시계열 자료를 이용한 시단위 댐 유입량 예측

Estimation of Hourly Dam Inflow using Time Series Data

엄장일* · 정관수**

Eom, Jangil*, and Jung, Kwansue**

Abstract

In order to determine the amount of dam discharge and the release schedule when a hydrological event such as flash flood occurs, a multiple linear regression analysis based on time series data was performed to estimate the inflow. The input variables used for the multiple linear regression analysis model were rainfall, discharge amount, and the preceding inflow amount during a period of 20 years (1998 ~ 2017) at the Seomjin River Dam. For the review of the time series data, a method to remove the random variation component of the inflow due to the limit of the existing inflow calculation method was proposed, and the inflow estimation method was compared and analyzed before and after the inflow estimation. The prediction of the inflow using the empirical model, namely the multiple linear regression analysis model, was good, and the prediction accuracy of the peak inflow was particularly good.

Key words : Estimation of Dam Inflow, Time Series, Multiple Linear Regression

요 지

홍수와 같은 긴급한 수문사상 발생시 댐 방류 규모 및 방류 시점 결정 등 댐 운영의 핵심 자료인 유입량 예측을 위하여 시계열 자료 기반의 다중선형회귀분석을 시행하였다. 다중선형회귀분석모델에 사용된 입력변수로는 섬진강댐의 20년(1998~2017) 시단위 강수량, 방류량 및 선행 유입량을 사용하였다. 특히 시계열 자료의 검토시 기존의 유입량 산정방법의 한계로 인한 유입량 산정시의 불규칙 변동 성분을 제거하는 방법을 제시하였으며, 유입량 예측시 보정전후 유입량을 비교 분석하여 유입량 재산정 방법의 적정성을 확인하였다. 경험적 모델인 다중선형회귀분석모델을 이용한 유입량 예측 결과 단기 유입량 예측의 정확성이 양호하였으며, 특히 첨두유입량 예측의 정확도가 우수하였다.

핵심용어 : 댐 유입량 예측, 시계열 자료, 다중선형회귀분석

1. 서 론

댐 운영의 안정성 확보 및 수자원의 효율적 관리를 위해 댐 유입량의 정확한 산정은 가장 기본이 되는 사항이다. 특히 근래 기후변화로 인한 기록적인 수문현상의 발생 빈도 증가는 재난관리 측면에서 많은 불확실성을 야기하고 있다. 이에 본 연구에서는 빅데이터 기반의 시계열 수문자료를 활용한 Data Driven Model을 이용하여 시단위의 단기 댐 유입량을 예측코자 하였다. 과거 댐 유입량 산정시 주로

활용된 모델로는 실제 유역의 수문학적 응답을 물리법칙에 적용하는 물리적 모델과 수학적 연산을 통해 현황을 분석하는 개념적 모델이 주로 사용되었으나, 최근 오랜 기간 축적된 수문자료를 활용하여 변수들간의 인과관계를 추론하는 경험적 모델이 머신러닝 기법의 개발에 힘입어 활발히 연구되고 있다.

이와 관련된 선행 연구로 Hwang et al. (2007) 등은 서울시 일급수량과 충주댐의 일유입량 예측을 위하여 다중선형회귀분석 모델과 SVM 모델, ANN 모델을 이용하여 분석하였

*정회원, ¹ 충남대학교 박사과정

² K-water 차장(E-mail: uji@kwater.or.kr)

Member, ¹ Ph.D. Program, Chungnam National University

² Senior Manager, K-water

**교신저자, 정회원, 충남대학교 토목공학과 교수(Tel: +82-42-821-5671, Fax: +82-42-825-0318, E-mail: ksjung@cnu.ac.kr)

Corresponding Author, Member, Professor, Department of Civil Engineering, Chungnam National University

