Supporting Information

Energy-Saving Scheduling for Multiple Water Intake Pumping Stations in Water Treatment Plants Based on Personalized Federated Deep Reinforcement Learning

Dongsheng Wang[a], Ao Li[a], Yicong Yuan[a], Tingjun Zhang[a], Liang Yu[a*], Chaoqun Tan[b*]

a College of Automation & College of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing 210023, China

b School of Civil Engineering, Southeast University, Nanjing 210096, China

Corresponding author contact details: Chaoqun Tan, Liang Yu
Email: tancq@seu.edu.cn, Liang.yu@njupt.edu.cn

This supporting information contains:
11 pages
4 Tables
2 Figures
12 Equations
3 Algorithms

List

Table S1 List of Abbreviations in Alphabetical Order

Description	Abbreviation
Average energy consumption.	AEC
Average reservoir level violation.	ALV
Average pressure violation.	APV
Average pressure variation violation.	APVV
Deep reinforcement learning.	DRL
Federated Learning.	FL
Long short-term memory.	LSTM
Personalized Federated Learning.	PFL
Personalized Federated Learning based LSTM.	PFL-LSTM
Personalized Federated Learning based MAAC.	PFL-MAAC
Personalized Federated Learning based Multi-Agent Attention Deep Reinforcement Learning.	PFL-MAADRL
Reinforcement learning.	RL
Multi-actor-attention-critic.	MAAC
Multi-agent Attention DRL.	MAADRL
Multi Water-intake pumping station.	MWIPS
Water Distribution Network.	WDN
Water Distribution Networks.	WDNs
Water-intake pumping station.	WIPS
Water-intake pumping stations.	WIPSs

Table S2 List of Parameter Description

Description	Parameter
Time slot index.	t
Time slot length (h).	τ
Water-intake pumping station index.	i
Total number of time slots per episode.	T
Total number of Water-intake pumping stations.	N
The upper limit of the reservoir level of WIPS i..	l_i^{max}
The lower limit of the reservoir level of WIPS i.	l_i^{min}
The upper limit of the water intake of WIPS i.	Q_i^{max}
The lower limit of the water intake of WIPS i.	Q_i^{min}
The limited value of main pipe pressure changes of WIPSs.	p_v^{max}
The upper limit of the main pipe pressure of WIPS i.	p_i^{max}
The lower limit of the main pipe pressure of WIPS i.	p_i^{min}
The WIPS energy consumption at slot t of WIPS i.	$\boldsymbol{\Phi_{i,t}}$
The weights of actor network.	θ
The weights of critic network.	ψ

Table S3 The main experimental parameters

Parameter	Value	Parameter	Value
p_v^{max}	0.10MPa	Y	50000
p_1^{max}	0.36MPa	K_{test}	1440h
p_1^{min}	0.13MPa	τ	1h
p_2^{max}	0.31MPa	T_{update}	1h
p_2^{min}	0.18MPa	Q_i^{min}	$5000\text{m}^3/\text{h}$
p_3^{max}	0.38MPa	b	32
p_3^{min}	0.08MPa	N_a	128
l_1^{min}	0.9m	N_c	128
l_1^{max}	2.9m	$buffer_{size}$	19200
l_2^{min}	1m	B_{size}	120
l_2^{max}	5m	φ	0.1
l_3^{min}	1m	ξ	0.001
α_a	0.001	l_3^{max}	4.2m
α_c	0.001	Q_i^{max}	$13000\text{m}^3/\text{h}$

Table S4 RB Schemes

Schemes	**Rule-1**	**Rule-2**
ϵ_i/m^3	200	400

Fig. S1 Performance details among all schemes for WIPS 3.

