



# Development of leakage detection model in water distribution networks applying LSTM-based deep learning algorithm

Lee, Chan Wook<sup>a</sup> · Yoo, Do Guen<sup>b\*</sup>

<sup>a</sup>Master of Engineering, Department of Civil Engineering, The University of Suwon, Hwaseong, Korea

<sup>b</sup>Assistant Professor, Institute of River Environmental Technology, Department of Civil Engineering, The University of Suwon, Hwaseong, Korea

Paper number: 21-042

Received: 31 May 2021; Revised: 16 June 2021; Accepted: 16 June 2021

## Abstract

Water Distribution Networks, one of the social infrastructures buried underground, has the function of transporting and supplying purified water to customers. In recent years, as measurement capability is improved, a number of studies related to leak recognition and detection by applying a deep learning technique based on flow rate data have been conducted. In this study, a cognitive model for leak occurrence was developed using an LSTM-based deep learning algorithm that has not been applied to the waterworks field until now. The model was verified based on the assumed data, and it was found that all cases of leaks of 2% or more can be recognized. In the future, based on the proposed model, it is believed that more precise results can be derived in the prediction of flow data.

**Keywords:** Water distribution networks, Leakage detection model, Deep learning

## LSTM 기반 딥러닝 알고리즘을 적용한 상수도시스템 누수인지 모델 개발

이찬욱<sup>a</sup> · 유도근<sup>b\*</sup>

<sup>a</sup>수원대학교 토목공학과 공학석사, <sup>b</sup>수원대학교 토목공학과 하천환경기술연구소 조교수

## 요 지

지하에 매설되어 있는 사회기반시설물 중 하나인 상수도시스템은 정수처리된 물을 수용가에게 수송 및 공급하는 기능을 가지고 있다. 최근들어, 계측능력이 향상됨에 따라 유량데이터에 의한 딥러닝기법을 적용한 누수 인지 및 탐지와 관련한 연구가 다수 수행되고 있다. 본 연구에서는 현재까지 상수도 분야에 적용되지 않은 LSTM 기반의 딥러닝 알고리즘을 활용하여 누수발생에 대한 인지 모형을 개발하였다. 가정한 데이터를 기반으로 모형에 대한 검증을 수행하였으며 2% 이상의 누수가 발생한 경우에 대하여 모두 인식이 가능한 것으로 나타났다. 향후, 제안된 모형을 토대로 유량 데이터 예측부분에 있어서 보다 정밀한 결과가 도출 될 수 있을것으로 판단된다.

**핵심용어:** 상수도시스템, 누수인지모형, 딥러닝

## 1. 서 론

상수관망시스템은 지하에 매설되는 사회기반시설물 중 하나로 정수처리된 물을 수용가의 수도꼭지까지 수송 및 공급하는 기능을 한다. 우리나라 전국 상수도 보급률은 2019년 기준 99.3% (ME, 2021)로 선진국 수준의 양적확대를 이루었으나

시설물 경과연수가 지속적으로 증가되고 있는 실정이다. 상수도 시스템은 지하에 매설되어 있기 때문에 지속적 또는 순간적으로 발생하는 누수와 같은 물 손실에 대하여 빠르게 인지하고 대처하는 것은 매우 어렵다. 전세계적으로도 이러한 물 손실을 방지하는 것은 지속가능한 수자원 관리에 있어 매우 중요한 요소중 하나로 정의되고 있다(Melgarejo-Moreno *et al.*, 2019).

상수관망시스템에서 발생하는 누수사고는 그 규모가 커서

\*Corresponding Author. Tel: +82-31-229-8676  
E-mail: dgyoo411@suwon.ac.kr (D. G. Yoo)

