# 정수처리공정에서의 빅데이터와 활용방안

2024. 12. 06.

K-water 연구원
상하수도연구소
형 진 석

K water

# CONTENTS


0 발표자 소개

I 정수처리공정과 빅데이터

II 정수처리공정에서의 빅데이터 활용

III 향후 연구 방향

0

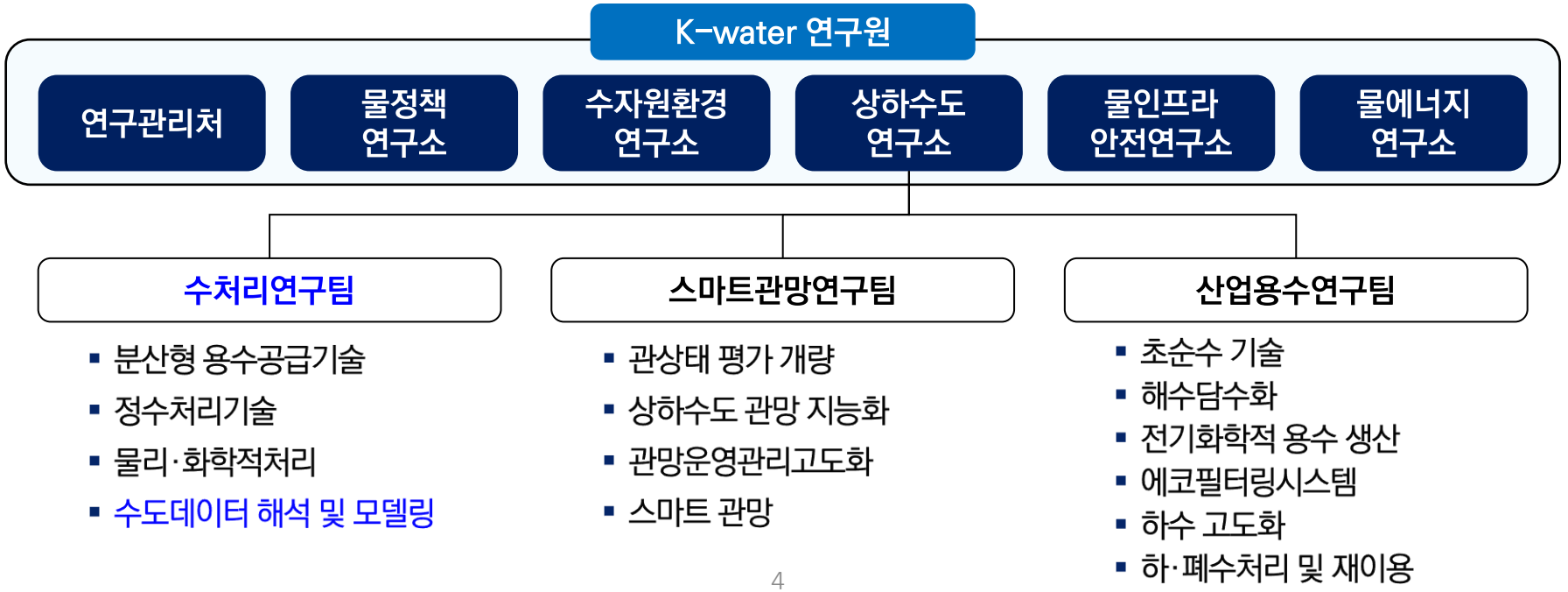
# 발표자 소개

K water

# 발표자 소개

## 형 진 석

- 서울시립대학교 환경공학부 공학사 (08학번)
- 서울시립대학교 일반대학원 환경공학과 공학석사 ('17.08)
  - 석사학위논문 : 정수지 소독능을 고려한 단기 물 수요예측 기반 송수펌프 최적운영기법 개발
- 서울시립대학교 일반대학원 환경공학과 공학박사 ('22.02)
  - 박사학위논문 : 빅데이터 기반 정수처리공정의 운영 진단 및 최적 의사결정 모델 개발
- 서울시립대학교 환경공학부 연구교수 ('22.03 ~ '24.07)
- K-water 연구원 상하수도연구소 선임연구원 ('24.07 ~ 현재)



I

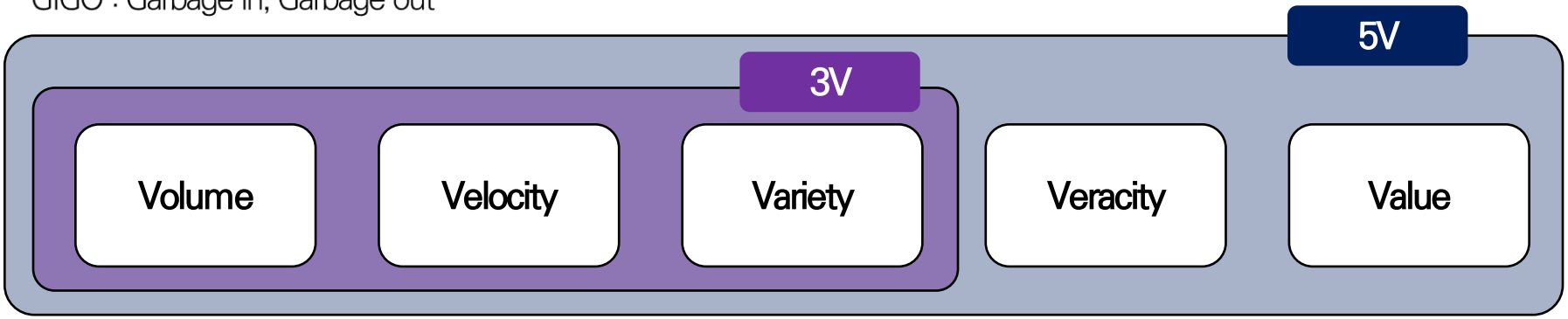
# 정수처리공정과 빅데이터

K water

# 빅데이터

## 빅데이터란?

- 빅데이터 란 "높은 통찰력, 의사결정, 프로세스 자동화를 위해 비용 효과가 높은 혁신적인 정보처리 과정을 요하며, 대용량의 데이터 규모(high-Volume), 빠른 속도(high-Velocity), 다양성(high-Variety) 을 지닌 정보자산" (Gartner, 2001)
  - Volume : 단순 저장되는 물리적 데이터양을 나타내며 빅데이터의 가장 기본적인 특징
  - Velocity : 데이터의 고도화된 실시간 처리 (데이터 생성, 저장, 시각화 등)
  - Variety : 다양한 형태의 데이터 (정형, 비정형 등)
    - 정형 데이터 : 행과 열 등으로 구조화되는 데이터(엑셀 스프레드 시트, 센서 등에 의해 계측된 수치화 데이터)
    - 비정형 데이터 : 구조화 되지 않는 데이터(오디오, 이미지, 동영상, 텍스트 데이터 등)
- 최근 정확성(Veracity), 가치(Value)가 빅데이터의 신규 구성요소로 대두되며, 5V로 표현되기도 함.
  - GIGO : Garbage in, Garbage out



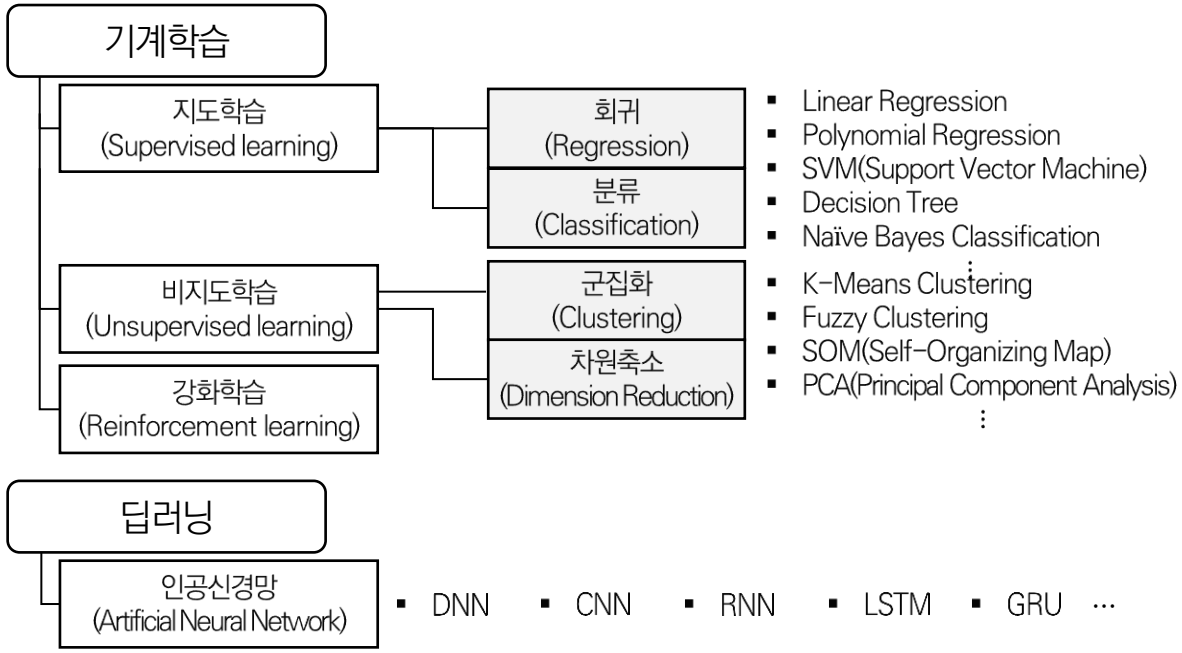

〈빅데이터의 구성요소〉

# 빅데이터

## 무엇을 할 수 있나?

- 빅데이터 분석 : 대규모 데이터 세트에서 유의미한 패턴, 트랜드, 통찰(Insight)를 찾아내는 것
  - 통계적 방법, 데이터 마이닝, 데이터 시각화 등 분석 방법론 활용
- 인공지능 : 인간의 인지적 작업을 모방하여 문제해결, 의사결정을 수행할 수 있는 기술
- 빅데이터 분석의 Tool로써 인공지능을 활용함으로써 데이터 수집, 분석 자동화가 실현되고 있음.

