문에서는 *RCC*의 성능을 평가하기 위해 4가지 자료계열과 1가지 자료기반 모형을 선정하였다. 자료계열은 자료자체의 선형성의 정도(입출력간의 선형성의 크기 및 입력자료 수 고려)에 따라 선형성이 강한 것부터 비선형성이 강한 것까지 총 4가지를 선정하였고, 모형은 비선형 모형 1가지를 선택하였다.

3.1 예측모형의 선정

선정된 모형은 SVM(Support vector machine)이다. SVM을 선택한 이유는 선형자료의 예측에서 비선형자료의 예측에 이르기까지 안정적이고 정확한 예측이 가능하다고 알려져 있기 때문이다. SVM을 간단히 설명하면, SVM은 신경망과는 다르게 철저히 통계적, 수학적 이론을 근거로 개발되었으며 그 단위 구조가 신경망과 유사하여 상당히 견고한 모형으로 평가받고 있다. SVM은 비선형적인 문제를 풀기위해 저차원의 입력공간(Input space)을 고차원의 특징공간(Feature space)으로 가정을 하게 된다. 다시 말해 단순 평면상에서의 비선형 자료도 적절한 개수의 임의의 차원을 추가하면 충분히 선형화시킬 수 있기 때문이다. SVM은 Vapnik(1995)에 의한 처음 고안되었으며 내부적으로 커널(Kernel) 이란 개념(신경망의 transfer function과 유사)을 적용하여 입력값을 적절한 출력값으로 계산한다. SVM의 장점은 이론적 기반이 두터워 결과에 대한 신뢰도가 높고 더 중요한 점은 가중치 매개변수를 결정하기 위한 함수가 2차식(quadratic function)으로 신경망의 최대단점중의 하나인 지역해(local mimia)에 빠질 우려가 없다는 점이다.

3.2 자료계열의 선정

자료계열은 자료기반 모형의 입출력간 선형성의 정도에 따라 4가지 선정하였다. 출력자료 기준으로, 선형성이 비교적 강한 자료로 St. Johns River의 Cocoa지점(USGS station number 02232400) 일평균 유출량 자료와 금강의 공주지점의 수위(상류의 금남과 공주간의 관계)를 선정하였다. 그리고 비선형성이 강한 자료로는 서울의 일급수량자료와 충주댐 일유입량 자료를 사용하였다. 여기서 언급한 자료의 선형성은 예측 모형 구현시 입력과 출력간의 상관관계를 의미한다. 표1은 본 논문에서 사용한 자료계열의 예측모형 구성을 위한 입출력 구성을 나타낸다.

Table 1. Data sets that were used as input and target data sets

Data sets		Unit	Length	AC	Lag of AC
(Case I) Daily mean stream flow					
Stream flow of Cocoa (Streamflow)	Total	ft ³ /s	12784	0.997	1 day
	Training	ft ³ /s	12418	0.997	1 day
	Forecasting	ft ³ /s	366	0.999	1 day
(Case II) Hourly water level					
Water level of GN (GN)	Total	m	22271	0.989	3 hour
	Training	m	20280	0.989	3 hour
	Forecasting	m	1991	0.984	3 hour
Water level of GJ (GJ)	Total	m	22271	0.992	3 hour
	Training	m	20280	0.992	3 hour
	Forecasting	m	1991	0.989	3 hour
(Case III) Daily water demand					
Water demand (WD)	Total	m ³	1827	0.855	1 day
	Training	m ³	1462	0.855	1 day
	Forecasting	m ³	365	0.840	1 day
Max. temperature (MT)	Total	℃	1827	0.957	1 day
	Training	℃	1462	0.957	1 day
	Forecasting	℃	365	0.960	1 day
Sunshine duration (SD)	Total	hour	1827	0.290	1 day
	Training	hour	1462	0.293	1 day
	Forecasting	hour	365	0.279	1 day
(Case IV) Daily mean inflow of dam					
Inflow (If)	Total	m ³ /s	6574	0.675	1 day
	Training	m ³ /s	6209	0.676	1 day
	Forecasting	m ³ /s	365	0.654	1 day
Rainfall (Rf)	Total	m ³ /s	6574	0.291	1 day
	Training	m ³ /s	6209	0.304	1 day
	Forecasting	m ³ /s	365	0.097	1 day

4. 적용 결과

그림2는 구성된 입출력 자료간의 선형성이 서로 다른 4가지 자료계열에 대해 SVM 모형으로 예측을 시행하여 실적치와 예측치 간의 *NSC*, *PI* 그리고 *RCC*를 산정한 결과이다. *AC*와

