Memetic Reinforcement Learning based Maximum Power Point Tracking Design for PV Systems under Partial Shading Condition

Xiaoshun Zhang [1], Shengnan Li[2], Tingyi He[2], Bo Yang [3,*], Tao Yu [4], Haofei Li [4], Lin Jiang [5], Liming Sun[6]

[1] College of Engineering, Shantou University, 515063 Shantou, China;
[2] Electric Power Research Institute of Yunnan Power Grid Co., Ltd., Kunming 650217, China;
[3] Faculty of Electric Power Engineering, Kunming University of Science and Technology, 650500 Kunming, China;
[4] College of Electric Power, South China University of Technology, 510640 Guangzhou, China;
[5] Department of Electrical Engineering & Electronics, University of Liverpool, Liverpool, L69 3GJ, United Kingdom;
[6] Guangzhou Shuimuqinghua Technology Co. Ltd., 510898 Guangzhou, China
* Correspondence: yangbo_ac@outlook.com, Tel.: +86-183-145-96103

Abstract: Solar energy has attracted significant attentions around the globe, while one of its most crucial task is to harvest the maximum available solar power under different weather conditions, also known as maximum power point tracking (MPPT). This paper proposes a novel memetic reinforcement learning (MRL) based MPPT scheme for photovoltaic (PV) systems under partial shading condition (PSC). In order to enhance the searching ability of MRL, the memetic computing structure is incorporated into reinforcement learning (RL). In particular, a virtual population is used for the global information exchange between different agents, such that the learning rate can be dramatically accelerated. Besides, a RL based local search is designed in each memeplex, which can effectively improve the optimum quality. Comprehensive case studies are undertaken, such as start-up test, step change of solar irradiation, ramp change of solar irradiation and temperature, and field atmospheric data of Hong Kong. The PV system responses are then evaluated and compared to that of seven typical MPPT algorithms.

Keywords: Solar energy; MPPT; partial shading condition; memetic reinforcement learning; virtual population

Nomenclature

Variables		***Abbreviations***	
V_{pv}	PV output voltage	**MPPT**	maximum power point tracking
I_{pv}	PV output current	**PV**	Photovoltaic
I_g	cell's photocurrent	**PSC**	partial shading condition
I_D	diode's photocurrent	**INC**	incremental conductance
I_S	cell's reverse saturation current	**MRL**	memetic reinforcement learning
I_{RS}	d-q components of the grid current	**GA**	genetic algorithm
T_c	cell's absolute working temperature, K	**PSO**	particle swarm optimization
T_{ref}	cell's reference temperature, K	**ABC**	artificial bees colony
S	total solar irradiation, W/m^2	**CSA**	Cuckoo search algorithm
E_g	bang-gap energy of the semiconductor used in the cell	**GWO**	grey wolf optimizer
N_p	number of panels connected in parallel	**TLBO**	teaching-learning-based optimization
N_s	number of panels connected in series	**GMPP**	global maximum power point
		LMPP	local maximum power point
PV system parameters		***The MRL parameters***	
q	electron charge, 1.60217733×10^{-19} Cb	n	number of memeplexes
A	p-n junction ideality factor, between 1 and 5	P_s	population size of each memeplex
k	Boltzman's constant, 1.380658×10^{-23} J/K	α	learning factor
k_i	cell's short-circuit current temperature coefficient	γ	discount factor
R_s	cell's series resistance	k_{max}	maximum iteration number
R_p	cell's parallel resistance	L	real-coded length

1. Introduction

Nowadays, the malignant drawbacks of fossil fuel energy, e.g., considerable greenhouse gases emission, water and air pollution, human health damage, have urgently driven the worldwide deployment of renewable energy [1-4]. Among various types of renewable energy, solar energy is one of most widely used technology while photovoltaic (PV) system installation has been dramatically increased thanks to the cost reduction of polysilicon in the past decade. Meanwhile, it is a mature and reliable technology with many elegant merits, including easy installation,

high safety, pollution-free, little maintenance and noise-free [5,6].

One of the most crucial task of PV system operation is to harvest the maximum available solar energy under different weather conditions, which is well known as maximum power point tracking (MPPT) [7]. Generally speaking, each PV cell has a limited efficiency when converting the solar energy into electricity. Moreover, the PV cell exhibits a nonlinear power-voltage (*P-V*) and current-voltage (*I-V*) characteristics, which leads to a single optimal operation point corresponding to the maximum power/efficiency. Particularly, the *P-V* and *I-V* curve will become quite complex and may exhibit multiple local maxima when several PV cells are connected in series and parallel in the presence of partial shading condition (PSC) [8-10].

