

# Digital twin for smart water and wastewater treatment works

Lijun Zhang | PhD, Pr Eng

Professor,

Department of Electrical and Electronic Engineering,

Department of Mechanical and Mechatronic Engineering,

Stellenbosch University

# Contents



Aims of being smart

DT for water and wastewater treatment overview

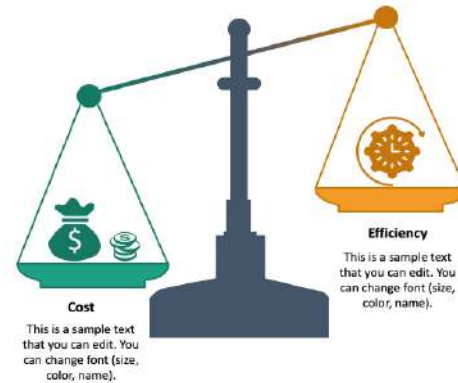
Core functions of the DT

Case studies

# Aims of smart water/sewer treatment works



**Environmental sustainability**



**Cost-effectiveness**

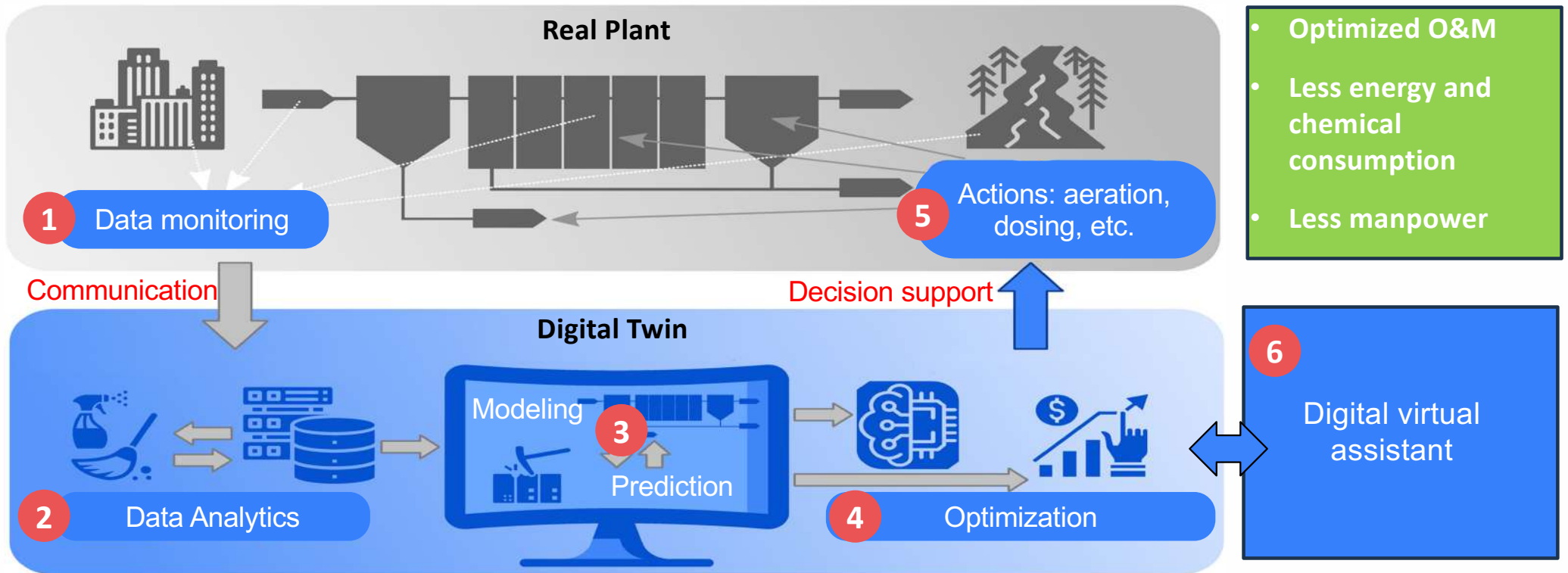


**Process digitalisation**



**Autonomous operation**

# DT for WWTP overview



# Data communication

## Features:

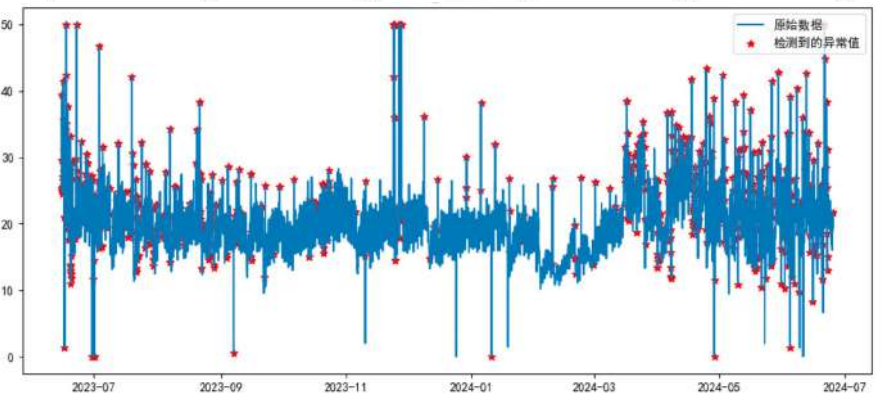
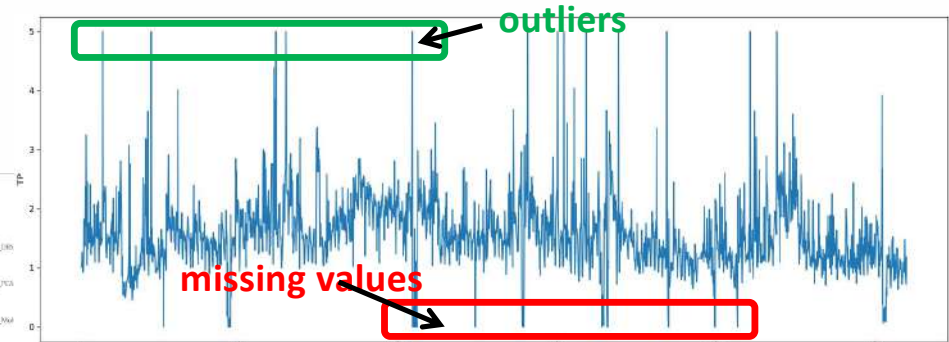
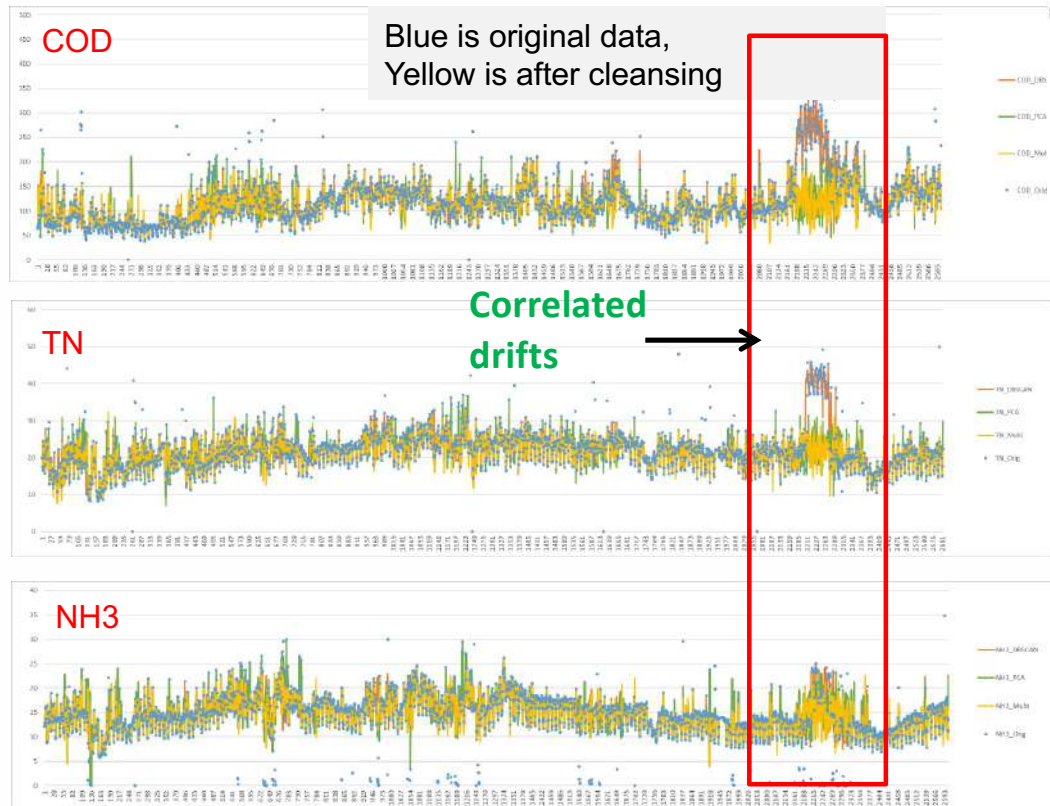
- OPC UA compliance real-time data communication
- Supports MQTT for IoT devices
- Alarm data monitoring and analytics



Real-time data communication with SCADA and IoT devices

# Data analytics — cleansing

**Problem:** Sensor blockage/malfunction, measurement drifting, communication errors



## Performance Indicators:

- False negative rate  $\leq 2\%$ , false alarm rate  $\leq 6\%$ , detection delay  $\leq 5s$ , accuracy  $\geq 90\%$ , and F1-Score  $\geq 90\%$ .

