Data flow in living contexts: a systemic analysis for an unsustainable model

Demand Response Scheduling Considering Pollutant Diffusion Uncertainty of Industrial Customers

Yingjun Wu, *Member, IEEE*, Zhiwei Lin, *Student Member, IEEE*, Yijun Xu, *Senior Member, IEEE*, Gianfranco Chicco, *Fellow, IEEE*, Tao Huang, *Senior Member, IEEE*, Junjie Shao, *Student Member, IEEE*, Zhaorui Chen, *Student Member, IEEE*

***Abstract*—The demand response scheduling scheme requires the consideration of both the industrial customers' economic cost and the environmental influences from pollutants. However, the diffusion process of the latter, although of paramount importance, is typically ignored in the existing literature. To address this issue, we propose a demand response scheduling scheme that not only precisely simulates the diffusion process through a spatio-temporal diffusion model, but incorporates the uncertainty into the diffusion trajectories via a Markov decision process. This enables the schedule-maker optimally select the industrial customers to participate in the demand response with a minimum cost while reducing the environmental influences of the pollutants simultaneously. Using it, a deep reinforcement learning approach is further advocated in the optimization procedure to improve the scalability of the proposed method. Simulation results on the modified IEEE-118 test system reveal the validity of the proposed method.**

***Index Terms*—demand response, pollutant diffusion, uncertainty, Markov decision process, deep reinforcement learning**

NOMENCLATURE

Indexes and sets

t	Index for time, $t \in \mathbf{T}$
k	Index for node connected to DN, $k \in \mathbf{B}_{DN}$
l	Index for production process, $l \in \mathbf{L}$
g	Index for pollutant emitted by IC, $g \in \mathbf{G}$
m	Index for fossil fuel, $m \in \mathbf{M}$
w	Index for atmospheric factor, $w \in \mathbf{W}$
$\mathbf{B}_{IC}$	Set of nodes connected to ICs
$\mathbf{S}$	Set of Markov state

Superscript and subscript

$\bullet_{dr}$	Corresponding parameter $\bullet$ for making DR scheduling
$\bullet_{ser}$	Corresponding parameter $\bullet$ for purchasing DR services from ICs
$\bullet_{com}$	Corresponding parameter $\bullet$ for compensating IC to participate in DR
$\bullet_{emi}$	Corresponding parameter $\bullet$ of environmental influences of pollutants from ICs
$\bullet_{in}, \bullet_{out}$	The input/output of corresponding parameter $\bullet$

Variables

ΔP	The responsive volume provided by ICs, $\Delta P \in \Delta\mathbf{P}$
q	The emission of pollutants from ICs, $q \in \mathbf{q}$
$e_{ene}, e_{fuel}, e_{oth}$	The consumption of energy consumption except electricity/fossil energy/other raw materials, $e_{ene} \in \mathbf{E}_{ene}, e_{fuel} \in \mathbf{E}_{fuel}, e_{oth} \in \mathbf{E}_{oth}$

Vectors

$\boldsymbol{r}_j^{t'}$	Position coordinate in the state $S_j^{t'}$, $\boldsymbol{r}_j^{t'} = \left(x_j^{t'}, y_j^{t'}, z_j^{t'}\right)$
$\boldsymbol{r}_0$	Position coordinate in the state S^0, $\boldsymbol{r}_0 = (x_0, y_0, z_0)$
$\boldsymbol{v}_j^{t'}$	Wind vector in the state $S_j^{t'}$, $\boldsymbol{v}_j^{t'} = \left(v\left(x_j^{t'}\right), v\left(y_j^{t'}\right), v\left(z_j^{t'}\right)\right)$

Scalars

C	The cost of schedule-maker
B_{rel}	The saved cost of power reserve capacity decline
ϑ	The unit cost
$\mathbb{R}(\bullet)$	The reserve capacity cost function
α, β	The output share of the production process flow/raw materials, $\alpha \in \boldsymbol{\alpha}, \beta \in \boldsymbol{\beta}$
c	The coefficient for describing the degree of pollutant emission, $c \in \mathbf{c}$
P	The electricity consumption
E_{dif}	The index for evaluating environmental influence of pollutants from ICs
D	The smoke cluster emitted from ICs

This work was supported by the Fundamental Research Funds for the Central Universities (B230201048), the National Natural Science Foundation of China under Grant 51877181 and the University Grants for Basic Research (56_RBA22HUT01). *(Corresponding author: Yijun Xu)*

Y. Wu, Z. Lin, J. Shao, and Z Chen are with College of Energy and Electrical Engineering, Hohai University, Nanjing, China (e-mail: yingjunwu@hotmail.com, lzw_222300@163.com, shaojunjie2021@163.com, and lhsxczr@163.com).

Y. Xu, is with the Electrical Engineering Department, Southeast University, Nanjing, Jiangsu 210096, China, (e-mail: yijunxu@seu.edu.cn).

G. Chicco is with the Department of Energy, Politecnico di Torino, Italy (e-mail: Gianfranco.chicco@polito.it).

T. Huang is with the Department of Energy, Politecnico di Torino, Italy and Xihua University, China (e-mail: tao.huang@polito.it).

Q	The pollutant concentration
G	The diffusion coefficient
l_j	The Euclidean distance between $\boldsymbol{r}_j^{t'}$
U	The voltage of the node
F_P, F_Q, F_S	The active/reactive/apparent power flow
r, x	The resistance, and reactance of the line
θ	The phase angle of the line
	Quantity
$\bullet_{\min}, \bullet_{\max}$	minimum and maximum (limits) of the corresponding quantity $\bullet$
	Acronyms
DN	Distribution network
DR	Demand response
IC	Industrial customer

I. Introduction

With the integration of renewable energy sources into the power systems, the power systems are facing some emerging challenges to address the associated uncertainty [1]. Facing it, DR serves as an effective tool to alleviate intermittent fluctuations [2]. The traditional DR program includes price-based DR and incentive-based DR. The former refers to the power demand adjustment when customers receive the pricing signal including time-of-use pricing [3], real-time pricing, and critical peak pricing [4]. This DR program aims to reduce the peak-to-valley difference through peaking-cutting and valley-filling using the electricity price adjustment [5]. The latter refers to the incentive-based DR, a.k.a. emergency power demand curtailment [6]. The manners for executing the incentive-based DR programs cover direct load control [7], interruptible load [8], demand-side bidding [9], emergency demand response [10], capacity market programs [11], and ancillary services programs [12], to cite a few. Here, the above references sorely focus on the economy of the DR programs.

Nowadays, with the decarbonization and cleanliness of modern power systems, DR has also evolved as an effective means of reducing carbon emissions [13]. The current studies usually apply DR in the electricity market or economic dispatch to mitigate the environmental influences of carbon emission. For example, Zheng *et al*. [14] presented an integrated methodology that considers RES and real-time-pricing-based DR as an option for planning distribution systems in a transition toward low-carbon sustainability. Fleschutz *et al*. [15] proposed a two-merit order-based method to approximate hourly marginal emission factors for the evaluation of the price-based DR on carbon emissions in European electricity markets. In [16], the DR was introduced into the traditional unit commitment strategy to alleviate the carbon emission on the generation side during both peak and valley load periods. In [17], carbon emissions abatement and incentive strategies in peaking shedding events were explored when facing pressure from both emissions tax and customer non-economic response. Dominguez *et al*. [18] further suggested an environmentally committed asset planning approach to remedy the existing issues to some extent. Stoll *et al*. [19] calculated the dynamic CO_2 signal utilized in the hourly DR programs.

Although references [14]-[19] utilize the DR technology to reduce carbon emission on the generation side, the main participants of the DR are industrial customers (ICs), whose main energy consumption relies on fossil fuels. The carbon emission of the ICs in the DR is rarely considered in the above studies. In addition, in the DR, apart from CO_2 emissions, there are also some pollutants such as SO_2, CO, and NO_x from incomplete ignition of fuels, as well as other greenhouse gases (fluorinated gases: hydrofluorocarbons, perfluorocarbons, SF_6, NF_3; CH_4) from chemical production [20],[21] that can heavily impact the environment. Therefore, the influences of pollutants released by the ICs cannot be neglected in the DR scheduling design.

Although some attempts have been initiated to evaluate the environmental influences in the DR [22],[23], they purely focus on the quantity of the pollutants. More specifically, these studies treat the pollutant as "static" stuff, instead of a "dynamic" diffusion process. However, the pollutants are non-static that will inevitably diffuse in the atmosphere. The greater the diffusion of the pollutants, the larger the environmental influence. Therefore, both the "emission quantity" and the "diffusion extent" should be considered in the evaluation of the environmental influence of pollutants. This statement is also advocated in some studies on the environmental economic dispatch of power systems. For example, Chen *et al.* [24] utilized an air pollutant dispersion model to simulate the pollutant diffusion process in power dispatch. In [25], an air pollutant dispersion model is also used in unit commitment to consider the environmental influence of pollutant diffusion. They generally utilize an oversimplified Gaussian plume model with a simple diffusion coefficient to simulate the diffusion process of the pollutants. However, the diffusion trajectories are indeed dominated by the wind, a key factor that not only determines the polluting intensity but directly guides the polluting direction. More importantly, the scattered distribution of ICs inherently meets different atmospheric conditions that lead to different wind conditions. Unfortunately, these critical factors are unable to be depicted in the traditional Gaussian plume model.

This motivates us to propose a turbulent diffusion theory-based modified Gaussian plume model employed in a novel DR scheduling scheme that can characterize the environmental influence of the pollutants from the ICs in a more realistic manner. As for the uncertainties in the diffusion process of the pollutants, we propose to model the DR scheduling scheme as a stochastic optimization program based on Markov Decision Process. This finally achieves a precise modeling of a stochastic diffusion process of the pollutants in the DR. The main contributions are as follows:

1) To realize the decarbonization and cleanliness of the power systems, a DR scheduling scheme with the proposed turbulent diffusion theory-based modified Gaussian plume model is proposed for the first time to simulate the influence of the pollutants of the ICs.

2) To obtain a precise evaluation result reflecting the influence of the pollutants, we further propose a novel assessment method that considers both the emission quantity

and the diffusion extent. In this method, the Markov Decision Process (MDP) is adopted to simulate the uncertainties of the diffusion trajectory of the pollutants.

3) Facing the curse of dimensionality arising from the MDP-based model that has to deal with a huge number of atmospheric factors and the intrinsic uncertainties in the diffusion of pollutants, a deep reinforcement learning-based MDP solution is further elaborated. Using it, the number of the Markov states of the diffusion trajectory can be reduced by deep reinforcement learning and dynamic approximate programming.

The rest of this paper is organized as follows. Section II presents the problem formulation. Section III elaborates on the proposed method. The case study is provided in Section IV. Finally, Section V concludes the paper.

II. Problem Formulation

A. Traditional DR scheduling scheme

In the traditional DR scheduling model, the objective function minimizes the total cost of purchasing DR services that select ICs for participation. The optimization model can be expressed as

$$\begin{cases} \min\limits_{\Delta \mathbf{P}^t}: \sum_{t\in\mathbf{T}}\sum_{k\in\mathbf{B}_{IC}} C_{dr,k}^{t}\left(\Delta P_k^t\right) \\ s.t. \qquad \boldsymbol{h}\left(\Delta \mathbf{P}^t\right)=0, \boldsymbol{g}\left(\Delta \mathbf{P}^t\right)\le 0 \end{cases}, \tag{1}$$

where $\boldsymbol{h}(\bullet)=0, \boldsymbol{g}(\bullet)\le 0$ are the equality and inequality constraints in the DR scheduling scheme (detailed model of constraints are listed in Section III.D). In (1), $C_{dr,k}^{t}\left(\Delta P_k^t\right)$ can be calculated by

$$C_{dr,k}^{t}\left(\Delta P_k^t\right)=C_{ser,k}^{t}+C_{com,k}^{t}-B_{rel,k}^{t}, \tag{2}$$

$$C_{ser,k}^{t}=\vartheta_{dr}^{t}\cdot\Delta P_k^t, \tag{3}$$

$$C_{com,k}^{t}=\vartheta_{com}^{t}\sum_{l\in\mathbf{L}_k}\sum_{m\in\mathbf{M}_k}\Delta e_{m,l}^{t}, \tag{4}$$

$$B_{rel,k}^{t}=\mathbb{R}\left(P_k^t\right)-\mathbb{R}\left(P_k^t-\Delta P_k^t\right). \tag{5}$$

Here, $\mathbb{R}(\bullet)$ represents the reserve capacity cost function, which can be expressed as a quadratic function [26].