다. 각 모델들의 분석 결과에 대한 상관계수 및 평균제곱근오차(RMSE) 등의 평가 결과 SVM 모델이 타 모델에 비해 안정적이고 정확한 예측 결과를 보여 주었다. Ryu et al. (2012) 등은 충주댐의 일유입량 예측을 위하여, 웨이블릿 변환(Wavelet Transform)을 적용한 ANN 모델 분석을 시행하였다. 원시자료는 1990~2007 사이의 일강우량 및 일유입량 자료를 사용하였으며, Lag time을 고려한 입력자료 선정을 위해 절대상관계수를 산정한 결과 하루전, 2일전, 3일전의 유입량 및 강우량을 입력변수로 선정하였다. 일반적인 시계열 자료에는 경향성 및 주기성이 포함되어 있기에 선정된 입력자료를 웨이블릿 변환을 통해 디노이징(Denoising) 하였으며, ANN 분석은 디노이징 이전 자료와 이후 자료 두 가지를 대상으로 시행하였다. 시행 결과는 디노이징 자료의 R^2 (결정계수), RMSE (평균제곱근오차) 및 RRMSE (평균제곱근상대오차) 값이 보다 우수하였다. Lee et al. (2018) 등은 최근 기후변화 및 유역개발로 인하여 수문환경이 급격히 변화하고 있는 메콩강 유역 주요 지점의 유출량 모의를 위하여 물리적 기반의 수문모형인 SWAT 모델과 딥러닝 알고리즘의 Long-Short term Memory (LSTM) 모델을 이용하였으며, 유출모의의 정확도 및 두 모델의 장단점을 분석하였다. 분석결과 NSE계수(Nash-Sutcliffe Efficiency)는 SWAT 모델은 0.90, LSTM 모델은 0.99의 정확도로 나타났으며, 메콩강과 같은 대유역의 수문 시계열 자료의 모의를 위해서는 다양한 입력 자료를 요구하는 물리적 수문모형 대신 선행 시계열 자료의 변동성을 학습하여 예측하는 LSTM 기법 등 데이터기반의 심층신경망 모형의 적용이 가능할 것으로 분석하였다. Nagesh Kumar et al. (2004) 등은 인도 Hemavathi 강의 월별 유출량을 ANN 알고리즘인 앞먹임 신경망(Feed Forward Neural Networks) 기법과 재현신경망(Recurrent Neural Networks, RNN) 기법을 이용하여 예측하였다. 모델 분석시 노이즈나 관측시의 에러가 없는 양질의 자료 확보를 위하여 시계열 자료 생성 모델인 Autoregressive Moving Average (ARMA) 모델을 이용하여 데이터를 정상화하여 학습시킨 후 이후 실제 관측데이터와 비교하였다. 두 모델 예측값에 대한 R^2 및 RMSE 통계분석 결과 FFM 모델은 R^2 0.629, RMSE 295 m^3/s 의 값을 RNN 모델은 R^2 0.870, RMSE 180 m^3/s 으로 분석되어 재현신경망이 유출량을 보다 잘 모의하는 것으로 분석되었다. Ghumman et al. (2011) 등은 역전과 ANN 알고리즘을 적용하여 파키스탄 Hub 강의 수문자료를 기반으로 한 월별 유출량을 예측하였고 또한 종래의 수학기반 개념모델로 분석한 결과와 비교하였다. 역전과 ANN 모델 분석은 훈련단계를 통해 월별 유출량 예측을 시행하였으며, 검증단계와 테스트단계를 수행한 결과 종래의 개념적 모델과 분석결과가 매우 유사하게 나왔으나, 훈련단계에서는 두 모델 모두 실제값과 유사한 결과를 도출하였으나 검증단계와 테스트단계에서는 두 모델 모두 실제값과는 다소 차이가 발생하였다. 또한 Meher (2014)는