〈인공지능, 기계학습, 딥러닝의 관계〉



# 정수처리공정에서의 빅데이터

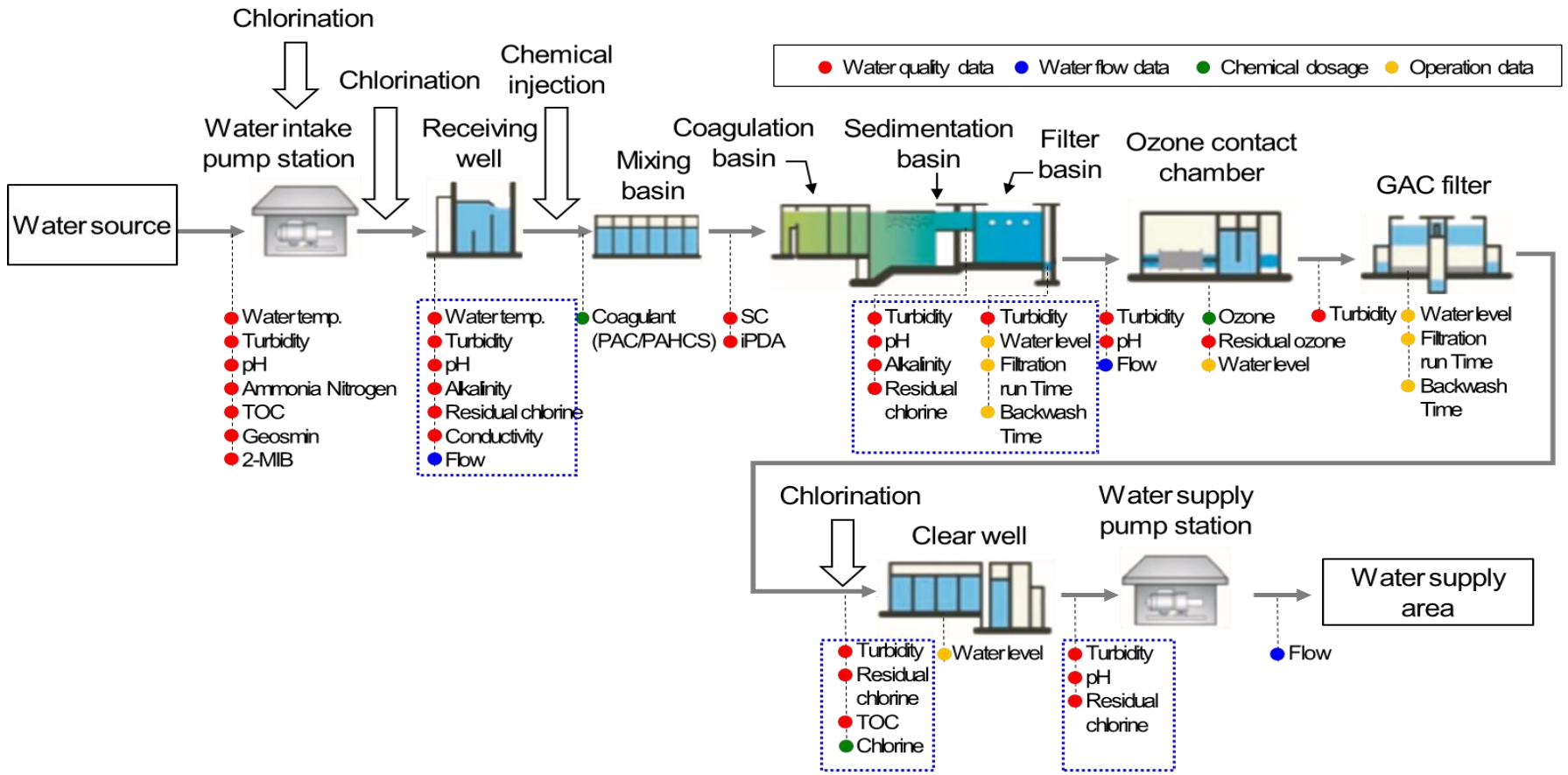
## 상수도 시스템 에서의 수질계측지점(상수도설계기준 해설편)

| 구 분 | | 수온 | pH | 탁도 | 전기 전도도 | 알칼 리도 | 잔류 염소 | 불소 | SCD | 입자 계수 | TOC | NH3 | 조류 측정 | COD (UV) | SS | 기타 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 취수장(수원) | | ○ | ○ | ● | ○ | ○ | | | | | ○ | ○ | ○ | | | ○ |
| 착수정 | | ● | ● | ● | ● | ● | ○ | | | | ○ | | | | | ○ |
| 혼화지 | | | ○ | | | ○ | ○ | | ○ | | | | | | | |
| 침전지 | | | ○ | ● | | | ○ | | | | | | | | | |
| 여과지 또는 정수지전단 | | | ○ | ● | | ○ | ○ | ○ | | ○ | ○ | | | | | |
| 정수지 | | ● | ● | ● | ○ | | ● | | | | | | | | | |
| 가압장 | | | | | | | ○ | | | | | | | | | |
| 배수지 | 10,000 $m^3$/일 이상 | ● | ● | ● | ○ | | ● | | | | | | | | | |
| | 10,000 $m^3$/일 미만 | ○ | ○ | ○ | ○ | | ○ | | | | | | | | | |
| 송배수 관로 | | ○ | ○ | ○ | ○ | | ○ | | | | | | | | | |
| 방류수 | | | ● | | | | | | | | | | | ○ | ○ | |

※ 1. 범례 : ● 기본설치항목, ○ 선택적 설치항목
2. 취수장 탁도계는 착수정 도착시간이 1시간 이상인 취수장 또는 광역취수장에 설치하며 취수장과 정수장이 인접된 경우는 설치 불필요
3. 여과지 "탁도"의 경우 통합 여과수 탁도계이며 여과지별 탁도계는 선택 사항 임
4. "기타"는 취수원의 수질특성상 온라인 감시(망간, 질산성질소 등)가 필요할 때 또는 수원의 목적상 온라인 감시가 필요할 때 다항목수질자동측정기를 설치운영
5. 송배수관로 상의 수질에 대한 온라인 감시가 필요할 경우 수질 자동측정기 설치를 검토할 수 있다.