눈으로 식별 가능한 중·대규모 누수와 특정 기술 또는 장비를 활용해 탐지할 수 있는 소규모 누수로 구분 가능하다. 중·대규모 누수의 경우 인근 주민들의 민원에 의해 사고가 발생한 위치를 인지할 수 있으나, 한번 사고가 발생되면 피해규모가 커질 수 있어 최대한 빠른 시간내에 사고를 인지할 수 있어야 한다. 소규모 누수의 경우 지하에 매설되어있는 상수관망시스템의 특성상 해당하는 위치를 발견하기까지의 지속기간이 길어 장기간의 기간동안 많은 양의 물 손실이 발생하게 된다. 따라서 소규모 누수의 경우 적극적인 누수탐사와 인지가 매우 중요하다 할 수 있다. 이와 같은 소·중·대규모 누수 모두 인지 과정을 토대로 위치파악, 주변관로 차단, 대응 등의 순으로 누수에 대한 보강을 수행하게 된다. 이러한 과정에서 시간을 가장 많이 줄일 수 있는 부분은 누수 인지 부분이라고 할 수 있다. 누수인지는 크게 누수량 산정, 누수위치 인지 및 탐지 그리고 누수제어의 과정으로 진행된다. 누수량 산정은 하향식 누수량 평가(Top-down leakage assessment)방식을 활용한 IWA 방법(Lambert and Hirner, 2000)과 야간 최소 유량(Minimum Night Flow, MNF) 분석에 의한 방법 등으로 구분하여 전통적으로 실무에 적용되고 있다. 이와 같은 누수량의 평가를 위해 Lambert (2001)는 수리학적으로 압력수두와 누수량의 관계를 통한 누수량 산정식인 FAVAD (Fixed and variable area discharges)를 제안하였다. 해당 식에 활용되는 매개변수에 대한 불확실성을 정량화하기 위하여 Lee *et al.* (2018)은 누수량 데이터를 통계적으로 분석한 바 있다. 이와 같은 연구는 상수관망 시스템의 전반적인 누수량 산정에 의해 상수관망시스템의 불확실성을 정량화하여 제시할 수 있다는 의의가 있으나 실시간적인 관점에서의 누수사고 인지에는 한계가 있다.

일반적으로, 누수에 대한 탐지방법은 장치기반 누수 탐지 기술(Hardware-based Techniques)과 모델기반 누수탐지 기술(Model-based Techniques)로 구분할 수 있다(Yoo *et al.*, 2018). 모델기반 누수탐지 기술은 다시 수리해석 기반 방법(Hydraulic model based techniques) 및 계측데이터 기반 방법(Hydraulic measurement based methods)으로 구분된다. 모두 각각의 장단점이 있으나 정확한 탐지 여부, 소요되는 인건비 등 다양한 방법별 한계점 역시 존재한다. 계측데이터 기반 방법은 계측시스템과 모델의 성능이 확보된다는 조건하에 인력이 최소화될 수 있는 방법이라 할 수 있다. 최근들어 계측데이터의 저장능력이 증대되고 이를 활용할 수 있는 다양한 방법들이 개발됨에 따라 상수도 분야에서도 이러한 데이터를 활용하는 것에 대한 연구가 활발히 진행되고 있다. 특히, 인공지능(Artificial Intelligence, AI)기법의 발달은 상수도시스템에서의 유량데이터로 하여금 특이치를 확인하는데 많은 활용

이 되고 있다(Mounce *et al.*, 2010; Romano *et al.*, 2014). 하지만, 상수도시스템의 경우 학습모형의 개발시 최신 딥러닝 기반 모형의 활용보다 인공지능망이 적용된 연구가 다수 수행된 바 있다(Mounce *et al.*, 2002; 2003; 2010).

본 연구에서는 딥러닝 알고리즘 중 하나인 순환신경망(RNN Recurrent Neural Network)의 경사도사라짐 문제를 개선한 LSTM (Long Short Term Memory Networks)기법을 활용하여 상수도시스템의 유량 예측능력을 향상시키고 보다 정확하고 신속한 누수인지모형을 개발하고자 하였다. 개발된 모형은 유량을 예측하고 결정된 임계값을 넘어설 경우 누수로 인지하는 예측-분류 기반의 누수인지모형으로 가정한 데이터를 기반으로 모형에 대한 검증을 수행하고 그 결과를 분석하였다.

## 2. 누수인지를 위한 딥러닝 모형

### 2.1 방법론

본 연구에서는 예측-분류 기반의 누수인지를 위하여 Fig. 1과 같은 방법론을 구축하였다. 먼저, 상수관망시스템에 설치되어있는 유량계측기에 의해 데이터를 확보한다. 확보된 데이터에 대한 전처리가 수행되고, 이를 기반으로 향후 발생하게 되는 유량을 예측한다. 상수관망시스템의 물 사용량은 주간/야간, 평일/주말 등 지속적으로 변동될 수 있다. 본 연구에서는 이에 대한 다양한 형상을 추출하여 유량 예측의 정확도를 향상시키고자 하였다. 다음으로 실제 계측값과 예측값간의 변동성을 확률적으로 산정하여 누수범위에 대한 임계값을 설정한다. 만일, 실제 계측된 유량값이 임계값을 넘는 경우 경보를 발생시키며, 상수관망시스템 어딘가에 물 손실이 크게 발생하였다는 것을 알리게 되고, 본 모형은 종료되게 된다.