# Data analytics — cleansing

## Solution:

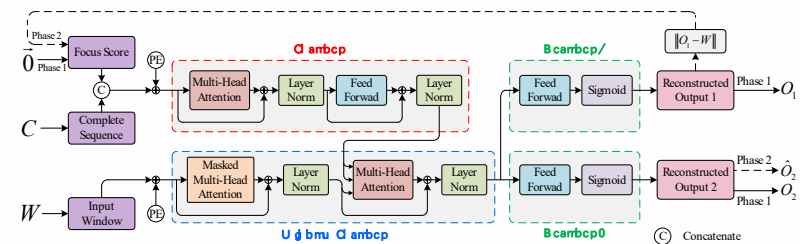
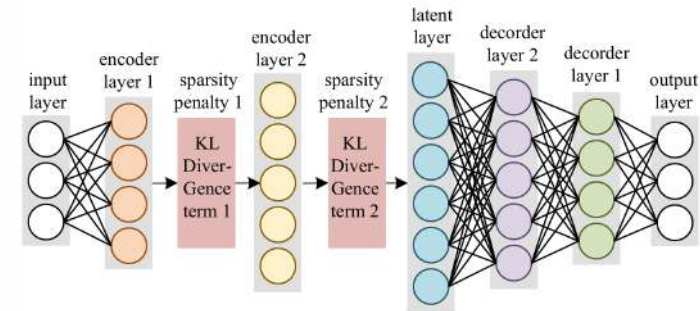
- Anomaly detection (identify anomalies)
- Data imputation (filling back)

## Features:

- Single variable + multi-variable algorithms
- Supports both online (clean as data arrives) and offline (batch clean historical data)
- Unsupervised machine learning to navigate unlabeled data (PCA, clustering, SAE, Transformer)

## Benefits:

- Continuous monitoring of data quality (raise alarm if many anomalies from a sensor)
- Quality data support for simulation and operation optimization
- No need for manual labelling of anomalies

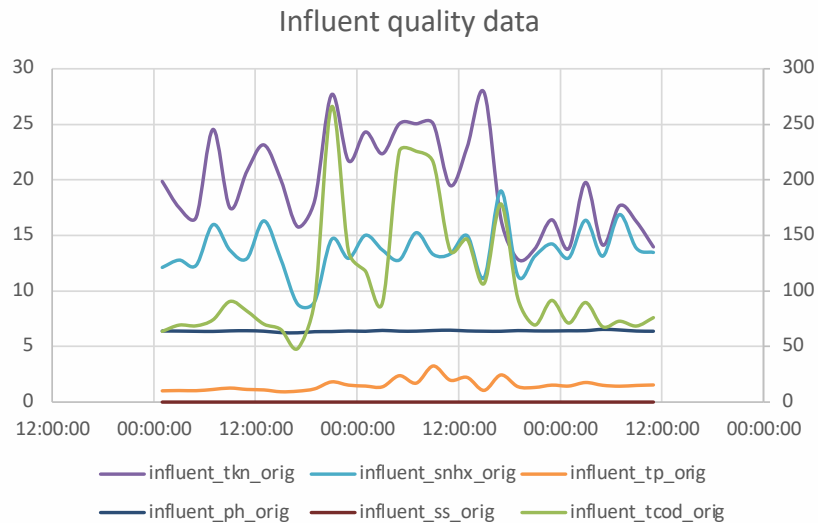


Multi-variable data cleansing combining unsupervised learning techniques

# Data analytics — forecasting

## Problem:

- Lacks proactive operating support using influent forecast feed forward
  - Inherent influent quality fluctuations
  - Sudden changes due to abnormal operating conditions such as heavy rain, high temperature, etc.
  - Instantaneous feedback control doesn't work that well due to large process delays.
- Lacks capability to support WWTP prescriptive analytics of “what-if” scenarios



Influent variation is common. How to deal with it?

Forecast Horizon	Parameter	Mean absolute percentage error (MAPE)
4-hour ahead	Influent flow rate	6%
	Influent COD	20%
	Influent Total Nitrogen	10%
	Influent Ammonia Nitrogen	10%
	Influent Total Phosphate	20%
	Influent pH	2%
	Influent temperature	1%
72-hour ahead	Influent flow rate	25%
	Influent COD	25%
	Influent Total Nitrogen	15%
	Influent Ammonia Nitrogen	15%
	Influent Total Phosphate	25%
	Influent pH	5%
	Influent temperature	2%

# Data analytics — forecasting

## Solution:

- Deep-learning neural network for multi-step ahead forecasting (customized Transformer, Informer, GRU-attention mechanism)

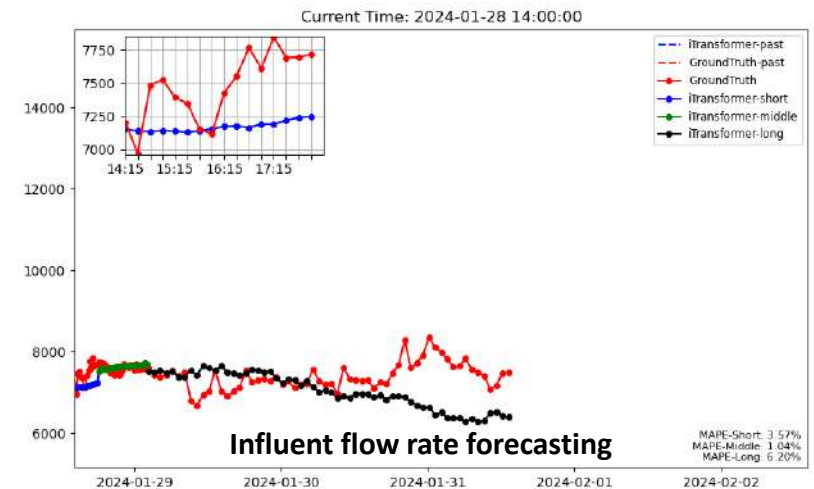
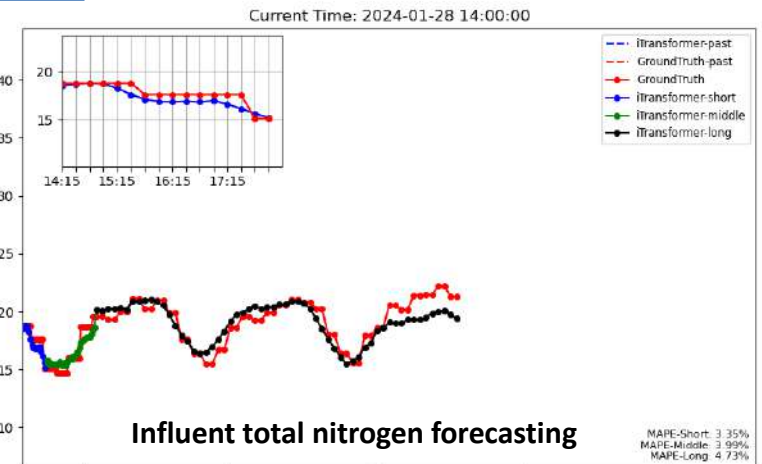
## Features:

- Supports long-, medium- and short-term forecasting
- Supports multi-variable forecasting

## Benefits:

- Better effluent compliance
- Better support for emergency planning and future upgrading (predictive analytics)

Deep learning multi-step, multi-resolution time series forecasting





# Twining — modelling

## Solutions:

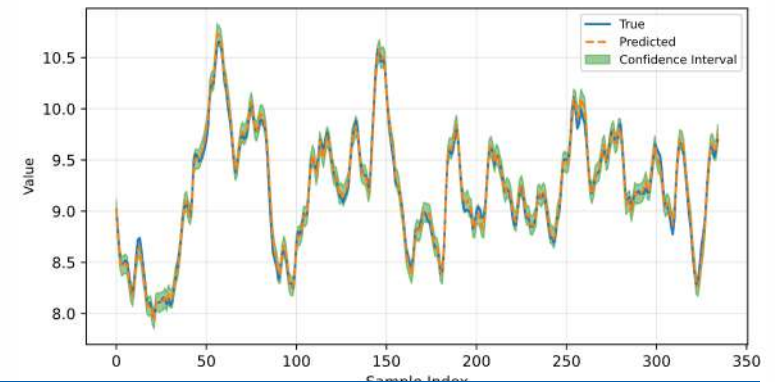
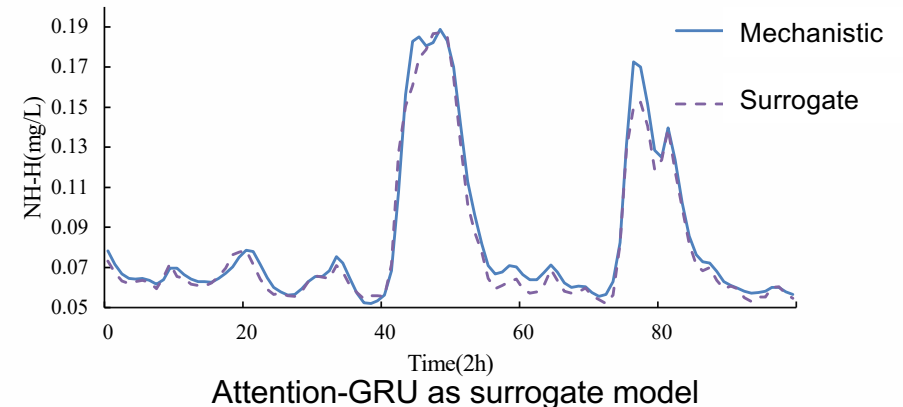
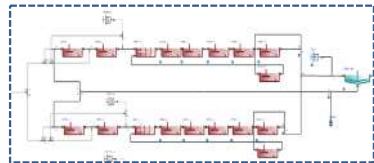
- Hybrid black-box white-box modelling for point and interval prediction
- Deep transfer learning based model self-calibration