Basically, MPPT algorithms attempt to maintain the operation point at the global maxima. Conventional MPPT techniques, e.g., incremental conductance (INC) [11], hill-climbing [12], perturb & observe (P&O) [13], have been widely adopted in practice thanks to their high reliability and structure simplicity. They can usually achieve an efficient MPPT under uniform solar irradiation scenario. However, a severe power oscillation may emerge when these algorithms converge closely to the maximum power point (MPP) due to the use of fixed step size. To handle this thorny obstacle, many adaptive/variable step size based schemes have been reported in [14,15]. These methods can dynamically increase the step size to improve the tracking efficiency when the MPP is still far while decrease the step size to reduce the oscillation when it approaches the MPP. Additionally, another inherent disadvantage of such strategies is the fact that they may be readily trapped at a local maximum power point (LMPP) in the presence of multiple MPPs, which inevitably leads to a low energy conversion efficiency [16].

In practice, PSC often occurs where the solar irradiation is unequally distributed among the PV module resulted from the shadows of buildings, trees, clouds, birds, dirt, etc. [17], by which there may exist multiple LMPPs and one single global MPP (GMPP). Normally, the aforementioned MPPT methods are inadequate to seek the GMPP under PSC. As a consequence, plenty of meta-heuristic algorithms have been proposed to effectively and efficiently seek the GMPP [18]. In reference [19], a P&O algorithm was integrated inside the genetic algorithm (GA) function and creates a single algorithm, such that the population size and the number of iterations are decreased. Thus, the MPP can be found in a shorter time. Moreover, a novel overall distribution MPPT algorithm was proposed to rapidly search the area near the GMPPs, which is further integrated with the particle swarm optimization (PSO) to improve the accuracy of MPPT [20]. Besides, literature [21] adopted artificial bee colony (ABC) algorithm to achieve MPPT under various weather conditions and PSC, which uses few parameters and its convergence is independent of the initial conditions. Meanwhile, work [22] applied Cuckoo search algorithm (CSA) via Levy flight to guarantee an efficient MPPT under PSC. Additionally, reference [23] presented a flower pollination algorithm (FPA) to mitigate PSC in buildings integrated with PV power systems, such that the solar energy conversion efficiency can be improved. Moreover, a human society inspired algorithm called teaching-learning-based optimization (TLBO) was employed to accurately track the GMPP under PSC, which structure is very simple with a fast convergence rate [24]. Meanwhile, animal behaviour based whale optimization algorithm (WOA) was utilized to achieve a rapid and oscillation-free MPPT under PSC, which just requires few parameters associated with a relatively low computational burden [25]. Furthermore, grey wolf optimizer (GWO) was used to resolve PSC, which owns a fast identification speed of GMPP [26]. In work [27], a dynamic leader based collective intelligence (DLCI) algorithm was proposed, which has a high accuracy in the searching of GMPP and can provide good dynamic performance and very quick convergence rate by dynamically switching between various sub-optimizer during the MPPT process. Further, memetic salp swarm algorithm (MSSA) algorithm was developed for PV systems affected by PSC, which owns the merits of fewer control parameters and independent convergence with the initial operation conditions [28].

However, the aforementioned meta-heuristic algorithms have two major drawbacks for MPPT of PV systems, as follows:

- *High convergence randomness*: under the same weather condition, the meta-heuristic algorithms may easily converge to different optimal solutions in different runs due to their inherent feature of random searching mechanisms. Hence, it might readily result in a larger power fluctuation of PV systems;
- *Difficulty in balancing the optimum quality and computation time*: in order to find GMPP among multiple LMPPs, the meta-heuristic algorithms generally require long computation time with a larger population size and more iterations in order to obtain a high-quality optimum. Nevertheless, the control cycle of MPPT is ultra-short, thus they have to reduce the population size and maximum iterations, which in turn often causes a low-quality LMPP.

Over the last few decades, reinforcement learning (RL) [29] has attracted an extensive study and a wide successful application thanks to its superior online learning ability. From the results reported in work [30]-[31], it has been revealed that RL has a higher convergence stability and a shorter computation time compared with that of traditional meta-heuristic algorithms for an optimization or control problem. As a consequent, RL is adequate to be a powerful tool for MPPT of PV systems under PSC. However, the conventional RL only employs a single agent

for knowledge learning, which may easily lead to a low learning efficiency. Therefore, this paper proposes a novel memetic reinforcement learning (MRL) for handling the MPPT of PV systems, which has the following two major improvements compared with the conventional RL, as follows:

1) A memetic computing framework [32] is introduced in MRL, which can achieve a more effective balance between the global search and local search. Hence, it can rapidly search a high-quality optimum to approximate GMPP more closely.

2) MRL not only adopts multiple groups of agents (i.e., multiple memeplexes) for an independent local search, but also employs a virtual population for an effective global information exchange between different agents. Hence, it can effectively improve the convergence stability with increasing diversity of searching agents [32], which will lead to a smaller power fluctuation of PV systems.