# 정수처리공정에서의 빅데이터

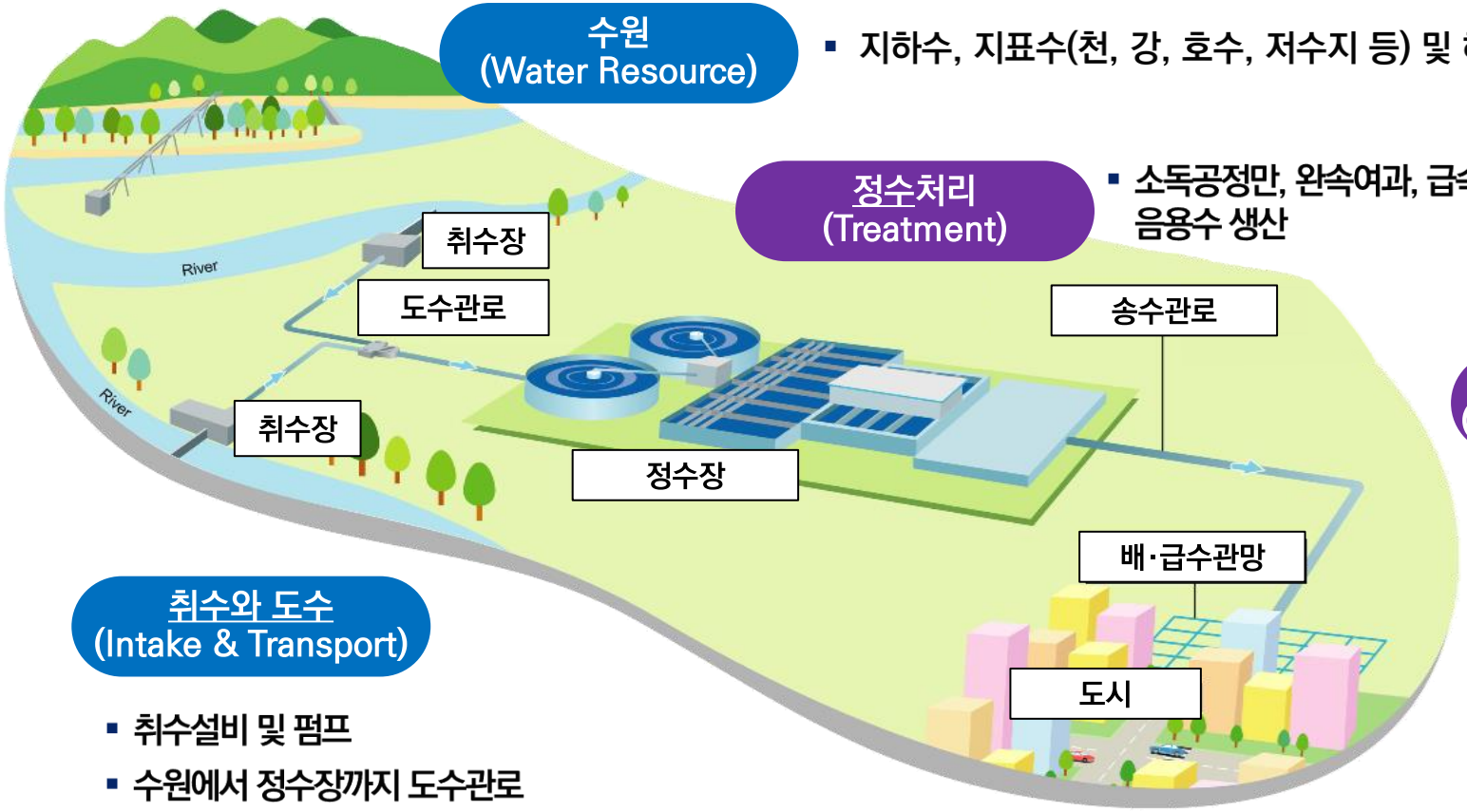
## 정수장 계측 지점 및 항목 예시 (A정수장 일보)




**정수장에서 축적되어 빅데이터화 되고 있는 운영 데이터와 AI 기술을 연계활용한 정수처리 공정 운영, 진단 방법에 대한 연구 추진 중**

# 상수도 시스템

## 상수도 시스템의 구성




**수원 (Water Resource)**

- 지하수, 지표수(천, 강, 호수, 저수지 등) 및 해수

**정수처리 (Treatment)**

- 소독공정만, 완속여과, 급속여과 등 다양한 공정의 조합으로 음용수 생산

**슬러지 처리 (Disposal of Residue)**

- 충분히 탈수되었을 때만 매립 가능
- 지자체(local authority) 승인이 있는 경우 하수처리장으로 운송

**취수와 도수 (Intake & Transport)**

- 취수설비 및 펌프
- 수원에서 정수장까지 도수관로
- 조류 및 미생물의 재성장을 막을 수 있는 소독 설비 포함

**저류, 송·배수, 급수 (Storage, Bulk Transport & Distribution)**

- 배급수관으로 소비자에게 전달
- 여러 관로가 모여 관망을 형성

# 정수시설

## 정수시설의 개념

# 정수시설

## 표준정수처리

- 입자성 물질(미생물 포함) 제거와 소독

• 수처리용 약품(응집제)을 신속하고 균일하게 수중에 확산시킴.

• 원수의 수위동요를 안정시키고, 원수유량을 조절하여 다음단계의 정수처리가 정확하고 쉽게 될 수 있도록 함.

착 수 정 | 혼 화 지

• 약품주입, 혼화 및 응집단계에서 성장한 Floc의 대부분을 분리작용에 의하여 제거

• 약품주입으로 Floc화한 탁질을 빠르게 침전되는 크기로 성장시킴.

응 집 지 | 침 전 지

**정수처리공정중 핵심 공정임**

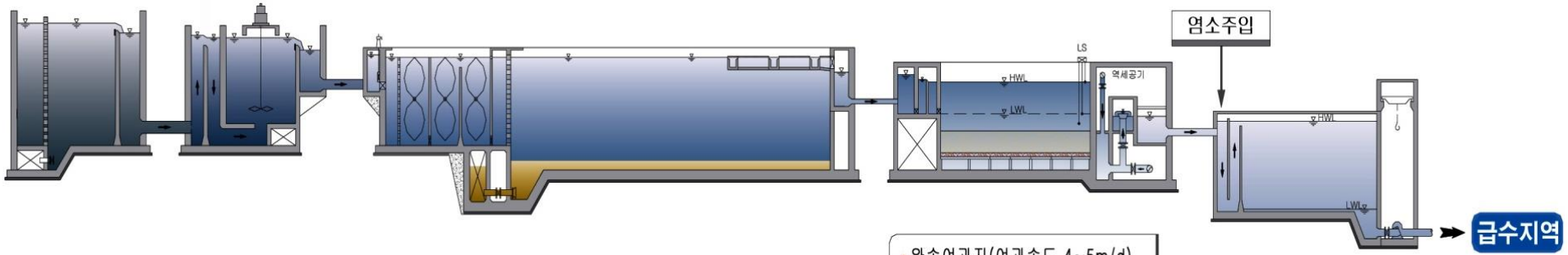
• 침전중의 현탁물질을 여재에 부착시키거나 여층의 체거름 작용에 의하여 제거.

여 과 지

• 처리된 물을 급수지역까지 송수함.

• 주입된 염소가 신속균등하게 물과 혼합되게하여 완전한 살균 및 소독을 함.

염소접촉지 | 정 수 지 | 송수펌프

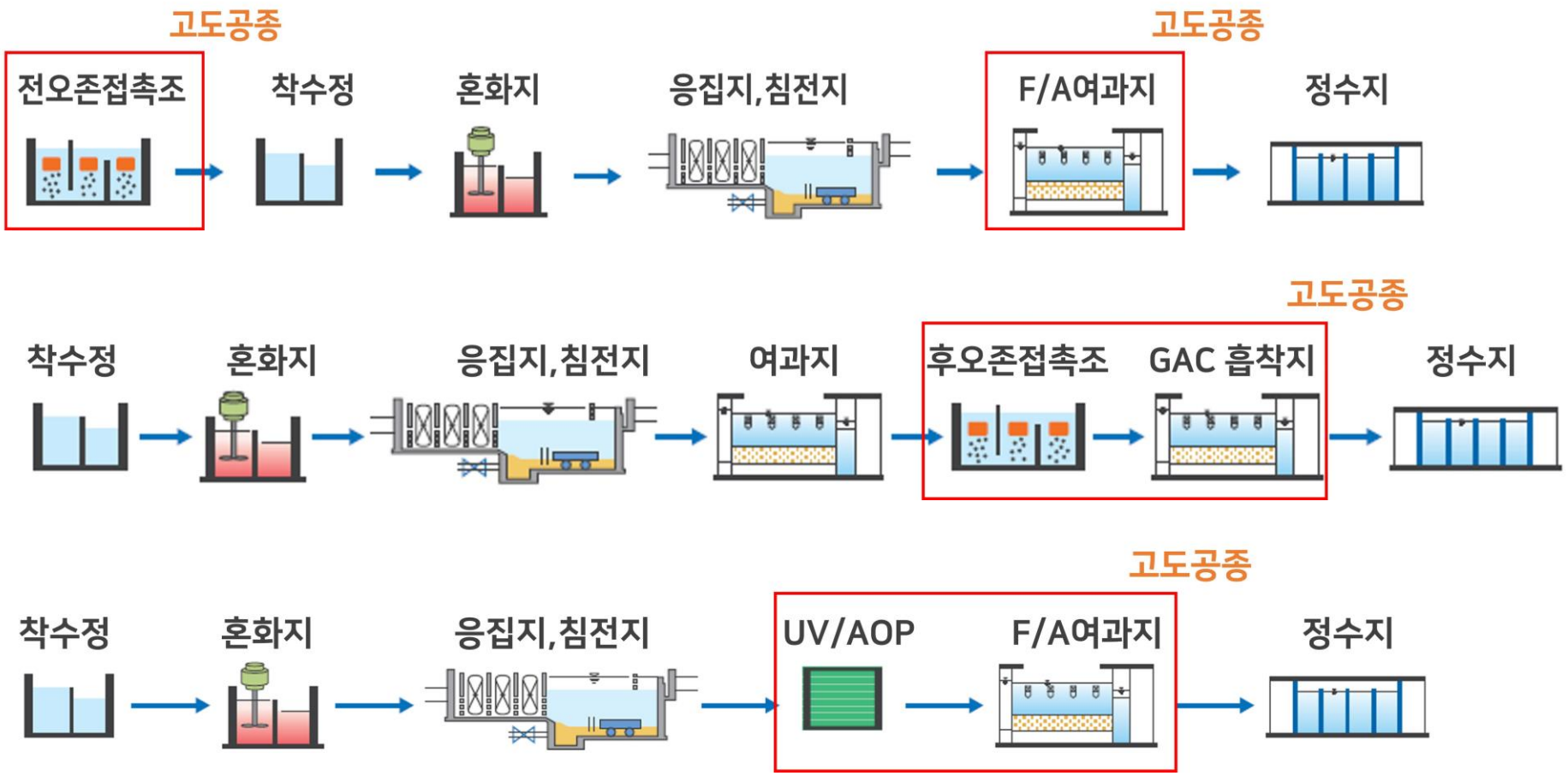



• 완속여과지(여과속도 4~5m/d)
• 급속여과지(여과속도 120m/d이상)