## Features:

- Works with mechanistic models
- Deterministic and stochastic predictions
- Surrogate model for fast simulation/optimization

## Benefits:

- Better long-term model accuracy + robust interval prediction
- Support for (fast) operation optimization + robust against uncertainties



Hybrid modelling + model auto-calibration for accurate long term prediction and fast optimisation

# Twining — prescriptive analytics

## Problems:

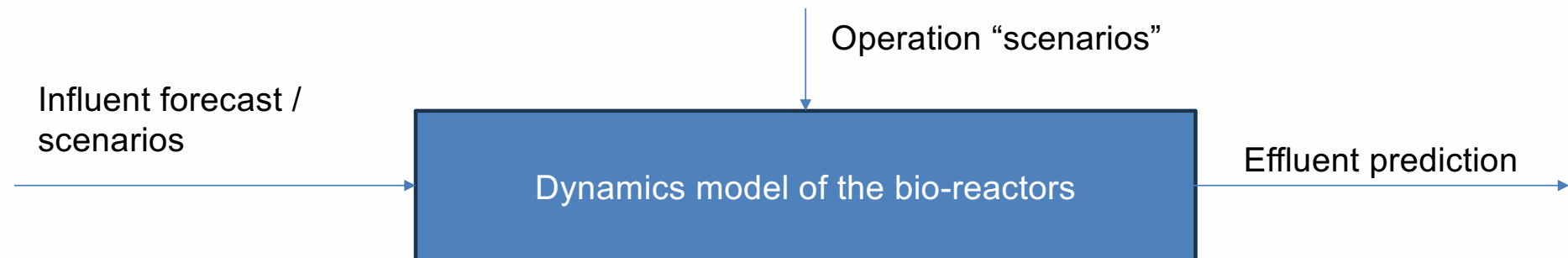
- Lacks accurate analysis/preview of plant operation scenarios (e.g., shock influent, equipment availability, etc.)

## Solution:

- Combines influent forecasting and accurate process model to support “what-if” scenario simulation

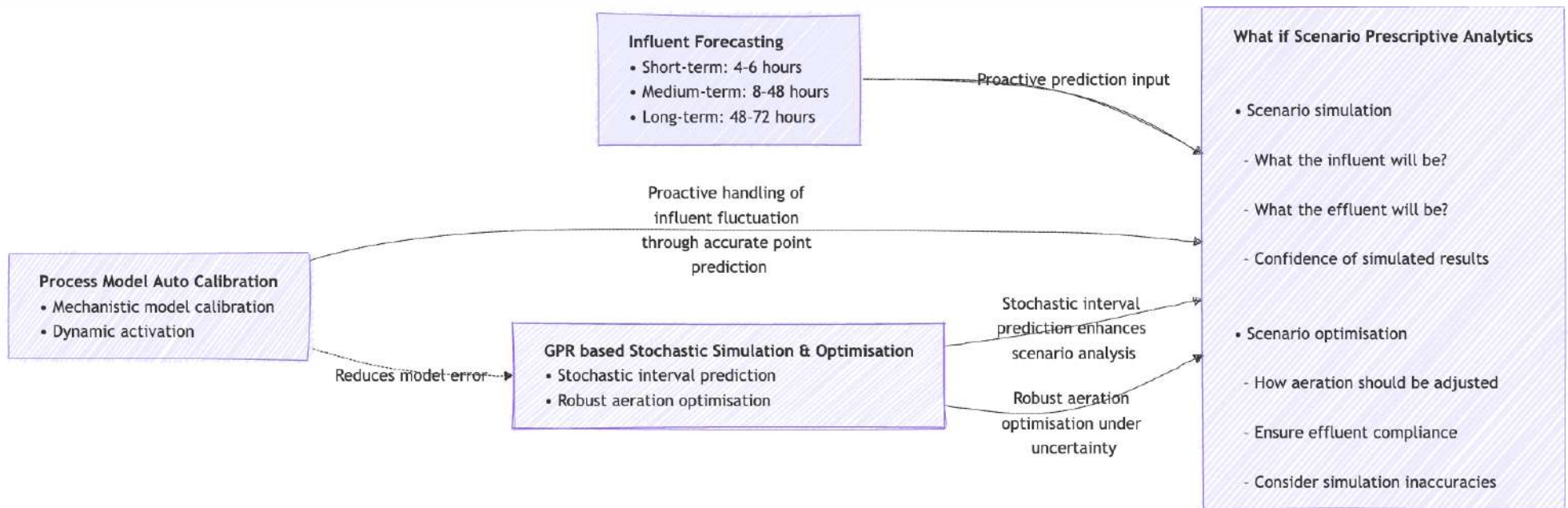
## Benefits:

- Better insights into different plant operating conditions supporting operations and plan upgrade



**“What-if” scenario analysis support**

# Twining — prescriptive analytics



“What-if” scenario analysis support

# Decision support — optimisation

## Problem:

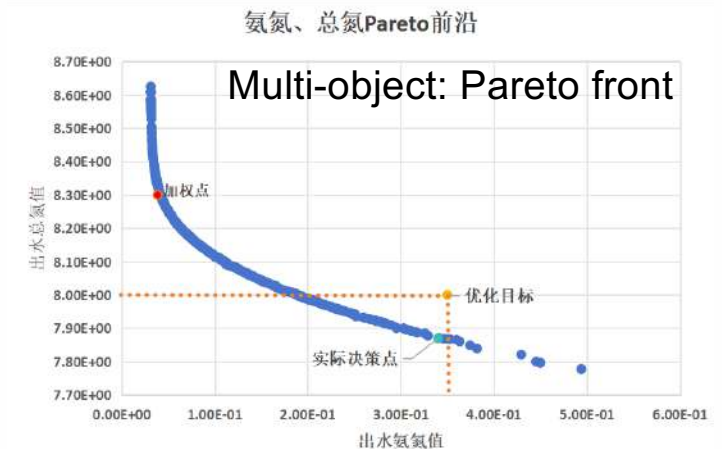
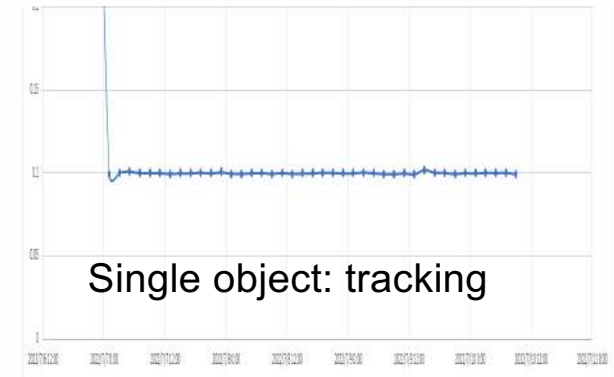
- The objectives of a WWTP usually include (conflicting goals):
  - Ensure effluent quality
  - Minimize operating costs
- The optimised decisions need to be robust against uncertainties

## Solution:

- AI and MPC empowered single and multi-objective optimization algorithms (combing meta-heuristics, swarm intelligence and optimal control theory)
- Robust optimization based on deep learning stochastic prediction

## Benefits:

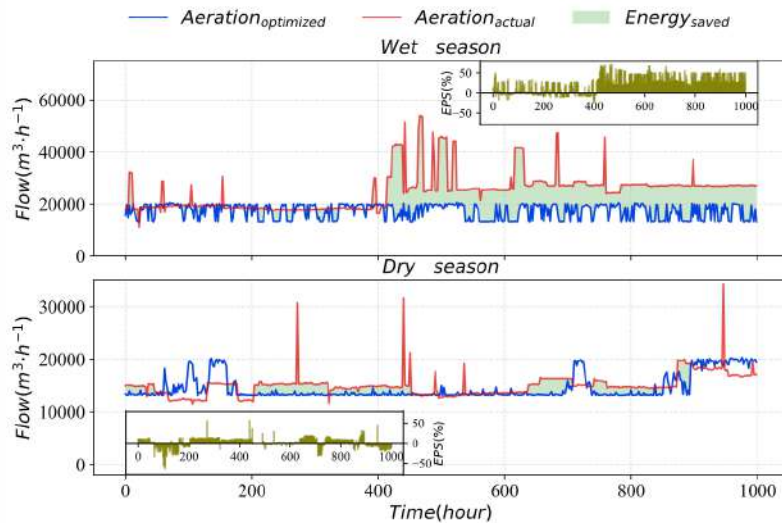
- Reduces energy and chemical costs
- Improves effluent compliance
- Supports robust optimisation
- Supports multi-objective Pareto optimization



Reduce human knowledge reliance, improve operation efficiency.