The remaining of this paper is organized as follows: Section 2 develops the PV systems model under PSC; In Section 3, MRL is described while its application on PV systems for MPPT under PSC is presented in Section 4; Furthermore, comprehensive case studies are undertaken in Section 5. Then, Section 6 provides the dSpace based HIL experiment results. At last, Section 7 summarizes the main findings/contributions of the paper.

2. PV Systems Modelling under PSC

2.1. PV cell modelling

A PV cell is generally a p-n semiconductor junction diode, by which the solar energy is converted into the electrical power. Figure 1 illustrates its equivalent circuit, which includes a light generated current source, a parallel diode, and a series resistance, respectively. Normally, PV cells are grouped together to form PV modules, which are combined in both series and parallel to provide a desired output power [33]. Denote the number of PV cells in series and in parallel to be N_s and N_p, respectively, the relationship between the output current and voltage can be described by [34,35]

$$I_{pv} = N_p I_g - N_p I_s \left(\exp\left[\frac{q}{AKT_c} \left(\frac{V_{pv}}{N_s} + \frac{R_s I_{pv}}{N_p} \right) \right] - 1 \right) \tag{1}$$

where the meaning of each symbol is given in Nomenclature.

The generated photocurrent I_g is determined by the solar irradiation, as follows

$$I_g = \left(I_{sc} + k_i (T_c - T_{ref}) \right) \frac{s}{1000} \tag{2}$$

Besides, the PV cell's saturation current I_s varies with the temperature according to the following relationship:

$$I_s = I_{RS} \left[\frac{T_c}{T_{ref}} \right]^3 \exp\left[\frac{qE_g}{Ak} \left(\frac{1}{T_{ref}} - \frac{1}{T_c} \right) \right] \tag{3}$$

The above equations (1)-(3) indicate that the current generated by the PV array is simultaneously relied on the solar irradiation and temperature.

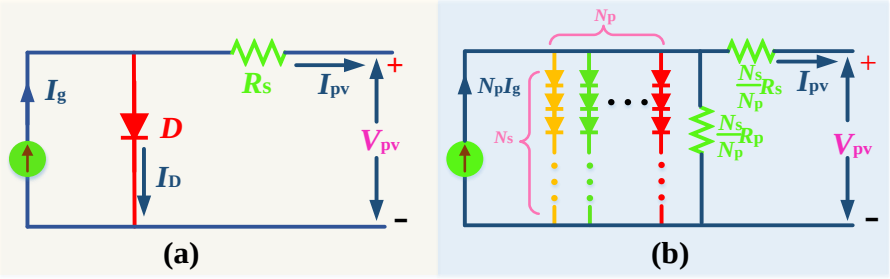


Figure 1. Equivalent circuit of PV cell/array circuit. (a) Single cell circuit; (b) PV array circuit with N_s in series and N_p in parallel.

2.2 PSC effect

Figure 2(a) illustrates a PV array in a typical series-parallel configuration. Here, the PV modules are connected in strings, with four modules per string. When one of the modules in the string exposes less illumination due to shading, e.g., dirt, trees, birds, its voltage will drop and behave as a load rather than a generator. Then, a hot spot will be emerged while generally a bypass diode is connected in parallel with each PV module to protect the shaded module from potential damage. Additionally, a blocking diode is also connected at the end of each string to provide the protection against reverse current resulted from the voltage mismatch between the parallel-connected strings [36].

Although the diodes are able to save the shaded PV cells from the above issue, they inevitably change the PV characteristics and produce a two-peak curve. Such phenomenon becomes quite complicated when several strings are connected in parallel to supply a higher current. Basically, various PV curves are generated by each string depending on the number of shaded PV cells [37]. These multi-peak PV curves are then combined due to the parallel connection, which then forms a multi-peak curve described by Fig. 2(b). Hence, the PV systems should

always operate at the GMPP in order to extract the maximum available solar energy from the PV array. Otherwise, considerable amount of power might be lost when operating at a LMPP.