B. DR scheduling scheme considering the pollutant emission of ICs

For realizing a low-carbon and environmental DR scheduling, some existing studies usually add an extra cost, which represents the influences of pollutants emitted by ICs, into the objective function. Thus, the corresponding optimization model can be reformulated as

$$\begin{cases} \min\limits_{\Delta \mathbf{P}^t,\Delta \mathbf{q}}: \sum_{t\in\mathbf{T}}\sum_{k\in\mathbf{B}_{IC}}\left[C_{dr,k}^{t}\left(\Delta P_k^t\right)+C_{emi,k}^{t}\left(\Delta P_k^t,\Delta q_k^{g,t}\right)\right] \\ s.t. \qquad \boldsymbol{h}\left(\Delta \mathbf{P}^t\right)=0, \boldsymbol{g}\left(\Delta \mathbf{P}^t\right)\le 0 \end{cases}. \tag{6}$$

Here, $C_{emi,k}^{t}\left(\Delta P_k^t,\Delta q_k^{g,t}\right)$ is evaluated by the quantity of the pollutants as

$$C_{emi,k}^{t}\left(\Delta P_k^t,\Delta q_k^{g,t}\right)=\sum_{g\in\mathbf{G}}\vartheta_{emi,k}^{g,t}\Delta q_k^{g,t}, \tag{7}$$

Generally speaking, pollutant emissions during industrial production arise from fossil fuel combustion, chemical reactions, and electricity consumption. Suppose that there are *L* process flows and *G* pollutants of the IC. Thus, the emission of pollutant, *g*, at time *t* can be written as

$$q^{g,t}=q_{fuel}^{g,t}+q_{elec}^{g,t}, \tag{8}$$

$$q_{fuel}^{g,t}=\sum\nolimits_{l\in\mathbf{L}}\left(\alpha_l^{g,t}\left(\sum\nolimits_{m\in\mathbf{M}}c_m^{g,t}e_{m,l}^{t}\right)\right), \tag{9}$$

$$q_{elec}^{g,t}=\sum\nolimits_{l\in\mathbf{L}}P_l^t\beta_l^{g,t}c_l^{g,t}. \tag{10}$$

Equation (8) describes the emission, $q^{g,t}$, of pollutant, *g*. Equations (9),(10) specify the formulation of $q_{fuel}^{g,t}$ and $q_{elec}^{g,t}$.

It is worth pointing out that when the IC participates in the DR, some other types of energy, such as natural gas, thermal energy, cooling energy, and raw materials used in process flows change accordingly to the variations of the power supply. These changes may affect the emission of pollutants. Thus, a generic model for describing the industrial process flow should be established. Considering the heterogeneity in ICs, we take the process flow as a black box and only concentrate on the relationship between the inputs and the outputs. For the inputs, energy consumption of the ICs in DR can be divided into electricity and other types of energy while the raw materials can also be divided into fossil fuel and other materials. For the outputs, let us suppose that there are four types of production from the industrial process, i.e., the main production, the sub-production, pollutants produced by fossil fuel combustion, and pollutants produced by electricity consumption. According to [27], the process flow can be generally expressed as

$$\begin{bmatrix} N_{l,pro}^{t} \\ \mathbf{N}_{l,sub}^{t} \\ \mathbf{q}_{l,fuel}^{t} \\ \mathbf{q}_{l,elec}^{t} \end{bmatrix}=\begin{bmatrix} \phi_{l,P}^{t,pro} & \boldsymbol{\Phi}_{l,E}^{t,pro} & \boldsymbol{\Phi}_{l,E}^{t,pro} & \boldsymbol{\Phi}_{l,E}^{t,pro} \\ \boldsymbol{\Phi}_{l,P}^{t,sub} & \boldsymbol{\Phi}_{l,E}^{t,sub} & \boldsymbol{\Phi}_{l,E}^{t,sub} & \boldsymbol{\Phi}_{l,E}^{t,sub} \\ 0 & \boldsymbol{\Phi}_{l,E}^{t,fuel} & \boldsymbol{\alpha}_l^t\mathbf{c}_{l,F}^{t} & 0 \\ \boldsymbol{\beta}_l^t\mathbf{c}_{l,E}^{t} & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} P_l^t \\ \mathbf{E}_{l,ene}^{t} \\ \mathbf{E}_{l,fuel}^{t} \\ \mathbf{E}_{l,oth}^{t} \end{bmatrix}. \tag{11}$$

Here, the middle matrix is the production coefficient matrix; $S_{l,pro}^{t}$ is the quantity of main production; $\mathbf{S}_{l,sub}^{t}$ is the set of sub-productions; $\mathbf{q}_{l,fuel}^{t},\mathbf{q}_{l,elec}^{t}$ are the sets of pollutants produced by fossil fuel combustion and electricity consumption, respectively.

Considering that the requirement of production is almost constant when participating in DR, the IC would convert other energy and materials into electricity. Thus, the conversion relationship between the energy and raw materials in the process flow can be expressed as

$$\begin{bmatrix} P_{out} \\ \mathbf{E}_{ene,out} \\ \mathbf{E}_{fuel,out} \\ \mathbf{E}_{oth,out} \end{bmatrix}=\begin{bmatrix} \eta_P & \boldsymbol{\eta}_{EP} & \boldsymbol{\eta}_{efP} & \boldsymbol{\eta}_{eoP} \\ \boldsymbol{\eta}_{PE} & \boldsymbol{\eta}_{ene} & \boldsymbol{\eta}_{efE} & \boldsymbol{\eta}_{efE} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{fuel}-\boldsymbol{\eta}_{efP}-\boldsymbol{\eta}_{efE} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{oth}-\boldsymbol{\eta}_{eoP}-\boldsymbol{\eta}_{efE} \end{bmatrix}\begin{bmatrix} P_{in} \\ \mathbf{E}_{ene,in} \\ \mathbf{E}_{fuel,in} \\ \mathbf{E}_{oth,in} \end{bmatrix}. \tag{12}$$

Here, the middle matrix represents the conversion efficiency matrix[1]. To give readers a better understanding, we employ

[1] Here, many detailed parameters of the IC operation should be accessible to evaluate the quantity of pollutant emission. These parameters could be obtained in multiple ways. First, much data can be available on a third-party

equations (11),(12) in a process flow of electrolytic aluminum, a typical IC. The details can be found in in Appendix A.

Thus, the DR schedule-maker should evaluate the environmental influence of the pollutant emission before and after the IC participation. The corresponding emission of the pollutant, g, can be calculated through

$$\begin{aligned}\Delta q^{g,t} &= \Delta q^{g,t}_{fuel} + \Delta q^{g,t}_{elec} \\ &= \sum\nolimits_{l\in\mathbf{L}}\left(\alpha_l^{g,t}\sum\nolimits_{m\in\mathbf{M}} c_m^{g,t}\Delta e_{m,l}^{t} + \Delta P_l^t \beta_l^{g,t} c_l^{g,t}\right)\end{aligned}. \tag{13}$$

Remark 1: *Here, it is worth pointing out that although the model in (6) outperforms the traditional DR model, as illustrated (1), by incorporating the pollutant emission of ICs, it purely focuses on the quantity of the pollutants while ignoring the influence of pollutants diffusion. Therefore, it calls for further improvement over the model in (6). This is a critical concern that we will elaborate next.*

III. The Proposed Method

In this section, we will present the proposed turbulent diffusion theory-based modified Gaussian plume model employed in the DR scheduling scheme. Also, a Markov decision process to incorporate the associated uncertainties and a deep reinforcement learning approach to ensure the scalability of the proposed method will be presented.

A. MDP-based pollutant diffusion uncertainty model

When evaluating the influence of pollutants from the ICs, the DR schedule-maker should not only consider the emission but also the diffusion. However, the modeling of pollutant diffusion is not trivial since its trajectory is highly related to lots of atmospheric factors, e.g., wind speed and wind direction, etc. Besides, the influence of these factors on the diffusion trajectory is sequential. All these lead to the stochasticity of the diffusion trajectory, as shown in Fig. 1.

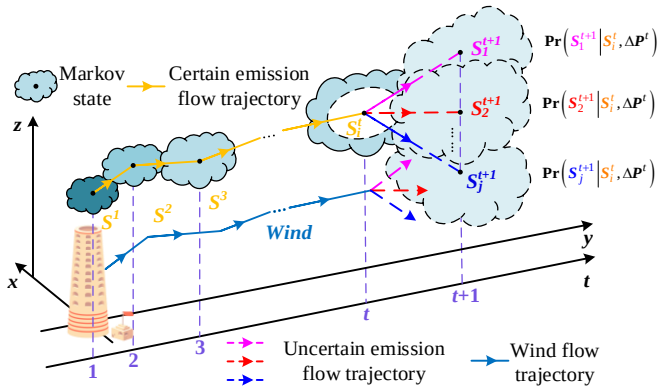

Fig. 1. Plot for the pollutant diffusion trajectory. There are two aspects that should be illustrated. First, the diffusion trajectories are indeed dominated by the wind, a key factor that not only determines the polluting intensity but directly guides the polluting direction. Second, the uncertainty of wind condition leads to the stochasticity of the diffusion trajectory.

platform that monitors energy and materials consumption. Besides, we can apply some state-of-art methods of parameter estimation to obtain some parameters of ICs which are hard to access. What is more, some parameters that rarely influence the results of DR scheduling could also be evaluated from experience.

To tackle the abovementioned uncertainty of the diffusion trajectory, an MDP-based pollutant diffusion uncertainty model is formed. In this model, the diffusion trajectory is modeled via Markov states. Here, let us define $\mathbf{H}^t$ as the set of positions for the original point of the pollutant diffusion at time, t. Define $\mathbf{F}_w^t$ as the set of positions of points under the influence of atmospheric factor, w, at time, t. Thus, the set of current positions of pollutant diffusion at time, t, can be expressed as

$$\mathbf{S}^t = \mathbf{H}^t \bigcup\nolimits_{w\in\mathbf{W}} \mathbf{F}_w^t, \tag{14}$$

On the diffusion trajectory, the current Markov state at time, t, has a probability of reaching the future Markov state at $t+1$. This probability is called *transition probability*, determined by atmospheric factors and the actions of the ICs. The latter can be viewed as responsive volumes due to the participation in DR. Hence, the transition probability can be calculated by

$$\Pr\left(S_j^{t+1}\middle|S_i^t,\Delta P^t\right) = \prod\nolimits_{w\in\mathbf{F}_w^{t+1}} \Pr\left(s_w^{t+1}\middle|s_w^t,\Delta P^t\right), \tag{15}$$

where $\Pr\left(S_j^{t+1}\middle|S_i^t,\Delta P^t\right)$ means the transition probability from the state S_i^t to the state S_j^{t+1} under the action ΔP^t, while $\Pr\left(s_w^{t+1}\middle|s_w^t,\Delta P^t\right)$ means the transition probability from the value of w^{th} atmospheric factor, s_w^t, to the value, s_w^{t+1}, under the action, ΔP^t.