# 정수시설

## 고도정수처리

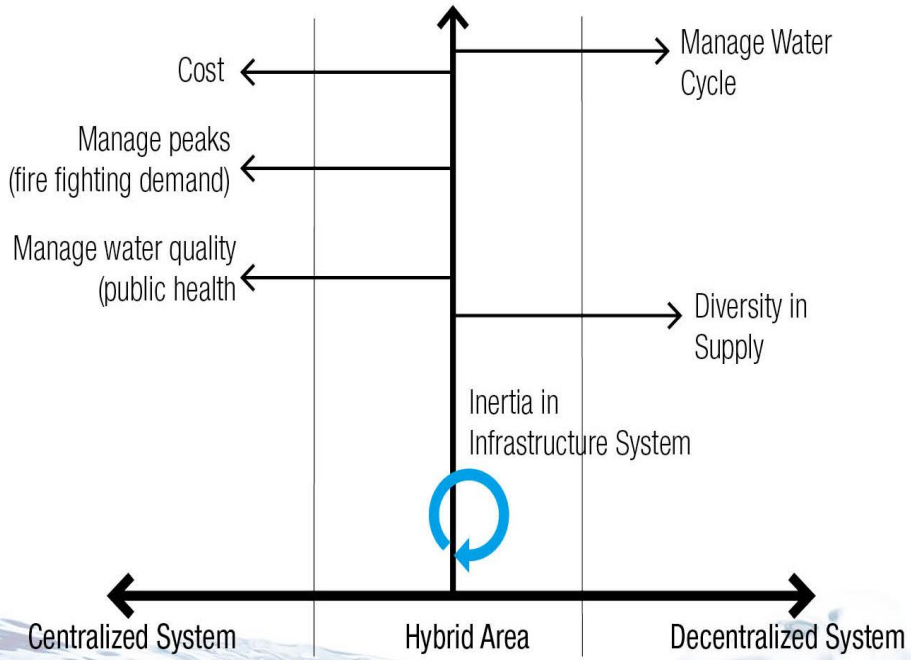
- 표준정수처리(응집/침전/여과/소독)으로 완전히 제거되지 않는 맛냄새 유발물질, 미량유기물질 등을 제거
- 대표적으로 활성탄 공정, 오존공정, UV 공정, 고도산화(AOP) 공정 등



# 분산형 용수공급 시스템

## 분산형 용수공급 시스템의 개념

- 목적에 맞는 용수공급 체계로서 개별, 수원별 물공급 또는 집중형과 연계된 용수공급 형태
  - 소규모로 수처리 시설 분산 배치
  - 처리된 수돗물의 공급과정 중 체류시간, 거리 최소화
  - 공급과정 중 2차 오염 최소화




| | | Degree of(De-) Centralization | | |
|---|---|---|---|---|
| | | Complete Decentralization | Semi-Centralization | Complete Centralization |
| perspectives | Technical Concept | Point-of-use/ point of Entry treatment | Few small treatment plants (cluster systems) | Centralized treatment plant |
| | Ownership | Homeowner | (private or public) Utilities, contractors, investors | (private or public) Utility, contractor, investor |
| | Operation | Homeowner | Utilities, contractors | Utility, contractor |

# 분산형 용수공급 시스템

## 스마트시티 스마트정수장 구축 시범사업

- 도시 내 소규모 최첨단 정수처리시설을 소비자 인근에 분산 설치하여, 관로 이송거리 최소화, 염소사용량 최소화(저염소, 0.1 ~ 0.2 mg/L) 를 통해 시민이 만족하는 고품질(병물 수준) 수돗물 공급

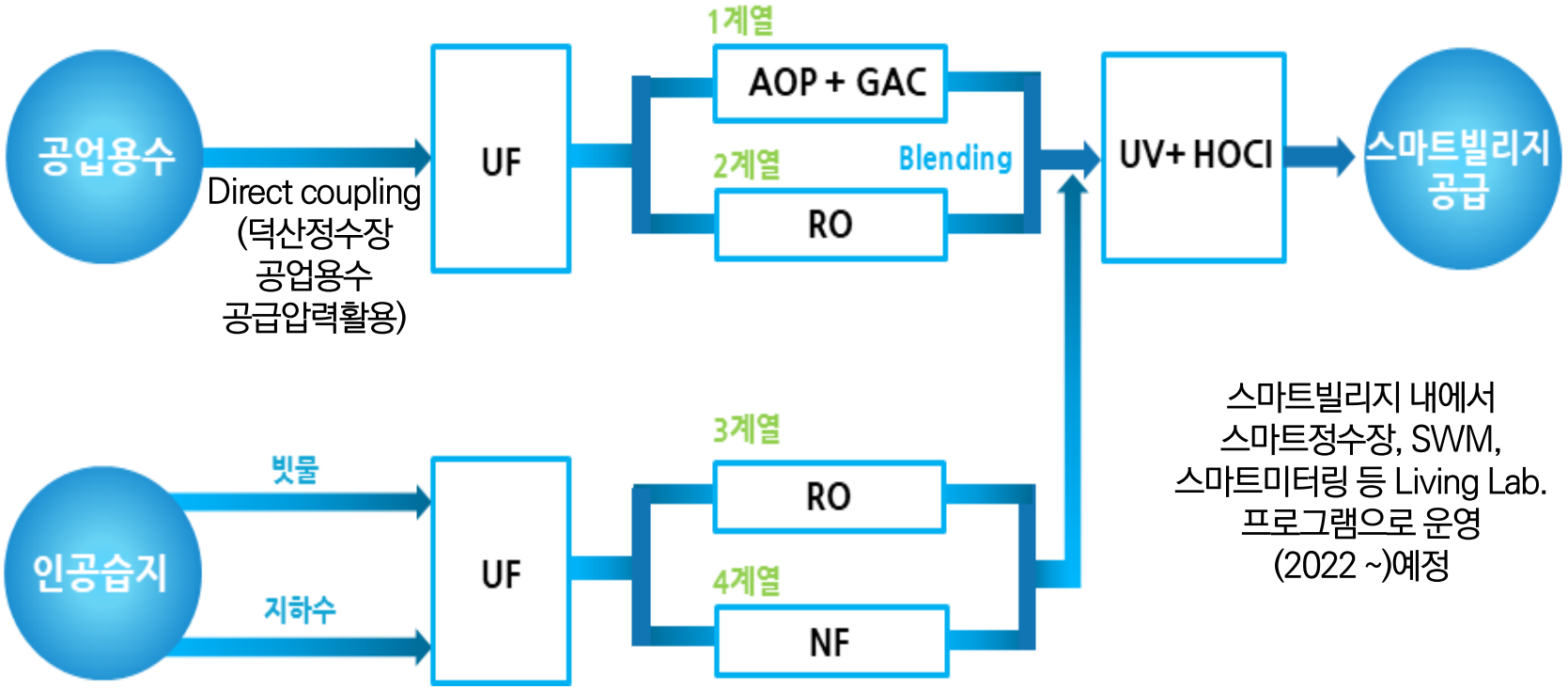
- 부산 EDC 스마트정수장 개요

| 구 분 | | 시 설 개 요 |
|---|---|---|
| 취수 시설 | 위 치 | · 부산광역시 강서구 명지동 3040-8번지 일원 |
| | 취수용량 | · Q=150㎥/일 (정수시설용량 × 50% 여과 회수율 고려) |
| | 집수매거 | · 집수매거 : D600, L=10m×2본 |
| | 취수원 | · 지하수, 빗물(인공습지)<br>· 덕산(정) 공업용수 |
| 정수 시설 | 위 치 | · 부산광역시 강서구 명지동 3005-1번지 일원 |
| | 시설용량 | · Q=100㎥/일 (신설) |
| | 처리공법 | · (공급)수질에 따라 1,2계열 생산수 Blending 처리<br>- 1계열 : 막여과(UF) + 고도정수처리(AOP+활성탄)<br>- 2계열 : 막여과(UF+RO)<br>· (보조)다수원 처리를 통한 연계 및 R&D 시설<br>- 3계열 : 막여과(UF+RO)<br>- 4계열 : 막여과(UF+NF) |
| | 건축개요 | · 대지면적 608.6㎡ / 지하1층~지상3층<br>· 연면적 661.0㎡(건폐율 37.3%, 용적률 83.2%)<br>· 에너지 절약계획(친환경 에너지)<br>- 지붕일체형 태양광집광판(BIPV)설치<br>- 냉난방설비 수열 열교환방식 적용 |