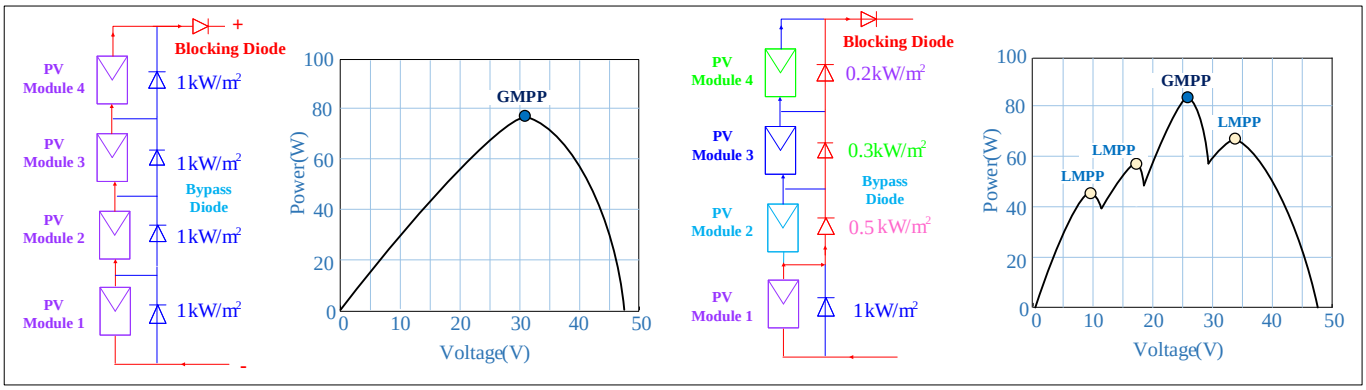


Figure 2. PSC effect. (a) *P-V* curve under uniform solar irradiation and temperature and (b) *P-V* curve under PSC.

3 Memetic Reinforcement Learning

3.1 Basic concept of memetic computing

According to the modern definition in reference [38], memetic computing is a broad subject which studies complex structures composed of interactive memes, which will effectively solve a specific optimization via continuous interactions and evolutions. Particularly, the main physical significance of the memetic algorithm is that both of searching efficiency and population diversity can be effectively improved via an efficient cooperation the population-based global search and multiple memeplexes based local search [32]. Under such framework, MRL is designed as a hybrid global-local heuristic searching method, in which the detailed definitions of the main terms are given as follows [39]:

- *Memetic*: which is originated from *meme* that is generally a unit of cultural evolution, and is analogous to the gene in GA.
- *Virtual population*: the traditional meta-heuristic algorithms are highly depended on the concept of population, which can evaluate the goodness of an individual by calculating its fitness value. In contrast, the population of MRL is only regarded as hosts of memes, in which the physical characteristics of each individual will not be changed. Hence, all the learning agents are taken as a virtual population.
- *Memeplex*: the virtual population will be divided into multiple parallel groups of agents, in which each group is called a memeplex.

Note that each memeplex is responsible for an independent local search, then the global information exchange can be implemented for the virtual population.

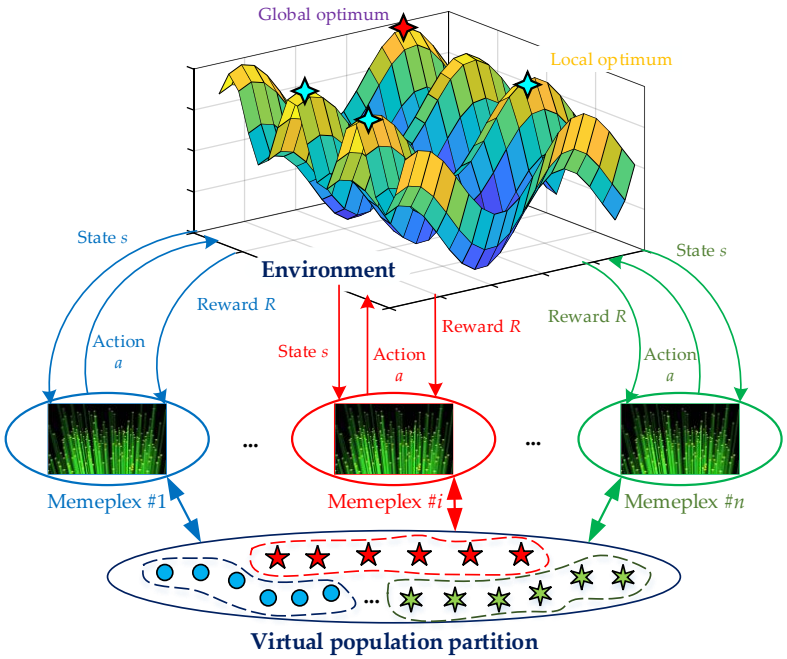


Figure 3. Optimization principle of MRL.

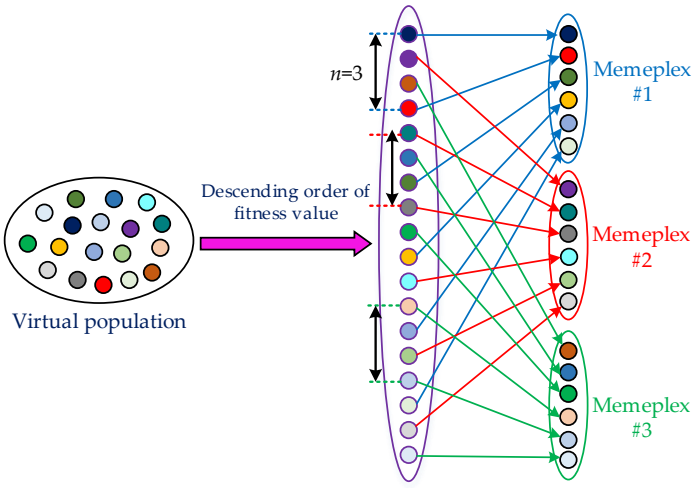


Figure 4. Partition principle of virtual population based on global information exchange.