B. Objective function

As described in Section III.A, the environmental influences of pollutants are not only determined by the total quantity but also by their diffusion processes driven by stochasticity. Accordingly, we propose an MDP-based DR scheduling scheme, in which the schedule-maker expects to minimize the scheduling cost and reduce the environmental influence of pollutants. Very importantly, the latter is characterized by both emission and diffusion. The associated objective function is, therefore, formulated as

$$\begin{cases}\min\limits_{\Delta\mathbf{P}^t,\Delta\mathbf{q}} : \sum_{t\in\mathbf{T}}\sum_{k\in\mathbf{B}_{IC}} C_k^t\left(S_{k,i}^t,\Delta P_k^t,\Delta\mathbf{q}_k^t\right) \\ s.t. \quad \boldsymbol{h}\left(\mathbf{S}^t,\Delta\mathbf{P}^t,\Delta\mathbf{q}^t\right)=0, \boldsymbol{g}\left(\mathbf{S}^t,\Delta\mathbf{P}^t,\Delta\mathbf{q}^t\right)\le 0\end{cases}, \tag{16}$$

Also, the Markov state transition probability in (15) reveals that the responsive volumes of the ICs allocated by the schedule-maker could not only influence the current states at time, t, but also future states at time, $t+1$. Thus, an MDP-based recursive objective function is proposed for coping with the state transition probability, which can be expressed as

$$\begin{aligned}&C_k^t\left(S_{k,i}^t,\Delta P_k^t,\Delta\mathbf{q}_k^t\right) = C\left(S_{k,i}^t,\Delta P_k^t,\Delta\mathbf{q}_k^t\right) + \\ &\sum\nolimits_{S_{k,j}^{t+1}\in\mathbf{S}_k^{t+1}} \Pr\left(S_{k,j}^{t+1}\middle|S_{k,i}^t,\Delta P_k^t\right) C_k^{t+1}\left(S_{k,j}^{t+1},\Delta P_k^t\right)\end{aligned}, \tag{17}$$

$$C\left(S_{k,i}^t,\Delta P_k^t,\Delta\mathbf{q}_k^t\right) = C_{dr}\left(S_{k,i}^t,\Delta P_k^t\right) + C_{dif}\left(S_{k,i}^t,\Delta P_k^t,\Delta\mathbf{q}_k^t\right), \tag{18}$$

where $C_{dr}\left(S_{k,i}^{t},\Delta P_k^t\right)$ is calculated by (2)-(5).

To consider the environmental influence of pollutant diffusion from ICs, we modify (7) into (19) by introducing a new index, $E_{dif,k}^{g,t}\left(S_{k,i}^{t},\Delta P_k^t\right)$, as

$$C_{dif}\left(S_{k,i}^{t},\Delta P_k^t\right)=\sum\nolimits_{g\in\mathbf{G}}\vartheta_{dif,k}^{g,t}E_{dif,k}^{g,t}\left(S_{k,i}^{t},\Delta P_k^t\right). \quad (19)$$

Considering the environmental influence of pollutants especially from ICs, we use the pollutant concentration (mg/L) to form the index $E_{dif,k}^{g,t}\left(S_{k,i}^{t},\Delta P_k^t\right)$. In order to calculate the pollutant concentration, we propose a turbulent diffusion theory-based modified Gaussian plume model that can simultaneously characterize the emission and diffusion of pollutants.

C. Environmental pollution evaluation method considering pollutant emission and diffusion

Now, let us elaborate the proposed turbulent diffusion theory-based modified Gaussian plume model that can simultaneously characterize the emission and diffusion of the pollutants. Here, similar to (17), the atmospheric factors and the action of the IC are coupled with both the current and the future states. Thus, an MDP-based recursive evaluation index is formed as

$$\begin{aligned}&E_{dif}^{t}\left(S_i^t,\Delta P^t\right)=D_i^t\left(S_i^t,\Delta P^t\right)+\\&\sum\nolimits_{S_j^{t+1}\in\mathbf{S}^{t+1}}\Pr\left(S_j^{t+1}\middle|S_i^t,\Delta P^t\right)E_{dif}^{t+1}\left(S_j^{t+1},\Delta P^t\right)\end{aligned}, \quad (20)$$

Now, let us further denote $Q_j^{t'}$ and $G_j^{t'}$ as the pollutant concentration and diffusion coefficients in the state $S_j^{t'}$, respectively. This leads to

$$D_i^t\left(S_i^t,\Delta P^t\right)=\sum\nolimits_{t'=0}^{t}\sum\nolimits_{S_j^{t'}\in\mathbf{S}^{t'}}Q_j^{t'}G_j^{t'}. \quad (21)$$

Here, according to Fick's law of diffusion [28], the pollutant concentration can be viewed as the partial differential equations of time and space positions. By defining $\boldsymbol{r}_j^{t'}=\left(x_j^{t'},y_j^{t'},z_j^{t'}\right)$ and $\boldsymbol{r}_0=\left(x_0,y_0,z_0\right)$ as position coordinate in the state $S_j^{t'}$ and S^0, respectively, $Q_j^{t'}$ is described as

$$\begin{cases}\dfrac{\partial Q_j^{t'}}{\partial t'}-G_j^{t'}\nabla^2Q_j^{t'}+\boldsymbol{v}_j^{t'}\cdot\nabla Q_j^{t'}=q_j^{t'}\cdot\ell\left(t'-t_0\right)\cdot\\\left(\delta\left(x_j^{t'}-x_0\right),\delta\left(y_j^{t'}-y_0\right),\delta\left(z_j^{t'}-z_0\right)\right)\\\nabla^2Q_j^{t'}=\dfrac{\partial^2Q_j^{t'}}{\partial\left(x_j^{t'}\right)^2}+\dfrac{\partial^2Q_j^{t'}}{\partial\left(y_j^{t'}\right)^2}+\dfrac{\partial^2Q_j^{t'}}{\partial\left(z_j^{t'}\right)^2}\\\nabla Q_j^{t'}=\dfrac{\partial Q_j^{t'}}{\partial\left(x_j^{t'}\right)}+\dfrac{\partial Q_j^{t'}}{\partial\left(y_j^{t'}\right)}+\dfrac{\partial Q_j^{t'}}{\partial\left(z_j^{t'}\right)}\end{cases}, \quad (22)$$

where $q_j^{t'}$ can be calculated by (8)-(10). Here, for the readers' convenience, we provide the detailed solution for the partial differential equations (22) in Appendix B[2]. Using it, for the diffusion trajectory in the state, $S_j^{t'}$, the associated concentration can be expressed as

$$\begin{aligned}Q_j^{t'}=&\frac{q_j^{t'}}{8\pi G_j^{t'}l_j^{t'}}\exp\left(l_j^{t'}\cdot\left|\boldsymbol{v}_j^{t'}\right|/2G_j^{t'}\right)\\&\left\{\exp\left(\left(\boldsymbol{r}_j^{t'}-\boldsymbol{r}_0\right)\cdot\boldsymbol{v}_j^{t'}/2G_j^{t'}\right)\cdot\operatorname{erfc}\left(l_j^{t'}\cdot/2\sqrt{G_j^{t'}\left(t'-t_0\right)}\right.\right.\\&\left.\left.+\left|\boldsymbol{v}_j^{t'}\right|\left(\left(t_j^{t'}-t_0\right)/4G_j^{t'}\right)^{-1/2}\right)+\exp\left(-\left(\boldsymbol{r}_j^{t'}-\boldsymbol{r}_0\right)\cdot\boldsymbol{v}_j^{t'}/2G_j^{t'}\right)\cdot\right.\\&\left.\operatorname{erfc}\left(l_j^{t'}\cdot/2\sqrt{G_j^{t'}\left(t'-t_0\right)}-\left|\boldsymbol{v}_j^{t'}\right|\left(\left(t_j^{t'}-t_0\right)/4G_j^{t'}\right)^{-1/2}\right)\right\}\end{aligned}, \quad (23)$$

Here, $\operatorname{erfc}(\bullet)=\left(2/\sqrt{\pi}\right)\int_{\bullet}^{\infty}\exp\left(-x^2\right)dx$ stands for the error compensation function [29].

To further demonstrate the influences of wind conditions on pollutant diffusion, a demo for the pollutant with an initial pollutant concentration of 10 mg/L under two different diffusion models is given in Fig. 2. The initial location of the pollutant is $\boldsymbol{r}_0=(6,7,10)$ in km. The simulation time is 1 hour with a step of 5×10^{-5} (s). The space size is 10×14 (km) with a simulation step of 0.01 (km). In this simulation, the wind conditions change every 15 minutes as depicted as t_1-t_4. The variations of wind directions and speeds are shown as Fig. 2.

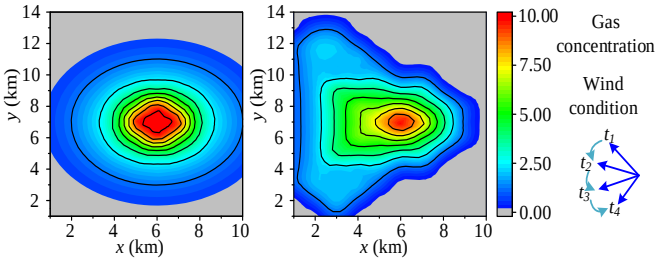

Fig. 2. Pollutant diffusion distribution with (a) Gaussian plume model; (b) our proposed model.

Remark 2: *Obviously, the pollutant diffuses in an ellipsoid under Gaussian plume model while, as a contrast, it diffuses in a shape like droplets along with wind direction under our proposed method. Therefore, it comes with no surprise that our proposed model characterizes the diffusion trajectory in a more precise and realistic manner. This is also a major contribution of our article.*

Apparently, using (23), the pollutant concentration under some specific wind conditions can be simplified and summarized in Table I.

[2] Here, one reviewer has pointed out that the difference between pollutant emission and diffusion should be clarified. This is, indeed, an important issue that needs to be addressed. First, in Fig. 2, the environmental influence of emission is viewed as evaluating the influence only in state S^1. However, the influence of diffusion is regarded as evaluating the influence from the state S to the state S^{t+1}. Furthermore, from the perspective of mathematical modeling, pollutant emission is only viewed as a quantity while pollutant diffusion is regarded as a partial differential equation.

TABLE I. POLLUTANT CONCENTRATION UNDER SOME SPECIFIC SCENARIOS

Condition	Pollutant concentration
$t_j^{t'} \le t_0$	$Q_j^{t'} = 0$
$t_j^{t'} \to \infty$	$Q_j^{+\infty} = \left(q_j^{t'} / 4\pi G_j^{t'} l_j^{t'}\right) \exp\left(\left(\left(\boldsymbol{r}_j^{t'} - \boldsymbol{r}_0\right) \cdot \boldsymbol{v}_j^{t'} - l_j^{t'} \cdot \left\|\boldsymbol{v}_j^{t'}\right\|\right) / 2G_j^{t'}\right)$
$\boldsymbol{v}_j^{t'} = \boldsymbol{0}$	$Q_j^{t'} = q_j^{t'} / 4\pi G_j^{t'} l_j^{t'}$
$\boldsymbol{v}_j^{t'} = \left(v\left(x_j^{t'}\right), 0, 0\right)$	$Q_j^{t'} = \left(q_j^{t'} / 4\pi G_j^{t'} l\right) \exp\left(v\left(x_j^{t'}\right)\left(\left(t' - t_0\right) - l\right) / 2G_j^{t'}\right)$