인도 중부 내륙에 위치한 Mahanadi 강 유역의 강우-유출량 분석을 위해 총 23개 강우관측소의 110년(1901~2010년)간의 월별 강우량 데이터를 주기성 시계열 분석모형인 Seasonal Autoregressive Integrated Moving Average (SARIMA) 모형과 ANN-MLP (Multi layer Perceptron) 모형을 이용하여 월별 강우량을 예측하고 SCS-CN 모델에 적용하여 월별 유출량을 예측하였다. 특히 ANN 분석시 과적합(over-fitting) 방지를 위해 분석데이터를 학습데이터(1901~1990, 관측자료수 1,080개), 훈련 데이터(1991~2002, 관측자료수 144개) 및 검증 데이터(2003~2010, 관측자료수 96개)로 나누어 분석하였다. 분석 결과 월별 강우량 예측은 SARIMA 모형과 ANN-MLP 모형의 성능이 비슷하였으나 유출량 예측은 ANN-MLP 모형이 SARIMA 모형보다 좋았지만 그다지 우수하지는 않았으며, ANN-MLP 모형의 정확도 향상을 위하여 보다 많은 입력자료와 알고리즘의 개발이 필요하다고 제안하였다.

인공신경망 모델의 경우 학습모델이 훈련 데이터에 지나치게 집중되어 신규 데이터의 적용시 분석 정확도가 떨어지는 과적합 문제, 지역해 수렴 문제 및 계산 절차를 파악하기 어려운 블랙박스모델이라는 제약사항이 있다. 본 연구에서는 전통적인 통계분석 모델인 다중회귀분석모델에 머신러닝 분석기법을 적용하여 유입량 예측을 시행하였다.

다중회귀분석모델은 아래와 같이 표현할 수 있다.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \epsilon \quad (1)$$

여기서, y 는 종속변수, x_i 는 독립변수들, β_0 와 β_i 는 회귀계수로 y 절편과 각 독립변수들의 가중치를 의미하며, ϵ 은 오차항이다. 다중회귀분석의 원리는 오차(ϵ)들의 제곱합이 최소가 되는 회귀계수 β_0 와 β_i 를 찾는 것으로 회귀계수의 추정기법으로는 최소자승법(least squares method), 최대우도추정법(maximum likelihood estimation) 등이 있다(Jung et al., 2018). 본 연구에서는 최소자승법을 이용하였으며, 회귀분석을 통한 수문자료의 분석관련 국내외 연구를 살펴보면, Jeong and lee (2010)는 대전광역시 도심하천을 대상으로 현재까지의 강우량, 수위 및 상류지역의 수위를 기반으로 최대 2시간 후의 수위변화를 예측할 수 있는 회귀모델을 구현하였으며, 평균 표준편차가 1~4 cm, R^2 의 값이 대부분 0.95 이상을 나타내어 예측시스템이 안정적으로 작동됨을 연구하였다. 또한 Patel et al. (2016) 등은 인도 중부 마디아 푸라데시주 발라켓 지역의 Wainganga 강의 14년(1994~2008) 간의 일강우-유출 자료를 이용하여 강우-유출에 관한 다중선형회귀분석모델을 개발하였다. 위 유입량 예측에 관한 기존 연구는 대부분 일단위의 예측으로 돌발홍수와 같은 급격한 수문변화시 시간위의 예측에 적용하기에는 다소 어려운 점이 있다.