3.2 RL based local search

3.2.1 Knowledge learning

As one of most frequently used RL, Q-learning is adopted for the local search of MRL. In general, Q-learning [40] only employs a single agent to interact with the environment, which can update its knowledge according to the feedback reward after making an action at the current state (See Fig. 3). Hence, it easily leads to a low learning efficiency. To handle this issue, a cooperative group of agents are employed to update the knowledge matrix of Q-learning [41], while the real-coded associative memory is used to decompose the original large-scale knowledge matrix into multiple small-scale knowledge matrices. Following the learning rule of Q-learning, the knowledge matrices of each memeplex can be updated as follows:

$$\begin{cases} \boldsymbol{Q}_{ij}^{l,k+1}\left(s_{ij}^{lp,k}, a_{ij}^{lp,k}\right) = \boldsymbol{Q}_{ij}^{l,k}\left(s_{ij}^{lp,k}, a_{ij}^{lp,k}\right) + \alpha\Delta\boldsymbol{Q}_{ij}^{l,k} \\ \Delta\boldsymbol{Q}_{ij}^{l,k} = R_{ij}^{lp,k}\left(s_{ij}^{lp,k}, a_{ij}^{lp,k}\right) + \gamma \max\limits_{a_{ij}^{l} \in \boldsymbol{A}_{ij}^{l}} \boldsymbol{Q}_{ij}^{l,k}\left(s_{ij}^{lp,k+1}, a_{ij}^{l}\right) - \boldsymbol{Q}_{ij}^{l,k}\left(s_{ij}^{lp,k}, a_{ij}^{lp,k}\right) \\ i = 1,2,\cdots,n; j = 1,2,\cdots,J; l = 1,2,\cdots,L; p = 1,2,\cdots,P_{\mathrm{s}} \end{cases} \tag{4}$$

where the subscripts i and j represent the ith memeplex and the jth controllable variable, respectively; the superscripts l, p, and k represent the lth real code, the pth agent, and the kth iteration, respectively; $\boldsymbol{Q}_{ij}^{l,k}$ is the knowledge matrix; $\Delta\boldsymbol{Q}_{ij}^{l,k}$ is the knowledge increment; $s_{ij}^{lp,k}$ is the state explored by the pth agent at the kth iteration; $a_{ij}^{lp,k}$ is the action explored by the pth agent at the kth iteration; α is the knowledge learning factor; γ is the discount factor; $R_{ij}^{lp,k}$ is the feedback reward obtained by the pth agent; $\boldsymbol{A}_{ij}^{l}$ is the action set of the lth real code; n is the number of memeplexes; J is the number of controllable variable; L is the real-coded length; and P_{s} is the population size.

3.2.2 Exploration and exploitation

In each memeplex, each agent will select an action for obtaining a higher quality solution according to the current knowledge matrices. During this process, it needs to balance the exploration and exploitation. In this paper, the ε-greedy rule [42] is used for selecting the action, as follows:

$$a_{ij}^{lp,k} = \begin{cases} \arg\max\limits_{a_{ij}^{l} \in \boldsymbol{A}_{ij}^{l}} \boldsymbol{Q}_{ij}^{l,k}\left(s_{ij}^{lp,k}, a_{ij}^{l}\right), & \text{if } q_0 < \varepsilon \\ a_{\mathrm{rand}}, & \text{otherwise} \end{cases} \tag{5}$$

where q_0 is a random value, which is uniformly distributed from 0 to 1; ε is the exploitation weight, i.e., the probability of choosing a greedy action; and a_{rand} is a random action selected from the action set, respectively.

3.3 Global information exchange

In order to avoid trapping at a low-quality optimum, the virtual population will be regrouped into different memeplexes according to all the agents' goodness. In this paper, the population partition rule of shuffled frog leaping algorithm (SFLA) [43] is employed for the global information exchange in MRL since SFLA is a classical memetic algorithm. Hence, the partition process can be achieved according to the descending order of fitness value for a maximum optimization. As illustrated in Fig. 4, the best solution will be assigned to memeplex #1, then the second best solution will be assigned to memeplex #2, and so on. Therefore, the re-organizing solution set of the ith memeplex can be described as follows [43]:

$$\boldsymbol{Y}^i = \left[\boldsymbol{x}_p, f_p \middle| \boldsymbol{x}_p = \boldsymbol{x}_{i+n(y-1)}, f_p = f_{i+n(y-1)},\ y = 1,2,\cdots,P_s\right],\ \ i = 1,2,\cdots,n \tag{6}$$

where $\boldsymbol{x}_p$ and f_p are the pth agent's solution and fitness value, respectively, which are determined by the descending order of the fitness value, i.e., $\boldsymbol{x}_1$ denotes the current best solution, p=1, 2, …, nP_s.