### 2.2 LSTM (Long-Short Term Memory networks)기반 유량 예측

LSTM (Long-Short Term Memory networks)은 일반적으로 자연어 처리(Natural Language Processing, NLP)와 같은 정적인 시계열데이터를 다루기에 최적화된 인공신경망(Artificial Neural Networks, ANN)중 하나이다. 기존 순환신경망(Recurrent Neural Networks, RNN)에서 발생하는 경사도사라짐 문제(Vanishing Gradient Problem)를 보완한 최신의 딥러닝 기법으로 상수관망시스템에서의 계절적, 시간적으로 변동되는 유량을 예측하기에 보다 적합할 것으로 판단되어 본 연구에서는 유량 예측모델로 LSTM을 활용하고자 하였다. LSTM은 Fig. 2와 같이 은닉층(Hidden Layer)에 메모리블록

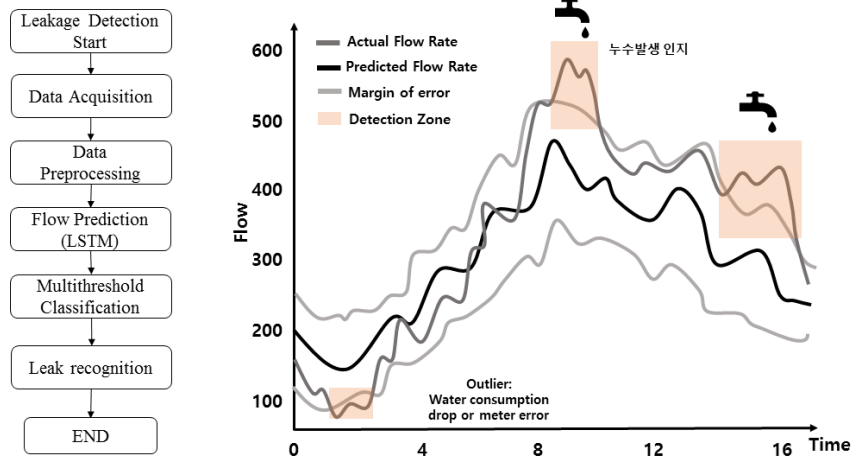


Fig. 1. Regional based drought resiliency assessment procedure

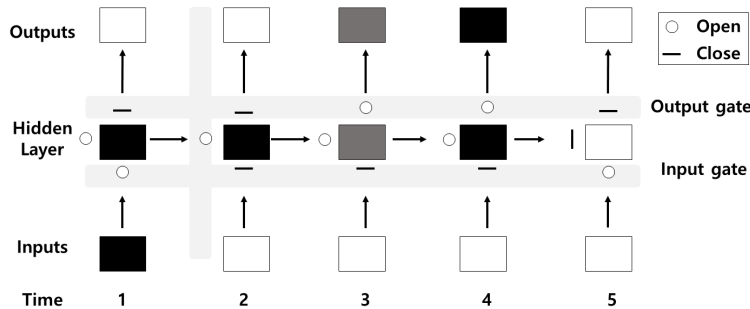


Fig. 2. Schematic diagram of LSTM for solving the vanishing gradient Problem

을 추가하여 장기적인 기억이 가능하도록 구성된 알고리즘으로 메모리블록이 열고 닫음을 실행하면서 시간“1”에 대한 시스템의 영향력을 꾸준히 유지하는 방법을 사용함으로써 경사도 사라짐 문제를 해결 하였다.

### 2.3 임계값 설정

일반적으로 시계열 데이터에 의한 탐지문제에는 관리도 (Control Charts)방법이 활용되곤 한다. 상수도시스템에서도 누수 인지를 위해서 다양한 관리도 방법이 적용된 바 있다. 관리도 방법은 데이터의 한계선을 설정한 그래프를 의미하며 본 연구에서는 계측된 유량자료의 상태를 활용하여 이상치를 식별하고자 하였다. 본 연구에서 활용한 슈하르츠 관리도는 Fig.3과 같이 표현될 수 있다. 중심선은 유량데이터의 시계열 자료를 의미하며 상하한의 임계값을 넘어가는 경우 특이치로 판별되게 된다. 슈하르츠 관리도(Shewhart, 1924)의 임계값은 일반적으로(관리하고자 하는 값(유량값))±3 × (관리하고자 하는 값의 표준편차)로 정의된다. 슈하르츠 관리도는 Eq. (1)과 같이 산출 가능하다.