D. Constraints

The equality and inequality constraints $\boldsymbol{h}(\bullet) = 0$, $\boldsymbol{g}(\bullet) \le 0$ can be expressed as follows:

$$\begin{cases} \Delta P_{k,\min}^t \le \Delta P_k^t \le \Delta P_{k,\max}^t, k \in \mathbf{B}_{IC} \\ \Delta P_k^t = 0, k \notin \mathbf{B}_{IC} \end{cases}, \tag{24}$$

$$q_k^{g,t} \le q_{k,\max}^{g,t}, g \in G, k \in \mathbf{B}_{IC}, t \in \mathbf{T}, \tag{25}$$

$$\begin{cases} \boldsymbol{e}_{l,\min}^t \le \boldsymbol{e}_l^t \le \boldsymbol{e}_{l,\max}^t \\ \boldsymbol{e}_{l,\,fuel,\min}^t \le \boldsymbol{e}_{l,\,fuel}^t \le \boldsymbol{e}_{l,\,fuel,\max}^t, l \in \mathbf{L}_k, t \in \mathbf{T} \\ \boldsymbol{e}_{l,oth,\min}^t \le \boldsymbol{e}_{l,oth}^t \le \boldsymbol{e}_{l,\,fuel,\max}^t \end{cases}, \tag{26}$$

$$U_k^t - U_{k'}^t \le 2\left(r_{kk'} F_{P,kk'}^t + x_{kk'} F_{Q,kk'}^t\right), k \in \mathbf{B}_{DN}, k' \in \mathbf{B}_k, t \in \mathbf{T}, \tag{27}$$

$$\begin{cases} F_{P,k}^t - \Delta P_k^t = U_k^t \sum_{k' \in \mathbf{B}_k^t} U_{k'}^t \left(g_{kk'} \cos\theta_{kk'}^t + b_{kk'} \sin\theta_{kk'}^t\right) \\ F_{Q,k}^t = U_k^t \sum_{k' \in \mathbf{B}_k^t} U_{k'}^t \left(g_{kk'} \cos\theta_{kk'}^t - b_{kk'} \sin\theta_{kk'}^t\right) \\ k \in \mathbf{B}_{DN}, t \in \mathbf{T} \end{cases}, \tag{28}$$

$$\left(F_{P,kk'}^t\right)^2 + \left(F_{Q,kk'}^t\right)^2 \le \left(F_{S,kk'}^t\right)^2, k \in \mathbf{B}_{DN}, k' \in \mathbf{B}_k, t \in \mathbf{T}, \tag{29}$$

$$\begin{cases} -F_{S,kk'}^t \le F_{P,kk'}^t \le F_{S,kk'}^t \\ -F_{S,kk'}^t \le F_{Q,kk'}^t \le F_{S,kk'}^t \\ -\sqrt{2}F_{S,kk'}^t \le F_{P,kk'}^t + F_{Q,kk'}^t \le \sqrt{2}F_{S,kk'}^t \\ -\sqrt{2}F_{S,kk'}^t \le F_{P,kk'}^t + F_{Q,kk'}^t \le \sqrt{2}F_{S,kk'}^t \end{cases}, k \in \mathbf{B}_{DN}, k' \in \mathbf{B}_k, t \in \mathbf{T}, \tag{30}$$

$$U_{k,\min}^t \le U_k^t \le U_{k,\max}^t, k \in \mathbf{B}_{DN}, t \in \mathbf{T}. \tag{31}$$

where $\mathbf{B}_k$ is the set of the nodes connected to node k; $g_{kk'}, b_{kk'}$ are the values of k-th row and k'-th column of the nodal admittance matrix. Equation (24) represents the responsive volume constraint of ICs. Equation (25) limits the maximum pollutant emission of ICs. Equation (26) represents raw material and energy constraints in the process flow of ICs. Equations (27)-(31) ensure the security of DN through power flow constraints, power balance constraints, line capacity constraints, and voltage constraints, respectively. To reduce the computation burden, the nonlinear constraints in (29) are relaxed to a group of linear constraints as described in (30) [30].

E. MDP-based DR optimization model and the solution

According to Sections III.A-B, we model the DR scheduling problem as an MDP and consider each IC as an agent. Now, let us define the Markov state space as the pollutant diffusion trajectory set $\mathbf{S}_k$, where $\mathbf{S}_k = \left\{S_{k,i}^t \middle| t \in \mathbf{T}, i \in \mathbf{I}_k^t\right\}$ and $\mathbf{I}_k^t$ is the set of the number of pollutant diffusion trajectory positions of IC$_k$ at time, t. The action space is defined as the responsive volumes that the schedule-maker allocates to all ICs, which is expressed as $\Delta\mathbf{P}_k = \left\{\Delta P_k^t \middle| t \in \mathbf{T}\right\}$. The state transition probability is related to current atmospheric factors and the action (see (15)). For simplicity, we use $\gamma_{k,ij}^{\Delta t}$ to express the state transition probability, as shown in

$$\gamma_{k,ij}^{\Delta t} = \Pr\left(S_{k,j}^{t+1} \middle| S_{k,i}^t, \Delta P_k^t\right). \tag{32}$$

Let us define the reward function, $r_k^t\left(S_{k,i}^t\right)$, as the opposite value of $C\left(S_{k,i}^t, \Delta P_k^t\right)$. We have

$$r_k^t\left(S_{k,i}^t\right) = -C\left(S_{k,i}^t, \Delta P_k^t\right). \tag{33}$$

Let us define the reward function $r_k^{\Delta t}\left(S_{k,j}^{t+1} \middle| S_{k,i}^t\right)$ of the transition from $S_{k,i}^t$ to $S_{k,j}^{t+1}$ as the opposite value of $C_{emis}\left(S_{k,i}^t, \Delta P_k^t\right)$. We have

$$r_k^{\Delta t}\left(S_{k,j}^{t+1} \middle| S_{k,i}^t\right) = -C_{emis}\left(S_{k,i}^t, \Delta P_k^t\right). \tag{34}$$

The value function $V_k^t\left(S_{k,i}^t\right)$ of the IC$_k$ in the state $S_{k,i}^t$ under the action ΔP_k^t is calculated as

$$\begin{aligned} V_k^t\left(S_{k,i}^t\right) &= \mathbb{E}\left[-C_{dr,k}^t\left(S_{k,i}^t, \Delta P_k^t\right)\right] \\ &= \mathbb{E}\left[r_k^t\left(S_{k,i}^t\right) + r_k^{\Delta t}\left(S_{k,j}^{t+1} \middle| S_{k,i}^t\right) + \sum_{S_{k,j}^{t+1} \in \mathbf{S}_k} \gamma_{k,ij}^{\Delta t} V_k^{t+1}\left(S_{k,j}^{t+1}\right)\right] \end{aligned}, \tag{35}$$

where $\mathbb{E}[\bullet]$ represents the expectation function.

Thus, the proposed model is reformulated into approximate dynamic programming with decision variables, $\Delta\mathbf{P}^t\left(\mathbf{S}^t\right)$, as

$$\begin{aligned} &\Delta\mathbf{P}^t = \arg\max_{\Delta\mathbf{P}^t} \sum_{k \in \mathbf{B}_{IC}^t} \mathbb{E}\left[r_k^t\left(S_{k,i}^t\right) + \sum_{S_{k,j}^{t+1} \in \mathbf{S}_k} \gamma_{k,ij}^{\Delta t} V_k^{t+1}\left(S_{k,j}^{t+1}\right)\right] \\ &s.t.\ (24) - (31) \end{aligned}. \tag{36}$$

The constraints $\boldsymbol{h}\left(\Delta\mathbf{P}^t\right) = 0, \boldsymbol{g}\left(\Delta\mathbf{P}^t\right) = 0$ (see (24)-(31)) could be added into the reward function via Lagrange relaxation method [31]. Then, the reformulated reward function with an augmented Lagrange function can be expressed as

$$L_k^t\left(\Delta P_k^t, \boldsymbol{\lambda}, \boldsymbol{\mu}\right) = r_k^t\left(S_{k,i}^t\right) - \boldsymbol{\lambda}^T \boldsymbol{h}\left(\Delta\mathbf{P}^t\right) - \boldsymbol{\mu}^T \boldsymbol{g}\left(\Delta\mathbf{P}^t\right), \tag{37}$$

where $\boldsymbol{\lambda}, \boldsymbol{\mu}$ are the set of Lagrange multipliers. Equation (36) can be reformulated as

$$\Delta\mathbf{P}^t = \arg\min_{\boldsymbol{\lambda}, \boldsymbol{\mu} \ge \mathbf{0}} \max_{\Delta\mathbf{P}^t} \sum_{k \in \mathbf{B}_{IC}^t} \mathbb{E}\left[\begin{aligned} &L_k^t\left(\Delta P_k^t, \boldsymbol{\lambda}, \boldsymbol{\mu}\right) \\ &+ \sum_{S_{k,j}^{t+1} \in \mathbf{S}_k} \gamma_{k,ij}^{\Delta t} V_k^{t+1}\left(S_{k,j}^{t+1}\right) \end{aligned}\right], \tag{38}$$

A framework to introduce the process of reformulation can be seen in Fig. 3. It describes the relationship between the optimization problem (16) and the proposed deep reinforcement problem (38).

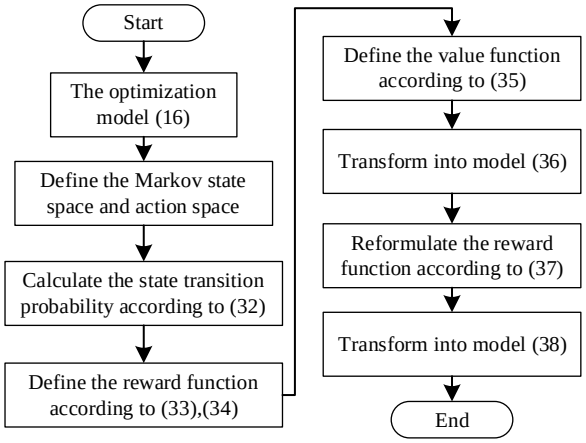

Fig. 3. The process of reformulation.

However, the numerous atmospheric factors with uncertainties on pollutant diffusion lead to an enormous number of states and actions, as well as complex state transitions in the MDP-based optimization model. Thus, the solution of the proposed optimization model (36) falls into the well-known "curse of dimensionality". To solve it, we advocate a deep reinforcement learning-based method combined with approximate dynamic programming. The key is to estimate the approximate future value function, $V_k^{t+1}\left(S_{k,j}^{t+1}\right)$. Commonly, temporal-difference (TD) prediction is an effective method for estimation [32], which can be viewed as

$$\tilde{V}_k^{t+1}\left(S_{k,j}^{t+1}\right) \approx \begin{array}{l} \tilde{V}_k^{t}\left(S_{k,i}^{t}\right)+\varphi\Big[r_k^{\Delta t}\left(S_{k,j}^{t+1}\middle|S_{k,i}^{t}\right) \\ +\gamma_{k,ij}^{\Delta t}\tilde{V}_k^{\Delta t}\left(S_{k,j}^{t+1}\middle|S_{k,i}^{t}\right)-V_k^{t}\left(S_{k,i}^{t}\right)\Big] \end{array}, \quad (39)$$

$$V_k^{\Delta t}\left(S_{k,j}^{t+1}\middle|S_{k,i}^{t}\right)=\tilde{V}_k^{t+1}\left(S_{k,j}^{t}\right)-V_k^{t}\left(S_{k,i}^{t}\right), \quad (40)$$

where $\tilde{V}_k^{t}\left(S_{k,i}^{t}\right)$ is the estimated value function of IC$_k$ in the state $S_{k,i}^{t}$; φ is the step size. In the TD prediction, $r_k^{\Delta t}\left(S_{k,j}^{t+1}\middle|S_{k,i}^{t}\right)+\gamma_{k,ij}^{\Delta t}\tilde{V}_k^{\Delta t}\left(S_{k,j}^{t+1}\middle|S_{k,i}^{t}\right)$ is called TD target value; $r_k^{\Delta t}\left(S_{k,j}^{t+1}\middle|S_{k,i}^{t}\right)+\gamma_{k,ij}^{\Delta t}\tilde{V}_k^{\Delta t}\left(S_{k,j}^{t+1}\middle|S_{k,i}^{t}\right)-V_k^{t}\left(S_{k,i}^{t}\right)$ is called the TD error that measures the difference between the estimated value TD target value.