댐 유입량 예측에 앞서 시계열 자료에 포함된 이상치 및 불규칙 변동 성분에 대한 검토를 수행하였다. 주요 검토

대상은 시계열 자료 중 유입량 자료로 기존의 댐 유입량 산정방식은 현실 여건상 유입량을 실시간 측정할 수 없기에 댐 방류량과 수위-저류용량 관계를 이용하여 유입량을 역으로 추산하는 방식(저류함수식)을 사용하고 있다. 즉, 수위가 변하지 않는 기간의 유입량은 방류량을 적용하고 수위가 최소독치인 1 cm 이상 변경시 댐의 수위-저류용량 관계를 만족시키기 위해 일시에 많은 유량이 유입되는 것으로 유입량을 산정하고 있으나, 실제 현상은 방류량을 초과하는 유입량에 의해 수위가 점진적으로 상승하게 된다. 이와 같은 유입량 산정방식의 특성으로 인하여 사전 강우나 방류량의 심한 변화가 없음에도 댐 수위 변동시 순간적으로 유입량이 과다하게 산정되는 문제가 발생하고 있다. 이러한 불규칙 변동 성분을 제거하기 위해 새로운 방식의 유입량 산정이 필요하다. 댐 유입량 산정 관련 연구로, Kim (2004)은 저류방정식에 의한 댐 유입량 산정시의 측정오차에 대한 개선을 위해 증발량을 고려한 유입량 산정방법, 상류 수위관측소 유량과의 상관관계를 분석하는 방법 및 평활화 기법을 적용하여 유입량의 측정오차를 줄이는 연구를 시행하였다. 본 연구에서는 시단위 유입량 산정시, 수위-저류용량, 댐 방류량 및 수위변화 시간을 고려한 유입량 산정 개선 방법을 제시하였다.

2. 대상 유역 및 연구 방법

2.1 대상 유역 현황

본 연구의 대상 유역인 섬진강댐은 우리나라 최초의 유역 변경식 다목적댐으로 연간 관개용수 및 생공용수로 435백만³을 공급하고 있으며, 칠보 수력발전소 및 하류 발전소를 통해 연간 188 GWh의 전력을 생산하고 있다.

Fig. 1. Dam Watershed

유역 현황으로는 유로연장은 35.8 km, 유역면적은 763 km², 저수면적은 26.5 km²이며, 행정구역상으로 임실군, 진안군, 순창군 및 정읍시에 분포하고 있다. 섬진강댐 강우관측소

는 Fig. 1과 같이 유역내 4개소가 있으며, 각 강우관측소의 관측개시일은 1990.12.15로 약 28년간의 수문관련 시계열자료를 확보하고 있다. 본 연구에서는 초기 시계열자료의 결측 등을 고려하여 1998~2017년간(20년)의 시계열자료를 대상으로 연구를 수행하였다.

2.2 연구 방법

시단위 댐 유입량 예측의 신뢰도 향상을 위해 시계열 자료에 대한 분석을 선행하였다. 강우량 자료는 계측시 자료의 검교정이 시행되고 있으며, 방류량의 경우 각 수문별 조건표에 의한 방류량 산정으로 특이한 이상치는 발견하지 못했으나, 댐 유입량의 경우 기존 저류함수법에 의한 유입량 산정방법의 한계로 불규칙 변동이 많이 발생하고 있다. 이에 유입량 자료의 변동성 개선을 위해 본 연구에서는 수위 상승시와 하강시를 구분하여 다음과 같이 유입량을 재산정하였다.

$$Inflow = O \pm S \frac{1}{3600} \frac{1}{\Delta t} \begin{cases} + \text{수위상승시} \\ - \text{수위하강시} \end{cases} \quad (2)$$

여기서, S 는 구간 저류량(m^3), Δt 는 수위변화시간(hr), O 는 방류량(m^3)이다. 즉, 각 수위별 저수용량과 수위변화 소요시간, 방류량을 이용하여 유입량을 재산정 하였으며, 재산정 유입량의 변동성 개선 효과 분석을 위해 통계분석을 시행하였다.

댐 유입량 예측을 위해 대표적인 통계분석 방법인 다중선형 회귀분석모델을 개발하였다. 입력자료로는 1998~2017년의 강우량, 방류량 및 선행 유입량을 이용하였으며, 모델의 신뢰도 확보를 위하여 시계열 자료를 training data (1998~2011), validation data (2012~2014), test data (2015~2017)로 분류하여 모델의 학습과 검증을 시행 후 최종적으로 개발 모델의 정확도를 테스트하였다. 유입량 예측은 섬진강댐 유역의 도달 시간(약 4.5시간)을 고려하여 최대 6시간까지로 한정하였으며, 1~6시간 후의 단기 댐 유입량을 예측을 위해 총 6개의 시나리오별로 모델 분석을 시행하였다. 모델 분석시 입력자료는 ‘강우량, 1시간에서 6시간 전 선행 유입량 및 방류량’으로 구성하였다. 시나리오별 모델 입력자료는 Table 1과 같다.

