

A field implementation of linear prediction for leak-monitoring in water distribution networks

1. Introduction

- Water distribution networks (WDN)에서 누수 탐지 방법.

1. water balance method

일반적으로 시스템의 물 흐름을 연구하기 위해 district metered areas (DMAs)을 도입하고 기본 유압 매개변수를 실시간 획득을 포함

2. acoustic method

3. non-acoustic method

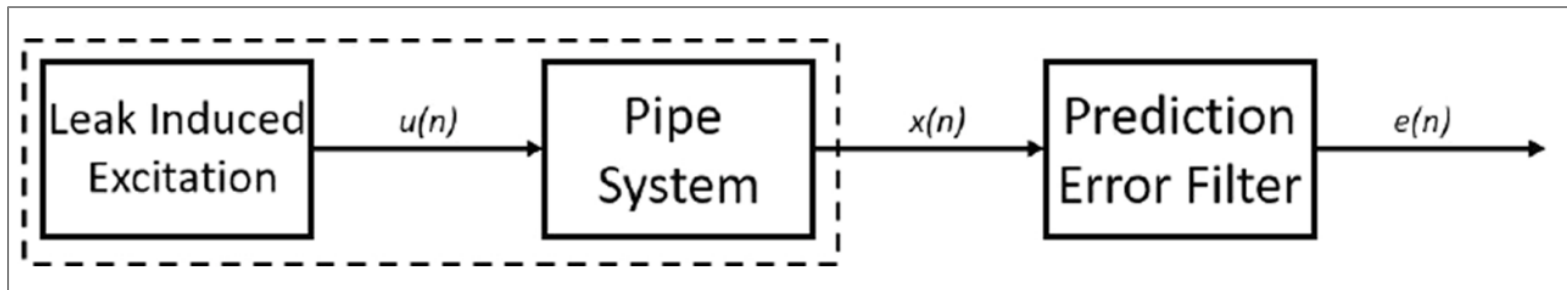
2번과 3번은 다양한 종류의 장비를 사용하여 측정을 하고 측정된 값을 분석하기 위한 지식이 필요하다.

따라서 본 논문은 누수 탐지를 위한 Linear Prediction(LP) 기반 방법을 제시한다.

누수를 탐지하는 순서: 전체적인 영역의 누수 여부 -> 누수의 정확한 위치

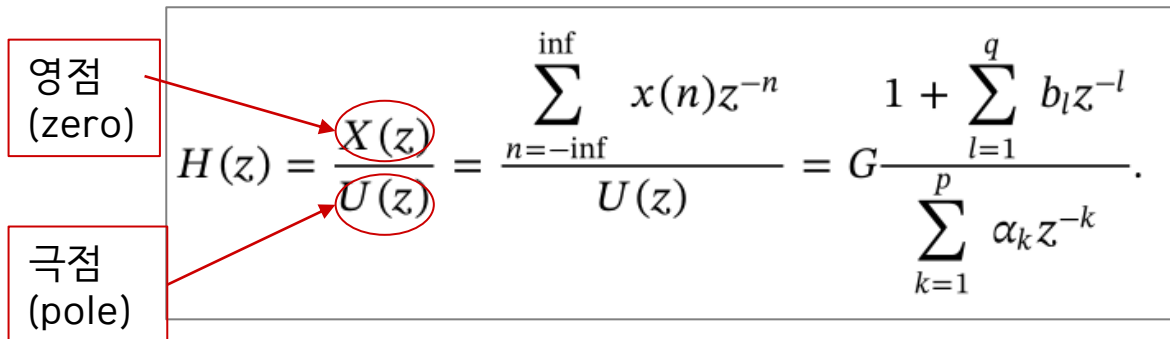
2. LP for leak-detection

- 누수는 좁은 주파수 대역에 포함된 에너지로 자극이 되는 것을 가정하며, 시스템의 전반적인 반응은 유체가 채워진 가압 파이프의 충격 반응으로 흥분 스펙트럼의 컨볼루션이다.
- 아래의 그림처럼 음향 누수 특징 $x(n)$ 은 $u(n)$ 에 의해 흥분된 짧은 시간 세그먼트 내에서 시간에 대해 불변하는 이산 시간 선형 시스템의 출력으로 모델링이 된다.



2. LP for leak-detection

- LP모형을 주파수 영역에서 autoregressive moving average (ARMA) 모델로 표현



The diagram shows the transfer function $H(z) = \frac{X(z)}{U(z)} = \frac{\sum_{n=-\infty}^{\infty} x(n)z^{-n}}{U(z)} = G \frac{1 + \sum_{l=1}^q b_l z^{-l}}{\sum_{k=1}^p \alpha_k z^{-k}}$. A red box labeled '영점 (zero)' has an arrow pointing to the numerator $X(z)$, which is circled in red. Another red box labeled '극점 (pole)' has an arrow pointing to the denominator $U(z)$, which is also circled in red.

- 이 모델에서 특별한 두 가지 경우가 있다.
 - $a_k = 0, 1 \leq k \leq p$ 인 경우를 all-zero 모델 혹은 moving average (MA) 모델이라고 부른다.
 - $b_l = 0, 1 \leq l \leq q$ 인 경우를 all-pole 모델 혹은 autoregressive (AR) 모델이라고 부른다.