4 MRL based MPPT Design of PV systems

As a meta-heuristic algorithm, MRL is also highly independent on the specific mathematical model of PV systems under PSC, e.g., the PV systems can be regarded as a black box for MRL. More specifically, MRL can be implemented for searching GMPP of PV systems only with the output voltage and current under different duty cycles of PWM, as shown in Fig. 5. In fact, the output voltage and current can be easily measured in practice by voltage/current meters, while the duty cycle variation of PWM is a practical method to control the voltage transfer of a boost DC-DC converter for the PV systems. Therefore, MRL has a high practical utility for MPPT of PV systems under PSC.

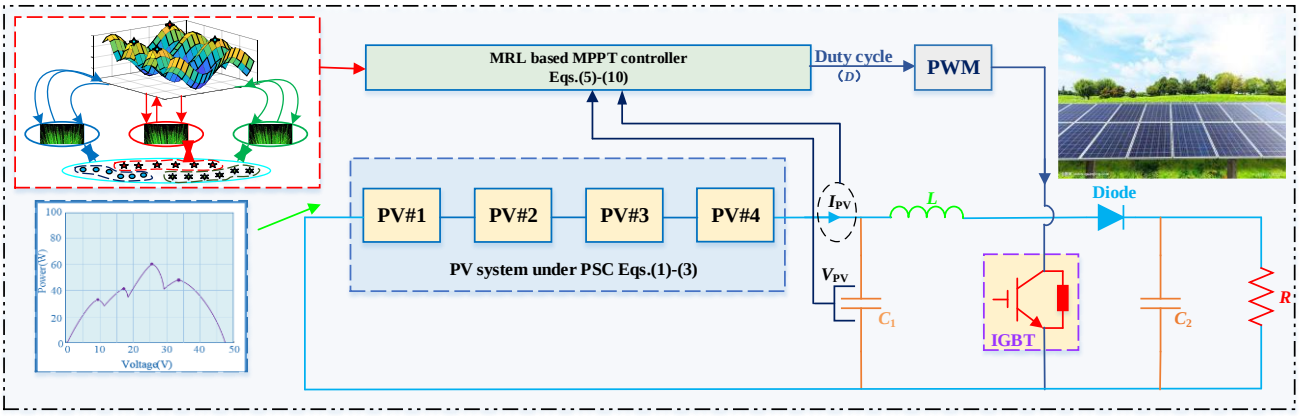


Figure 5. Overall structure of MRL based MPPT of PV system under PSC.

4.1 Optimization model for MPPT under PSC

As shown in Fig. 5, the MPPT can be achieved by regulating the PV system output voltage V_{pv} under PSC. Hence, the output voltage V_{pv} can be regarded as the optimization variable, while it should be limited within its lower and upper bounds. By taking the maximization of output active power as the objective function, the optimization model for MPPT under PSC can be written by

$$\max f\left(V_{\mathrm{pv}}\right) = P_{\mathrm{out}}\left(V_{\mathrm{pv}}\right) = V_{\mathrm{pv}} * I_{\mathrm{pv}}\left(V_{\mathrm{pv}}\right) \tag{7}$$

$$\text{s.t. } V_{\mathrm{pv}}^{\min} \le V_{\mathrm{pv}} \le V_{\mathrm{pv}}^{\max} \tag{8}$$

where P_{out} is the output active power of PV systems; $V_{\mathrm{pv}}^{\min}$ and $V_{\mathrm{pv}}^{\max}$ are the lower and upper bounds of the output voltages, respectively.

4.2 Transformation from actions to solutions

By considering the lower and upper bounds limits in (8), the agent's solution can be transformed based on the selected actions, as follows:

$$\boldsymbol{x}_{ip} = V_{\mathrm{pv}}^{\min} + \frac{\sum_{l=1}^{L} 10^{l-1} \times \left(a_{ij}^{lp} - 1\right)}{10^{L} - 1} \times \left(V_{\mathrm{pv}}^{\max} - V_{\mathrm{pv}}^{\min}\right),\quad i = 1,2,\cdots,n; p = 1,2,\cdots,P_s; j = 1 \tag{9}$$

where $\boldsymbol{x}_{ip}$ denotes the solution obtained by the pth agent in the ith memeplex.