Table 1. Input Data by Scenarios

Scenario	Input data
# 1	rainfall + 1hr preceding inflow + outflow
# 2	rainfall + 2hr preceding inflow + outflow
# 3	rainfall + 3hr preceding inflow + outflow
# 4	rainfall + 4hr preceding inflow + outflow
# 5	rainfall + 5hr preceding inflow + outflow
# 6	rainfall + 6hr preceding inflow + outflow

또한, 당초 시계열 유입량과 재산정 시계열 유입량에 의한 유입량 예측값의 변화 비교를 위해 두 가지 케이스로 검토하였다. CASE 1은 원본 시계열 유입량 자료를 이용한 예측이며, CASE 2는 원본자료의 불규칙 변동 성분을 제거한 재산정 유입량을 이용한 예측이다.

3. 분석결과 및 고찰

3.1 시계열 자료 분석

기존 유입량 산정방식은 수위 변화가 없는 시간은 방류량을 유입량으로 간주하고, 수위 변화시에는 수위-저류용량 관계를 만족시키기 위해 유입량이 일시적으로 과다하게 산정하는 방식으로 Table 2 및 Fig. 2와 같이 유입량이 요동치는 현상을 볼 수 있다. 이러한 유입량의 노이즈는 댐 방류 규모 및 방류 시기의 결정에 있어 잘못된 의사결정을 초래할 수 있다.

Table 2. Time Serise Data (2017)

time	dam_level	rainfall	inflow	outflow
01-01 6:00	183.45	0	2.66	2.66
01-01 7:00	183.46	0	42.39	2.67
01-01 8:00	183.46	0	2.68	2.68
01-01 9:00	183.46	0	2.65	2.65
01-01 10:00	183.46	0	2.68	2.68
01-01 11:00	183.47	0	42.36	2.64
01-01 12:00	183.47	0	2.68	2.68
01-01 13:00	183.47	0	2.69	2.69
01-01 14:00	183.47	0	2.67	2.67
01-01 15:00	183.48	0	42.67	2.67
01-01 16:00	183.48	0	2.69	2.69

Fig. 2. Raw Inflow Plot (2017)

이에 2.2장에서 언급한 바와 같이 유입량을 재산정하였다. 2017년 시자료를 대상으로 원자료 유입량과 재산정 유입량을 통계분석한 결과 Table 3과 같이 최대 유입량과 평균

유입량은 매우 유사한 결과를 보여 주었으나, 변동은 488.64 → 417.21로 15%가 줄어들었으며, 표준편차는 22.11 → 20.43으로 8%가 저감되었다.

Table 3. Inflow Analysis (2017)

Item	Raw inflow	Recalculated inflow
max	294.91 m ³ /s [100%]	294.88 m ³ /s [100%]
median	3.88 m ³ /s	6.97 m ³ /s
Average	11.96 m ³ /s [100%]	12.25 m ³ /s [102%]
variance	488.64 [100%]	417.21 [85%]
standad deviation	22.11 [100%]	20.43 [92%]

이와 같이 유입량 재산정 결과 최대 유입량과 평균 유입량에서는 큰 변화 없이 유입량의 불규칙 변동이 제거되었음을 확인할 수 있다.

Fig. 3은 재산정 유입량 차트로 Fig. 2의 원본 유입량 차트와 비교시 불규칙 변동이 많이 제거되었고, 특히 저유량 구간의 유입량 트렌드를 보다 잘 표현하고 있다.

Fig. 3. Recalculated Inflow Plot (2017)

3.2 유입량 예측

입력 자료를 훈련데이터, 검증데이터, 테스트데이터로 분류하였다. 훈련데이터를 이용하여 회귀분석모델을 개발한 후 검증데이터에 적용하여 모형의 적합성을 판단하고 테스트데이터에 적용하여 모델을 평가하였다.