Furthermore, the artificial neural network with parameters, θ, is used to approximate $V_k^{t+1}\left(S_{k,j}^{t+1}\right)$. To guarantee the robustness of the training process against errors, there are two techniques. One is to use a target network to calculate the temporal-different (TD) target value [33]. Within it, we initially create a copy of the neural network $V_k^{t+1}\left(S_{k,j}^{t+1},\theta\right)$, denoted as $V_k^{t+1}\left(S_{k,j}^{t+1},\hat{\theta}\right)$. Network $V_k^{t+1}\left(S_{k,j}^{t+1},\hat{\theta}\right)$ is called the target network and original network $V_k^{t+1}\left(S_{k,j}^{t+1},\theta\right)$ is called the online network. During the training process, the parameters θ of online network are updated every step while the parameter $\hat{\theta}$ of target network are updated every C steps (making it equal to the current θ). With the target network, TD target value can be calculated as

$$\hat{y}\left(S_{k,j}^{t+1}\middle|S_{k,i}^{t}\right)=\begin{cases} r_k^{\Delta t}\left(S_{k,j}^{t+1}\middle|S_{k,i}^{t}\right) & S_{k,j}^{t+1}\text{ is terminal} \\ r_k^{\Delta t}\left(S_{k,j}^{t+1}\middle|S_{k,i}^{t}\right)+ & \\ \gamma_{k,ij}^{\Delta t}V_k^{\Delta t}\left(S_{k,j}^{t+1}\middle|S_{k,i}^{t}\right) & S_{k,j}^{t+1}\text{ is non-terminal} \end{cases}. \quad (41)$$

The other technique used to ensure the robustness of the neural network against errors as well as improve data efficiency is the experience replay mechanism [34]. More specially, we maintain a replay buffer $\mathbf{D}$ of capacity D for ICs. After receiving the DR scheduling assignment, each IC$_k$ will store the transition $\left(S_i, r\left(S_i\middle|S_j\right)\right)$ in the replay buffer $\mathbf{D}$. At each step of training, we randomly choose M transitions from $\mathbf{D}$ to form a mini-batch $\mathbf{M}$. With the mini-batch, we can calculate the loss function, $\ell(\theta)$, in a mean square error (MSE) manner as

$$\ell(\theta)=\frac{1}{M}\sum_{\left(S_i^t, r\left(S_i^t\middle|S_j^t\right)\right)\in\mathbf{M}}\left[\hat{y}\left(S_{k,j}^{t+1}\middle|S_{k,i}^{t}\right)-V_k^{t+1}\left(S_{k,j}^{t+1},\theta\right)\right]^2, \quad (42)$$

Then, the parameters are updated by gradient descent to minimize the loss function

$$\theta \leftarrow \theta-\eta\nabla\ell(\theta). \quad (43)$$

Here, η is the learning rate. The flow chart of the proposed solution is given in Fig. 4.

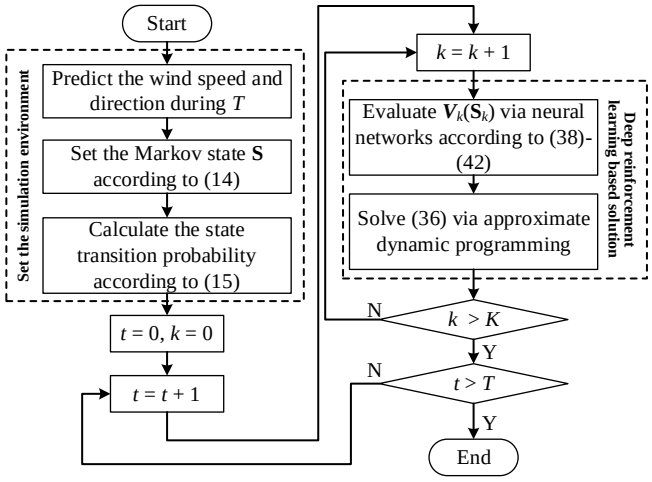

Fig. 4. Flow chart of the proposed solution.

IV. Case Study

In this section, we test the performance of the proposed method in a modified IEEE 118-node distribution system.

A. *Experiment Setup*

This section constructs the city region with the modified IEEE 118-node distribution test system, where the locations of ICs, commercial customers, and residential customers are presented in Fig. 5 [35]. Fig. 6 shows a typical day load curve. In the case study, the DR period is set from 9:00 to 12:00 and the schedule-maker is supposed to make the scheduling every 15 minutes[3]. Table II gives the DR service and incentive compensation unit cost in each period.

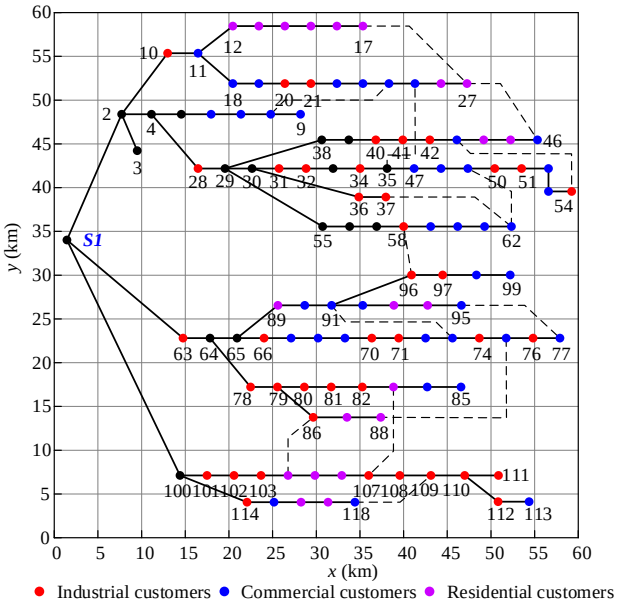

Fig. 5. The modified IEEE-118 node test system.

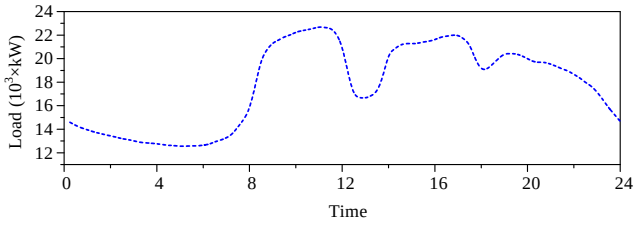

Fig. 6. Typical load curve.

In the simulation system, there are 4 types of ICs participating in the DR, i.e., iron and steel ICs (ISIC), electrolytic aluminum IC (EAIC), coking IC (CIC), and cotton manufacturing IC (CMIC). Their corresponding process flow, raw materials, and energy supplies come from [36],[37],[38],[39].

TABLE II. UNIT COST IN DIFFERENT PERIODS

Period	ϑ_{dr}^{t} (kWh/$)	ϑ_{com}^{t} (kWh/$)
Valley (23:00-7:00)	0.2	0.3
Flat (7:00-10:00,14:00-17:00)	0.6	0.7
Peak (10:00-14:00,17:00-23:00)	1.1	1.15

ISICs are connected to nodes 34, 36, 37, 40, 41, 42, 50, 51, 54, 58, 97, respectively. EAICs are connected to nodes 70, 71, 74, 76, 82, 96, 107, 108, 109, 110, respectively. CICs are connected to 63, 66, 78, 79, 80, 81, 82, 86, 101, 102, respectively. CMICs are connected to 10, 20, 21, 28, 31, 32, 103, 111, 112, 114, respectively.

The neural network used to approximate the state value function is plotted in Fig. 7 with 3 hidden layers as 128, 64, and 32 neurons, respectively. We apply rectified linear unit to each layer. At the input layer, the discrete state variables (i.e., time step t and location $\boldsymbol{r}$) are represented with one-hot encoding. The Adam optimizer is applied to learn the neural network parameters with a learning rate $\eta = 0.001$. The other hyperparameters during training are as follows: the number of episodes is 50, the capacity of replay buffer is 5×10^{5}, the discount factor is 0.9, and the decay step size of the exploration probability is 1.5×10^{-4}.

The proposed problem is written and solved in Python with Gurobi, and the neural networks are trained in Python with Pytorch [40], an open-source deep learning platform. All experiments are carried out on a computer with a 10-core 3.70 GHz Intel CoreTM i9-10900X processor and 32 GB of RAM.

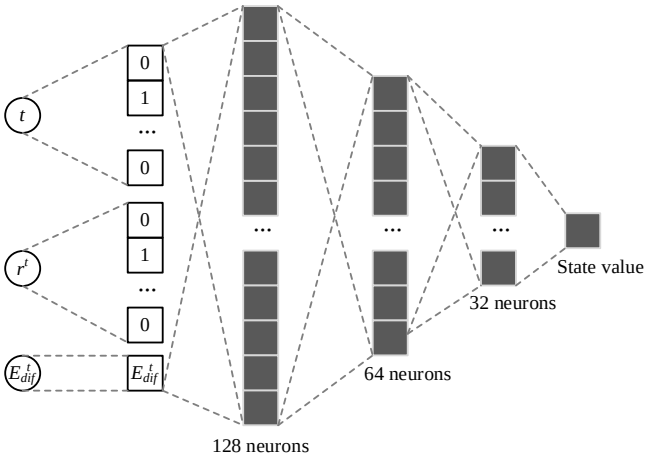

Fig. 7. Neural network architecture for estimating state value. The network has 3 hidden layers as 128, 64, and 32 neurons, respectively. Rectified linear unit is applied to each layer. At the input layer, the discrete state variables (i.e., time step t and location $\boldsymbol{r}$) are represented with one-hot encoding.

B. *Base Case*

We demonstrate the validity of the proposed model in this paper via comparing with three other DR scheduling models. The models are defined as follows:

- M_1: DR scheduling model without pollutant emission
- M_2: DR scheduling model with pollutant emission, but without pollutant diffusion

[3] Here, we would like to emphasize that 15 minutes are for ICs to do the demand response but not to simulate the diffusion process. In the demand response simulation, 15 minutes is not a standard value, but a quite common choice in the existing literature, e.g., [42], [43]. Of course, we can also select other interval for the demand response interval, such as 20 minutes or 30 minutes. In our diffusion process simulation, we utilize Partial Differential Equation Toolbox™, one of the toolboxes for solving partial differential equation via MATLAB. In our experiment, we also found that this time interval is 5×10^{-5} (s), space interval on X-axis, Y-axis, and Z-axis is 0.01 (km), at least, suitable for the data background of the manuscript.

- M_3: DR scheduling model with pollutant emission and certain diffusion
- M_4: DR scheduling model with pollutant emission and uncertain diffusion (our proposed model)

Fig. 8 gives the total DR scheduling cost under M_1-M_4. In each responsive period, the total cost of M_1 is the highest while that of M_2-M_4 decreases. The total cost of M_4 is the lowest among all four models. Fig. 9(a)-(b) gives the breakdown of the schedule-maker for the DR scheduling cost. The cost for purchasing DR services from ICs is similar under M_1-M_4. The cost for incentivizing IC to participate in the DR of M_2 ranks first, followed by M_3, M_4, and M_1 accordingly. Furthermore, the equivalent benefit for DN performance enhancement under M_1 is the highest but its equivalent cost of influences is also the highest. The schedule-maker via M_2-M_4 prefers to reduce the influence of pollutants at the expense of some economic benefits. It shows that the pollutant emission from ICs has a huge influence on DR scheduling.

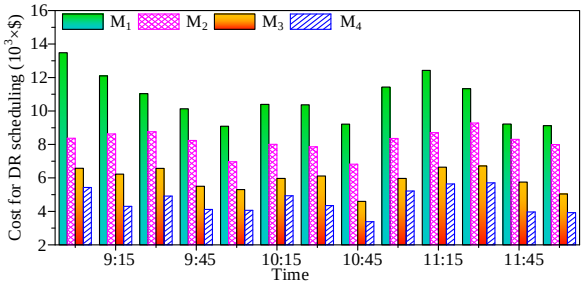

Fig. 8. Total cost for the DR scheduling.