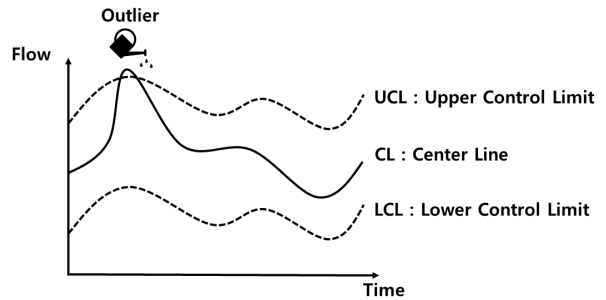


Fig. 3. Set of thresholds for burst recognition

$$\bar{\bar{X}} \pm A_3 \bar{S} \tag{1}$$

여기서,  $\bar{\bar{X}}$ 는 하위그룹의 평균 그리고,  $\bar{S}$ 는 하위그룹 표준편차의 평균을 의미한다. 이때,  $A_3$ 는 자료의 수(n)에 따른 관리도 계수이다. 이를 기반으로 상한과 하한의 임계구간을 설정하게 되고 임계구간을 넘어가는 수치(값)의 경우 특이치로 식별되게 된다.

### 3. 적용 및 결과

#### 3.1 가상 데이터 생성

본 연구에서 구축한 상수도시스템 누수인지 딥러닝 모형을 학습하고 성능을 평가하기 위해서는 실제와 유사한 계측 유량데이터를 생성할 필요가 있다. 이에 따라, 우리나라 D시에서 실제로 운용중인 상수관망시스템 블록중 하나에 대한 유량 자료를 활용하고자 하였다. D시 상수도시스템은 Fig. 4와 같이 1개의 배수지에서 285개의 절점에 물을 공급하고 있으며 373개의 관로로 구성되어 있다. 유량이 계측되는 지점은 배수지 앞단인 #D501로 가정하였으며 해당 배수지는 일 57,843톤의 물을 공급하고 있는 것으로 나타났다.

D시 상수도시스템의 계측지점에서 확보된 유량 데이터는 Table 1과 같다. 이는 EPANET2.0 (Rossman, 2000)기반의

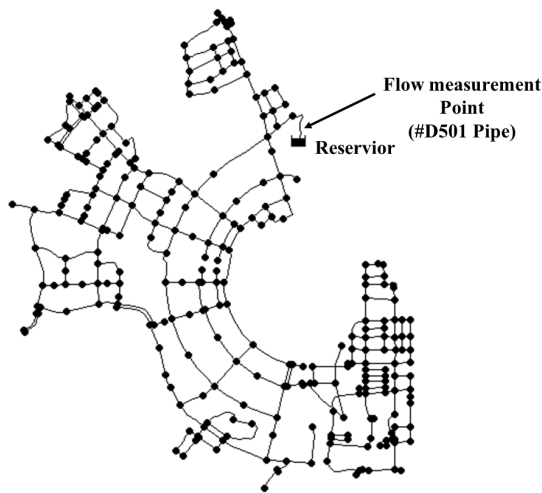


Fig. 4. Target water distribution network for inflow data generation

Table 1. Flow time-series data based on 24-hrs demand pattern

Time (Hours)	Flow (CMD)	Time (Hours)	Flow (CMD)
0:00	50902.64	12:00	64785.18
1:00	45118.25	13:00	64785.18
2:00	41647.62	14:00	61314.55
3:00	39333.86	15:00	59000.79
4:00	38755.42	16:00	58422.35
5:00	38755.42	17:00	60736.11
6:00	43382.93	18:00	64785.18
7:00	57843.91	19:00	68255.82
8:00	75775.52	20:00	67098.94
9:00	71148.02	21:00	63049.86
10:00	69412.7	22:00	61892.98
11:00	67677.38	23:00	54373.28
Average Flow			57,843.90

Table 2. Assumed demand multiplier for 24-hrs

Time (Hour)	Demand Multiplier
0:00	1.001
0:05	1.004
0:10	1.005
0:15	0.999
0:20	0.996
0:25	0.995
0:30	1.002
0:35	1.007
0:40	1.008
0:45	0.993
0:50	0.992
0:55	0.998