유입량 예측을 위해 개발된 다중선형회귀분석 모델은 시나리오별로 총 6개가 개발되었으며(Table 4), 대부분의 독립변수에서 p-value가 0.05이하였으며, 회귀식에 대한 F-statistic의 p-value는 모두 0.05이하로 개발된 다중회귀분석모델이 통계적으로 의미가 있는 것으로 판단된다. 각 시나리오별 예측 결과에 대한 통계량 분석(CC 상관계수, R² 결정계수, Relative Peak Error (RPE) 상대첨두치오차) 결과는 Tables 5, 6, 7과 같다.

Table 4. Multiple Regression Analysis Model (x_1 rainfall, x_2 outflow, $x_3 \sim 8$ preceding inflow)

Item	CASE 1	CASE 2
# 1	$y = -0.41 - 2.83x_1 + 0.12x_2 + 0.94x_3$	$y = -0.57 - 2.14x_1 + 0.10x_2 + 0.96x_3$
# 2	$y = -0.72 - 4.48x_1 + 0.12x_2 + 0.66x_3 + 0.30x_4$	$y = -0.60 - 2.51x_1 + 0.10x_2 + 0.89x_3 + 0.07x_4$
# 3	$y = -0.61 - 3.81x_1 + 0.11x_2 + 0.69x_3 + 0.37x_4 - 0.10x_5$	$y = -0.56 - 2.18x_1 + 0.10x_2 + 0.90x_3 + 0.13x_4 - 0.07x_5$
# 4	$y = -0.46 - 3.05x_1 + 0.11x_2 + 0.67x_3 + 0.42x_4 - 0.01x_5 - 0.14x_6$	$y = -0.47 - 1.67x_1 + 0.10x_2 + 0.88x_3 + 0.15x_4 + 0.07x_5 - 0.15x_6$
# 5	$y = -0.37 - 2.69x_1 + 0.11x_2 + 0.66x_3 + 0.41x_4 - 0.03x_5 - 0.07x_6 - 0.10x_7$	$y = -0.42 - 1.54x_1 + 0.10x_2 + 0.88x_3 + 0.16x_4 + 0.09x_5 - 0.08x_6 - 0.09x_7$
# 6	$y = -0.34 - 2.63x_1 + 0.11x_2 + 0.65x_3 + 0.42x_4 - 0.04x_5 - 0.06x_6 - 0.08x_7 - 0.03x_8$	$y = -0.42 - 1.53x_1 + 0.10x_2 + 0.87x_3 + 0.16x_4 + 0.09x_5 - 0.07x_6 - 0.09x_7 - 0.01x_8$

Table 5. Result of Statistical Analysis with Training Data (1998~2011)

Scenario	CASE 1			CASE 2		
	CC	R ²	RPE(%)	CC	R ²	RPE(%)
# 1 (1hr-prediction)	0.95	0.91	4.06	0.97	0.94	2.83
# 2 (2hr-prediction)	0.96	0.92	2.50	0.97	0.94	3.29
# 3 (3hr-prediction)	0.96	0.92	1.18	0.97	0.94	3.20
# 4 (4hr-prediction)	0.96	0.92	0.49	0.97	0.94	2.13
# 5 (5hr-prediction)	0.96	0.92	0.94	0.97	0.94	1.80
# 6 (6hr-prediction)	0.96	0.92	1.04	0.97	0.94	1.75
Average	0.96	0.92	1.70	0.97	0.94	2.50

Table 6. Result of Statistical Analysis with Validation Data (2012~2014)

Scenario	CASE 1			CASE 2		
	CC	R ²	RPE(%)	CC	R ²	RPE(%)
# 1 (1hr-prediction)	0.91	0.84	6.42	0.96	0.92	6.78
# 2 (2hr-prediction)	0.93	0.86	5.68	0.96	0.92	7.08
# 3 (3hr-prediction)	0.93	0.86	5.18	0.96	0.92	6.43
# 4 (4hr-prediction)	0.93	0.86	4.68	0.96	0.92	5.45
# 5 (5hr-prediction)	0.93	0.87	4.31	0.96	0.92	4.97
# 6 (6hr-prediction)	0.93	0.87	4.17	0.96	0.92	4.92
Average	0.93	0.86	5.07	0.96	0.92	5.94