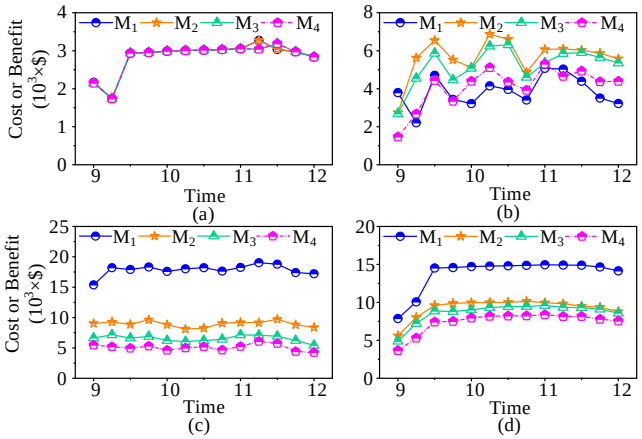

Fig. 9. Breakdown of the schedule-maker for DR scheduling cost with (a) cost for buying DR services from ICs; (b) Cost for compensating ICs to participate in the DR; (c) equivalent cost of the influences of the pollutants on the environment (d) equivalent benefit for promoting performances of the DN.

For insight comparisons among M_2-M_4, Fig. 10 and Fig. 11 show the influence of pollutant diffusion. Fig. 10(a)-(b) shows that under M_1 and M_2, the schedule-maker mainly selects ISICs and EAICs with high responsive potential to participate in the DR. It leads to the pollution space ratio up to approximately 22%. Fig. 10(c) reveals that under M_3, the schedule-maker tends to choose CICs and CMICs with middle responsive potential but small pollutant diffusion. The pollution space ratio under M_3 decreases by nearly 7% and 3% compared with M_1 and M_2, respectively. Fig. 10(d) shows that the schedule-maker prefers to select more scattered ICs with middle or even small potential to reduce the influence of pollutant diffusion. The pollution space ratio under M_4 descends by approximately 4.5%. In short, the proposed model is more beneficial to the influence reduction of pollutants and realizes a low-carbon and environmentally-friendly DR. Furthermore, the schedule-maker should consider not only the influences of the pollutant emission but also the influences of the pollutant diffusion from the DR participants.

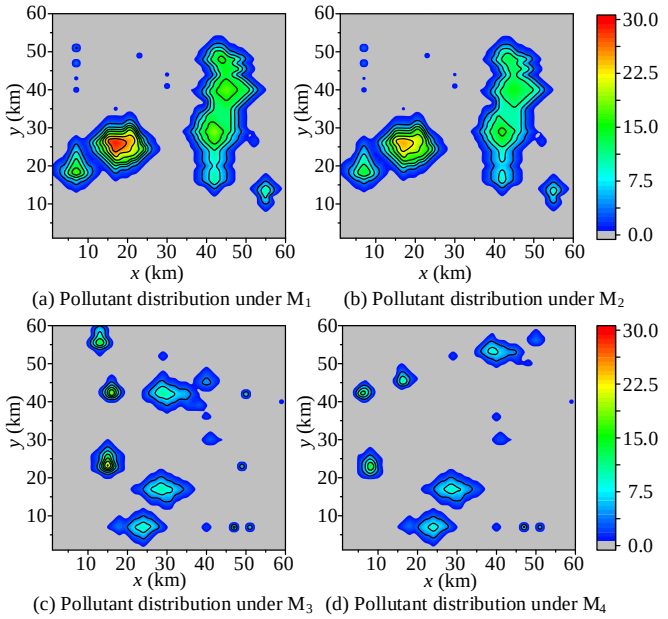

Fig. 10. Pollutant distribution under different DR scheduling with (a) M_1;(b) M_2;(c) M_3;(d) M_4.

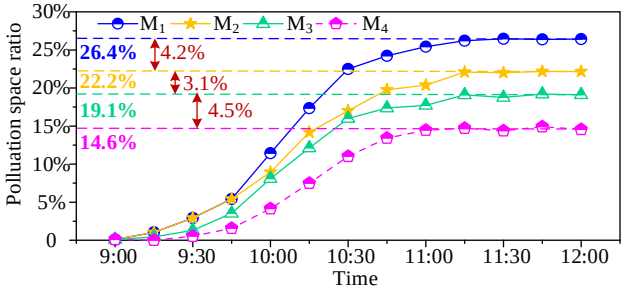

Fig. 11. Pollution space ratio under different DR scheduling.

C. *Analysis on impacting factors*

Many factors can directly affect pollutant diffusion, and, therefore, indirectly influence the results of the DR scheduling. Therefore, we further evaluate the impact of some factors on the proposed method through two insight cases.

1) Weather condition

Solar radiation and cloud thickness under different weather conditions play important roles in pollutant diffusion. Hence, this paper sets five scenarios with different weather conditions to investigate the impacts of these factors on DR scheduling. The scenarios are set as follows:

- S_1: Sunny day with high solar radiation
- S_2: Sunny day with middle solar radiation
- S_3: Sunny day with low solar radiation

- S_4: Cloudy day with thin clouds
- S_5: Cloudy day with thick clouds

Fig. 12 gives the variation of the breakdown of the DR scheduling cost, pollution space ratio, and diffusion coefficient under S_1-S_5. It is easily found that with the decrease in solar radiation and the increase in cloud thickness, the diffusion coefficient gradually descends. The pollution space ratio declines correspondingly and the total DR scheduling cost decreases as well. Therefore, the schedule-maker should select the ICs with small pollutant diffusion under the weather conditions beneficial for pollutant diffusion. Further, when the weather conditions are unconducive to pollutant diffusion, the schedule-maker can choose the ICs with high responsive potentials and middle pollutant emission.

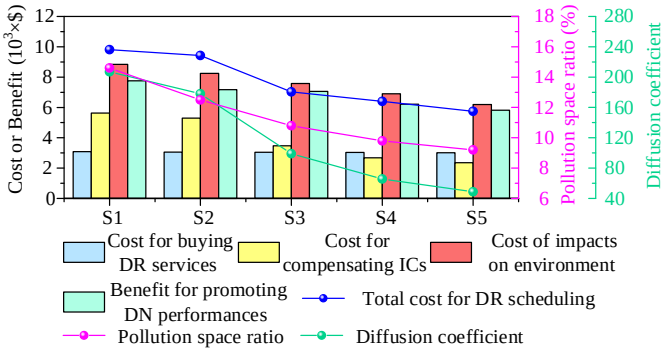

Fig. 12. Results under different weather conditions.

2) *Uncertainty of the wind*

The uncertainty of the wind is an essential factor that leads to the curse of dimensionality in the solution of the proposed model. Thus, to illustrate the effectiveness of the proposed deep reinforcement learning-based solution method, the cumulative rewards and computational time to reach convergence of the proposed method with different uncertainty levels, from 10%- 50%, of the wind are tested.

Fig. 13 gives episodic average cumulative rewards for the proposed method under different uncertainty factors of wind. With the increase in the uncertainty factor, the number of episodes increases gradually. Table III shows the computation time to reach convergence of the proposed method with different uncertainty factors of wind.

TABLE III. COMPUTATION TIME TO REACH CONVERGENCE OF THE PROPOSED METHOD WITH DIFFERENT UNCERTAINTY FACTORS OF WIND

Uncertainty factor	Total time (min)	Number of episodes	Average time per episode (min)
10%	277.56	8	34.70
20%	312.88	14	22.35
30%	489.33	25	19.57
40%	512.29	30	17.08
50%	675.82	40	16.90

Although the total solution time increases with the increase in the uncertainty factor, the average time per episode decreases gradually. It demonstrates that the proposed solution method can efficiently alleviate the curse of dimensionality in the solution of the MDP-based model.

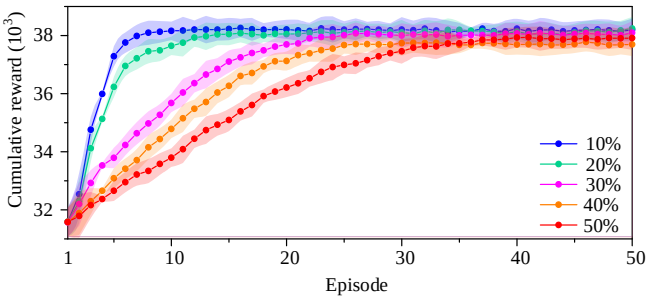

Fig. 13. Episodic average cumulative reward for the proposed method under different uncertainty factors of wind.

3) *Robustness of training*

This subsection analyses the effectiveness of the target network (defined as M_1) and experience replay mechanism (defined as M_2) to guarantee the stability of training in the neural network. Fig. 14 shows the impact of M_1 and M_2 on the proposed method. The absence of M_2 decreases the average cumulative reward by 35.94% while the absence of M_1 decreases it by just 4.62%. What is worse, the average cumulative reward drops by 39.15% with the proposed method without both of them. Thus, it is a must to keep these techniques to maintain a stable training process. It also proves that our proposed method combined with M_1 and M_2 can ensure robustness against errors in neural network training.

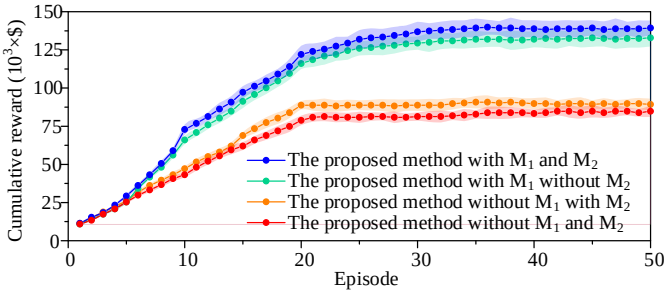

Fig. 14. Impact of M_1 and M_2 on the proposed method. Error bars are the 95% confidence intervals across 6 random seeds.

V. CONCLUSION

This paper proposes a DR scheduling considering the pollutant diffusion uncertainty of ICs. In the proposed model, an MDP-based evaluation method is proposed considering both pollutant emission and diffusion of ICs. To tackle the uncertainty of pollutant diffusion, the pollutant diffusion trajectory is modeled as the Markov state and a precise space-time diffusion model is formed. Further, considering the pollutant diffusion uncertainty, an MDP-based recursive DR optimization model is proposed with the objective of minimizing the scheduling cost and reducing the environmental influences of pollutants. In addition, the presence of numerous atmospheric factors with high uncertainty on pollutant diffusion leads to enormous states and actions, as well as complex state transitions in the MDP-based optimization model. A deep reinforcement learning-based solution method has then been proposed to overcome the curse of dimensionality.

The case studies executed on a modified IEEE-118 node distribution test system with 40 ICs demonstrate the validity of

the proposed method. The simulation results also compare the proposed method with three other DR scheduling methods and show that the proposed method can effectively reduce the environmental influences of pollutant emission and diffusion.

Furthermore, some insight studies on the factors that influence the proposed method have given valuable suggestions about the DR participant selection under different weather conditions to the schedule-maker. In addition, in order to illustrate the effectiveness of the deep reinforcement learning-based solution method, we have also presented the results of the simulations carried out under different uncertainty factors of wind. The simulation results are very useful for the schedule-maker to develop a reasonable DR scheduling plan.

APPENDIX A.

For example, Fig. 15 depicts the process flow of electrolytic aluminum, a typical IC. Its process flow can be expressed as Fig. 15.

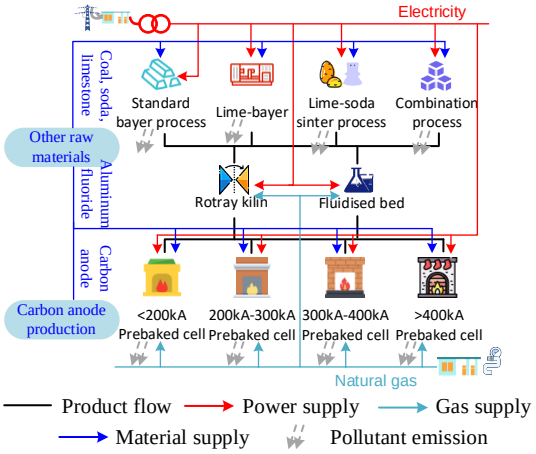

Fig. 15. The process flow of electrolytic aluminum. The electrolytic aluminum industry consists of five subsectors: bauxite mining, alumina refining, primary aluminum smelting, anode production and casting. Each subsector has its own distinct characteristics, such as production processes, technologies and energy demand.