2. LP for leak-detection

- LP모델을 주파수 영역에서 autoregressive moving average (ARMA) 모델로 표현

영점 (zero) → $X(z)$

극점 (pole) → $U(z)$

$$H(z) = \frac{X(z)}{U(z)} = \frac{\sum_{n=-\infty}^{\infty} x(n)z^{-n}}{\sum_{k=1}^p \alpha_k z^{-k}} = G \frac{1 + \sum_{l=1}^q b_l z^{-l}}{\sum_{k=1}^p \alpha_k z^{-k}}.$$

- 이 모델에서 '스펙트럼 포락선(spectral envelope)'이 목적이다.
- 따라서 이산 시간 선형 시스템에서 all-pole 모델로 표현한다.

$$H(z) = \frac{G}{1 - \sum_{k=1}^p \alpha_k z^{-k}},$$

2. LP for leak-detection

$$H(z) = \frac{G}{1 - \sum_{k=1}^p \alpha_k z^{-k}},$$

- 위의 식을 시간 영역에서 표현하면 아래와 같다.

$$x(n) = \sum_{k=1}^p \alpha_k x(n-k) + Gu(n),$$

- 또한 예측 모델은 아래와 같이 표현한다

$$\hat{x}(n) = \sum_{k=1}^p a_k x(n-k),$$

2. LP for leak-detection

- a_k 을 추정하기 위해 $x(n)$ 와 $\hat{x}(n)$ 의 차이를 아래와 같이 표현한다.

$$E_n = \sum_n (x(n))^2 - \sum_{k=1}^p a_k \sum_n x(n)x(n-k).$$

- Auto-correlation 방식을 사용하여 식을 간소화하였다.

$$E_n = R(0) - \sum_{k=1}^p a_k R(k),$$

where,

$$R(i) = \sum_{n=-\infty}^{\infty} x(n)x(n-i).$$

3.2. Leak-localization

- Correlation 기반한 localization 기법은 보통 오랜 기간동안 얻어진 고해상도의 원본 데이터가 요구되며, 여러가지 통계적인 계산이 필요하다.
- 따라서 센서의 개수가 많아지게 되면 데이터 전송에 대한 오버헤드가 발생할 수 있다.
- 논문에서는 필터링 기법을 사용하여 전송에 대한 부담을 줄일 것이다.
- 1단계: 일반적인 해상도에서 localization을 한다.
- 결과는 상관관계 분석에 필요한 가장 적절한 센서 스트림과 누수의 일반적인 영역을 결정하는 것이다.
- 2단계: 1단계에서 식별된 센서들을 가지고 교차 상관관계를 수행한다.
- 이를 통해 모든 센서들에 대해 전송을 할 필요가 없어져서 비용 절감이 된다.

3.2. Leak-localization

- Step 1: coarse localization
 1. 평균 baseline RMS는 시뮬레이션된 누수 이전에 수집된 baseline 데이터를 사용하여 계산한다.
 2. 시스템에서 전기적인 소음과 펌프와 같은 일반적인 주기적 노이즈를 제거하기 위해 데이터를 필터링한다.
 3. 모든 경우 RMS 값으로 설명되는 각 센서 위치에 대한 음향 신호의 강도는 sensor location과 관련된 baseline RMS를 빼서 계산하고 정규화를 한다.
 4. 그런 다음 이 값들은 가장 큰 값에서 가장 작은 값으로 순위가 매겨진다.

3.2. Leak-localization

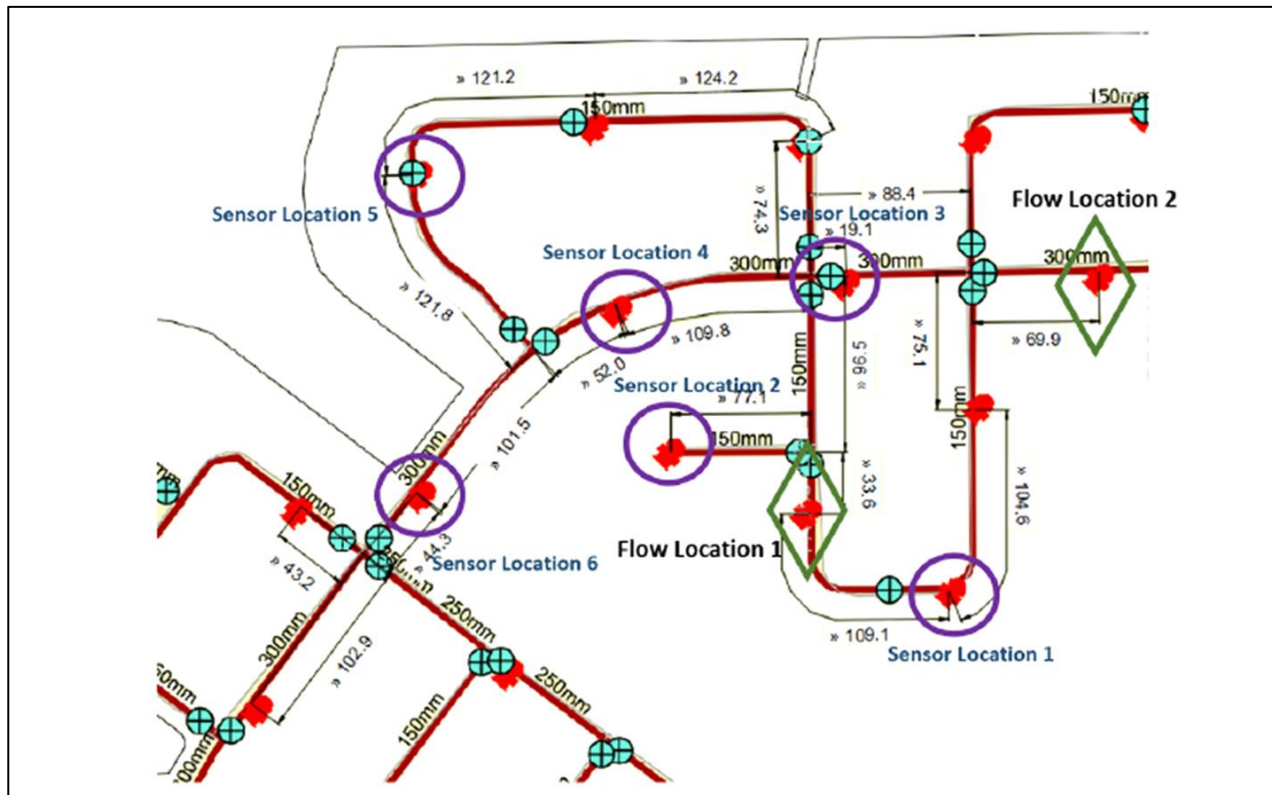
- Step 2: fine localization
- 1단계에서 계산을 통해 얻어진 누수에서 가장 가까운 두 위치의 센서가 있다.
- 기준 센서로부터 D_1 거리에 있는 누수 위치는 다음과 같이 계산된다.

$$D_1 = \frac{D - c\tau}{2},$$

- D 는 두 센서 위치 사이의 거리이고 c 는 파이프 내의 소리 확산 속도이다.
- τ 는 교차 상관의 정점에 해당하는 지연이다.