Table 7. Result of Statistical Analysis with Test Data (2015~2017)

Scenario	CASE 1			CASE 2		
	CC	R ²	RPE(%)	CC	R ²	RPE(%)
# 1 (1hr-prediction)	0.63	0.39	5.83	0.86	0.74	4.28
# 2 (2hr-prediction)	0.71	0.50	6.09	0.87	0.75	4.44
# 3 (3hr-prediction)	0.70	0.49	5.60	0.87	0.75	4.17
# 4 (4hr-prediction)	0.70	0.50	6.27	0.87	0.75	4.94
# 5 (5hr-prediction)	0.71	0.50	4.65	0.87	0.76	3.26
# 6 (6hr-prediction)	0.71	0.51	4.66	0.87	0.76	3.29
Average	0.69	0.48	5.52	0.87	0.75	4.06

훈련데이터를 이용한 분석 결과는 Table 5와 같으며, CASE 1의 분석 결과 CC는 0.96, R²는 0.92, RPE는 1.70%대로 예측 결과가 양호하였으며, CASE 2의 분석 결과 CC는 0.97, R²는 0.94, RPE는 2.50%로 예측 결과가 매우 양호하였다. 전반적으로 재산정 유입량을 이용하여 분석한 케이스 2의 분석 결과가 좀 더 우수하였다.

검증데이터를 이용한 유입량 예측 결과는 Table 6과 같이 CASE 1의 분석 결과 CC는 0.93, R²는 0.86, RPE는 5.07%대로 예측 결과가 비교적 양호하였으며, CASE 2의 분석 결과 CC는 0.96, R²는 0.92, RPE는 5.94%로 예측 결과가 양호하였다. 전반적으로 검증데이터 또한 재산정 유입량을 이용하여 분석한 케이스 2의 분석 결과가 양호하였다.

테스트데이터를 이용한 분석 결과는 Table 7과 같이 CASE 1의 분석 결과 CC는 0.69, R^2 는 0.48, RPE는 5.52%대로 예측 결과는 좋지 않았으며, CASE 2의 분석 결과 CC는 0.87, R^2 는 0.75, RPE는 4.06%로 예측 결과는 비교적 양호하였다. 테스트데이터 또한 재산정 유입량을 이용하여 분석한 케이스 2의 분석 결과가 더욱 우수하였다.

전반적인 유입량 예측 결과, 재산정 유입량을 활용한 케이스 2의 분석 결과가 원본 유입량을 적용한 케이스 1 분석보다 예측 성능이 우수하였다. 다만 테스트데이터의 예측 결과와 상대적으로 저조하였다. 이에 대한 원인으로 테스트데이터 기간의 최대 유입량이 훈련데이터 및 검증데이터 기간에 비해 현저히 작아 분석 모델이 다소 낮게 유입량을 예측한 결과로 보이며, 이에 대한 대책으로 유입량 규모별 분석 모델을 개발하여 유입량에 따라 선택적으로 적용하는 방안도 향후 연구할 필요가 있을 것으로 사료된다.

- Training data 기간 최대 유입량: 3,970 m^3
- Validation data 기간 최대 유입량: 1,685 m^3
- Test Data 기간 최대 유입량: 492 m^3

예측시간에 따른 유입량 예측결과로는 예측시간이 길어 질수록 두 케이스 모두 예측 정확도가 다소 높아지는 현상이 발생하였다. 이는 입력자료의 추가로 모델의 학습 성능이 개선되는 현상으로 판단된다.