For the IC of the electrolytic aluminum, S_{pro} is the production of aluminum; $\mathbf{S}_{sub}=\{S_{Ald}\}$, where S_{Ald} is the production of aluminum dust; $\mathbf{q}_{fuel}=\left[q_{fuel}^{XF}\quad q_{fuel}^{SO2}\quad q_{fuel}^{Al2O3}\quad q_{fuel}^{petro}\right]^T$, where $q_{fuel}^{XF},q_{fuel}^{SO2},q_{fuel}^{Al2O3},q_{fuel}^{petro}$ represents the emission of fluorinated gases (e.g., CF_4, C_2F_6, and HF), SO2, alumina dust, and petroleum dust from fossil fuel, respectively; $\mathbf{q}_{ele}=\left[q_{ele}^{CO2}\quad q_{ele}^{XF}\right]^T$, where $q_{ele}^{CO2},q_{ele}^{XF}$ represent the emission of carbon dioxide and fluorinated gases from electricity usage, respectively.

In addition, the whole process flow needs three types of energy, i.e., electricity, natural gas, and coal, and four types of raw materials, i.e., aluminum fluoride, soda, limestone, and carbon anode. Thus, $\mathbf{E}_{ene}=\{e_{gas}\}$, where e_{gas} is the consumption of natural gases; $\mathbf{E}_{fuel}=\{e_{coal}\}$, where e_{coal} is the consumption of coal; $\mathbf{E}_{oth}=\left[e_{sod}\quad e_{lst}\quad e_{AlF}\quad e_{Can}\right]^T$, where $e_{sod},e_{lst},e_{AlF},e_{Can}$ represent the consumption of some raw materials including soda, limestones, aluminum fluorides, and carbon anodes, respectively. Equation (11) can be reformulated into (44).

$$\begin{bmatrix} S_{pro} & S_{Ald} \\ q_{fuel}^{XF} & q_{fuel}^{SO2} \\ q_{fuel}^{Al2O3} & q_{fuel}^{petro} \\ q_{ele}^{CO2} & q_{ele}^{XF} \end{bmatrix}^T = \mathbf{\Phi}_{IC} \begin{bmatrix} P & e_{gas} \\ e_{gas} & e_{sod} \\ e_{lst} & e_{AlF} \\ e_{Can} & \end{bmatrix}^T \tag{44}$$

Here, $\mathbf{\Phi}_{IC}$ represents the production coefficient matrix of the IC of the electrolytic aluminum.

Besides, there are three types of energy or materials that can be converted each other, i.e., electricity, natural gas, and coal. Therefore, equation (12) can be reformulated into (45).

$$\begin{bmatrix} P_{out} \\ e_{gas,out} \\ e_{coal,out} \end{bmatrix} = \begin{bmatrix} \eta_P & \eta_{gP} & \eta_{cP} \\ \eta_{Pg} & \eta_{gas} & \eta_{cg} \\ 0 & 0 & 1-\eta_{cg}-\eta_{cP} \end{bmatrix} \begin{bmatrix} P_{in} \\ e_{gas,in} \\ e_{coal,in} \end{bmatrix} \tag{45}$$

Here, $\bullet_{out},\bullet_{in}$ are the output and input of energy or raw materials; the middle matrix represents the conversion efficiency matrix.

APPENDIX B.

For simplicity, let us omit the superscript t' and the subscript j in the following descriptions. According to Fick's Law of Diffusion [41], define $\boldsymbol{r}=(x,y,z)$, $\boldsymbol{r}_0=(x_0,y_0,z_0)$, and then we have

$$\frac{\partial Q}{\partial t}=G\cdot\nabla^2 Q-(\boldsymbol{v}\cdot\nabla Q), \tag{46}$$

$$\frac{\partial Q}{\partial t}=G\cdot\left(\frac{\partial^2 Q}{\partial x^2}+\frac{\partial^2 Q}{\partial y^2}+\frac{\partial^2 Q}{\partial z^2}\right)-\left(v_x\frac{\partial Q}{\partial x}+v_y\frac{\partial Q}{\partial y}+v_z\frac{\partial Q}{\partial z}\right), \tag{47}$$

where $\boldsymbol{v}=(v_x,v_y,v_z)$ represents the wind vector. The position coordinates $\boldsymbol{r}=(x,y,z)$ in $Q(\boldsymbol{r},t)$ can be transformed into $w(\boldsymbol{r})$ via Fourier Transformation theory, i.e., $Q(\boldsymbol{r},t)\to Q(w(\boldsymbol{r}),t)$, $w(\boldsymbol{r})=(w_x,w_y,w_z)$.

Then, equation (47) can be written as

$$\frac{\partial Q}{\partial t}=\left[-G\left(w_x^2+w_y^2+w_z^2\right)-j\left(w_xv_x+w_yv_y+w_zv_z\right)\right]Q, \tag{48}$$

where j is the imaginary unit.

Equation (48) can be integrated on both sides into

$$Q(w(\boldsymbol{r}),t)=Q(w(\boldsymbol{r}),t_0)\cdot \exp\left\{\left[-G\left(w_x^2+w_y^2+w_z^2\right)-j\left(w_xv_x+w_yv_y+w_zv_z\right)\right](t-t_0)\right\}. \tag{49}$$

Let us suppose that the pollutant emission of the IC q is located at $\boldsymbol{r}_0=(x_0,y_0,z_0)$, and the pollutant concentration

can be written in $Q(\boldsymbol{r},t_0)=q\delta(x-x_0)\delta(y-y_0)\delta(z-z_0)$. Then,

$$Q\big(w(\boldsymbol{r}),t_0\big)=q\exp\Big[-j\big(w_x x_0+w_y y_0+w_z z_0\big)\Big]. \quad (50)$$

To obtain $Q(\boldsymbol{r},t)$, two steps are required: 1) a Fourier inverse transformation for (49); 2) an integration over time t.

Step-1: The Fourier inverse transformation is applied in (49), and the result can be denoted as $\tilde{Q}(\boldsymbol{r},t)$, which can be calculated by

$$\tilde{Q}(\boldsymbol{r},t)=\frac{q}{\big(4\pi G(t-t_0)\big)^{3/2}}\exp\left(-\frac{\big|\boldsymbol{r}-\boldsymbol{v}(t-t_0)-\boldsymbol{r}_0\big|^2}{4G(t-t_0)}\right). \quad (51)$$

Step-2 : Perform an integration transformation for (51) over $t_0\to t$. The pollutant concentration $Q(\boldsymbol{r},t)$ can be expressed as

$$\begin{aligned}Q(\boldsymbol{r},t)=q\big(8\pi Gl\big)^{-1}&\exp\Big(\big(\big((\boldsymbol{r}-\boldsymbol{r}_0)\cdot\boldsymbol{v}\big)+l|\boldsymbol{v}|\big)\big/2G\Big)\\&\Big[\operatorname{erfc}\Big(l\big(G(t-t_0)\big)^{-1/2}\big/2+|\boldsymbol{v}|\big((t-t_0)/4G\big)^{1/2}\Big)\\&+\exp\Big(\big(\big((\boldsymbol{r}-\boldsymbol{r}_0)\cdot\boldsymbol{v}\big)+l|\boldsymbol{v}|\big)\big/2G\Big)\cdot\\&\operatorname{erfc}\Big(l\big(G(t-t_0)\big)^{-1/2}\big/2-|\boldsymbol{v}|\big((t-t_0)/4G\big)^{1/2}\Big)\Big]\end{aligned} \quad (52)$$

REFERENCES

[1] J. Hu, X. Liu, M. Shahidehpour, and S. Xia, "Optimal Operation of Energy Hubs With Large-Scale Distributed Energy Resources for Distribution Network Congestion Management," *IEEE Trans. Sustain. Energy*, vol. 12, no. 3, pp. 1755–1765, 2021, doi: 10.1109/TSTE.2021.3064375.

[2] J. Leithon, S. Sun, and T. J. Lim, "Demand Response and Renewable Energy Management Using Continuous-Time Optimization," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 991–1000, 2018, doi: 10.1109/TSTE.2017.2771359.

[3] F. Wang *et al.*, "Multi-Objective Optimization Model of Source–Load–Storage Synergetic Dispatch for a Building Energy Management System Based on TOU Price Demand Response," *IEEE Trans. Ind. Appl.*, vol. 54, no. 2, pp. 1017–1028, 2018, doi: 10.1109/TIA.2017.2781639.

[4] E. Mahboubi-Moghaddam, M. Nayeripour, J. Aghaei, A. Khodaei, and E. Waffenschmidt, "Interactive Robust Model for Energy Service Providers Integrating Demand Response Programs in Wholesale Markets," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 2681–2690, 2018, doi: 10.1109/TSG.2016.2615639.

[5] Y. J. Wu, X. Y. Liang, T. Huang, Z. W. Lin, Z. X. Li, and M. F. Hossain, "A hierarchical framework for renewable energy sources consumption promotion among microgrids through two-layer electricity prices," *Renew. Sustain. Energy Rev.*, vol. 145, p. 111140, 2021, doi: https://doi.org/10.1016/j.rser.2021.111140.

[6] S. Zheng *et al.*, "Incentive-Based Integrated Demand Response for Multiple Energy Carriers Considering Behavioral Coupling Effect of Consumers," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3231–3245, 2020, doi: 10.1109/TSG.2020.2977093.

[7] E. Shahryari, H. Shayeghi, B. Mohammadi-ivatloo, and M. Moradzadeh, "An improved incentive-based demand response program in day-ahead and intra-day electricity markets," *Energy*, vol. 155, pp. 205–214, 2018, doi: https://doi.org/10.1016/j.energy.2018.04.170.

[8] M. Yu, S. H. Hong, Y. Ding, and X. Ye, "An Incentive-Based Demand Response (DR) Model Considering Composited DR Resources," *IEEE Trans. Ind. Electron.*, vol. 66, no. 2, pp. 1488–1498, 2019, doi: 10.1109/TIE.2018.2826454.

[9] Y. Wu, Z. Lin, C. Liu, Y. Chen, and N. Uddin, "A Demand Response Trade Model Considering Cost and Benefit Allocation Game and Hydrogen to Electricity Conversion," *IEEE Trans. Ind. Appl.*, vol. 58, no. 2, pp. 2909–2920, 2022, doi: 10.1109/TIA.2021.3088769.

[10] H. Zhong, L. Xie, and Q. Xia, "Coupon Incentive-Based Demand Response: Theory and Case Study," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1266–1276, 2013, doi: 10.1109/TPWRS.2012.2218665.

[11] F. Wang *et al.*, "Smart Households' Aggregated Capacity Forecasting for Load Aggregators Under Incentive-Based Demand Response Programs," *IEEE Trans. Ind. Appl.*, vol. 56, no. 2, pp. 1086–1097, 2020, doi: 10.1109/TIA.2020.2966426.

[12] N. G. Paterakis, O. Erdin, A. G. Bakirtzis, and J. P. S. Catalo, "Qualification and Quantification of Reserves in Power Systems Under High Wind Generation Penetration Considering Demand Response," *IEEE Trans. Sustain. Energy*, vol. 6, no. 1, pp. 88–103, 2015, doi: 10.1109/TSTE.2014.2359688.

[13] T. Broeer, F. K. Tuffner, A. Franca, and N. Djilali, "A demand response system for wind power integration: greenhouse gas mitigation and reduction of generator cycling," *CSEE J. Power Energy Syst.*, vol. 4, no. 2, pp. 121–129, 2018, doi: 10.17775/CSEEJPES.2018.00500.