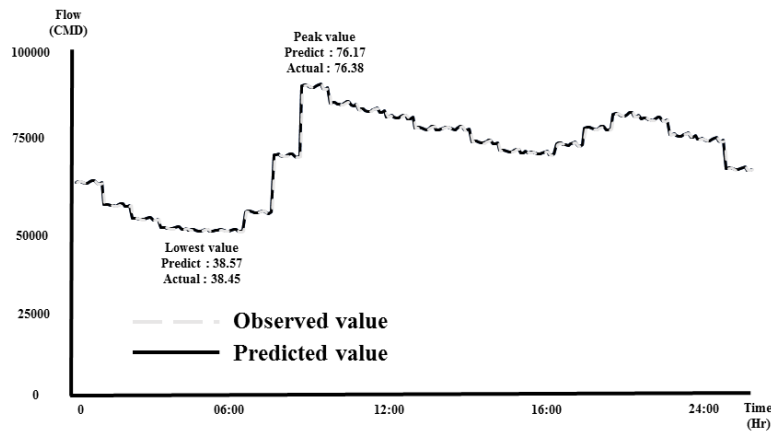


Fig. 5. Flow prediction results based on developed model

수리해석에 의한 결과이다. 기존 데이터의 경우 1시간 단위의 데이터로만 추출이 가능하였다. 본 연구의 목적은 계측에 의한 실시간적인 누수인지이므로 해당 데이터를 5분 단위의 시계열 데이터로 구축하였으며 각 시간단위 별로 평균이 1인 12개의 패턴을 적용하였으며 패턴의 값은 Table 2와 같다. 총 유량값의 변화는 발생되지 않도록 구축하였으며 구축한 계측 유량 데이터의 시계열 그래프는 Fig. 5와 같다. 기존 1시간 단위의 데이터에 의하면, 물 공급이 가장 피크(Peak)가 되는 시간은 08:00분에 발생하였으나 5분 단위 시계열로 변환 후 08:40분에 76,381톤의 물이 소비되는 것으로 최고점이 변경되었으며, 4:50분 및 5:50분에 약 38,445톤의 물이 소비되는 형태로 최저점 발생 시간이 변경되었다.

본 모형을 학습시키기 위해 앞서 가정한 5분 단위 데이터를 1월 1일부터 2월 28일까지 총 59일의 데이터를 만들었으며 2월 27일까지의 데이터는 LSTM 예측모형을 학습시키는데 활용하고 2월 28일의 데이터를 토대로 검증을 수행하고자 하였다.

### 3.2 누수발생 시나리오

가정된 데이터에 누수 사고 시나리오를 Table 3과 같이 수립하였다. 앞선절에서 설명한 것처럼, 대상시스템의 유입지점에 유량계측이 진행되며, 대상관망 내의 일부 지점에서 누수가 발생된 상황을 가정하였다. 첫 번째 시나리오는 0:00분

Table 3. Pipe burst Scenarios

Case Number	Time of Burst Occurrence	Flowrate (1,000 CMD)	
		Normal Condition	Added leakage flowrate
1	Random		1 ~ 15%
2	08:40 (Peak point)	Original flow	14% = 8.10
			12% = 6.94
			10% = 5.78
			6% = 3.47
			4% = 2.31
	04:50 (Peak point)		2% = 1.16

Table 4. Types of features for developed model

Feature	Input data
First period (daily interval sequence of t-1)	t-1-(24hr) × 7, t-1-(24hr) × 6, t-1-(24hr) × 5, t-1-(24hr) × 4, t-1-(24hr) × 3, t-1-(24hr) × 2, and t-1-(24hr) × 1
Second period (daily interval sequence of t-1)	t+1-(24hr) × 7, t+1-(24hr) × 6, t+1-(24hr) × 5, t+1-(24hr) × 4, t+1-(24hr) × 3, t+1-(24hr) × 2, and t+1-(24hr) × 1
Third period (daily interval sequence of t-1)	t-(24hr) × 7, t-(24hr) × 6, t-(24hr) × 5, t-(24hr) × 4, t-(24hr) × 3, t-(24hr) × 2, t-(24hr) × 1
Fourth period (daily interval sequence of t-1)	t-35min, t-30min, t-25min, t-20min, t-15min, t-10min, t-5min

부터 23:55분까지 기존 유량측정값 대비 1.01 ~ 1.15배의 누수가 발생하는 임의의 누수사고 10건을 두 번 발생 시켰다. 두 번째 시나리오는 08:40분과 04시 50분에 Table 3과 같은 누수사고를 가정하였다.

### 3.3 누수인지 모형 적용(가정 데이터 기반) 결과

누수인지 모형의 일부인 LSTM 모형의 경우 노드의 수를 128, 64, 48개로 구성하였으며 활성화 기능은 일반적으로 활용되고 있는 tanh를 활용하였다. 다음으로, 최적 학습률은 0.002, batch size는 60, epoch는 120을 적용하여 가중치를 총 7,200번 업데이트 하도록 하였다. 59일 동안의 데이터는 49일간의 학습 데이터 및 10일간의 테스트 데이터로 구분하였다. 시간별로 변화하는 유량의 변동성을 학습시키기 위하여 4가지의 시퀀스를 LSTM 모형에 추가로 적용하였으며 1주일 단위의 물 사용량 패턴을 고려하였다. Table 4의 첫 번째부터 두 번째 및 세 번째 형상은 평일 및 주말을 고려하고자 하는 식으로 t시점 기준으로 앞, 뒤 시점의 인접 시퀀스를 고려하게 된다. 네 번째 형상은 t+1 시점 예측시 t 시점을 기준으로 인접한 5분 단위 간격의 다양한 시점에서의 데이터를 활용하기 위한 식을 의미한다.