# 분산형 용수공급 시스템

## 스마트시티 스마트정수장

- 입자상물질 처리 : 컴팩트 한 수처리 시스템 구성을 위해 한외여과막(UF)으로 선정
- 고도정수처리 : 동일 수계 고도처리 공정으로 적용 사례가 많은 전·후오존($O_3$-AOP) 활성탄(GAC) 공정을 감안 처리효율 유사, 소요면적이 적으며, 설비가 단순/운영비용이 적게 드는 UV-AOP, 활성탄(GAC) 공정 적용
- 초고도처리(고압막 설비) : 유기물(TOC, 과불화합물, 신종오염물질 등) 및 염분 제거를 위하여 고압막여과 공정은 역삼투(RO) 공정을 도입하되, 보조계열과 NF(나노여과), RO 공정 설치를 통해 연계 및 다중수원 처리를 위한 R&D 시설로 활용




스마트빌리지 내에서
스마트정수장, SWM,
스마트미터링 등 Living Lab.
프로그램으로 운영
(2022 ~)예정

II

# 정수처리공정에서의 빅데이터 활용

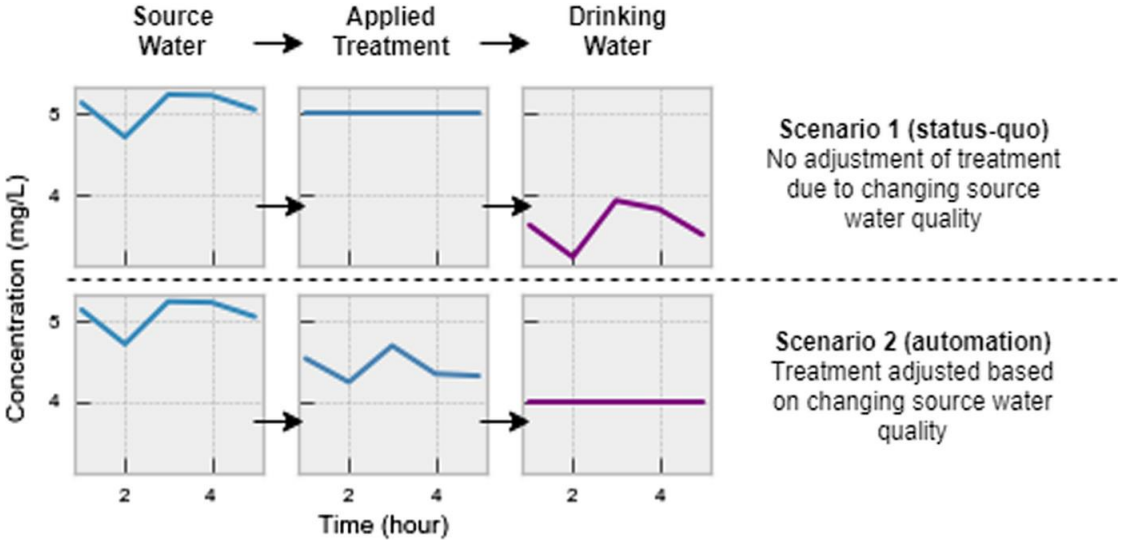
K water

# 정수처리공정에서의 도전 과제

## 기후변화, 사회적 이슈 등에 따른 정수처리공정에서의 도전과제

- 정수처리공정은 인간의 보편적인 활동에 필수적인 수돗물을 생산하는 과정으로 항상 안전한 수돗물을 안정적으로 생산할 수 있어야 함.
- 기후변화 영향으로 집중호우, 가뭄 빈도와 강도가 증가하고, 집중호우 시 하천 유출 특성, 비점오염물질의 유입 정도가 변화하거나, 가뭄에 의한 특정 수질인자 농도 증가 등 급격한 원수 수질변화가 발생하고 있음(Cho et al., 2014; Cho and Kim, 2016; MOE, 2020; Lee et al., 2016).
- 급격하게 원수 수질이 변화하는 경우, 정수처리공정 운영 난이도가 증가하며, 최종 정수수질이 안정적이지 않을 수 있음.
- 최근 다양한 수돗물 관련 이슈로 상수도 분야에 대한 관심 증대

〈수돗물 유충사고(연합뉴스 보도)〉

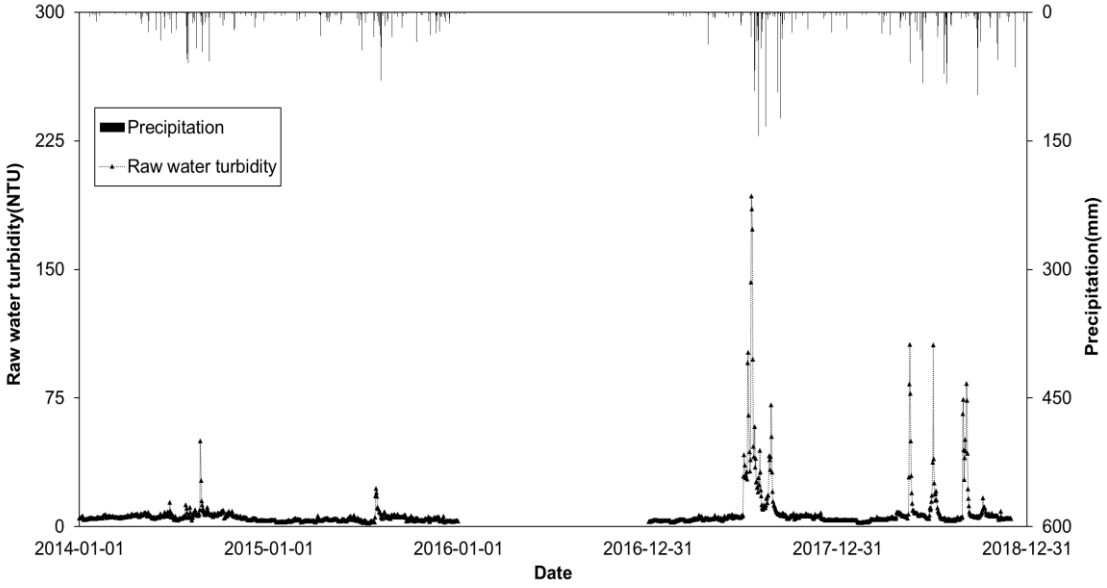



〈약품주입량 제어를 통한 정수 품질 변화 개념 예시(Aliasharafi et al., 2021)〉

# 활용 예시① 유입원수 특성 분석

## A정수장 원수 탁도 유입 현황

- 2014년, 2015년 2017년, 2018년 기상, 원수수질 데이터를 활용하여 연구대상 정수장 유입 원수의 탁도를 분석함.
- 유입 탁도 변동이 매우 크므로 고탁도 원수에 대응하기 위하여 응집제 종류, 보관기간 등을 고려하여 사전대비가 필요함.
- 2017년은 데이터 수집 기간 중 가장 많은 양의 강우량을 보인 일자가 포함되어 있으며, 정수장 유입 탁도 또한 가장 높았던 것으로 나타남.
- 데이터 수집 기간 중 탁도 값 분포의 상위 25%에 해당하는 3분위수는 6.60 NTU로 해당 탁도 이상을 보인 일수는 4개년간 358일 이며, 이 일 중 당일 강우가 발생한 일수는 127일로 35.5%를 차지함.
- 강우가 집중되는 시기에 탁도가 급증하는 경향을 보이나, 항상 강우가 발생한 당일 급증하지는 않았음.