# Decision support — optimisation

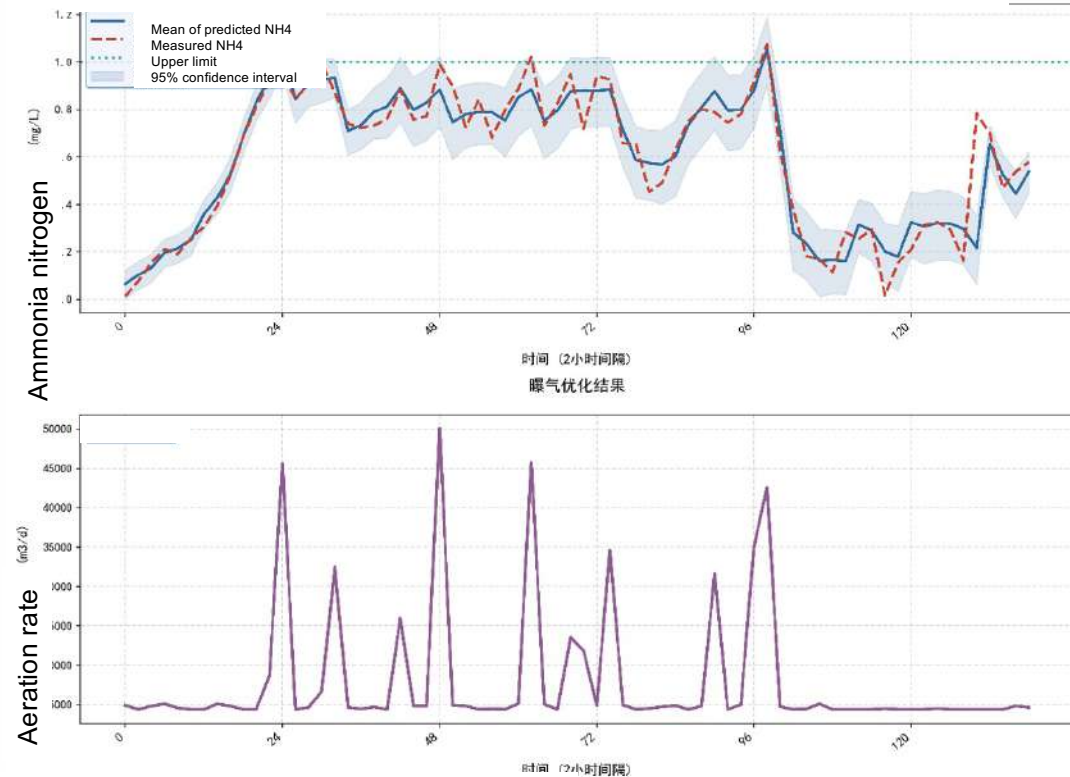
Illustrative example of energy savings achieved through AI aeration optimization



- Option 1: DO setpoint optimization
- Option 2: Direct aeration optimization

~20%-40% energy savings, 97% effluent compliance

Robust aeration optimization: minimize energy usage



Reduce human knowledge reliance. Reduce energy usage while keeping effluent compliant.

# Real-time control — controllers

## Problem:

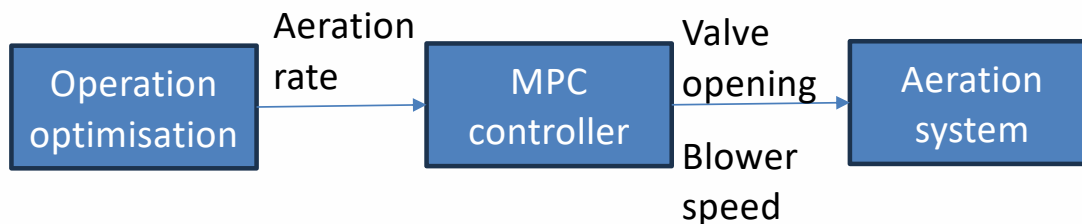
- PID control of the blower and valves lacks performance optimization

## Solution:

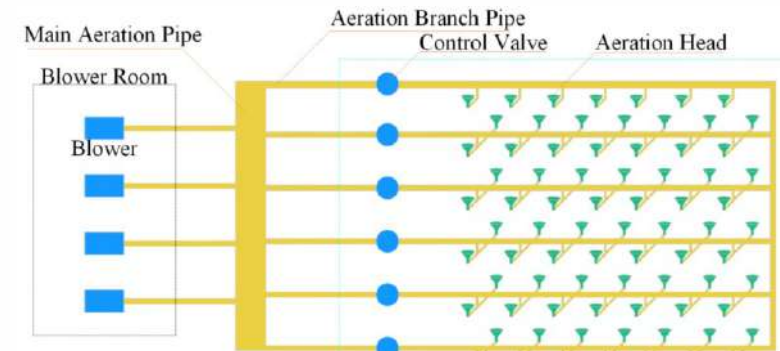
- Advanced optimal controller for blower speed and valve opening for precision aeration

## Benefits:

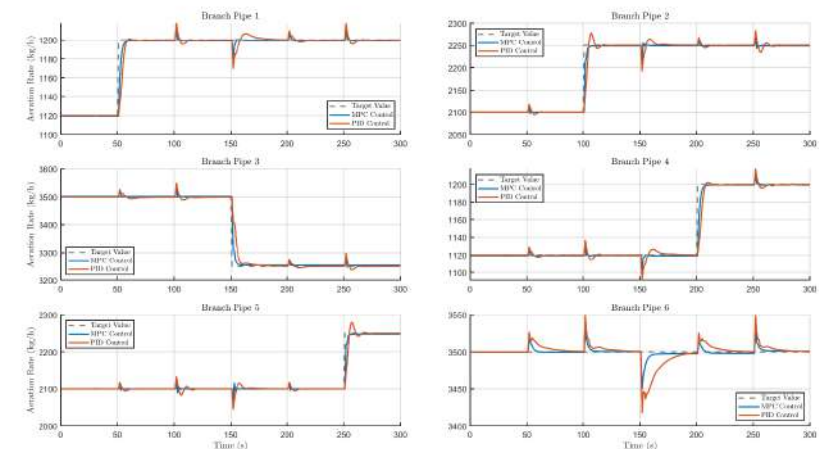
- Better tracking of decision from the optimization layer
- Further reduces energy consumption (~2%-5%)



**Optimal actuator control for closed-loop implementation**



System configuration



Performance comparison with PID

# Digital assistant — integration with DT

## Problem:

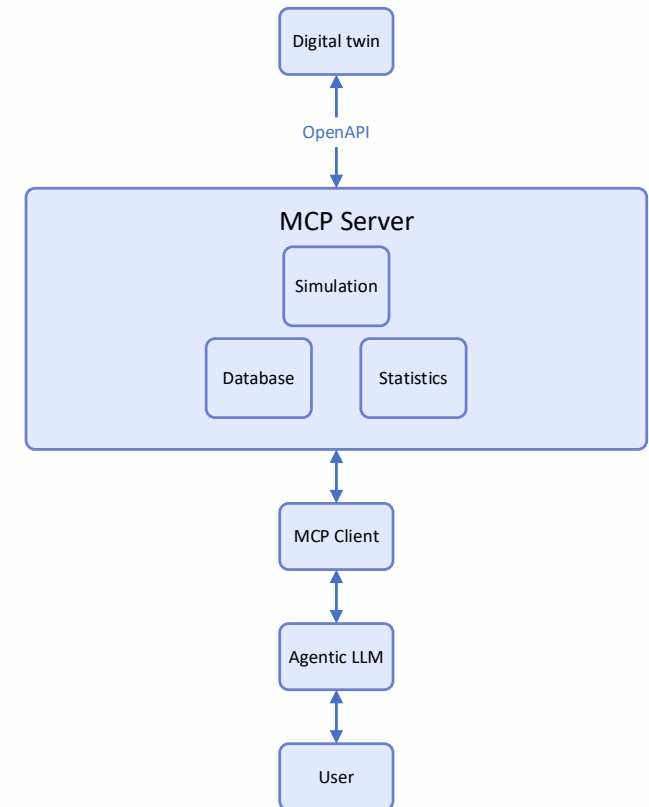
- Lacks natural language interface to gain insights into the plant's operation
- Cannot utilize expert knowledge effectively

## Solution:

- Integration of a DVA empowered by local knowledge base and access to DT functionalities through MCP servers

## Benefits:

- Better operating knowledge management
- Real-time access to DT's functions, including prescriptive analytics, operation optimization, etc.
- Integration with root cause analysis of operating conditions (such as effluent violations) for better decision making
- An innovative training platform



Natural language interface for access to DT functions, more versatile utilization of the DT

# Digital assistant — integration with DT

- **Start Time:** 2024-05-15 01:00:00
- **End Time:** 2024-05-15 03:00:00
- **Data Source:** 1 (cleaned data)
- **Frequency:** 120
- **Forecast Step:** 5

## How to Proceed

Let me know the details for your simulation, and I'll execute it for you. For example:

- "I want to run an AAO offline simulation from 2024-05-15 01:00:00 to 2024-05-15 03:00:00 with cleaned data and a forecast step of 5."
- Would you like to proceed with a specific simulation?

You: I want to run an AAO offline simulation from 2024-05-15 01:00:00 to 2024-05-16 01:00:00 with cleaned data

AI:

The AAO offline simulation has been successfully completed for the time range from 2024-05-15 01:00:00 to 2024-05-16 01:00:00 using cleaned data.

## Next Steps:

- If you'd like to query the simulation results from the database, let me know the specific variables or data you're interested in.
  - If you need further analysis or visualization of the results, just ask!
- Would you like to proceed with any of these options?