4.3 Feedback reward

Generally speaking, an action with a larger fitness value will obtain a higher feedback reward for a maximization optimization. In order to accelerate the learning rate, the feedback reward is designed with the cooperative mechanism of ant colony [44], which can be described as

$$R_{ij}^{lp,k}\left(s_{ij}^{lp,k}, a_{ij}^{lp,k}\right) = \begin{cases} \max\limits_{p=1,2,\cdots,P_s} f_p, & \text{if } \left(s_{ij}^{lp,k}, a_{ij}^{lp,k}\right) \in \boldsymbol{SA}_{ij}^{\text{best}} \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

where represents the best agent's station-action pair set of the jth controllable variable in the ith memeplex.

When the weather condition varies, the P-V and I-V output characteristics of the PV systems will be changed automatically as well. Therefore, the feedback rewards in (10) of different control strategies will be directly influenced, then the proposed algorithm will rapidly search a high-quality optimum to approximate the new GMPP of PV systems according to the updated feedback rewards. Hence, the proposed MRL can effectively handle the weather variability during different times of the day as well as in different seasons.

4.4 Execution procedure

To this end, the overall execution procedure of MRL based MPPT of PV systems under PSC is illustrated in

Fig. 6, where k_{max} represents the maximum iteration number.

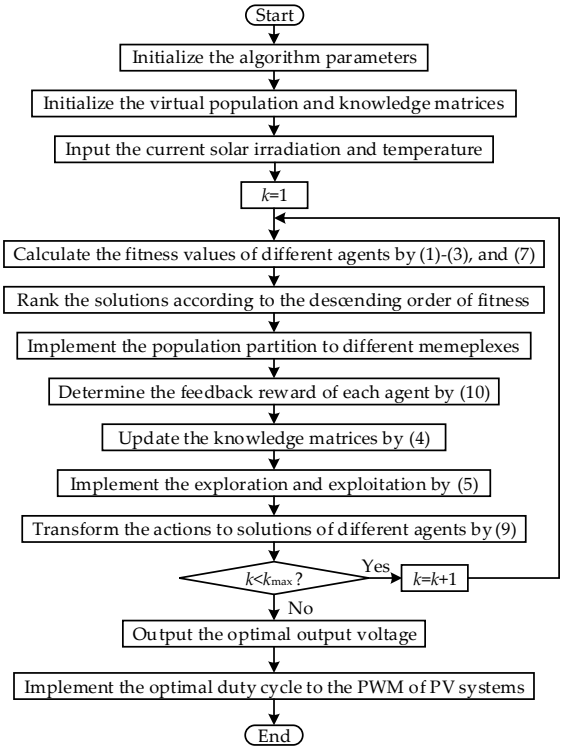


Figure 6. Overall execution procedure of MRL based MPPT of PV systems under PSC.

Table 1. The MRL parameters

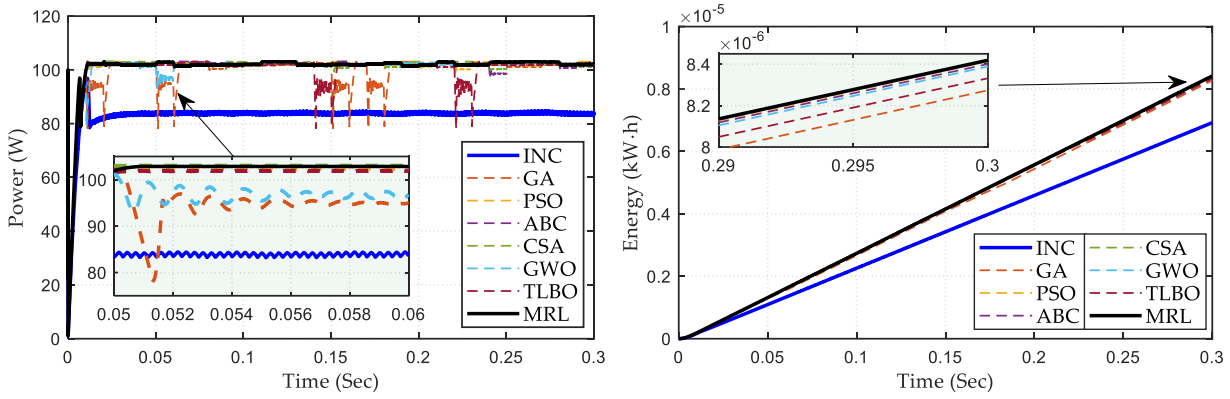
Parameter	Range	Value
n	$J>1$	3
P_s	$P_s>1$	10
α	$0<\alpha<1$	0.1
γ	$0<\gamma<1$	0.001
q_0	$0< q_0 <1$	0.9
k_{max}	$k_{max}>1$	10
L	$L>1$	4

5 Case Studies

In order to evaluate the MPPT performance under PSC of MRL, this section carries out four cases, e.g., start-up test; step change in solar irradiation; ramp change in both solar irradiation and temperature; and field atmospheric data of Hong Kong. The MPPT performance is compared to that of INC [11], GA [19], PSO [20], ABC [21], CSA [22], GWO [26], and TLBO [24], respectively. In addition, the MRL parameters are presented in Table 1. The simulation is executed on Matlab/Simulink 2016a. Moreover, the rated solar irradiation and temperature are set to be 1000 W/m^2 and 25℃, respectively.