Fig. 5는 2월 28일 테스트 데이터에 대한 관측 값과 예측 값을 그래프화 한 것으로 구축된 LSTM 모형의 민감도를 분석하고자 10번의 학습을 수행하였으며 관측 값과 예측 값 간의 MAPE는 평균 0.098로 나타나 관측값에 의한 예측이 대체적으로 잘 이루어 지는 것으로 나타났다.

#### 3.3.1 (Case 1)적용 결과

모형의 성능 평가를 위해서 딥러닝 모형의 평가지표로 많이 활용되고 있는 TPR, FPR 등 Table 5와 같은 7개의 지표를 활용하였다. 각각의 지표는 TP (참양성 비율), TN (참부정 비율), FP (부정 참 비율), FN (부정 음성 비율)에 의해 산정될 수 있다.

Case 1의 경우 일평균 유량을 기반으로 1.01 ~ 1.15배의 누수량을 랜덤으로 적용 시킨 상황이다. 임의로 10건의 누수사고가 발생한다고 가정하였다. 개발된 모형의 적용 결과는

Table 5 및 Fig. 6과 같다. 가장먼저, 실제로 누수가 발생한 것을 정확하게 식별한 비율을 의미하는 TPR의 경우 70%로 10건중 3건은 오측을 하는 경우가 발생하였다. 이는 누수발생빈도가 매우 적고, 누수량이 비교적 적은 것이 원인으로 판단된다. 정확도를 의미하는 MAPE는 0.08로 100번의 누수사고건수 중 8번은 오류인 것으로 나타났다. 정상상황에서 누수로 식별된 비율인 FPR의 경우 0.73으로 도출되었으며 실제로 누수 사고가 아닌 상황을 아니라고 예측한 FNR의 경우 30%로 나타났다. 이는 TPR의 반대 의미로서 앞선 결과와 마찬가지로 적은 누수량이 적용된 것이 원인으로 판단된다. 적용된 모형의 감지 능력을 평가할 수 있는 지표인 ACC는 98.24%로 전체 시점중에서 누수가 발생한 시점을 위와 같은 확률로 인지할 수 있는 것으로 나타났다. 다만, 정확한 누수사고를 인지할 수 있는지에 대한 정밀도 지표인 PPV의 경우 77.78%로 정확도와는 상이한 결과가 도출되었다.

3.3.2 (Case 2)적용 결과

Case 2의 경우 두 가지의 시나리오로 구분될 수 있다. 첫 번째 시나리오는 하루 동안의 물 공급량 패턴 중 가장 적은 양의 물이 공급되는 시점, 두 번째는 첫 번째와 반대로 가장 많은 양의 물이 공급되는 시점(8:40)에서 누수량을 적용하여 모의하

는 것이다. 개발 모형 적용에 따른 성능평가 결과는 Table 6과 같다. 물 사용량이 가장 적은 시점(4:50)에서 실제 누수를 정확하게 식별한 데이터의 비율(TPR)은 평균 93.33%로 가장 물 사용량이 많은 오전시간(8:50)에 비해 20% 이상 정확하게 인지하는 것으로 나타났다. 이는 물 사용량이 적을 경우 변동성이 많은 시간대에 비해 적음으로 나타나는 형상으로 판단된다. 정상상황에서 누수로 식별한 오측의 비인 FPR의 경우 물 사용량이 적은 새벽시간에 더 낮은 0.02%의 비율을 보였다. 이외에도 FNR, TNR의 지표에서 모두 물 사용량이 적은 시점에서 더 좋은 누수인지 효율을 보였다. 개발된 누수인지 모형의 정확도는 물 사용량에 따른 시점과 무관하게 99%이상의 값이 도출되었다. 정밀도를 의미하는 ACC지표에서도 99% 이상의 신뢰도를 보이는 것으로 나타났다. 따라서 개발된 누수인지 모형의 성능 평가 결과 물 사용량이 비교적 적은 시점에서 발생하는 누수사고에 대한 식별 능력이 더 좋은 것으로 나타났으며 물 사용량이 많은 지점에서의 누수인지 효율은 비교적 떨어지는 것을 확인 할 수 있었다.