4. Description of the field test site and data

- 이 시스템은 설치 비용과 에너지 소비량을 최소화하면서 WDN에서 이벤트 탐지를 위한 편리하고 경제적인 해결책을 제시한다.
- 1. 비전문가인 사용자를 위해 데이터 기반의 누수 탐지 및 localization 시스템 제공. 이를 통해 DMA 네트워크에 영향을 미치는 누수를 효과적으로 찾아내고 물 손실 및 검사 비용에 대해 절감할 수 있다.
- 2. DMA의 동일한 영역에 전개되는 센서 시스템의 통합에 기초한 자동 누수 감지 및 localization 분석의 기능을 제공한다.
- 3. DMA에 설치되었을 때 장치의 복잡한 설치와 초기화가 필요하지 않을 뿐만 아니라 네트워크가 정상 상태로 변경될 때 시스템이 사용자에게 해석될 수 있도록 하기 위해서다.
- 4. 네트워크를 통해 모든 장치에서 음향 데이터에 쉽게 접근할 수 있기 위해서다.
- 5. 도시 직원이 해석할 수 있는 직관적인 방법으로 결과를 제시하는 인터페이스를 제공하기 위해서다.



4.2. Sensors and data-acquisition system description

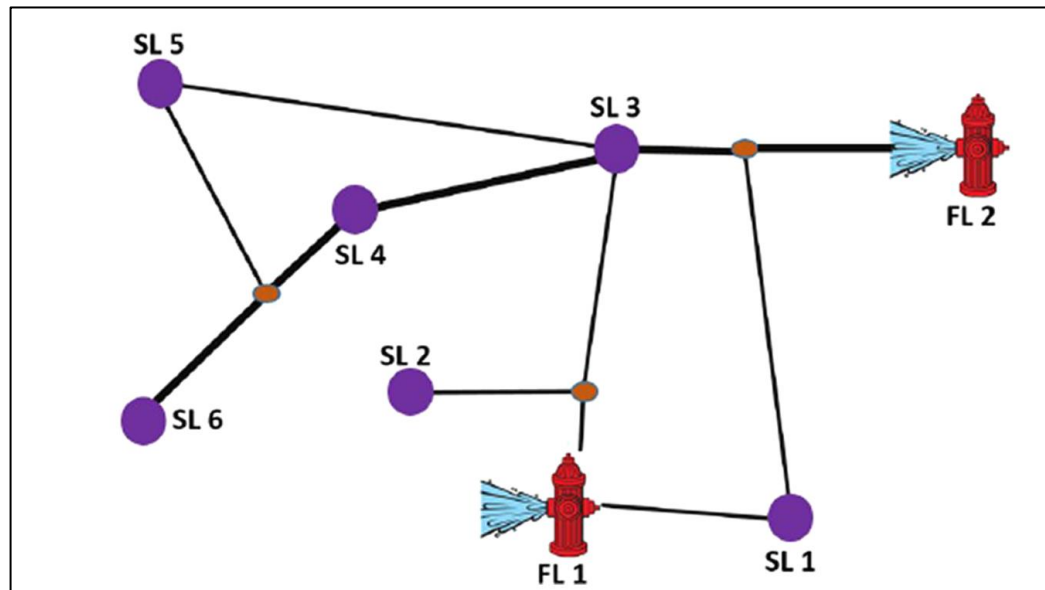
- 사용하는 센서들은 4가지이다.
- (1) hydrophone, (2) pressure sensor, (3) accelerometer, (4) temperature sensor
- 데이터 수집 시스템은 데이터 전송 오버헤드를 최소화하고 배터리 수명을 최대화하기 위해 센서 위치에서 기본적인 신호 처리가 가능하도록 설계되었다.

4.3. Data collection and description

- 2018년 가을 누수 시뮬레이션 테스트에서 3개의 날씨가 선택되었다.
- 10월 13일, 10월 21일, 11월 3일에 자정부터 새벽 4시 사이에 시뮬레이션되었다.
- 각 흐름 위에서 4개의 흐르는 양을 시뮬레이션했다.
- 200, 100, 50 and 25 (L/min)이고 3일 내내 지속되었다.
- 이 시험일마다 자정부터 오전 3시까지 거의 3시간 동안 연속적으로 두 수전들이 흐르게 되었다.