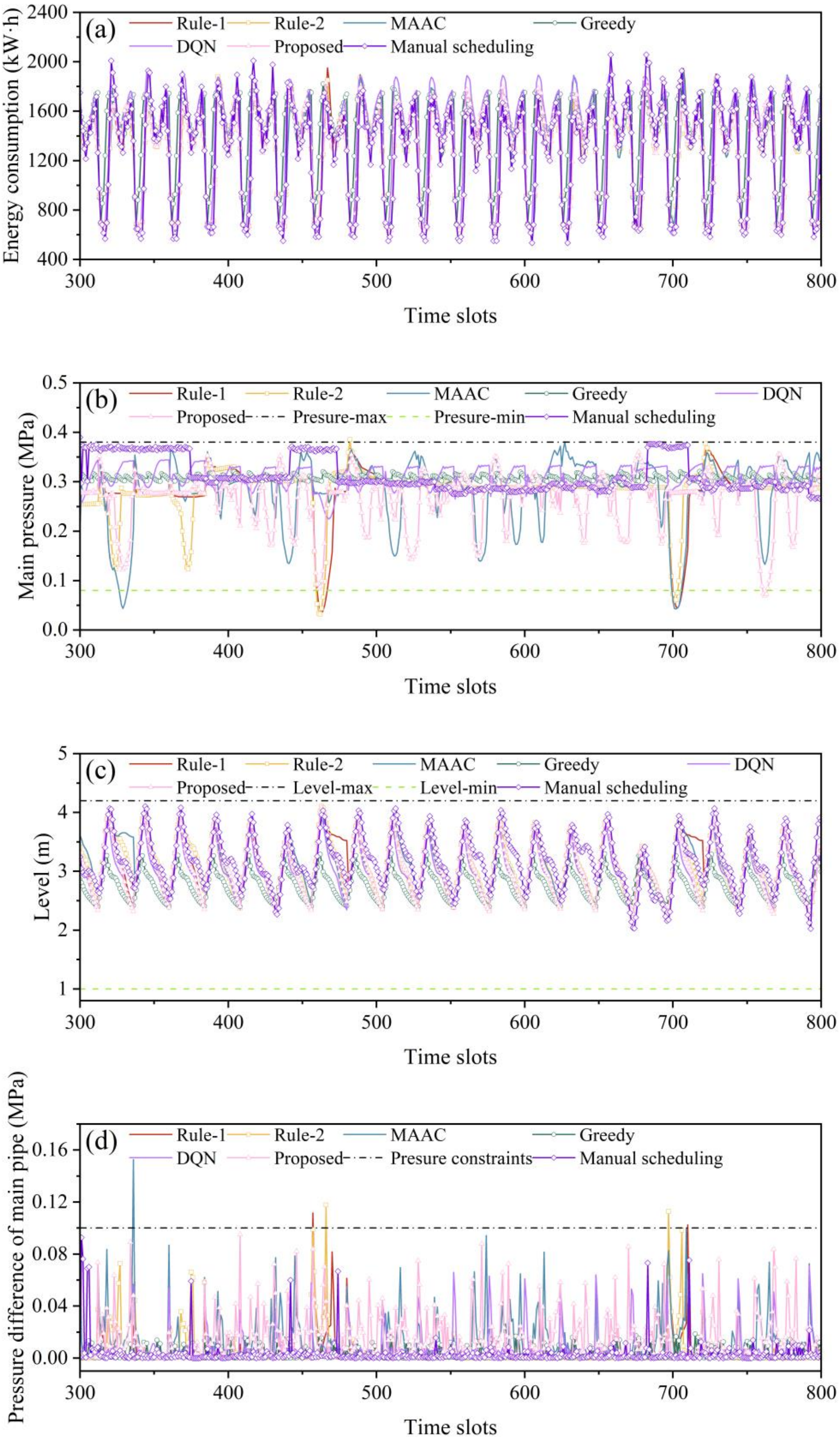
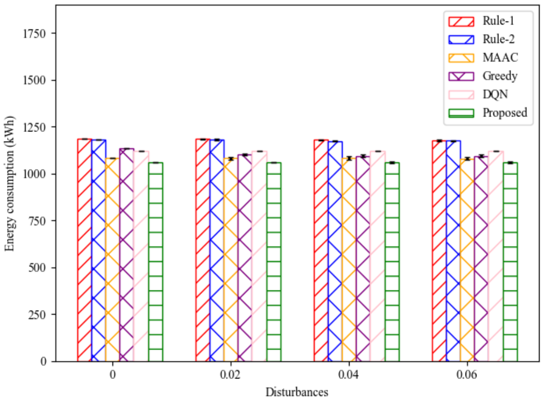
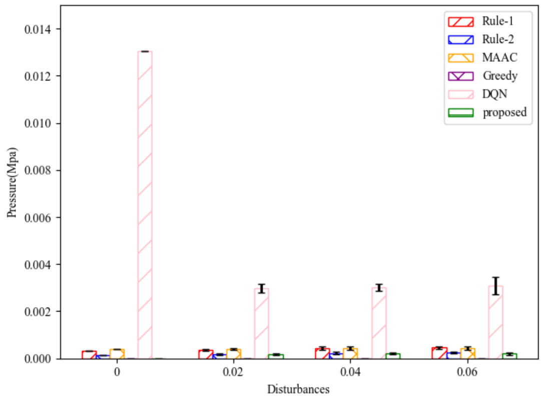


Fig. S1 Performance details of among all schemes for WIPS 3.