앞서 적용된 가정은 동일한 누수량에 대한 분석으로 두 시점에 대해 어느 정도의 누수량까지 식별 가능한지에 대한 분석을 수행하였다. 적용된 누수량의 변화에 따른 누수인지 여

Table 5. Results of performance indicators for Case 1

Performanc indicators (Unit: %)	Evaluation results
TPR (True Positive Rate)	70.00
FPR (False Positive Rate)	0.73
FNR (False negative rate)	30.00
TNR (True negative rate)	99.27
ACC (Accuracy)	98.24
PPV (Positive predictive value)	77.78
MAPE (Mean Absolute Percentage Error)	0.08

Table 6. Results of performance indicators for Case 2

Performanc indicators (Unit)	Evaluation results	
	04:50 (Peak time)	08:40 (Lowest time)
TPR (%)	93.33	73.33
FPR (%)	0.02	0.21
FNR (%)	6.67	26.67
TNR (%)	99.98	99.79
ACC (%)	99.95	99.70
PPV (%)	93.33	55.00
MAPE (%)	0.11	0.08

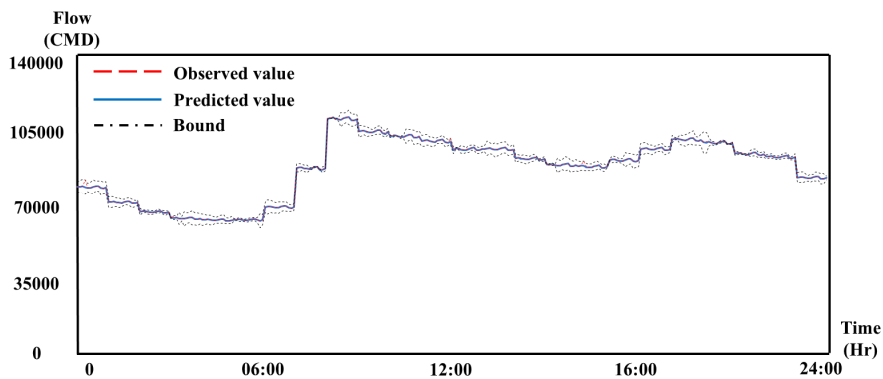
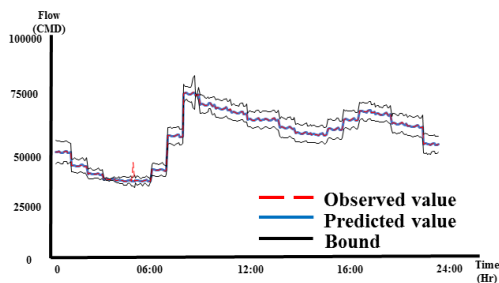


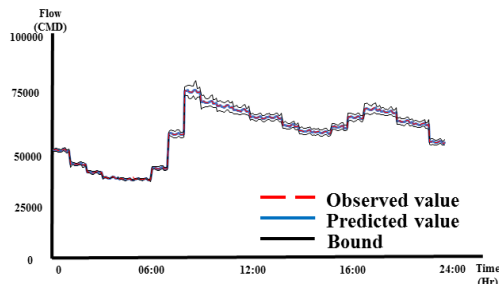
Fig. 6. Flow time-series results (observed & predicted) for Case 1

**Table 7.** Results of burst detection and false alarm according to the amount of leakage

Leakage rate (%)	Results of burst detection and false alarm (O/X)			
	Peak flow (at 8:40)		Lowest flow (at 4:50)	
	Burst detection	False alarm	Burst detection	False alarm
14%	O	X	O	X
12%	O	X	O	X
10%	O	X	O	X
6%	O	X	O	X
4%	X	O	O	X
2%	X	X	O	X



(a) Observed and predicted flow data for amount of 2% water leakage (Lowest Point)



(b) Observed and predicted flow data for amount of 14% water leakage (Peak Point)

**Fig. 7.** Results of performance indicators for Case 2

부에 대한 결과는 Table 7과 같다. 물 사용량이 많은 시점(8:40)에서는 4%이하에 대한 누수인지는 불가하였으며 오측의 횟수도 더 많은 것으로 나타났다. 물론, 2%에 대한 누수량을 잡아낼 수 있었지만, 이는 물 사용량이 많은 시점에서는 동일한 누수량이더라도 총량적인 관점에서의 비율은 작은 것이 원인으로 판단된다. 반면에 물 사용량이 적은 시점(4:50)에서는 2%의 누수량 까지 식별할 수 있었다. 향후 평균적으로 물 사용량이 많은 시점에 대한 정확한 누수 인지를 위해서는 보다 정밀도 높은 분석이 필요할 것으로 사료된다. 다만, 학습데이터의 수치가 거의 일정함으로 실제 운용되는 상수관망시스템에서 발생하는 유량데이터와는 상이성이 존재할 수 있으므로

향후 이에대한 보완이 필수적이라고 판단된다. Fig. 7(a)는 사용량이 적은 시점(4:50)에 대한 누수사고를 발생시킨 경우에 대한 결과이다. 그림에서 보이듯이 저점에서는 적절한 임계값 설정에 의해 사고를 인지할 수 있었다. 결과적으로 누수량이 큰 대규모 누수사고의 경우 Fig. 7(b)와 같이 충분히 인식이 가능한 것으로 나타났다.