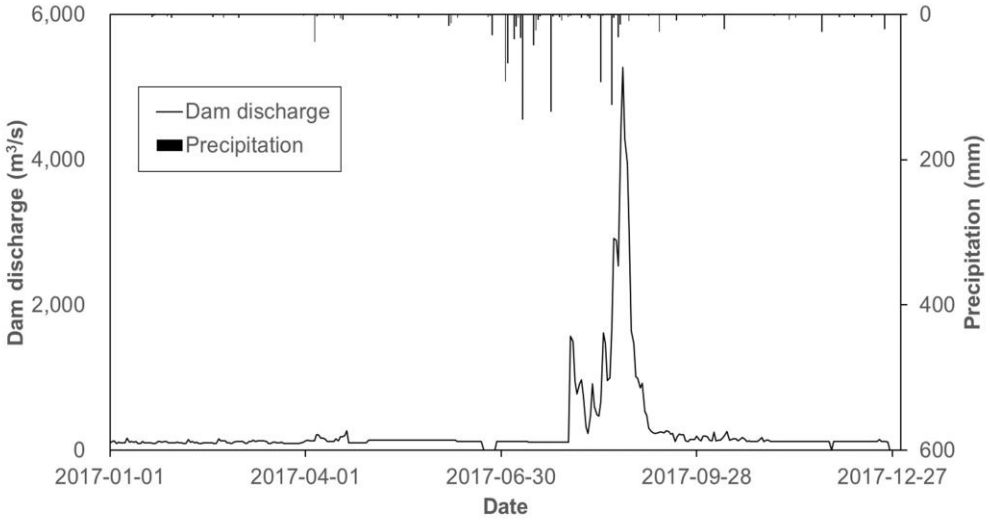


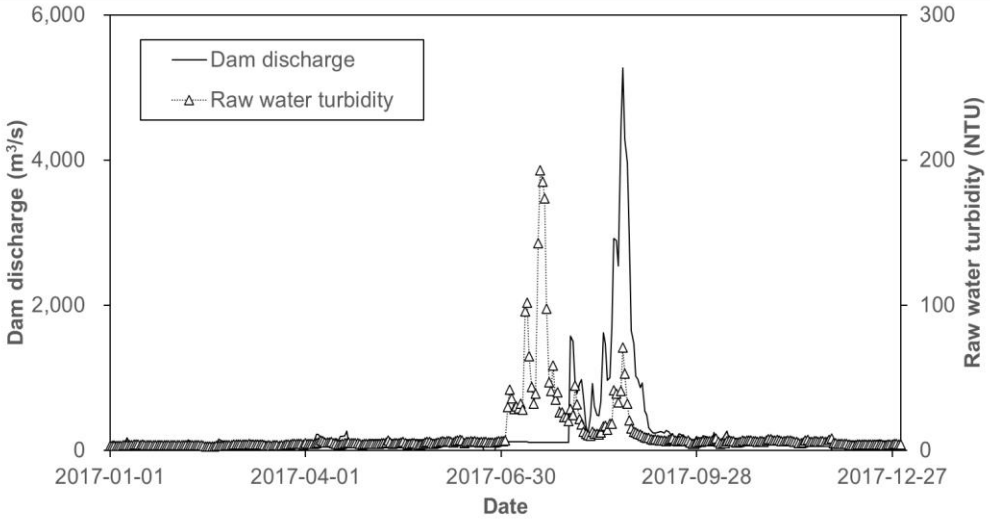

| Year | The number of days with turbidity greater than 6.60 NTU | | The number of rainy days | Max. daily precipitation (mm) | Ave. daily precipitation (mm) |
|---|---|---|---|---|---|
| | Total | rainy days on the day | | | |
| 2014 | 124 | 37 | 134 | 59.5 | 6.04 |
| 2015 | 21 | 8 | 129 | 80.0 | 6.14 |
| 2017 | 107 | 46 | 144 | 144.5 | 8.56 |
| 2018 | 106 | 36 | 135 | 96.5 | 9.65 |
| Total | 358 | 127 | 542 | 144.5 | 9.65 |

# 활용 예시① 유입원수 특성 분석

## A정수장 원수 탁도 유입 원인 검토

- 데이터 수집기간 중 가장 높은 탁도와 일일 강우량을 갖는 기간인 2017년에 대하여 연구대상 정수장의 취수원으로부터 약 25 km 상류에 위치한 P 댐의 방류량 데이터를 수집하여 일일 강우량, 원수 탁도의 관계를 분석함.
- P 댐에서는 수위와 저수량을 고려하여 강우가 집중적으로 발생한 기간 이후 방류량을 증가시킴.
- 강우가 발생하지 않았으나, 댐 방류량이 5,270.4 $m^3/s$로 가장 많았던 2017년 8월 25일에 원수 탁도는 71 NTU로 주변 일자의 탁도와 비교하여 첨두값을 나타냄.
- 강우 발생이 원수 탁도 변화에 영향을 미치는 것으로 판단되며, 추가적으로 집중호우 발생 이후 댐 방류 등에 따라서도 원수 탁도가 영향을 받을 수 있을 것으로 판단됨.
- 응집제 비축 등 고탁도 원수 유입 사전대비를 위하여 강우 발생 뿐 아니라 댐 방류 가능성 또한 고려할 필요가 있음.

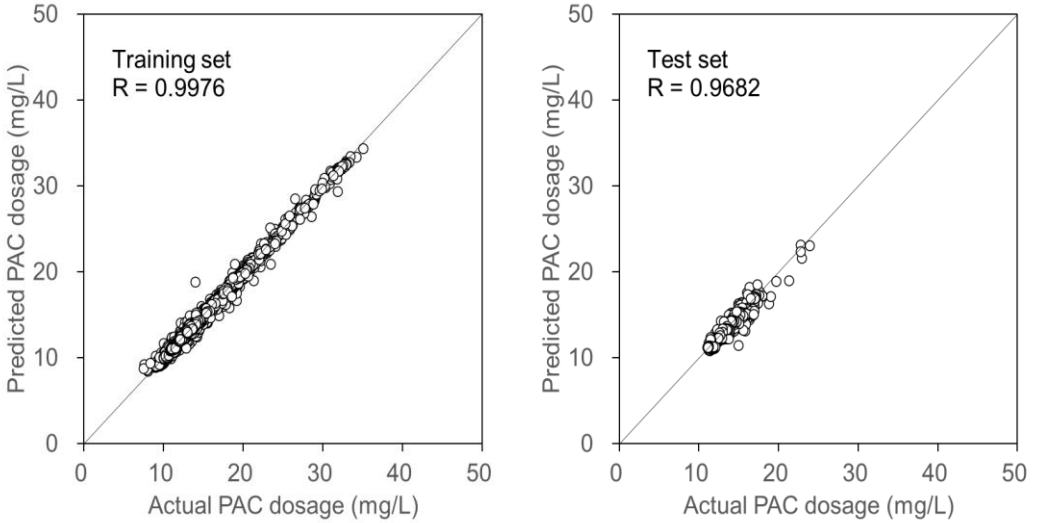


# 활용 예시② 약품 주입률 결정

## 약품 주입률 결정 모델 개발

- 집중호우 등으로 탁도 급증 시 적절한 응집제 주입률 결정 어려움.
- SVM, RF, MLP, GRU 모델을 활용하여 원수 수질(탁도, 수온, 알칼리도, 전기전도도) 기반 PAC 주입률 결정 모델을 개발하였으며, RF 모델의 성능이 가장 우수
  - 상관계수는 모두 매우 높은 값으로 나타남에 따라 검증 데이터에 대한 오차 지표를 기준으로 모델 성능을 비교함.



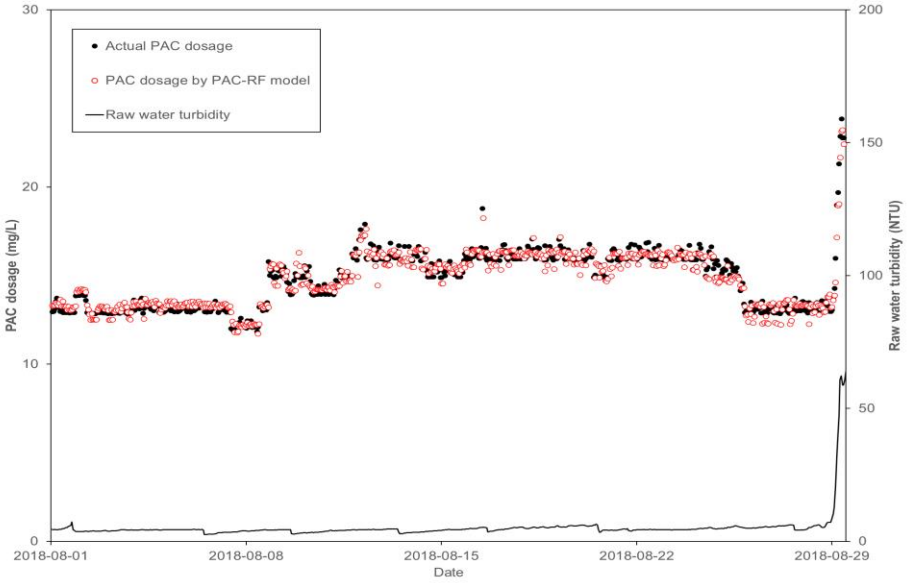

| Model | PI | Model performance | |
|---|---|---|---|
| | | Training | Test |
| PAC -SVM | R | 0.9731 | 0.9226 |
| | MAE | 0.70 mg/L | 0.53 mg/L |
| | MAPE | 6.22 % | 3.86 % |
| PAC -RF | R | 0.9976 | 0.9682 |
| | MAE | 0.13 mg/L | 0.30 mg/L |
| | MAPE | 1.02 % | 2.08 % |
| PAC -MLP | R | 0.9928 | 0.9025 |
| | MAE | 0.21 mg/L | 0.61 mg/L |
| | MAPE | 1.66 % | 4.15 % |
| PAC -GRU1 | R | 0.9913 | 0.9873 |
| | MAE | 0.65 mg/L | 0.41 mg/L |
| | MAPE | 5.33 % | 2.90 % |