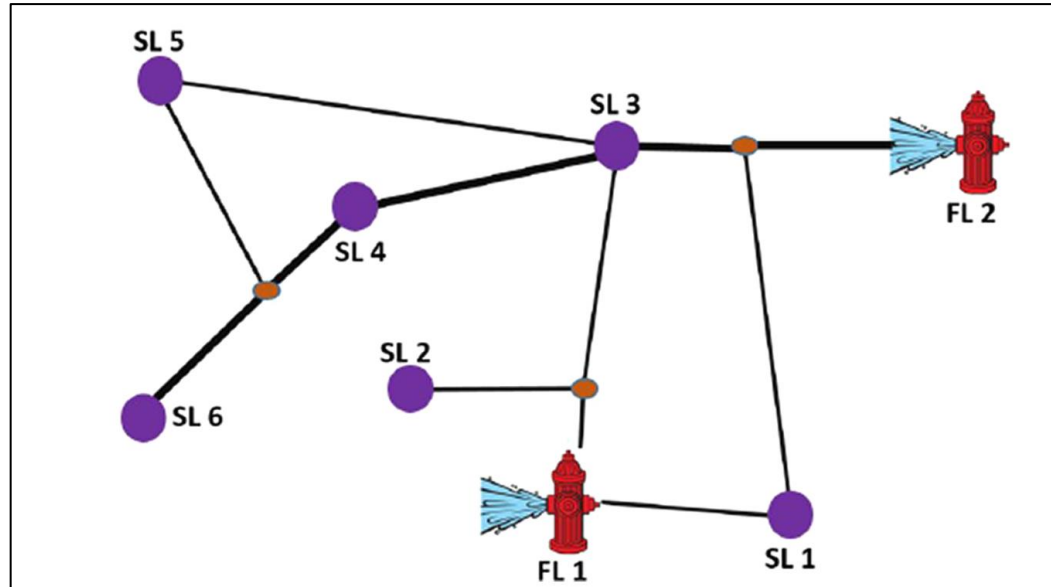
4.3. Data collection and description

- 첫번째 테스트 데이터는 SL 1, 2, 4, 6에서 수집되었고, 두번째 테스트 데이터는 SL 1, 2, 4, 5, 6에서 수집되었으며, 마지막으로 셋째 날에는 SL 1, 2, 4, 3, 5, 6에서 수집되었다.



- 또한 데이터를 수집하는 순서는 (1) leak-free data, (2) 200 L/min leak, (3) 100 L/min leak, (4) 50 L/min leak, (5) 25 L/min leak, (6) leak-free data로 된다.

4.3. Data collection and description

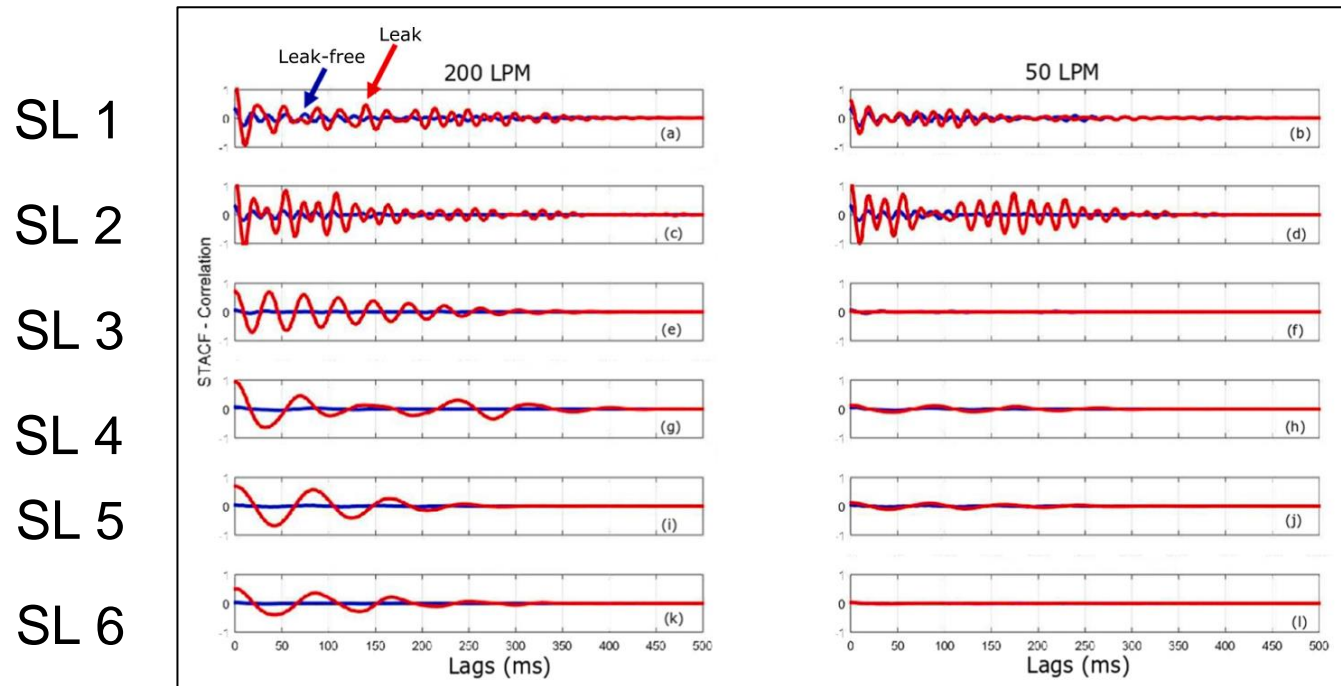


- 누수는 FL 1, 2 두 곳에서 발생을 한다.
- 데이터 세트의 충분한 가변성을 보장하기 위해, 음향 신호는 다른 날짜와 다른 흐름 크기로 얻는다.

데이터는 4kHz의 샘플링 주파수로 수집되었다.