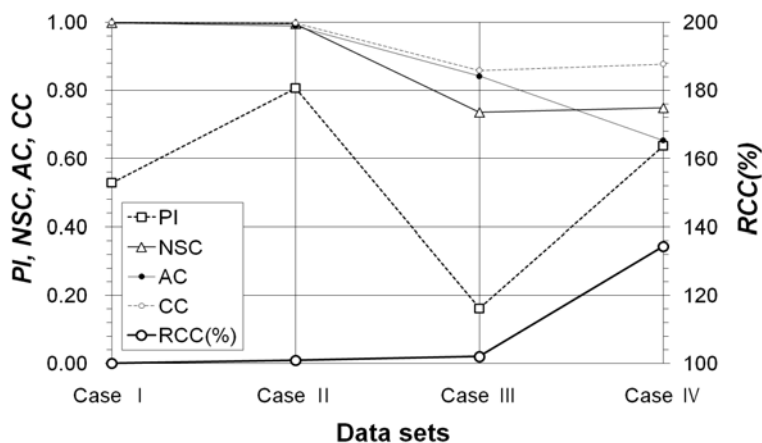


Fig. 2 The goodness-of-fit of RCC compared with PI and NSC as a validation test method to estimate the nonlinear forecasting efficiency with consistency

CC는 각각 대상 실측치의 자기상관계수와 예측치와 실측치 간의 상관계수를 의미한다. 그림의 결과를 보면 AC가 Case I에서 Case IV로 갈수록 작아지는 것을 볼 수 있다. 그러나 PI의 경우는 자료계열의 AC에 상관없이 변동하고 있음을 볼 수 있다. 즉, 자료계열의 자기상관 정도를 거의 고려하지 못하고 있다. 이는 PI의 경우는 대조모델의 기준치로 이전 시차의 값을 사용하기 때문에 정확도나 효율을 양적으로 측정할 수 있는 절대적인 표준치가 없

기 때문이다. 따라서 PI의 경우는 자료계열에 의존하는 상대적인 값이기 때문에 오차량 측정 방법과 마찬가지로 서로 다른 계열간의 직접적인 비교는 어렵다. 그리고 NSC의 경우도 AC가 작은 경우 자기상관 정도를 고려하지 못하고 있음을 볼 수 있다. NSC의 경우는 CC와 유사하게 거동하는데 이는 식 자체의 한계에서 나온 결과로 NSC의 경우 예측결과가 실측치와 어느 정도 부합하는지를 판별하는 데는 매우 훌륭한 성능을 보여준다. 단, AC를 고려한 결과가 아니라는 점이다. 따라서 NSC의 경우는 일반적인 수문모형의 예측의 정확도를 측정하는 데는 유용하나 자기회귀 모형의 예측의 효율을 측정하는 데는 적절하지 못하다. 이에 비해 RCC는 식 자체에서 AC와 CC를 고려하기 때문에 자기상관 정도 대비 예측의 정확도를 비교적 객관적으로 보여주고 있다. 그리고 100%를 기준으로 실제 자기상관에 대한 실측과 예측의 상관의 비를 나타내 주므로 직관적으로 자료계열의 특성과 모형의 예측성능을 동시에 RCC를 통해서 파악할 수 있다. 물론 RCC도 상한계가 없다는 점 등은 추가적인 논의가 필요할 것으로 보인다.

5. 결론

본 논문에서는 자료기반 수문 예측모형의 성능을 보다 객관적으로 평가할 수 있는 새로운 모형 성능평가 기법인 상대상관계수(Relative Correlation Coefficient; RCC) 제시하였다. RCC는 자기상관계수에 대한 관측치와 예측치간의 상관계수의 비로 산정되며, 자기상관정도에 따라 예측성능을 평가할 수 있다. 자기상관이 서로 다른 선형, 비선형 자료계열에 대해 비선형 자료기반 수문모형인 SVM을 적용하여 예측모형을 구성하였고 그 예측 결과를 이용하여 기존 모형 평가 기법의 한계를 제시하였다. 그리고 기존의 성능평가 기법과 RCC를 비교분석하여 자료기반 수문예측모형의 성능평가에 있어 RCC가 보다 객관적이고 일관성 있는 성능평가가 가능함을 보였다. 단, 모형의 성능평가기법은 개발된 목적에 따라 고유의 장점을 가지고 있기 때문에 모형의 성능평가에 있어 절대적이고 범용적인 기법은 없다. 따라서 객관적인 모형의 성능평가를 위해서는 여러 성능평가기법을 사용하여 종합적으로 판단하는 것이 중요하다.

6. 참고문헌

1. Nash JE, Sutcliffe JV. 1970. River flow forecasting through conceptual models: Part I—a discussion of principles. *Journal of Hydrology* 10(3): 282-290.
2. Vapnik, V. N. 1995 *The Nature of Statistical Learning Theory*. Springer Verlag, New York.