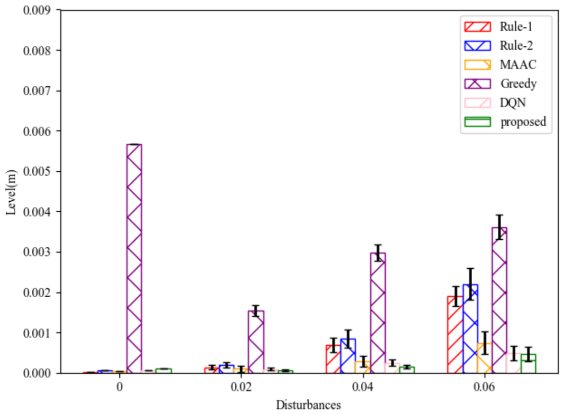
Fig. S2 Robustness performance of the proposed algorithm for MWIPS.



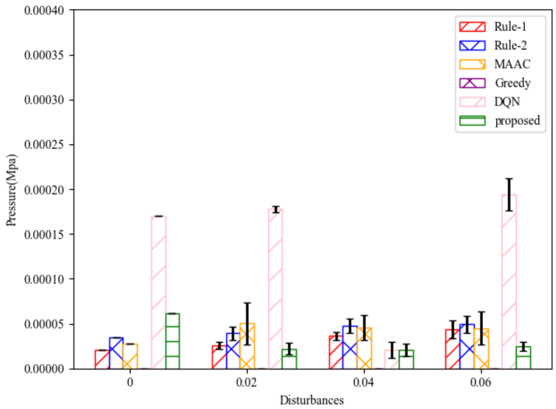
(a) AEC



(b)APV



(c)ALV



(d)APVV

Fig. S2 Robustness performance of the proposed algorithm for MWIPS.

Equation. S1

$$r_{1,i,t}(s_{t-1}, a_{t-1}) = \Phi_{i,t} \qquad (S1)$$

Equation. S2

$$r_{2,i,t}(s_t) = \left\{\left[l_{i,t} - l_i^{max}\right]^+ + \left[l_i^{min} - l_{i,t}\right]^+\right\} \qquad (S2)$$

Equation. S3

$$r_{3,i,t}(s_t) = \left\{\left[p_{i,t} - p_i^{max}\right]^+ + \left[p_i^{min} - p_{i,t}\right]^+\right\} \qquad (S3)$$

Equation. S4

$$r_{4,i,t}(s_{t-1}, a_{t-1}) = \left[\left|p_{i,t} - p_{i,t-1}\right| - p_v^{max}\right]^+ \qquad (S4)$$

Equation. S5

$$R_{i,t} = -\left(r_{1,i,t}(s_{t-1}, a_{t-1}) + \alpha_{i,1} r_{2,i,t}(s_t) + \alpha_{i,2} r_{3,i,t}(s_t) + \alpha_{i,3} r_{4,i,t}(s_{t-1}, a_{t-1})\right) \qquad (S5)$$

Equation. S6

$$\mathcal{L}_Q(\psi) = \sum\nolimits_{i=1}^{N} E_{(o,a,\tilde{o},r)\sim D}\left[\left(Q_i^{\psi}(o,a) - y_i\right)^2\right] \qquad (S6)$$

Equation. S7

$$y_i = r_i(o,a) + \gamma E_{\tilde{a}\epsilon\pi_{\bar{\theta}}}\left[-\varphi \log \pi_{\bar{\theta}_i}(\tilde{a}_i|\tilde{o}_i) + Q_i^{\bar{\psi}}(\tilde{o},\tilde{a})\right] \qquad (S7)$$

Equation. S8

$$\nabla_{\theta_i} J(\theta) = E_{o\sim D, a\sim\pi}\left[\nabla_{\theta_i} \log\left(\pi_{\theta_i}(a_i|o_i)\right) \rho_i\,(o_i, a_i)\right] \qquad (S8)$$

Equation. S9

$$\rho_i(o_i, a_i) = -\varphi \log\left(\pi_{\theta_i}(a_i|o_i)\right) + Q_i^{\psi}(o,a) - b\left(o, a_{\backslash i}\right) \qquad (S9)$$

Equation. S10

$$AEC = \frac{1}{K_{test}} \sum_{t=1}^{K_{test}} \Phi_{i,t} \tag{S10}$$

Equation. S11

$$APV = \frac{1}{K_{test}} \sum_{t=1}^{K_{test}} \left\{ \left[p_{i,t} - p_i^{max} \right]^+ + \left[p_i^{min} - p_{i,t} \right]^+ \right\} \tag{S11}$$

Equation. S12

$$APVV = \frac{1}{K_{test}} \sum_{t=1}^{K_{test}} \left[\left| p_{i,t} - p_{i,t-1} \right| - p_v^{max} \right]^+ \tag{S12}$$

Equation. S13

$$ALV = \frac{1}{K_{test}} \sum_{t=1}^{K_{test}} \left\{ \left[l_{i,t} - l_i^{max} \right]^+ + \left[l_i^{min} - l_{i,t} \right]^+ \right\} \tag{S13}$$