[14] B. Zeng, J. Zhang, X. Yang, J. Wang, J. Dong, and Y. Zhang, "Integrated Planning for Transition to Low-Carbon Distribution System With Renewable Energy Generation and Demand Response," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1153–1165, 2014, doi: 10.1109/TPWRS.2013.2291553.

[15] M. Fleschutz, M. Bohlayer, M. Braun, G. Henze, and M. D. Murphy, "The effect of price-based demand response on carbon emissions in European electricity markets: The importance of adequate carbon prices," *Appl. Energy*, vol. 295, p. 117040, 2021, doi: https://doi.org/10.1016/j.apenergy.2021.117040.

[16] R. Ma, K. Li, X. Li, and Z. Qin, "Economic and low-carbon day-ahead Pareto-optimal scheduling for wind farm integrated power systems with demand response," *J. Mod. Power Syst. Clean Energy*, vol. 3, no. 3, pp. 393–401, 2015, doi: 10.1007/s40565-014-0094-7.

[17] J. Lin, J. Dong, D. Liu, Y. Zhang, and T. Ma, "From peak shedding to low-carbon transitions: Customer psychological factors in demand response," *Energy*, vol. 238, p. 121667, 2022, doi: https://doi.org/10.1016/j.energy.2021.121667.

[18] O. D. Melgar-Dominguez, M. Pourakbari-Kasmaei, M. Lehtonen, and J. R. Sanches Mantovani, "An economic-environmental asset planning in electric distribution networks considering carbon emission trading and demand response," *Electr. Power Syst. Res.*, vol. 181, p. 106202, 2020, doi: https://doi.org/10.1016/j.epsr.2020.106202.

[19] P. Stoll, N. Brandt, and L. Nordström, "Including dynamic CO2 intensity with demand response," *Energy Policy*, vol. 65, pp. 490–500, 2014, doi: https://doi.org/10.1016/j.enpol.2013.10.044.

[20] M. Okubo, Y. Hiroyasu, and T. Kuroki, "Ion Cluster Formation by Nonthermal Plasma Induced by Pulse Corona Discharge Toward Indoor Air Cleaning," *IEEE Trans. Ind. Appl.*, vol. 56, no. 5, pp. 5480–5488, 2020, doi: 10.1109/TIA.2020.3010703.

[21] G. Hunt, J. J. Heiszwolf, and M. Sewell, "Enhanced Hydrated Lime—A Simple Solution for Acid Gas Compliance," *IEEE Trans. Ind. Appl.*, vol. 54, no. 1, pp. 796–807, 2018, doi: 10.1109/TIA.2017.2740389.

[22] N. Zhang, Z. Hu, D. Dai, S. Dang, M. Yao, and Y. Zhou, "Unit Commitment Model in Smart Grid Environment Considering Carbon Emissions Trading," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 420–427, 2016, doi: 10.1109/TSG.2015.2401337.

[23] J. Yang, N. Zhang, Y. Cheng, C. Kang, and Q. Xia, "Modeling the Operation Mechanism of Combined P2G and Gas-Fired Plant With CO_2 Recycling," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 1111–1121, 2019, doi: 10.1109/TSG.2018.2849619.

[24] Y. Chen, T. Yu, B. Yang, X. S. Zhang, and K. Qu, "Many-Objective Optimal Power Dispatch Strategy Incorporating Temporal and Spatial Distribution Control of Multiple Air Pollutants," *IEEE Trans. Ind. Informatics*, vol. 15, no. 9, pp. 5309–5319, 2019, doi: 10.1109/TII.2019.2896968.

[25] S. Lei, Y. Hou, X. Wang, and K. Liu, "Unit Commitment Incorporating Spatial Distribution Control of Air Pollutant Dispersion," *IEEE Trans. Ind. Informatics*, vol. 13, no. 3, pp. 995–1005, 2017, doi: 10.1109/TII.2016.2631572.

[26] X. Guan, S. Guo, and Q. Zhai, "The conditions for obtaining feasible solutions to security-constrained unit commitment problems," *IEEE Trans. Power Syst.*, vol. 20, no. 4, pp. 1746–1756, 2005, doi: 10.1109/TPWRS.2005.857399.

[27] H. Deng, K. Yang, Y. Liu, S. Zhang, and Y. Yao, "Actively Exploring Informative Data for Smart Modeling of Industrial Multiphase Flow Processes," *IEEE Trans. Ind. Informatics*, vol. PP, no. 99, p. 1, 2020.
[28] M. A. Melehy, "Diffusion forces and diffusion potentials," *Proc. IEEE*, vol. 51, no. 7, p. 1028, 1963, doi: 10.1109/PROC.1963.2384.
[29] Y. Jiang, G. Zhang, P. Chen, D. Li, X. Tang, and W. Huang, "Systematic Error Compensation Based on a Rational Function Model for Ziyuan1-02C," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3985–3995, 2015, doi: 10.1109/TGRS.2015.2388700.
[30] X. Chen, W. Wu, and B. Zhang, "Robust Restoration Method for Active Distribution Networks," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 4005–4015, 2016, doi: 10.1109/TPWRS.2015.2503426.
[31] S. Paternain, M. Calvo-Fullana, L. F. O. Chamon, and A. Ribeiro, "Safe Policies for Reinforcement Learning via Primal-Dual Methods," *IEEE Trans. Automat. Contr.*, vol. 68, no. 3, pp. 1321–1336, 2023, doi: 10.1109/TAC.2022.3152724.
[32] R. S. Sutton, "Learning to predict by the methods of temporal differences," *Mach. Learn.*, vol. 3, no. 1, pp. 9–44, 1988, doi: 10.1007/BF00115009.
[33] Y. Liang, Z. Ding, T. Ding, and W.-J. Lee, "Mobility-Aware Charging Scheduling for Shared On-Demand Electric Vehicle Fleet Using Deep Reinforcement Learning," *IEEE Trans. Smart Grid*, vol. 12, no. 2, pp. 1380–1393, 2021, doi: 10.1109/TSG.2020.3025082.
[34] L.-J. Lin, "Reinforcement Learning for Robots Using Neural Networks," Carnegie Mellon University, USA, 1992.
[35] Y. Wu *et al.*, "Resilience enhancement for urban distribution network via risk-based emergency response plan amendment for ice disasters," *Int. J. Electr. Power Energy Syst.*, vol. 141, p. 108183, 2022, doi: https://doi.org/10.1016/j.ijepes.2022.108183.
[36] I. and S. Production, "Electricity's Progress in the Iron and Steel Industry," *Trans. Am. Inst. Electr. Eng.*, vol. XLIV, pp. 693–696, 1925, doi: 10.1109/T-AIEE.1925.5061154.
[37] W. Zhang, H. Li, B. Chen, Q. Li, X. Hou, and H. Zhang, "CO2 emission and mitigation potential estimations of China's primary aluminum industry," *J. Clean. Prod.*, vol. 103, pp. 863–872, 2015, doi: https://doi.org/10.1016/j.jclepro.2014.07.066.
[38] L. Wang and K. Chen, "Performance Analysis of Induction Motors for Driving Coke-Transfer Cars of a Coke-Oven Plant in a Practical Iron-Making Factory," *IEEE Trans. Ind. Appl.*, vol. 50, no. 1, pp. 60–67, 2014, doi: 10.1109/TIA.2013.2270451.
[39] A. Müezzinoğlu, "Air pollutant emission potentials of cotton textile manufacturing industry," *J. Clean. Prod.*, vol. 6, no. 3, pp. 339–347, 1998, doi: https://doi.org/10.1016/S0959-6526(98)00013-4.
[40] A. Paszke *et al.*, "PyTorch: An Imperative Style, High-Performance Deep Learning Library," in *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, Red Hook, NY, USA: Curran Associates Inc., 2019.
[41] C. Gong, D. Tao, K. Fu, and J. Yang, "Fick's Law Assisted Propagation for Semisupervised Learning," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 26, no. 9, pp. 2148–2162, 2015, doi: 10.1109/TNNLS.2014.2376963.
[42] P. McNamara and S. McLoone, "Hierarchical Demand Response for Peak Minimization Using Dantzig–Wolfe Decomposition," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2807–2815, 2015, doi: 10.1109/TSG.2015.2467213.
[43] B. Mota, P. Faria, and Z. Vale, "Residential load shifting in demand response events for bill reduction using a genetic algorithm," *Energy*, vol. 260, p. 124978, 2022, doi: https://doi.org/10.1016/j.energy.2022.124978.

Yingjun Wu (M'13) received the B.S. degree from Nanchang University, Nanchang, China, in 2007, the M.S. degree from Southeast University, Nanjing, China, in 2009, and the Ph.D. degree from Politecnico di Torino, Torino, Italy, in 2013, all in electrical engineering.Since September 2013 to June 2018, he was an Assistant Professor with the College of Automation, Nanjing University of Posts and Telecommunications, Nanjing, China. He is currently an Associate Professor with the College of Energy and Electrical Engineering, Hohai University, Nanjing, China. His current research interests include active distribution grids, power system economics, and cyber-physical systems.

Zhiwei Lin (S'22) received a bachelor's degree in electrical engineering and the automatization from Xihua University, China, in 2019. He is currently a master student at Hohai University, China. His research interests include active distribution grids and power system economics.

Yijun Xu (SM'21) is a professor at Southeast University, Nanjing, China. He received his Ph.D. degree from the Bradley Department of Electrical and Computer Engineering at Virginia Tech, Falls Church, VA, in December 2018. He worked as a research assistant professor at Virginia Tech-Northern Virginia Center, Falls Church, VA, in 2021. He was a postdoc associate at the same institute from 2019 to 2020. He did a computation internship at Lawrence Livermore National Laboratory, Livermore, CA, and a power engineer internship at ETAP – Operation Technology, Inc., Irvine, California, in 2018 and 2015, respectively. His research interests include power system uncertainty quantification, uncertainty inversion, and decision-making under uncertainty. Dr. Xu is currently serving as an Associate Editor of the IET GENERATION, TRANSMISSION & DISTRIBUTION and an Associate Editor of the IET RENEWABLE POWER GENERATION. He is the co-chair of the IEEE Task Force on Power System Uncertainty Quantification and Uncertainty Aware Decision-Making.

Gianfranco CHICCO (M'98, SM'08, F'18) holds a Ph.D. in Electrotechnics Engineering and is a Full Professor of Electrical Energy Systems at Politecnico di Torino (POLITO), Italy. He received the title of Doctor Honoris Causa from the Universities Politehnica of Bucharest and "Gheorghe Asachi" of Iasi (Romania) in 2017 and 2018, respectively. He is the Chair of the IEEE R8 Italy Section. He is the Scientific Responsible of the research group on Power and Energy Systems at POLITO, and the Responsible of the Torino unit of the Italian Consortium ENSIEL. He is the Editor-in-Chief of Sustainable Energy Grids and Networks. His research topics include Power System Analysis, Distribution System Analysis and Optimization, Electrical Load Management, Energy Efficiency and Environmental Impact of Multi-Energy Systems, Data Analytics Applied to Power and Energy Systems, and Power Quality.

Tao Huang (M' 18, SM'23) received his Ph.D. degree from Politecnico di Torino, Turin, Italy in 2011. He is currently an associate professor with the Department of Energy, Politecnico di Torino, Italy. He is also adjunct professor with Xihua University and Polytechnic University of Henan, China. His research interests include vulnerability assessment, electricity markets, smart grids, active distribution network, and artificial intelligence in power systems.

Junjie Shao (S'23) received a bachelor's degree in School of Mechanical Engineering and Rail Transit from Changzhou University, China, in 2021. He is currently a master student at Hohai University, China. His research interests include renewable energy technology, frequency control, and the stability.

Zhaorui Chen (S'23) received a bachelor's degree in electrical power system and its automatization from Nanjing Institute of Technology, China, in 2022. She is currently a master student at Hohai University, China. Her research interests include active distribution grids and power system economics.