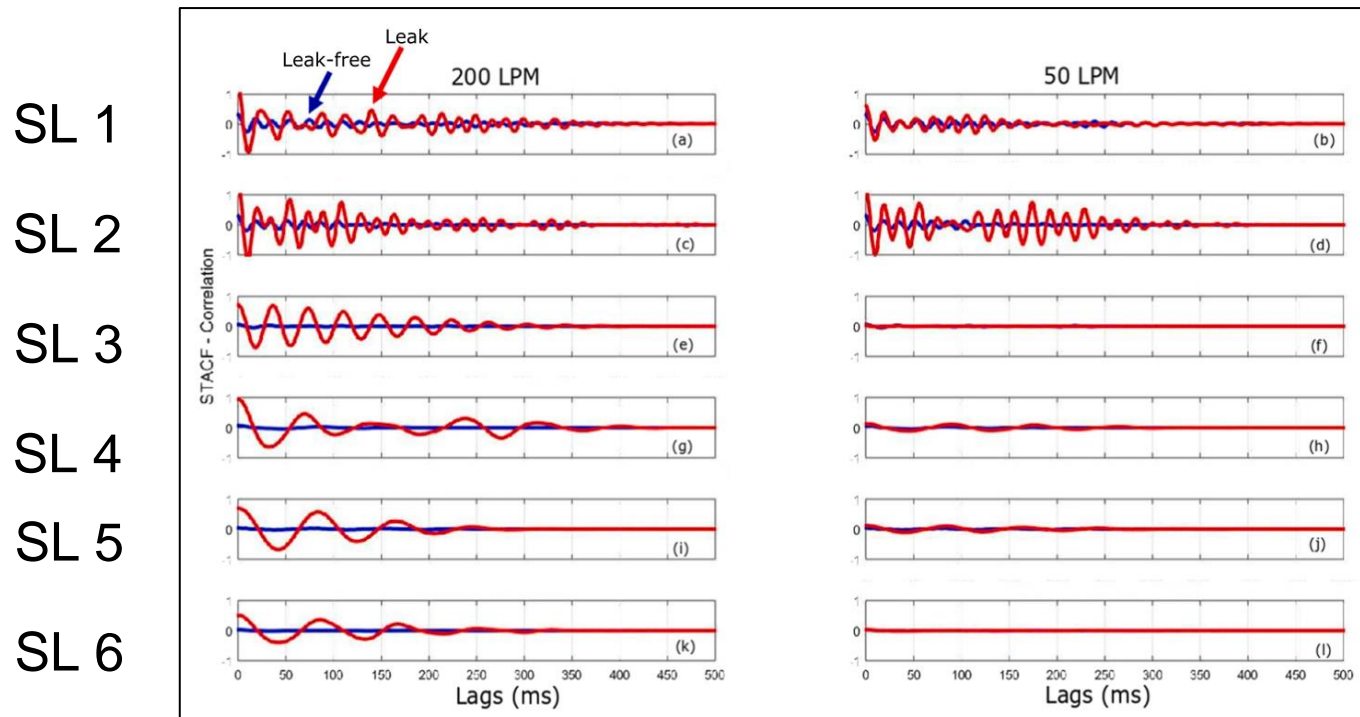
5.1. Leak characterization

- 모든 센서 위치에서 200 L/min과 50 L/min의 두 경에 대해 FL 1에 해당하는 경우



- 누수 에너지의 크기 차이는 주로 스펙트럼의 저주파 영역에서 볼 수 있다.

5.1. Leak characterization



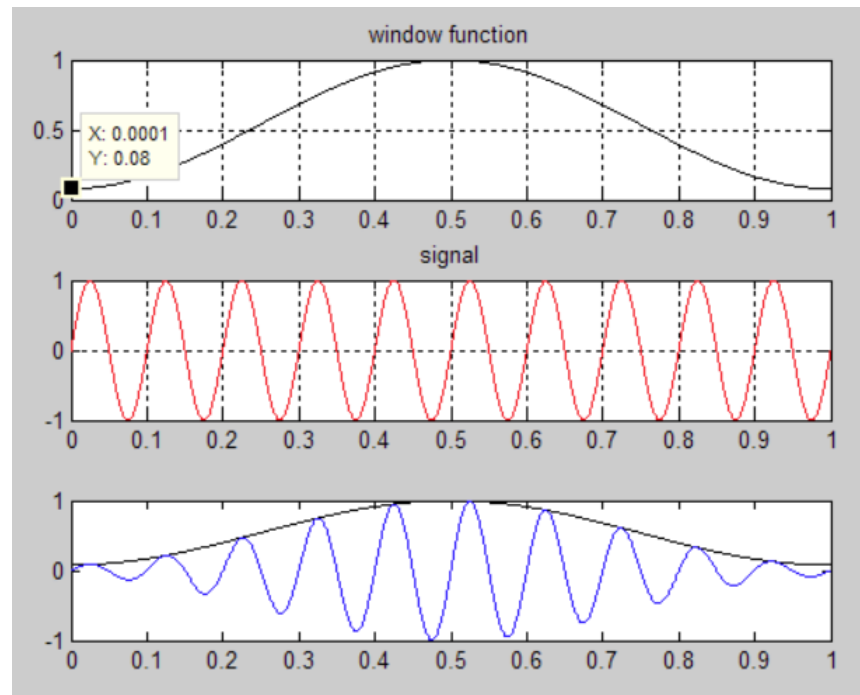
- 200 L/min인 경우에는 먼 경우에도 많은 영향을 받는 반면, 50 L/min인 경우에는 먼 경우에는 거의 차이를 확인할 수가 없다.

5.2. Leak-detection

- 5.2.2. Anomaly detection using multivariate GMM
- 데이터 준비 프로세스, LP 계수 추정 및 이상 사건 검출을 위한 세부사항은 아래와 같다.
 1. 누수가 없거나 누수인 케이스에 대한 음향 신호는 우선 약 1초 시간 프레임으로 분할되며, 인접 프레임 간에 50% 중복된다. 각 프레임은 특징 행렬의 표본을 하나씩 생성한다. 따라서 분석을 위한 표본 수는 프레임 수와 동일하다.
 2. LP 특징은 누수가 없거나 있는 경우 둘 다 $p = 50$ 의 모델 순서를 사용하여 추정한다. Hamming window는 LP 계수를 추출하기 전에 각 프레임에 적용된다. 이들 50개 계수의 3가지 주요 성분을 각 표본에 대해 취한다.
 3. 샘플의 40%는 학습에 사용되며, 나머지는 시험에 사용된다.
 4. 이후 성능을 측정하고 결과를 검토한다.

