Spatio-temporal wind field reconstruction method based on deep reinforcement learning algorithm

Z. Luo [a], Q. Yang [a] , K. Li [a,b,*]

[ab] School of Civil Engineering, Chongqing University, Chongqing 400045, China
[b] Key Laboratory of New Technology for Construction of Cities in Mountain Area (Chongqing University), Ministry of Education, Chongqing 400045, Chinamail address

SUMMARY:
With the gradual improvement of the research and development tasks in the field of wind power engineering, higher requirements are also put forward for the pneumatic test. Due to people's understanding of turbulence is not mature enough, the existing passive wind tunnels and multi-fan active wind tunnels cannot reproduce the wind speed samples by using the traditional PID control method, and can only achieve frequency domain reproduction. As an innovation, this paper combines the deep learning algorithm with the traditional flow field control problem to carry out a preliminary study on the turbulence time history reconstruction. Specifically, the DDPG algorithm is used to build the recursive model of the flow field, and then the CFD program is used to build the numerical flow field to extract and process the training data. Finally, a closed-loop control system of target wind speed and control method is constructed on 1 and 2 dimensions, and the error obtained can be controlled within 2.2% and 10% respectively. The results show that the DRL method has a good application prospect in the reconstruction of spatio-temporal wind field.

Keywords: Target tracking; deep reinforcement learning; spatio-temporal change wind field

1. INTRODUCTION

Wind tunnel is the most effective and practical aerodynamic test tool at present. It can generate and regulate air flow in its tubular space, so as to study the aerodynamic effects of gas flowing through objects. The multi-fan active control wind tunnel can control the wind speed of the fan at the inlet of the flow field to change the internal operation law of the flow field and reach the artificial standard. With the deepening of wind engineering research, higher and higher requirements are put forward for the aerodynamic test, as well as for the control accuracy and effect.

Teunissen first obtained two-dimensional turbulence with almost any desired average velocity distribution by actively controlling multiple jets that can independently control the jet velocity. Cao and others have successfully reproduced the non-stationary fluctuation of wind speed in the multi-fan wind tunnel. Wang introduced more control variables to study the wind speed direction of a single fan and tried to further explore the simulation potential of multi-fan wind tunnels.

Although most of the existing methods provide effective methods to simulate the flow field, the existing multi-fan wind tunnels can only match the wind frequency within a period of time due to the complex hydrodynamic relationship and nonlinear interaction between the fluids. This paper attempts to explore a higher goal, that is, to achieve the matching turbulence of time series and the recurrence of time-history wind. In order to achieve this goal, this paper uses the deep reinforcement learning algorithm to build a recursive model of the flow field, and then uses the CFD method to build a numerical flow field to extract and process the training data. Finally, satisfactory results are obtained in the multi-dimensional numerical flow field simulation, and the future research directions are discussed.

2. DESIGN OF SPATIO-TEMPORAL WIND FIELD RECONSTRUCTION MODEL BASED ON DEEP REINFORCEMENT LEARNING

As a powerful function fitting tool, neural network can be used to explore the relationship between target wind speed and control method. To achieve closed-loop control, the interaction between the model and the environment will occur many times during the training process of the model, which involves a large number of network parameters and training data. It is not feasible to simply manually adjust the neural network. Therefore, reinforcement learning is used as an automatic method for training neural network parameters.

Figure 1 shows the interaction between the environment and agents in the time-history wind reproduction problem based on reinforcement learning. At time t, the numerical flow field is considered as the environment , and the wind speed U_{fan} of the fan at the entrance is regarded as the action a_t. As the state s_t, the flow field information includes the wind speed U_t^{obs} at some observation points in the flow field and the target wind speed U_{t+1}^{target} at the next time at the target location.

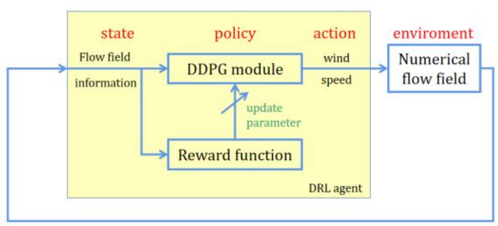

Figure 1. Flow field control framework based on deep reinforcement learning.

When the environment sends the state to the agent, the agent will update the parameters of the neural network to obtain a more accurate control method. Then the agent will give the output of the neural network under the current parameters as the wind speed of the fan, and the action will affect the environment. The return $r_t = -\sum \left|U_t^{obs} - U_t^{target}\right|$ will also be recorded. This is the key scalar to guide the learning process. During the above cycle, the parameters are corrected. Data is stored in the form of (s_t, a_t, r_t, s_{t+1}).

In this paper, DDPG algorithm is used to train neural network parameters to find closed-loop control strategy. The DDPG algorithm is composed of four neural networks, namely, strategy network $A(s|\theta^A)$, target strategy network $A'(s|\theta^{A'})$, value network $C(s,a|\theta^C)$ and target value network $C'(s,a|\theta^{C'})$, where θ is the parameter of each network and is distinguished by superscript.The parameters update of target network depends on the parameters update of strategy network and value network.For any segment of trajectory $(s_t^i, a_t^i, r_t^i, s_{t+1}^i)$, the target network is used to guide the calculation of value y_i, which is an important parameter to guide parameter update.

$$y_i = r_t^i + \gamma C'(s_{t+1}^i, A'(s_{t+1}^i|\theta^{A'})|\theta^{C'}) \tag{1}$$

In this way, when N trajectories $(s_t^i, a_t^i, r_t^i, s_{t+1}^i)$ are extracted, the parameters of value network can be updated by the gradient descent method of learning rate η^C.The parameters of the strategy network can be updated by the gradient rise method of the learning rate η^A.

$$\theta^C = \theta^C - \eta^C \frac{1}{N}\sum_{i=1}^{N} \nabla_{\theta^C}[y_i - C(s_t^i, a_t^i|\theta^C)]^2 \tag{2}$$

$$\theta^A = \theta^A + \eta^A \frac{1}{N}\sum_{i=1}^{N} \nabla_{\theta^A} C[s_t^i, A(s_t^i|\theta^A)|\theta^C]^2 \tag{3}$$

With the increase of r, the fitting ability of the model is also continuously improved, so as to realize the spatio-temporal wind field reconstruction.