#### 4. 결론

본 연구에서는 최신의 딥러닝 알고리즘중 하나인 LSTM 모형을 활용하여 기존 데이터 기억 능력의 저하로 인한 단점을 보완하여 데이터 기반의 예측능력이 향상된 누수인지 모형을 개발하였다. 개발된 모형을 가정된 누수사고 데이터에 의해 학습 및 검증을 수행하였으며, 약 99%의 정확도를 보이는 것으로 나타났다. 이와 같은 결과는 학습데이터 구성시 거의 일정한 데이터에 의한 학습과 비교적 적은 양의 데이터를 적용함에 따른 결과일 수 있으므로, 향후에는 실제 누수사고데이터를 다수의 지점에서 확보하고, 추가적인 모델의 검증이 필요할 것으로 판단된다. 즉, 보다 다양한 운영조건의 상수관망에 대한 적용을 통한 제안 방법론의 검증과 일반화가 필요하며, 다수 지점의 계측데이터를 활용한 누수의 발생인지 방법론 고도화 및 누수발생 지역 및 지점을 예측하는 연구 등이 이에 해당한다고 할 수 있다. 개발된 모형은 누수탐지를 위한 기초적인 연구자료가 될 수 있을것으로 기대되며, 제안된 모형을 토대로 유량데이터 예측부분에 있어서 보다 정밀한 결과가 도출 될 수 있을것으로 판단된다.

#### 감사의 글

본 결과물은 환경부의 재원으로 한국환경산업기술원의 지능형 도시수자원 관리사업의 지원을 받아 연구되었습니다 (2019002950002). 이에 감사드립니다.

#### References

Lambert, A. (2001). "What do we know about pressure-leakage relationships in distribution systems." *Proceedings of International Water Association International Specialised Conference: System Approach to Leakage Control and Water Distribution Systems Management*, Brno, Czech Republic.

- Lambert, A., and Hirner, W. (2000). *Losses from water supply systems: A standard terminology and recommended performance measures*. International Water Association, UK.
- Lee, C., Jun, J., Joo, J.G., and Yoo, D.G. (2018). "Uncertainty analysis for leakage of water supply networks using pressure dependent leakage analysis." *Journal of Korean Society of Hazard Mitigation*, Vol. 18, No. 7, pp. 129-136.
- Melgarejo-Moreno, J., López-Ortiz, M.I., and Fernández-Aracil, P. (2019). "Water distribution management in South-East Spain: A guaranteed system in a context of scarce resources." *Science of the Total Environment*, Vol. 648, pp. 1384-1393.
- Ministry of Environment (ME) (2021). *Water supply statistics*.
- Mounce, S.R., Boxall, J.B., and Machell, J. (2010). "Development and verification of an online artificial intelligence system for detection of bursts and other abnormal flows." *Journal of Water Resources Planning and Management*, Vol. 136, No. 3, pp. 309-318.
- Mounce, S.R., Day, A.J., Wood, A.S., Khan, A., Widdop, P.D., and Machell, J. (2002). "A neural network approach to burst detection." *Water Science and Technology*, Vol. 45, No. 4-5, pp. 237-246.
- Mounce, S.R., Khan, A., Wood, A.S., Day, A.J., Widdop, P.D., and Machell, J. (2003). "Sensor-fusion of hydraulic data for burst detection and location in a treated water distribution system." *Information Fusion*, Vol. 4, No. 3, pp. 217-229.
- Romano, M., Kapelan, Z., and Savić, D.A. (2014). "Automated detection of pipe bursts and other events in water distribution systems." *Journal of Water Resources Planning and Management*, ASCE, Vol. 140, No. 4, pp. 457-467.
- Rossman, L.A. (2000). *EPANET 2: Users manual*, U.S. Environmental Protection Agency, Washington, D.C., U.S.
- Shewhart, W.A. (1924). "Some applications of statistical methods to the analysis of physical and engineering data." *Bell System Technical Journal*, Vol. 3, No. 1, pp. 43-87.
- Yoo, D.G., Choi, D.Y., and Kim, K.P. (2018). "Trend and improvement direction of data-based medium-to-large leak recognition model." *Water for Future*, Vol. 51, No. 11, pp. 20-33.