You: I want to query the variables aao\_influent\_1\_1\_tp\_ted and aao\_influent\_1\_1\_fsrhx\_tkn\_ted

AI:

Here are the results for the variables aao\_influent\_1\_1\_tp\_ted (Total Phosphorus) and aao\_influent\_1\_1\_fsrhx\_tkn\_ted (Free Ammonia Nitrogen) from the AAO offline simulation:

## Preview of Data

Timestamp	TP (mg/L)	Free Ammonia Nitrogen (mg/L)
2024-05-15 01:00:00	3.476	0.577
2024-05-15 03:00:00	3.624	0.603
2024-05-15 05:00:00	3.772	0.578
2024-05-15 07:00:00	3.919	0.608
2024-05-15 09:00:00	2.138	0.379

**Total Rows: 12**

## Next Steps:

- The full data is saved in a JSON file at: `static\ab_results\db_offline_simulation_data_aao_2024-05-15_01-00-00_2024-05-16_01-00-00.json`.
  - Would you like to:
    1. **Visualize this data** (e.g., plot trends over time)?
    2. **Perform statistical analysis** (e.g., fit a model to the data)?
    3. **Query additional variables?**
- Let me know how you'd like to proceed!

You: Visualize this data and Perform statistical analysis

AI:

Here's the **visualization and statistical analysis** of the variables aao\_influent\_1\_1\_tp\_ted (Total Phosphorus) and aao\_influent\_1\_1\_fsrhx\_tkn\_ted (Free Ammonia Nitrogen) from the AAO offline simulation:

## 1. Statistical Analysis

### Total Phosphorus (TP)

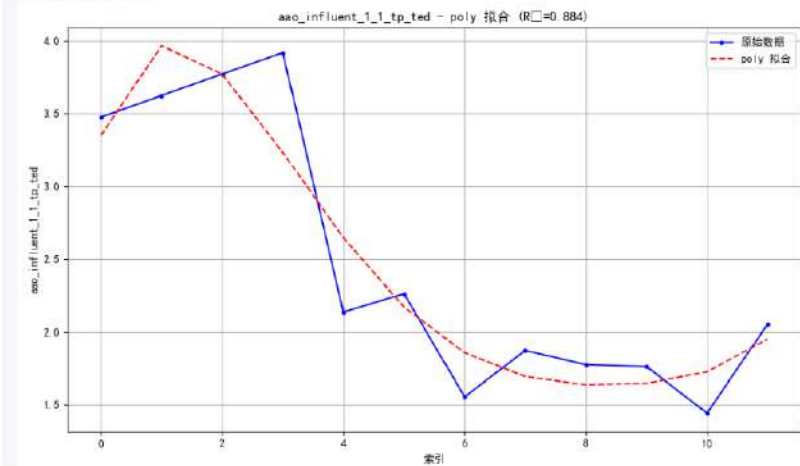
- **Max:** 3.919 mg/L
- **Min:** 1.443 mg/L
- **Mean:** 2.471 mg/L
- **Variance:** 0.882
- **Standard Deviation:** 0.939
- **Skewness:** 0.646 (slightly right-skewed)
- **Kurtosis:** -1.493 (platykurtic distribution)
- **Best Fit Model:** Polynomial (5th degree)
- **R<sup>2</sup>:** 0.884 (highly predictive)

### Free Ammonia Nitrogen (TKN)

- **Max:** 0.735 mg/L
- **Min:** 0.379 mg/L
- **Mean:** 0.612 mg/L
- **Variance:** 0.008
- **Standard Deviation:** 0.091
- **Skewness:** -1.342 (left-skewed)
- **Kurtosis:** 3.599 (leptokurtic distribution)
- **Best Fit Model:** Polynomial (5th degree)
- **R<sup>2</sup>:** 0.650 (moderately predictive)

## 2. Trend Plots

- **Total Phosphorus (TP):**



# Digital assistant — O&M support

## Problem:

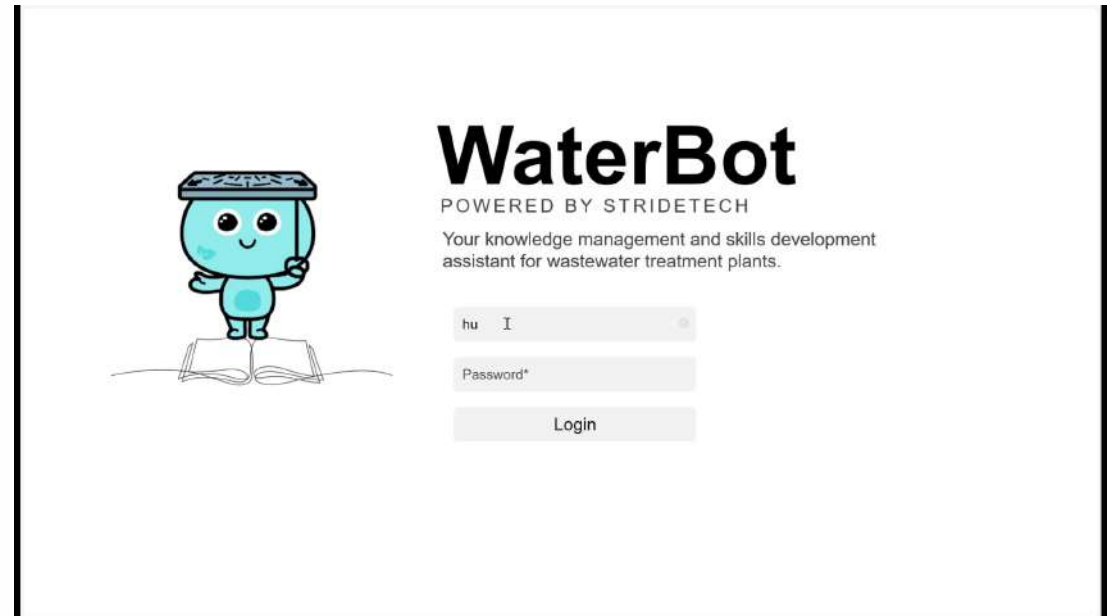
- Difficulty in effective identification and utilization of O&M manuals for O&M support

## Solution:

- DVA empowered by agentic RAG technique

## Benefits:

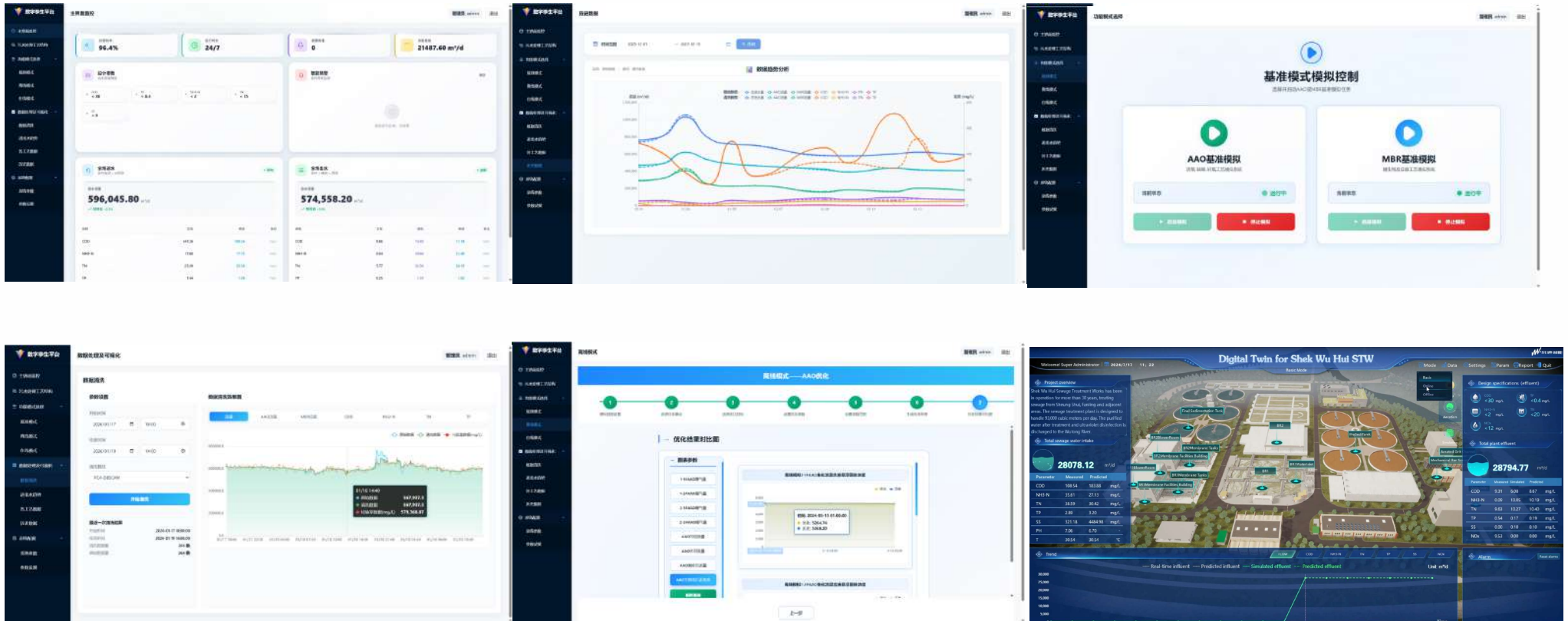
- All O&M documents can be integrated into the DVA digitally
- Instance access and search through all O&M documents for relevant solutions to O&M problems
- Great personnel training platform