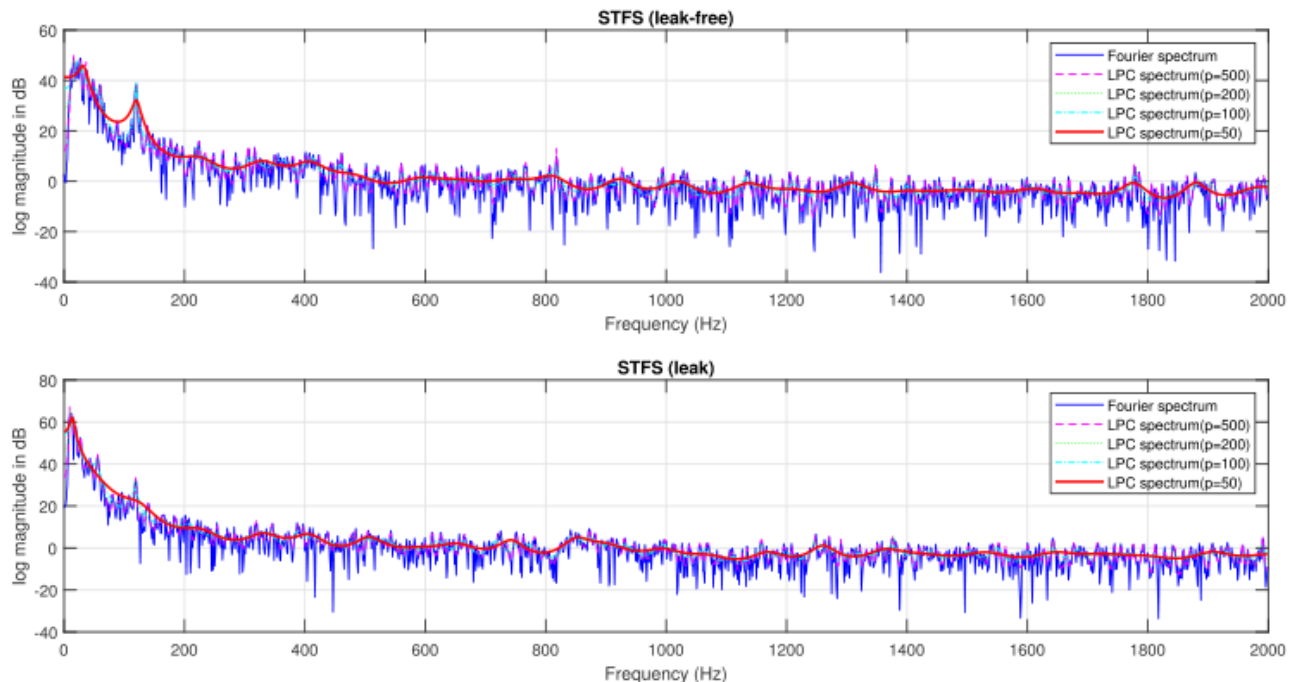
5.2. Leak-detection

- Hamming window



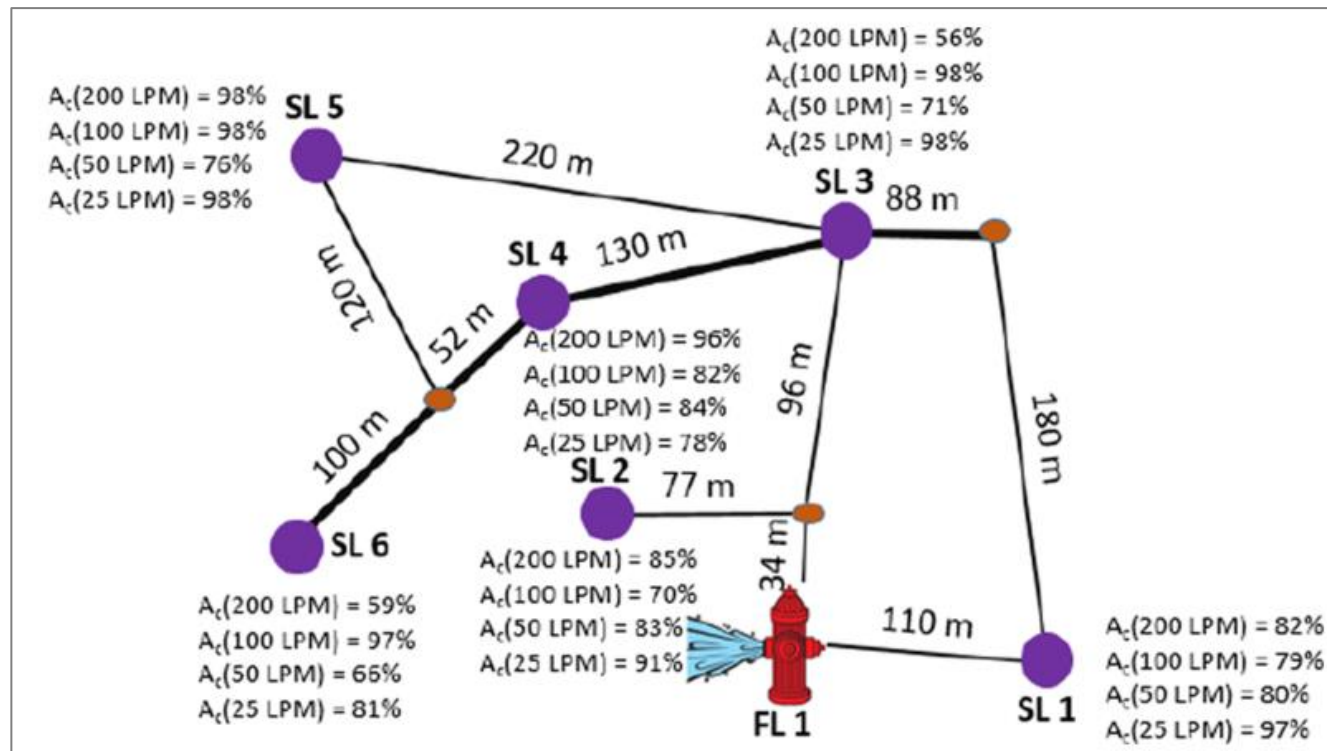
5.2. Leak-detection

- p 의 크기에 따른 short term 푸리에 스펙트럼 값
- $p = 50$ 인 경우는 전체적인 경향을 파악하고, $p = 500$ 인 경우에는 short term 푸리에 스펙트럼 값이 실제 스펙트럼 값과 일치해진다.