4. 결 론

댐 운영에 있어 가장 중요한 요소라 할 수 있는 유입량 예측의 경우, 이수측면에서는 일 또는 월단위의 유입량 예측으로도 댐 운영에는 지장이 없으나, 치수 차원을 고려시에 시간단위의 유입량 예측모델의 개발이 필요한 실정이다. 이에 본 연구에서는 통계분석기법인 다중선형회귀분석을 통해 유입량 예측모델을 개발코자 하였다. 과거 20년간의 섬진강 댐 시자료를 이용하여 모델의 입력자료를 구성하였으며, 예측모델 개발전 시계열 자료에 대한 사전 검토를 시행하여 유입량에 포함된 불규칙 변동 성분의 저감을 위해 수위변화 시간과 수위별 저수용량, 방류량을 고려한 새로운 댐 유입량 산정방법을 제시하였으며, 유입량 예측모델은 섬진강댐의 도달시간을 고려하여 1~6시간 후의 단기 유입량 예측모델을 개발 후 통계분석을 통해 모델의 적정성을 검토하였다.

연구 결과, 시계열 자료 분석과정에서 새로운 방법으로 재산정한 유입량의 경우 기존 유입량 자료와 비교시 최대 유입량과 평균 유입량의 변화량은 미미한 반면, 변동 및 표준편차 분석을 통해 유입량의 불규칙 변동 성분이 많이 제거되었음을 확인하였으며, 특히 저유량 구간의 유입량 트렌드를 보다 잘 분석하였다. 또한 유입량 예측 결과로는 재산정 유입량을 입력자료로 사용한 케이스2의 경우가 결정 계수 0.94~0.75, 상대점두오차 4.06~2.50% 수준으로 우수

한 분석결과를 보여 주었다. 본 유입량 예측에 관한 연구 결과는 댐 운영의 안정성을 확보하고 나아가 재해 관리 차원에서 활용 가능할 것으로 판단된다.

References

- Ghumman A.R., Ghazaw, Y.M., Sohail, A.R., and Watanabe, K. (2011). Runoff forecasting by artificial neural network and conventional model. *Alexandria Engineering Journal*, Vol. 50, No. 4, pp. 345-350.
- Hwang, S.H., Kim, J.H., and Jung, S.W. (2007). Hydrologic time series forecasting using SVM. *Proceedings of 2007 Conference*, Korea Water Resources Association, pp. 1972-1976.
- Jeong, D.K., and Lee, B.H. (2010) Development of urban flood water level forecasting model using regression method. *J. Korea Water Resour. Assoc.*, Vol. 43, No. 2, pp. 221-231.
- Jung, S.H., Lee, D.E., and Lee, K.S. (2018). Prediction of river water level using deep-learning open library. *J. Korean Soc. Hazard Mitig.*, Vol. 18, No. 1, pp. 1-11.
- Kim, D.I. (2004). *The improvement of inflow estimation in Yongdam Dam*. Master's thesis, Chungnam National University.
- Lee, G.H., Jung S.H., and Lee, D.E. (2018). Comparison of physics-based and data-driven models for streamflow simulation of the Mekong river. *J. Korea Water Resour. Assoc.*, Vol. 51, No. 6, pp. 503-514.
- Meher, J. (2014). *Rainfall and runoff estimation using hydrological models and ANN techniques*. Ph.D. dissertation, National Institute of Technology, India.
- Nagesh Kumar, D., Srinivasa Raju, K., and Sathish, T. (2004). River flow forecasting using recurrent neural networks. *Water Resources Management*, Vol. 18, No. 2, pp. 143-161.
- Patel, S., Hardaha, M.K., Seetpal, M.K., and Madankar, K.K. (2016). Multiple linear regression model for stream flow estimation of Wainganga River. *American Journal of Water Science and Engineering*, Vol. 2, No. 1, pp. 1-5.
- Ryu, Y.J., Shin, J.Y., Nam, W.S., and Heo, J.H. (2012). Forecast of the daily inflow with artificial neural network using wavelet transform at Chungju Dam. *J. Korea Water Resour. Assoc.*, Vol. 45, No. 12, pp. 1321-1330.

Received	September 27, 2018
Revised	October 1, 2018
Accepted	October 11, 2018