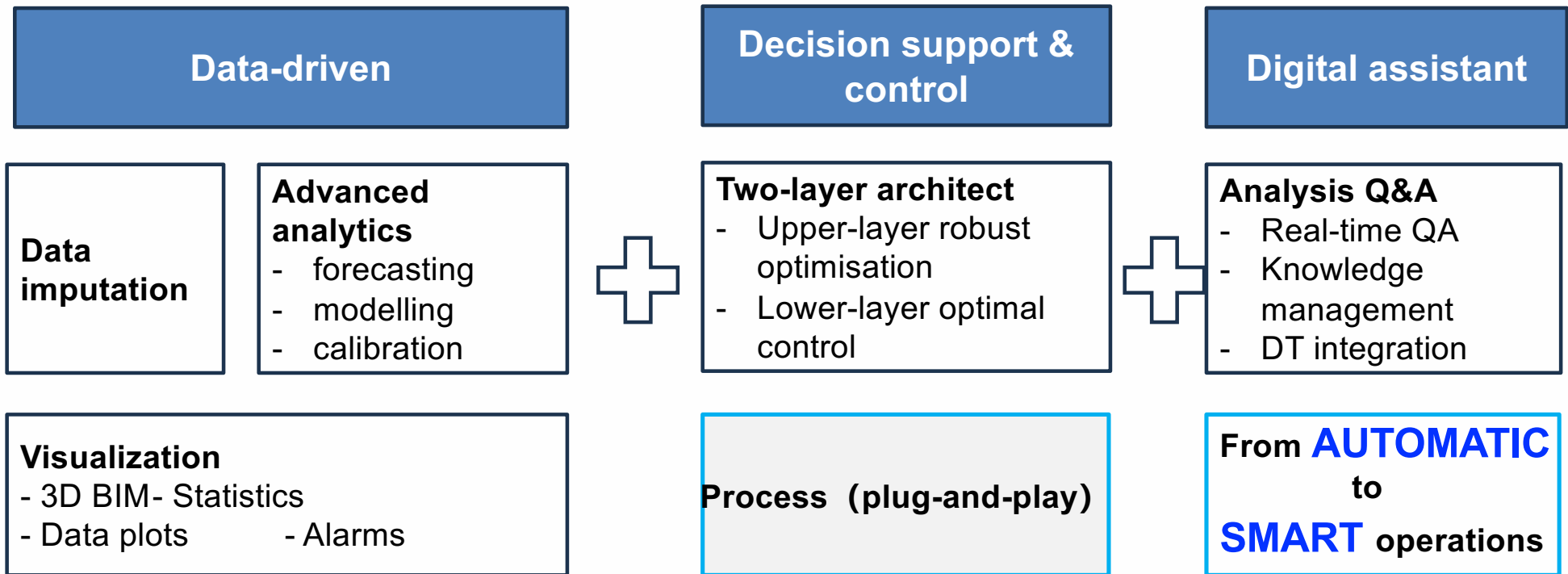
- Autonomous agents for smart evidence gathering and recommendation.
- Predefined workflows for stability and reliability.

**Better knowledge management, personnel training, and real-time support**

# Customizable UI: 2D or 3D



# Tech summary

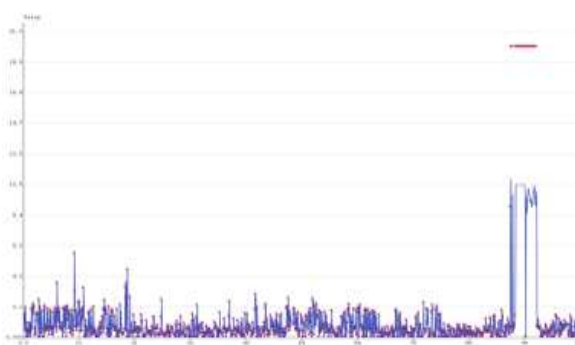


**digital + intelligent platform**

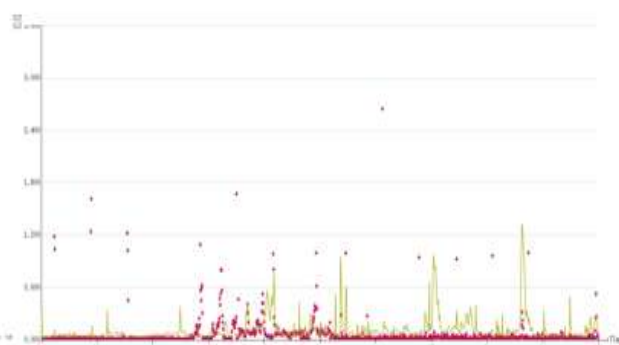
Empowering data-driven operation optimization and intelligent control of industrial processes and smart city applications

# Wuhan case study

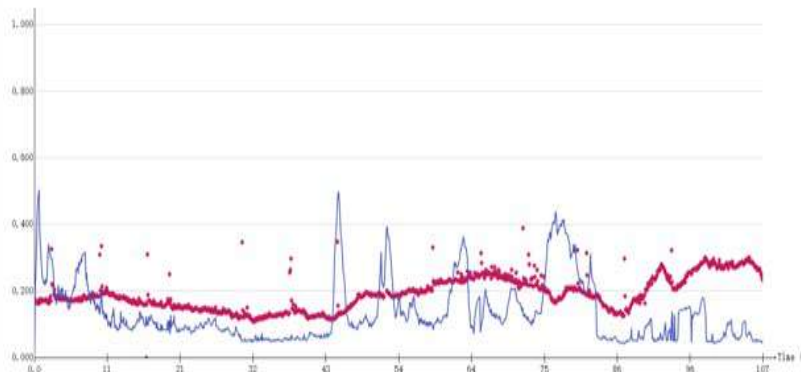
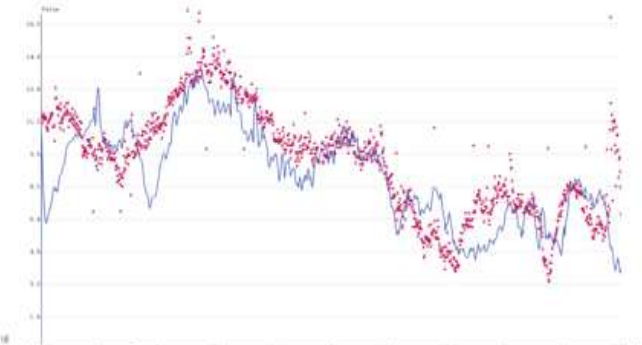
Modeling results using data from January 1 2022 to April 17 2023



Effluent SS, Mean Absolute Error: 0.01



Effluent NH4-N, Mean Absolute Error : 0.05



Effluent TP, Mean Absolute Error : 0.06



MLSS, Mean Absolute Percentage Error: 3.3%

# Wuhan case study

## The plant:

A WWTP in Wuhan  
1#AAO (200k m<sup>3</sup>/day)

## Requirements:

Effluent NH<sub>4</sub><1.5mg/L

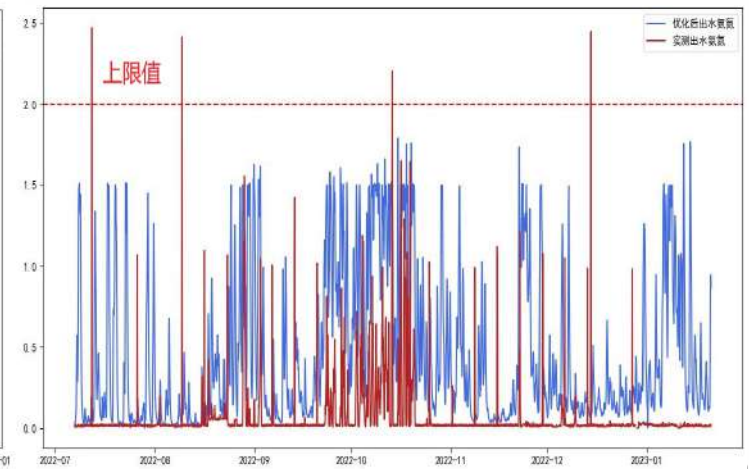
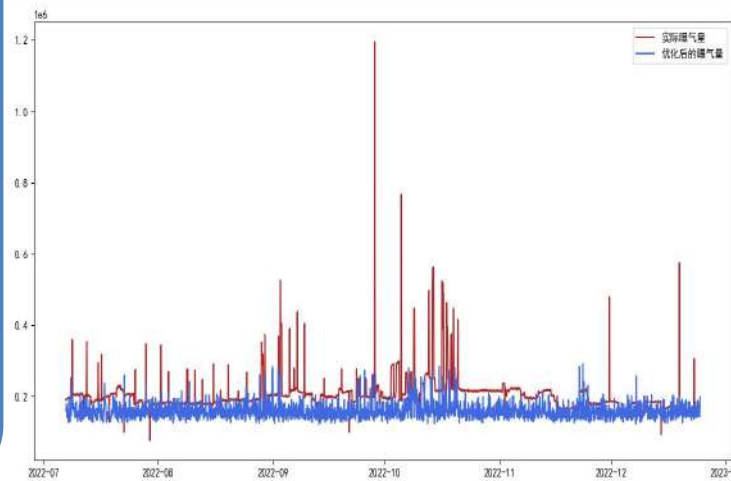
## Decision variable:

Aeration rates

## Duration:

2022.07.07-2023.01.20 (A  
total of 198 days)

	Aeration (m <sup>3</sup> /d)	Savings (%)
Plant actual	481334201.654	-
Optimised	377047473.249	<b>21.67</b>