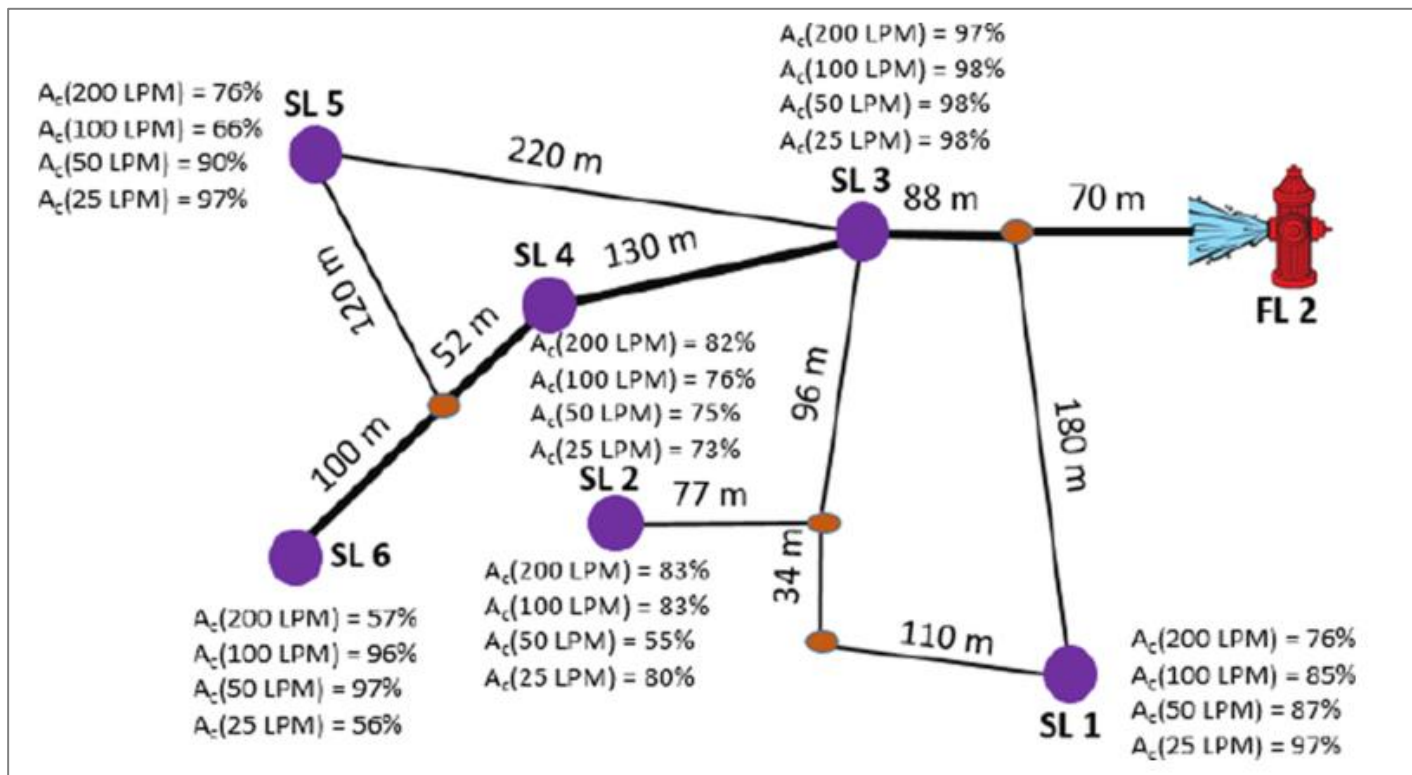
5.2. Leak-detection

- FL 1에서 누수가 일어났을 경우의 각 센서에서의 검출 정확도
- 유량의 차이가 검출 정확도에서는 큰 차이가 없다.