Algorithm. S1 PFL Algorithm

Algorithm. S1 PFL Algorithm

Input: N clients, each holds a set of private training dataset
Output: Personalized models and global model
Initialize the global model W_{global}
for each round t= $1,2,\cdots,\mathrm{F}$ **do**
    Initialize the personalized model parameters W_{local_i}
    for each client i in parallel **do**
        $\left(W_{local_i}, W_{personal_i}\right) \longleftarrow \boldsymbol{ClientUpdate}(i, W_{global})$
    end for
    $W_{global} \longleftarrow \boldsymbol{Aggregate}(N, \{W_{local_i} | i \in N\})$
end for
$\boldsymbol{ClientUpdate}(i, W_{global})$
$W_{local} \longleftarrow Copy(W_{global})$
for local iteration $i = 1,2,\cdots P$ **do**
    for data batch $b \in$ client dataset$_i$ **do**
        $W_{local_i} \longleftarrow W_{local_i} - \alpha.\nabla loss(W_{local_i}, b)$
    end for
end for
$W_{personal_i} \longleftarrow \boldsymbol{Personalize}(W_{local_i},$ client dataset$_i)$
return $W_{personal_i}$
$\boldsymbol{Aggregate}(N, W_{local_i})$
$W_{global\_new} \longleftarrow 0$
for client $i \in N$ **do**
    $W_{global\_new} = \frac{1}{N}\sum_{i \in N} W_{local_i}$
end for
return W_{global_new}

Algorithm. S2 Training Process of MAADRL-based Energy Scheduling Algorithm for MWIPS

Algorithm. S2 Training Process of MAADRL-based Energy Scheduling Algorithm for MWIPS

Input: The mainline pressure, reservoir water level and supplied water demand, personal model $W_{personal_i}$

Output: The weights of actor network and critic network (i.e., θ and ψ)

Initialize MWIPS environment with N agents

Initialize the capacity of experience replay buffer D

Initialize the weights of the actor network and the critic network are denoted by: π_i^{θ} and Q_i^{ψ}

Initialize the weights of the target network i.e., $\pi_i^{\bar{\theta}}$ and $Q_i^{\bar{\psi}}$

Initialize the weights of the personal model $W_{personal_i}$

for episode = $1,2,\cdots Y$ **do**

  Reset environment and receive initial observation $o_{i,1}$ for all agents

  for $t = 1,2,\cdots 24$ **do**

    Select actions $a_{i,t} \sim \pi_i^{\theta}(\cdot\,|o_{i,t})$ for each agent i and execute them

    Each agent i gets new state $o_{i,t+1}$ and reward $r_{i,t+1}$ from the MWIPS scheduling simulation environment

    Store transitions $(o_t, a_t, o_{t+1}, r_{t+1})$ in D

    if $K_{memory} \geq B_{size}$ and mod (t, T_{update}) **do**

      Sample B_{size} transitions $(o, a, \tilde{o}, r)$ from D

      Calculate $Q_i^{\psi}(o_i^c, a_i^c)$ for all agents, o_i^c and a_i^c denote the cth o_i and a_i of the B_{size} transition, $1 \leq \text{c} \leq B_{size}$

      Calculate $\tilde{a}_i^c \sim \pi_i^{\bar{\theta}}(\tilde{o}_i^c)$ and $Q_i^{\bar{\psi}}(\tilde{o}_i^l, \tilde{a}_i^l)$ for observation and action tuple c and agent i

      Minimize the joint regressive loss function (S6) to update critic network

      Calculate $a_i^c \sim \pi_i^{\bar{\theta}}(o_i^c)$ and $Q_i^{\psi}(o_i^l, a_i^l)$ for observation and action tuple c and agent i

      Update policies using (S8)

      Update the weight parameters of target networks:

      $\bar{\psi} \leftarrow \psi\xi + (1-\xi)\bar{\psi}, \bar{\theta} \leftarrow \theta\xi + (1-\xi)\bar{\theta}$

    end if

  end for

end for

Algorithm. S3 Testing Process of MAADRL-based Energy Scheduling Algorithm for MWIPS

Algorithm. S3 Testing Process of MAADRL-based Energy Scheduling Algorithm for MWIPS

1: **Input:** Load the weight parameters of the actor network, i.e., θ

2: **Output:** Action $a_{i,t}$ for each agent at the time slot t

3: Initialize environment for each agent, i.e., WIPS agent

4: Initialize each agent local observation $o_{i,1}$

5: **for** $t = 1,2,\cdots 24$ **do**

6: All agents select the action $a_{i,t}(1 \le i \le N)$: $a_{i,t} \sim \pi_\theta(\cdot\,|o_{i,t})$

7: Each agent i takes action $a_{i,t}$ under the environment, which affects the operation of the WIPSs

8: Each agent i obtains new observation $o_{i,t+1}$

24: **end for**