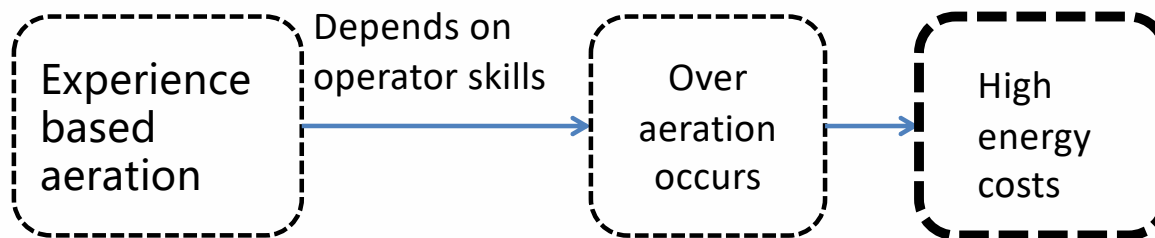




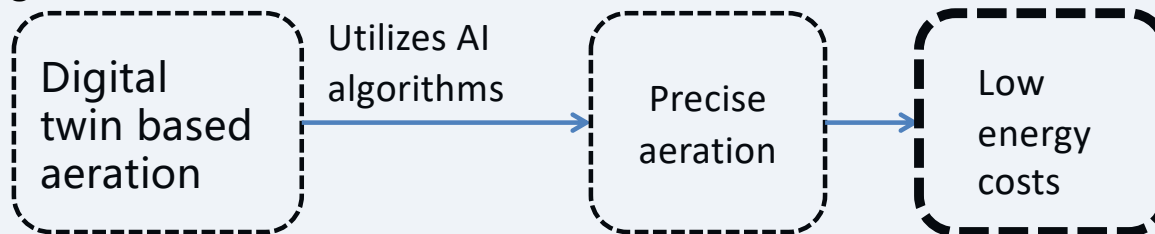
# Hong Kong case study

- Operation optimisation

## Aeration Optimization



### Digital twin



- Decision support
- Supervised operation
- ~10% energy savings

# Digital Twin for Shek Wu Hui STW

**User login**

User  
admin123

Password  
.....

Login

Quit

# Digital Twin for Shek Wu Hui STW



Welcome! Super Administrator | 2024/7/17 11: 22

Basic Mode

Mode | Data | Settings | Param | Report | Quit

## Project overview

Shek Wu Hui Sewage Treatment Works has been in operation for more than 30 years, treating sewage from Sheung Shui, Fanling and adjacent areas. The sewage treatment plant is designed to handle 93,000 cubic meters per day. The purified water after treatment and ultraviolet disinfection is discharged to the Wutong River.

## Total sewage water intake

**28078.12** m<sup>3</sup>/d

Parameter	Measured	Predicted	
COD	108.54	183.88	mg/L
NH3-N	35.61	27.13	mg/L
TN	38.59	30.42	mg/L
TP	2.89	3.20	mg/L
SS	321.18	4484.98	mg/L
PH	7.06	6.70	-
T	30.54	30.54	°C



- Basic
- Online
- Offline

## Design specifications (effluent)

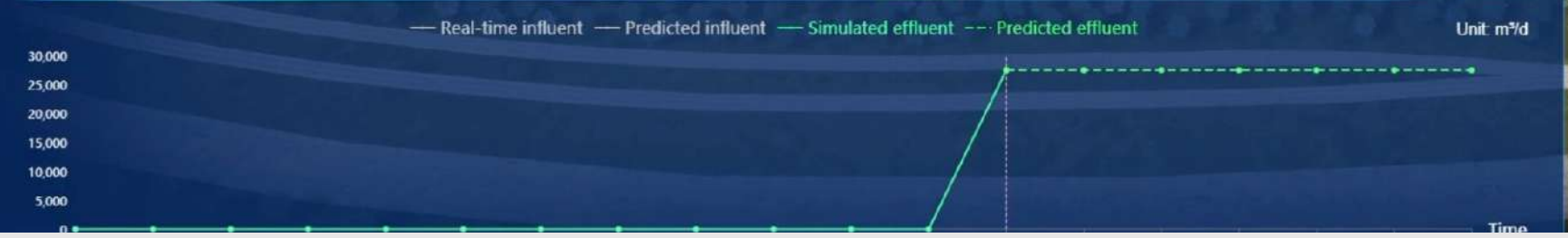
COD	<30 mg/L	TP	<0.4 mg/L
NH3-N	<2 mg/L	TN	<20 mg/L
NOx	<12 mg/L		

## Total plant effluent

**28794.77** m<sup>3</sup>/d

Parameter	Measured	Simulated	Predicted	
COD	9.31	6.08	8.67	mg/L
NH3-N	0.09	10.05	10.19	mg/L
TN	9.63	10.27	10.40	mg/L
TP	0.54	0.17	0.19	mg/L
SS	0.00	0.10	0.10	mg/L
NOx	9.53	0.00	0.00	mg/L

## Trend



## Alarm

Reset alarms

# Digital Twin for Shek Wu Hui STW

Welcome! Super Administrator | 2024/7/17 11: 22

Online Mode

Mode Data Settings Param Report Quit

## Project overview

Shek Wu Hui Sewage Treatment Works has been in operation for more than 30 years, treating sewage from Sheung Shui, Fanling and other areas. The sewage treatment plant is designed to handle 93,000 cubic meters per day. The treated water after treatment and ultraviolet disinfection is discharged to the Wutong River.

## Total sewage water intake



28078.12

Parameter	Measured	Predicted
COD	108.54	183.88
NH3-N	35.61	27.13
TN	38.59	30.42
TP	2.89	3.20
SS	321.18	4484.98
PH	7.06	6.70
T	30.54	30.54

## Trend

30,000  
25,000  
20,000  
15,000  
10,000  
5,000  
0

## Online Simulation



Aeration parameters (Aeration control)

Dissolved oxygen parameters (DO control)

Use raw data

Aerobic Zone 1:   
Aerobic Zone 2:   
Aerobic Zone 3:   
Aerobic Zone 4:

mg/L  
mg/L  
mg/L  
mg/L

Simulation in progress

Dosage

PAC dosage:  m<sup>3</sup>/h

RAS (return activated sludge) flow rate - Q:  L/s

MLR (mixed liquid return) flow rate - Q:  L/s

SAS (surplus activated sludge) flow rate - Q:

SAS (surplus activated sludge) flow rate - Q:  L/s

Start Simulation

## Design specifications (effluent)

DO	>30 mg/L	TP	<0.4 mg/L
NH3-N	<2 mg/L	TN	<20 mg/L
DO	>12 mg/L		

## Plant effluent

28794.77 m<sup>3</sup>/d

Measured	Simulated	Predicted	
9.31	6.08	8.67	mg/L
0.09	10.05	10.19	mg/L
9.63	10.27	10.40	mg/L
0.54	0.17	0.19	mg/L
0.00	0.10	0.10	mg/L
9.53	0.00	0.00	mg/L

Reset alarms

1:31.59

Time

# Digital Twin for Shek Wu Hui STW

Welcome! Super Administrator | 2024/7/17 11:22

Online Mode

Mode Data Settings Param Report Quit

## Project overview

Shek Wu Hui Sewage Treatment Works has been in operation for more than 30 years, treating sewage from Sheung Shui, Fanling and other areas. The sewage treatment plant is designed to handle 93,000 cubic meters per day. The water after treatment and ultraviolet disinfection is discharged to the Wutong River.

## Total sewage water intake



28078.12

Parameter	Measured	Predicted
COD	108.54	183.88
NH3-N	35.61	27.13
TN	38.59	30.42
TP	2.89	3.20
SS	321.18	4484.98
PH	7.06	6.70
T	30.54	30.54

## Trend

30,000  
25,000  
20,000  
15,000  
10,000  
5,000  
0

## Online Optimization



Select the decision variable to be optimized

Aeration rate  OR DO setpoint

SAS (surplus activated sludge) flow rate - Q  OR MLSS

MLR (mixed liquid return) flow rate - Q

RAS (return activated sludge) flow rate - Q

PAC (Polymeric Aluminum Chloride) dosage

Swing Zone Aeration/Anoxic State

Previous step

Stop online optimization

1:08.45

## Design specifications (effluent)

DO	30 mg/L	TP	<0.4 mg/L
NH3-N	2 mg/L	TN	<20 mg/L
Ox	12 mg/L		

## Plant effluent

28794.77 m<sup>3</sup>/d

Measured	Simulated	Predicted	
9.31	6.08	8.67	mg/L
0.09	10.03	10.17	mg/L
9.63	10.25	10.38	mg/L
0.54	0.17	0.19	mg/L
0.00	0.10	0.10	mg/L
9.53	0.00	0.00	mg/L

Reset alarms

# Digital Twin for Shek Wu Hui STW

Welcome! Super Administrator | 2024/7/17 11:22

Offline Mode

Mode Data Settings Param Report Quit

## Project overview

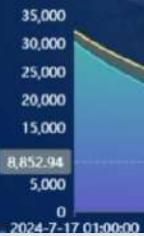
Shek Wu Hui Sewage Treatment Works has been in operation for more than 30 years, treating sewage from Sheung Shui, Fanling and other areas. The sewage treatment plant is designed to handle 93,000 cubic meters per day. The water after treatment and ultraviolet disinfection is discharged to the Wutong River.