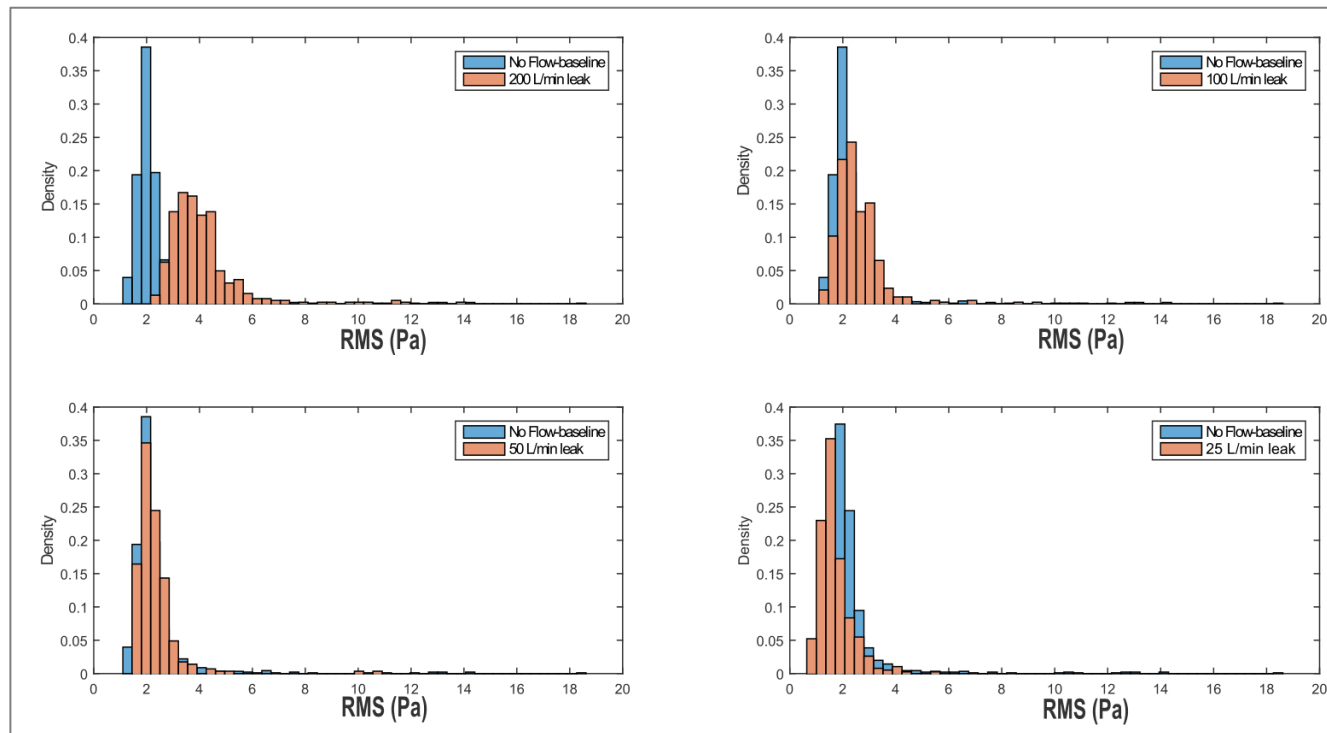
5.2. Leak-detection

- FL 2에서 누수가 일어났을 경우의 각 센서에서의 검출 정확도
- 가장 먼 곳을 제외하고 대부분 높은 성능을 보여준다.



5.3. Leak localization

- FL 1과 SL 2의 데이터를 사용하였다.
- 물의 흐름이 많은 경우에는 유의미한 값이 나온다.
- RMS를 통해 정확한 위치를 찾는 것은 어렵지만, 대략적인 위치 파악에는 사용할 수 있다.



5.3.1. Flow Location 1

- FL 1에서 각 SL과의 영향에 따른 순위

Table 2

Average location rank as a percentage for *FL 1* (true order: SL1/SL2, SL3, SL4, SL6, SL5) of all flow cases. The selected rank for each sensor location is in **bold**.

<i>Rank</i>	SL 1	SL 2	SL 3	SL 4	SL 5	SL 6
1 st	18.4	45.1	19.1	1.1	6.7	6.8
2 nd	42.2	14.0	8.0	5.4	5.4	16.1
3 rd	1.9	3.7	28.8	24.6	23.1	29.8
4 th	1.7	3.5	16.3	25.2	25.8	32.1
5 th	19.5	7.5	9.0	18.9	28.9	15.2
6 th	16.3	26.3	18.8	24.7	10.0	0.0

5.3.1. Flow Location 1

- FL 1에 제일 많이 영향을 주는 센서는 SL 2 → SL 1 → SL 3 순서이다.
- SL 1과 SL 2을 통해 계산된 오차와 SL 1과 SL 3을 통해 계산된 오차

Table 3

Average D_1 (in meters), μ_{D_1} , for each flow amount using *SLs 1 and 2*, and *SLs 1 and 3*. In both cases D_1 is taken as the distance from *SL 1* to the leak location, thus for both cases, the true distance is approximately $D_1 = 109$ m. Percentage errors are reported in parenthesis.

Flow Amount	$\mu_{D_1(SL1-2)} (m)$	$\mu_{D_1(SL1-3)} (m)$
200 L/min	95.99 (11.9%)	136.03 (24.8%)
100 L/min	107.91 (1.0%)	131.19 (20.4%)
50 L/min	100.40 (7.9%)	117.26 (7.5%)
25 L/min	99.59 (8.6%)	100.62 (7.7%)

5.3.2. Flow Location 2

- FL 2에서 각 SL과의 영향에 따른 순위

Table 4

Average location rank as a percentage for *FL 2* (true order: SL3, SL1, SL4, SL2, SL6, SL5). The selected rank for each sensor location is in **bold**.

<i>Rank</i>	SL 1	SL 2	SL 3	SL 4	SL 5	SL 6
1 st	33.1	22.6	52.7	1.1	2.5	10.8
2 nd	27.4	18.4	17.8	9.3	2.0	27.3
3 rd	11.1	8.2	8.6	34.4	5.7	26.2
4 th	8.7	8.7	7.6	24.1	20.7	24.2
5 th	13.6	33.0	3.2	11.3	16.7	<i>11.4</i>
6 th	6.1	9.1	0.0	19.7	52.3	0.0

5.3.2. Flow Location 2

- FL 2에 제일 많이 영향을 주는 센서는 SL 3 → SL 1 → SL 4 순서이다.
- SL 3과 SL 4을 통해 계산된 오차

Table 5

Average D_1 (in meters), μ_{D_1} , for each flow amount using *SLs* 3 and 4. D_1 is taken as the distance from *SL* 4 to the leak location, thus for both true $D_1 = 268$ m. Percentage errors are shown in the parenthesis.

Flow Amount	$\mu_{D_1(SL3-4)} (m)$
200 L/min	261.88 (2.3%)
100 L/min	230.85 (13.9%)
50 L/min	222.47 (17.0%)
25 L/min	247.90 (7.5%)

A field implementation of linear prediction for leak-monitoring in water distribution networks

감사합니다