3. SPATIO-TEMPORAL VARIATION IN NUMERICAL FLOW FIELD

In order to test the predictive effectiveness of the agent model, the closed-loop control is used to generate the control method and the control sample of the target area speed, and then the model is trained and optimized in the way of open-loop control. The training of model parameters is based on the Tensorflow library in Python environment.

In one-dimensional model, we use completely random control method to verify the fitting performance of the model in any case. Figure 3 shows the tracking of the new time series wind speed curve by the trained model, and the error of the model can be controlled within 2.2%.

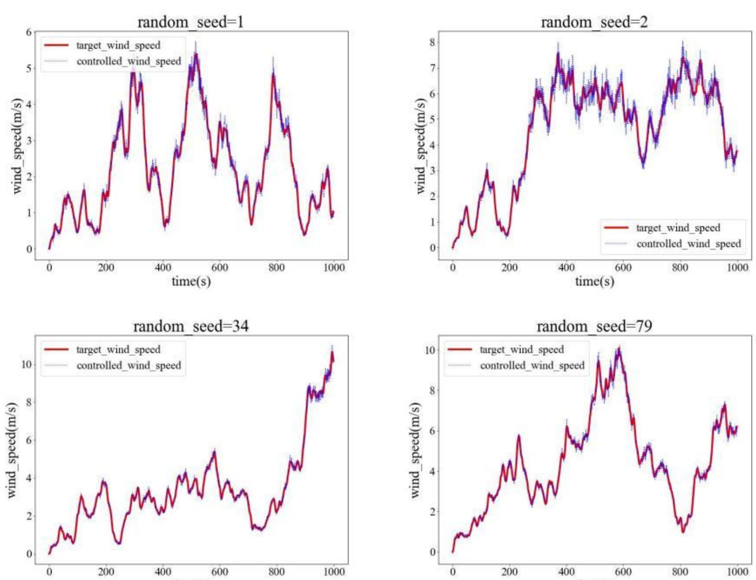

Figure 2. Fitting effect of one-dimensional spatio-temporal variation wind field.

In the two-dimensional model, we use the longitudinal arrangement of four fans as the flow controller. The increase of dimension is a challenge to the training of the model, but the parameters can still converge well in enough time. Figure 4 shows the simulation effect of four fans in a two-dimensional case.

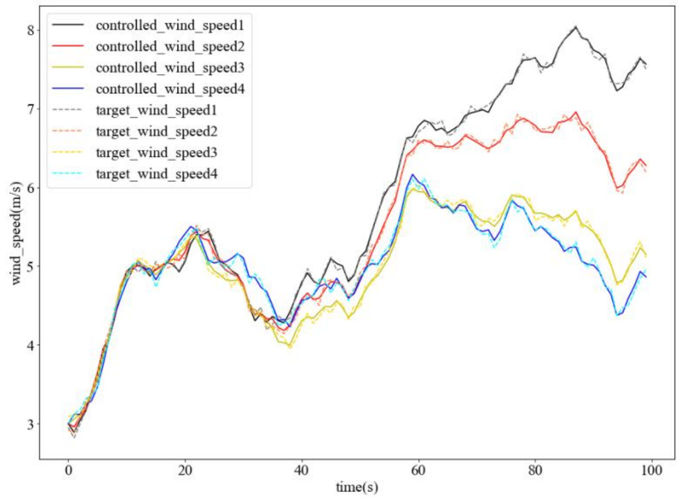

Figure 3. Fitting effect of two-dimensional spatio-temporal variation wind field

The statistical results show that the average relative error of the model in reproducing the wind speed time history generated from the training data is less than 5%, while the error in the target wind speed time history that has not been learned can be controlled within 10% (average 10 groups of test data)

4. CONCLUSIONS

In this paper, an accurate and effective control scheme based on depth RL is proposed to actively simulate the transient wind field in the multi-fan flow field. Numerical examples in one-dimensional and two-dimensional environments show that the proposed control scheme based on DRL performs well in reproducing the spatio-temporal wind field. It makes theoretical preparation for the construction of time domain reappearance wind tunnel in the future.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the supports from Key Laboratory for Wind Resistance Technology of Bridges (Shanghai),Ministry of Transport, PRC(KLWRTBMC22-01).

REFERENCES

Wang Jia-Ying et al. A multiple-fan active control wind tunnel for outdoor wind speed and direction simulation.[J]. The Review of scientific instruments, 2018, 89(3) : 035108.

Teunissen H W . Simulation of the planetary boundary layer in a multiple-jet wind tunnel[J]. Atmospheric Environment, 1975, 9(2):145-174.

Timothy P. Lillicrap et al. Continuous control with deep reinforcement learning.[J]. CoRR, 2015, abs/1509.02971.

Y. LeCun et al. Backpropagation Applied to Handwritten Zip Code Recognition[J]. Neural Computation, 1989, 1(4) : 541-551.

Li Shaopeng and Snaiki Reda and Wu Teng. Active Simulation of Transient Wind Field in a Multiple-Fan Wind Tunnel via Deep Reinforcement Learning[J]. J. Eng. Mech, 2021, 147(9): 04021056.