## Total sewage water intake



Parameter	Measured
COD	183.76
NH3-N	33.17
TN	32.68
TP	2.67
SS	48.83
PH	6.62
T	30.89

## Trend

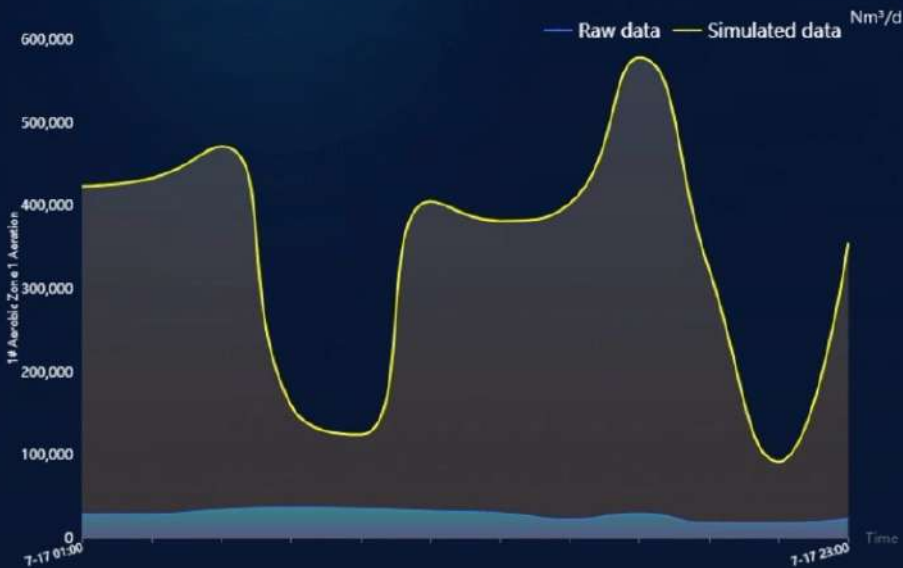


## Offline Optimization

- Optimization
- 1. Selection data
- 2. Selection algorithm
- 3. Set objectives
- 4. Select variable
- 5. Optimization results

### Variable

- Aeration
- DO
- Reflux
- PAC
- Sludge



时间: 2024-07-17 13:00  
● Historical influent: 28600  
● Historical effluent: 28412.4

## Effluent specifications (effluent)

DO	30 mg/L	TP	<0.4 mg/L
NH3-N	2 mg/L	TN	<20 mg/L
DOx	12 mg/L		

## Plant effluent

29948.02 m<sup>3</sup>/d

Measured	Simulated	Unit
104.31	12.29	mg/L
0.06	5.36	mg/L
3.80	5.85	mg/L
0.69	1.95	mg/L
0.00	0.08	mg/L
3.75	0.00	mg/L

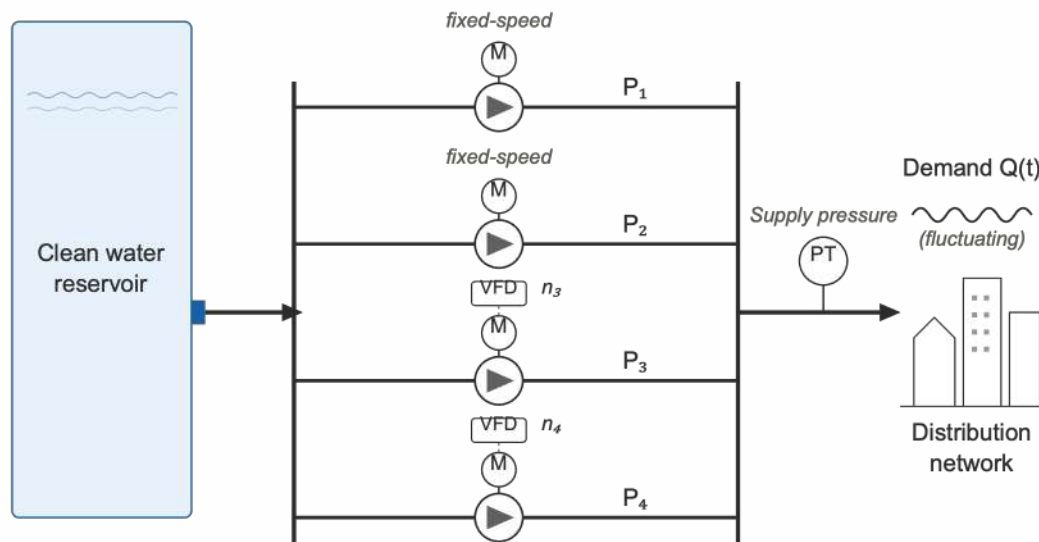
TP SS NOx

Unit: m<sup>3</sup>/d

Time 2024-7-17 23:00:00

# Water treatment plant optimisation

## Multi-pump operation scheduling



Objective: choose pump speeds  $n_1, n_2, n_3, n_4$  to supply the fluctuating demand  $Q(t)$  while holding the supply pressure at the setpoint  $P_{set}$

### System

- Mixed fixed-speed and VFD pumps
- Combined discharge supplies the distribution network

### Control objective

- Match the fluctuating customer demand
- Hold supply pressure at the setpoint
- Minimise total energy consumption

### Decision variables

- On/off status of fixed-speed pumps
- Continuous speeds of VFD pumps

### Constraints

- Flow balance
- Pump characteristics

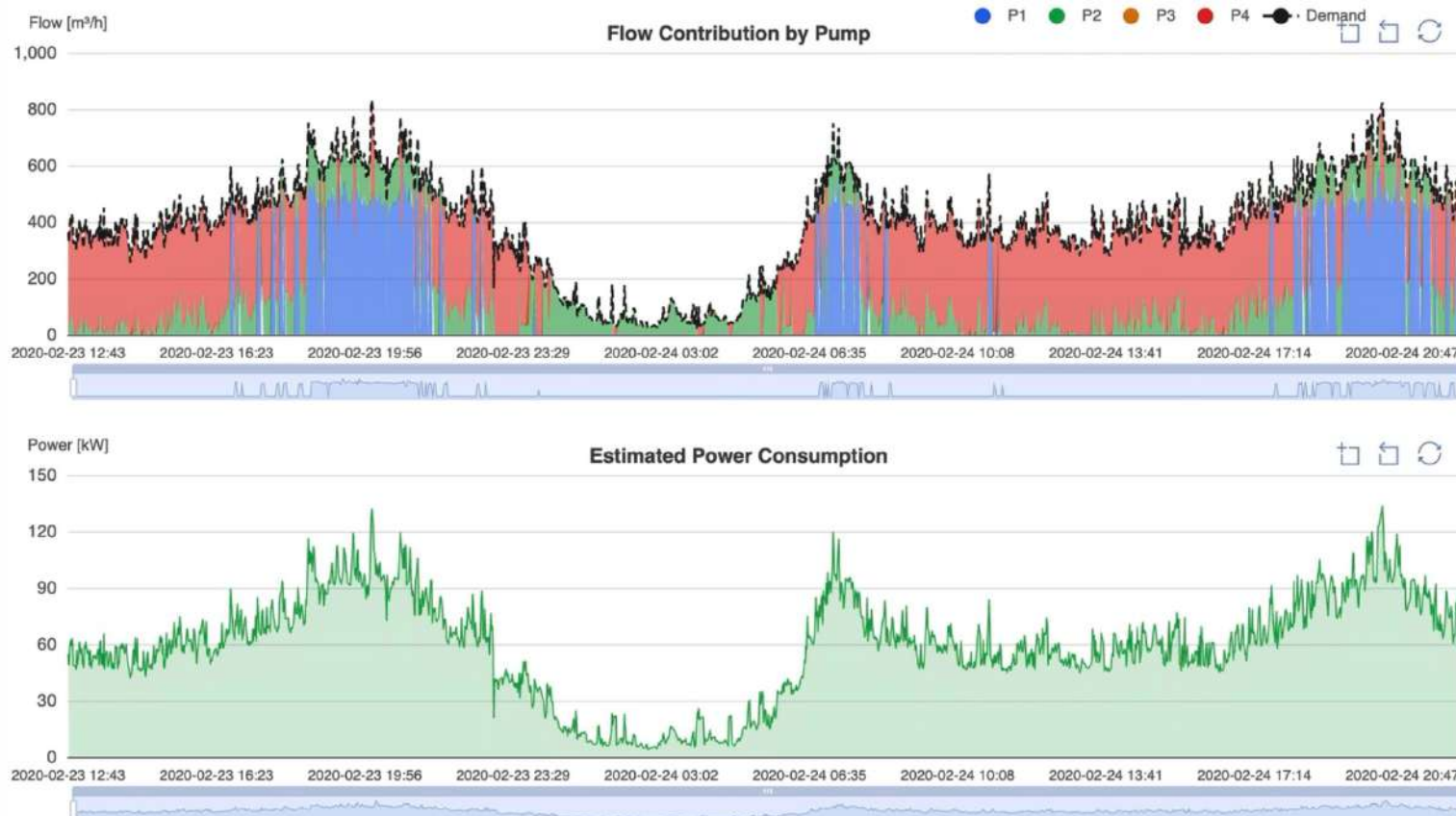
### Why it's hard

- Mixed-integer nonlinear problem
- Multiple local optima
- Continuous demand fluctuation

# Water treatment plant optimisation

The screenshot displays a web application interface for 'Pump Station Optimization'. The interface is divided into a left sidebar and a main content area. The sidebar, titled 'OPERATIONS', contains a list of menu items: '1 Upload', '2 Screen', '3 Fit Curves', '4 Schedule', '5 Real-Time', and 'Settings'. The 'Upload' item is currently selected. The main content area is titled 'Upload SCADA Data' and features a large dashed-line box. Inside this box, there is a plus sign (+) and the text 'Click to upload CSV' followed by 'Station\_data.csv from SCADA export'.

# Water treatment plant optimisation



- More balanced utilization of pumps
- 8-10% energy savings
- Autonomous operation
- Allows manual take-over when necessary

# Thank you !

## Acknowledgement