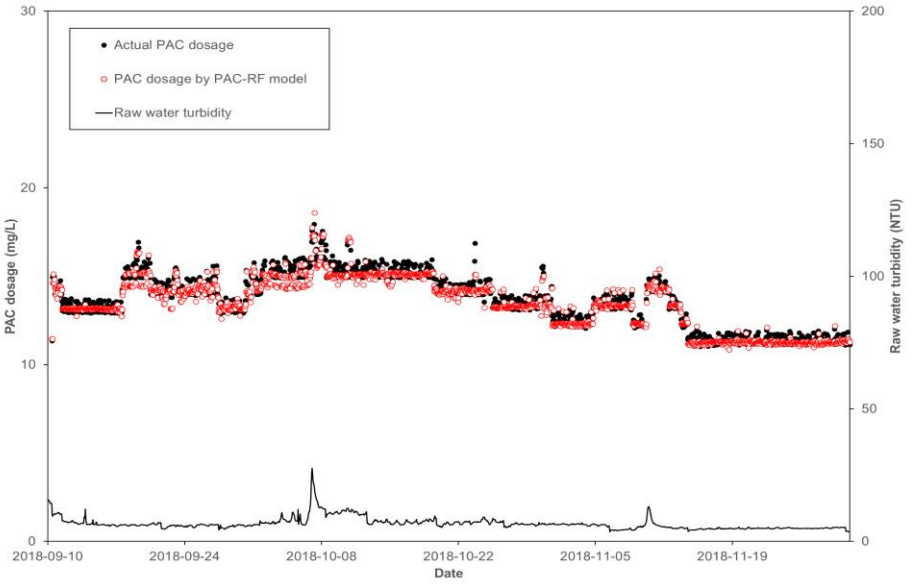
# 활용 예시② 약품 주입률 결정

## 약품 주입률 결정 모델의 활용

- PAC 주입률 결정 모델의 학습 및 검증에 활용되지 않은 예측기간 111일(2018.08.01 ~ 2018.08.29, 2018.09.10. ~ 2018.11.30)에 대하여 실제 주입률과 PAC-RF 모델에 의해 결정된 주입률을 비교함.
- 예측기간동안 실제 PAC 주입량과 PAC-RF 모델에 의해 결정된 PAC 주입량은 0.6 % 차이를 나타내므로, 실제 응집제 주입률을 결정할 때의 의사결정과정이 잘 모델화 된 것으로 판단됨.
- 유입 원수 수질을 고려하여 약품주입률 결정시 활용 가능

| Operation method | | Period | | |
|---|---|---|---|---|
| | | 08.01 ~ 08.29 | 09.10 ~ 11.30 | Total |
| Actual operation | PAC dose (kg) | 112,954 | 273,346 | 386,300 |
| PAC-RF model | PAC dose (kg) | 112,803 | 271,116 | 383,919 |
| | Difference from actual dose | -151 kg | -2,230 kg | -2,381 kg |
| | | -0.1 % | -0.8 % | -0.6 % |

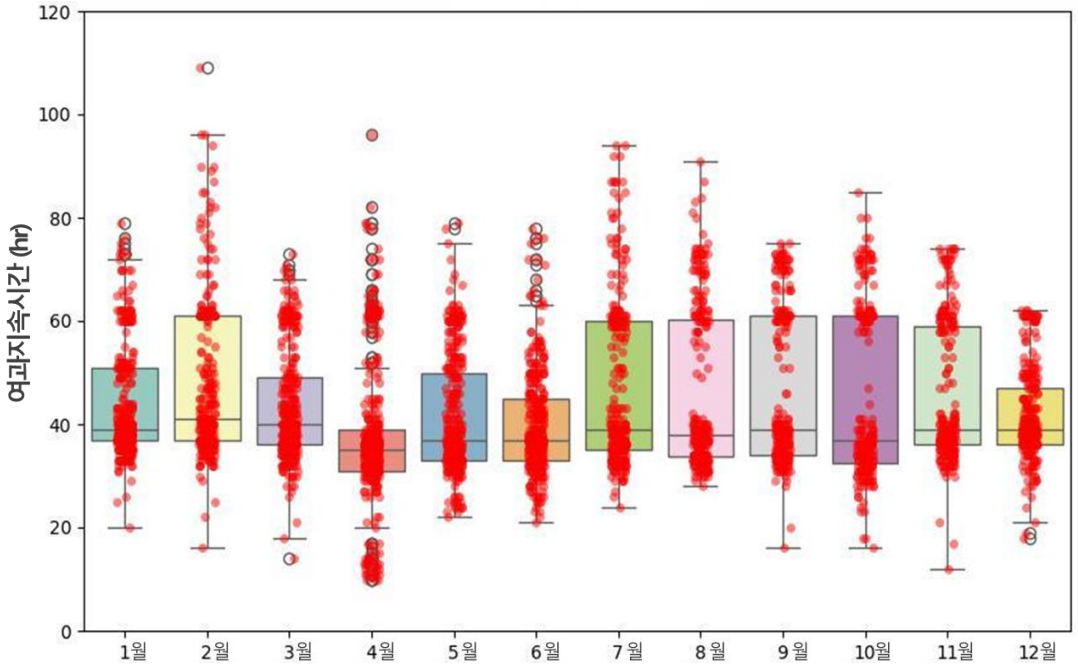


# 활용 예시③ 여과지 운영 현황 분석

## 여과지 운영 분석

- B 정수장의 4년간의 여과지 운영 자료를 활용하여 분석
- 계절별 원수 수질이 여과지 운영에 영향을 미칠 것 으로 판단하여 여과지속시간에 대하여 월별로 월별 Box-plot 작성
- Box 구간을 기준으로 통계적 기법에 의해 일반적인 분포를 벗어나는 여과지속시간이 많고, 짧은 여과지속시간으로 운영된 4월의 경우, 다른 시기보다 여과지 운영에 어려움이 있었을 것으로 판단됨.
  → 실제 시네드라 유입으로 여과지 폐색 가속화에 따라 여과지속시간 짧게 운영




| 구분 | | 1월 | 2월 | 3월 | 4월 | 5월 | 6월 | 7월 | 8월 | 9월 | 10월 | 11월 | 12월 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 분석대상 여과주기 | | 340 | 285 | 333 | 385 | 357 | 360 | 289 | 256 | 243 | 243 | 245 | 282 |
| 여과지 속시간 (시간) | 최대 | 79 | 109 | 73 | 96 | 79 | 78 | 94 | 91 | 75 | 85 | 74 | 62 |
| | 최소 | 20 | 16 | 14 | 10 | 22 | 21 | 24 | 28 | 16 | 16 | 12 | 18 |
| | 평균 | 43.7 | 47.8 | 43.6 | 37.1 | 41.3 | 39.8 | 46.6 | 45.6 | 46.2 | 45.1 | 45.7 | 41.7 |

# 활용 예시④ 여과지 운영 의사결정

## 여과수 탁도 예측 모델 개발

- 정수처리공정 운영 시, 수도법 상 정수처리기준을 만족하여야 하며,

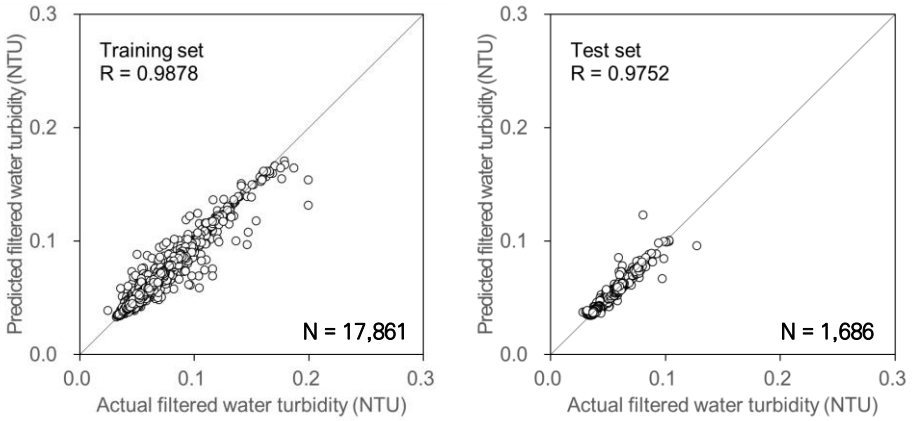
  정수처리기준에서는 탁도를 기준으로 정수처리기준 만족여부 판정

  – 통합 여과수에 대하여 1일 6회 이상(최소 4시간 간격)으로 탁도 측정

  – 급속여과 방식은 모든 측정 시료에 대하여 1.0 NTU 이하를 만족하며,

  매월 측정 시료의 95% 이상이 0.3 NTU 이하가 되어야 함.

- SVM, RF, MLP, GRU 모델을 활용하여 수질 자료(침전지 유출수의 탁도 및 pH) 유입유량, 개별 여과지 운영 조건(여과 개시 후 운영 시간, 누적 탁도) 기반 통합 여과수 탁도 예측 모델을 개발하였으며, RF 모델의 성능이 가장 우수




| Model | PI | Model performance | |
|---|---|---|---|
| | | Training | Test |
| FT-SVM-2 | R | 0.8869 | 0.7606 |
| | MAE | 0.0109 NTU | 0.0193 NTU |
| | MAPE | 22.66 % | 44.03 % |
| FT-RF-2 | R | 0.9878 | 0.9752 |
| | MAE | 0.0009 NTU | 0.0014 NTU |
| | MAPE | 1.56 % | 2.91 % |
| FT-MLP-2 | R | 0.9808 | 0.9725 |
| | MAE | 0.0012 NTU | 0.0070 NTU |
| | MAPE | 2.18 % | 15.31 % |
| FT-GRU1-2 | R | 0.9987 | 0.9174 |
| | MAE | 0.0007 NTU | 0.0052 NTU |
| | MAPE | 1.24 % | 9.72 % |

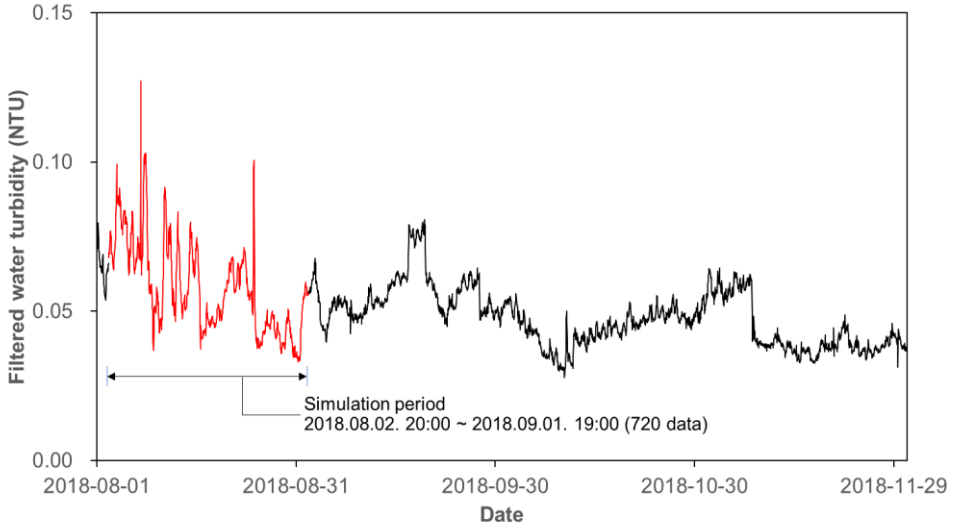
# 활용 예시④ 여과지 운영 의사결정

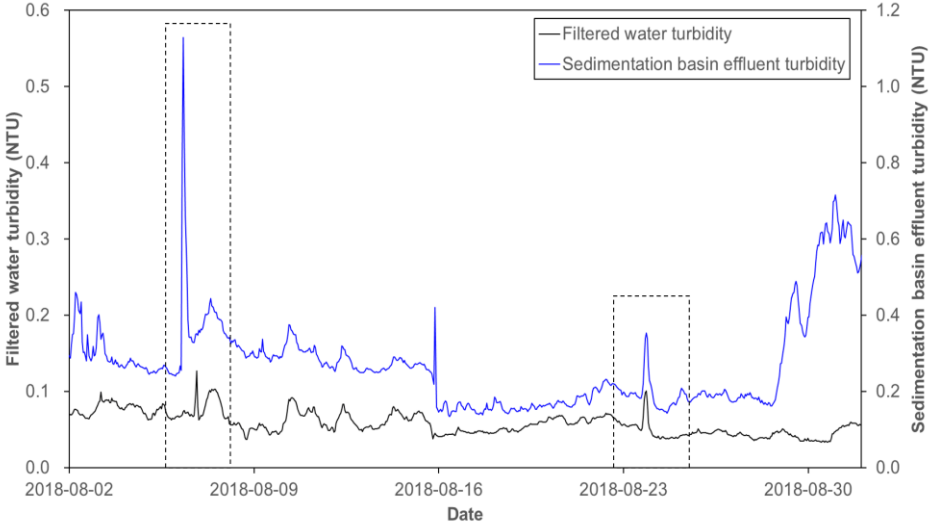
## 여과수 탁도 예측 모델의 활용

- 개별 여과지 유입 탁도 누적값을 고려하는 모델을 활용하여 여과지속시간 변화에 따른 여과수 탁도 예측을 수행하여 모델의 활용성을 검토함.

1) 모의 기간 및 조건 선정

- 모델 학습 및 검증에 사용하지 않은 2018. 08.01 ~ 11.30. 까지의 기간을 대상으로 30일(720시간)에 대한 시간단위 여과수 탁도의 변동계수(평균 대비 표준편차)를 산정한 결과, 2018. 08. 02. 20:00 ~ 09.01. 19:00 구간의 변동계수는 0.2636으로 가장 크게 나타남.
- 변동계수가 가장 큰 구간을 모의 기간으로 선정하였으며, 해당 기간은 여과수 탁도가 0.1 NTU를 8회 초과하며, 개별 여과지는 평균 11회의 여과주기를 갖고, 평균 63.5 시간(최소 41.0 시간, 최대 77.7 시간)의 여과지속시간을 갖음.
- 여과지속시간 시나리오는 최소 40시간부터 최대 80시간까지 10시간 간격으로 설정하였으며, 각 시나리오별 통합 여과수 탁도를 예측함.



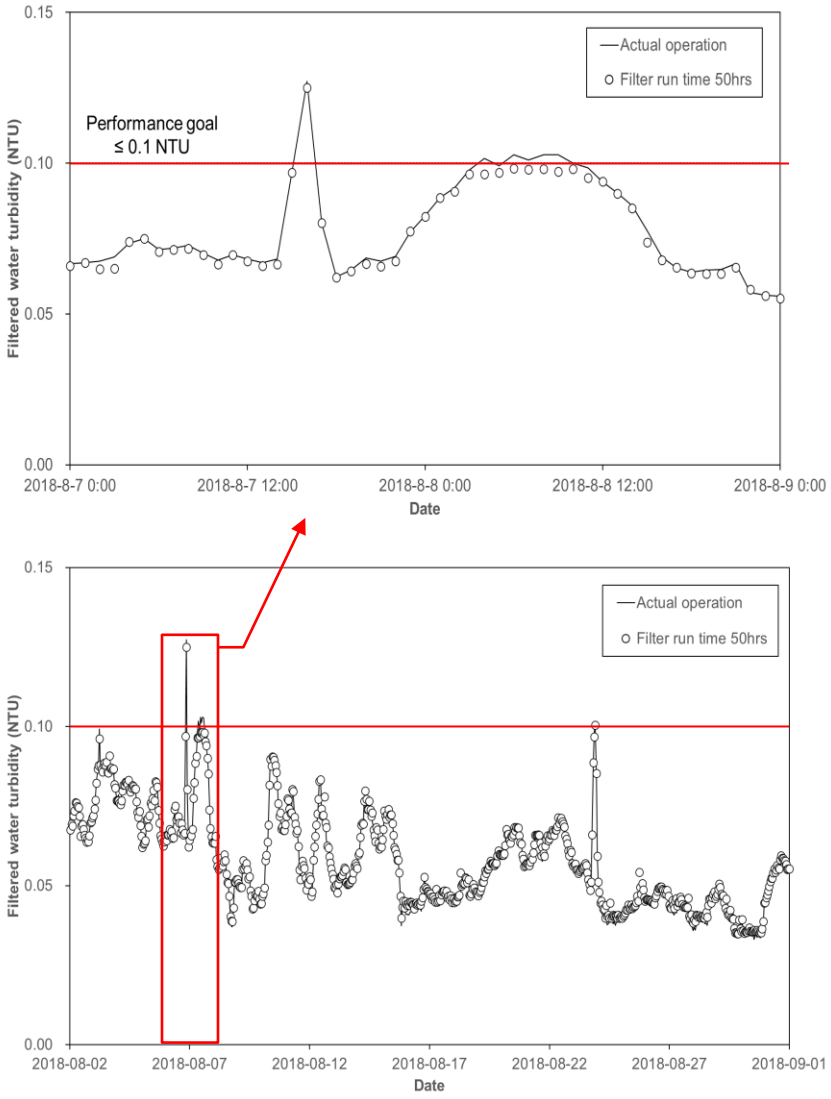
# 활용 예시④ 여과지 운영 의사결정

## 여과수 탁도 예측 모델의 활용

2) 시나리오 모의 결과

- 여과수 탁도 0.1 NTU 초과 횟수가 가장 적은 경우는 여과지속시간 50시간, 80시간 일 때 인 것으로 나타났으나, 하절기 고탁도 원수 유입을 대비하여 여과지속시간을 50시간으로 조정하는 것이 바람직할 것으로 판단됨.
- 여과수 탁도 예측 모델을 활용하여 여과 공정 운영 진단, 적정 여과공정 운영 계획 수립 등에 활용 가능
- 응집제 주입률, 침전수 탁도 예측과 연계하여 여과수 탁도 예측 모델의 운영 시 좀더 중장기적으로 모의 가능

| Filtration operation | Coefficient of variation | The number of hours exceeded 0.1 NTU | Filtered water turbidity (NTU) | | |
|---|---|---|---|---|---|
| | | | Max. | Ave. | Min. |
| Actual operation(63.5hrs) | 0.2636 | 8 | 0.1272 | 0.0580 | 0.0332 |
| Filter run time 40 hrs | 0.2564 | 7 | 0.1287 | 0.0583 | 0.0348 |
| Filter run time 50 hrs | 0.2536 | 2 | 0.1250 | 0.0582 | 0.0350 |
| Filter run time 60 hrs | 0.2569 | 9 | 0.1292 | 0.0583 | 0.0350 |
| Filter run time 70 hrs | 0.2539 | 6 | 0.1209 | 0.0582 | 0.0351 |
| Filter run time 80 hrs | 0.2531 | 2 | 0.1297 | 0.0582 | 0.0353 |



감사합니다
K water