5.1 Start-up test

The first test attempts to investigate the MPPT performance at start-up (from zero point) under PSC. The solar irradiation ranges from 400 W/m^2 to 1000 W/m^2, which is set to be 800 W/m^2, 600 W/m^2, and 400 W/m^2 for each PV array, respectively. Fig. 7 depicts the MPPT performance of different methods. One can readily find that INC can only converge to a LMPP since its final solution is completely determined by the initial solution under PSC. In contrast, other meta-heuristic algorithms can make the PV system generate more active power due to their enhanced global searching ability. Moreover, the energy obtained by MRL is the highest among all the methods, while its power fluctuation is the smallest. This verifies that MRL can not only converge to a high quality optimum for MPPT via a RL based local search, but also guarantee the convergence stability by the global information exchange based population partition.



(a) (b)

Figure 7. PV system responses of seven methods obtained on the start-up test. (a) Power and (b) Energy.

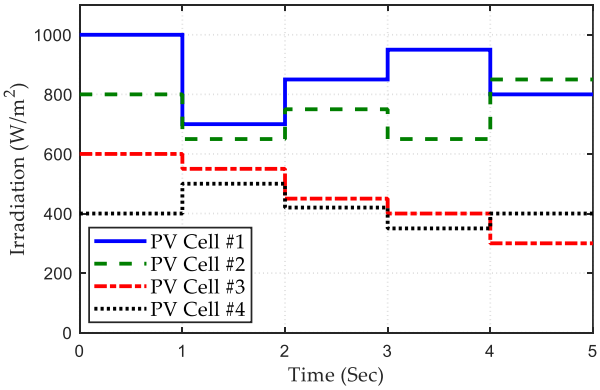
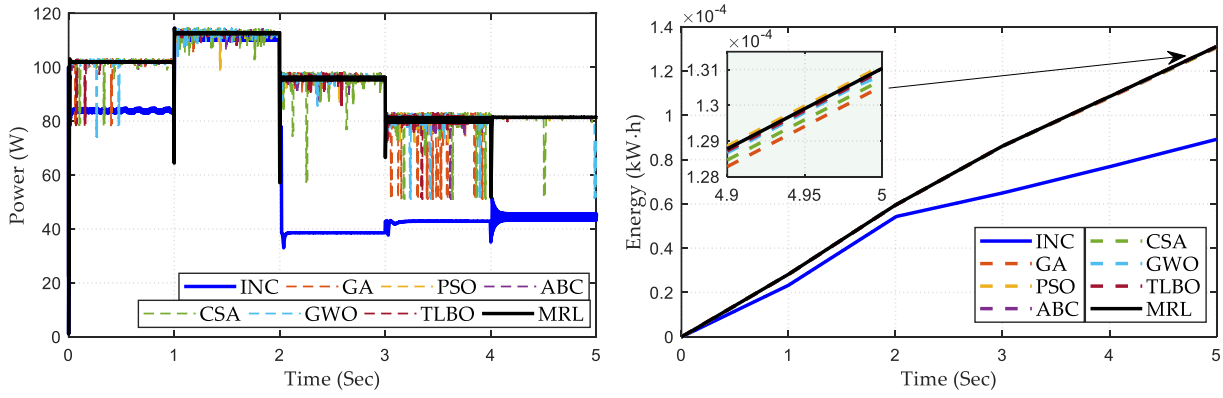


Figure 8. Step change of solar irradiation with PSC.

5.2 Step change in solar irradiation

To mimic the effect when a cloud rapidly passes over a PV array, a set of solar irradiation steps are applied on the PV array. The solar irradiation ranges from 300 W/m^2 to 1000 W/m^2, which varies at every second interval, as shown in Fig. 8. Throughout this test, the temperature keeps at 25℃.

Fig. 9 shows the online optimization results of different methods for MPPT under step change in solar irradiations. Similarly, all the meta-heuristic algorithms can outperform INC as they could generate much higher energy under the same weather conditions. Moreover, the energy output obtained by MRL is the highest, which can increase 46.98% of output energy compared with that of INC. Moreover, as the solar irradiation suddenly changes, all of the output current, voltage, and power obtained by the meta-heuristic algorithms excluding MRL present some oscillations.



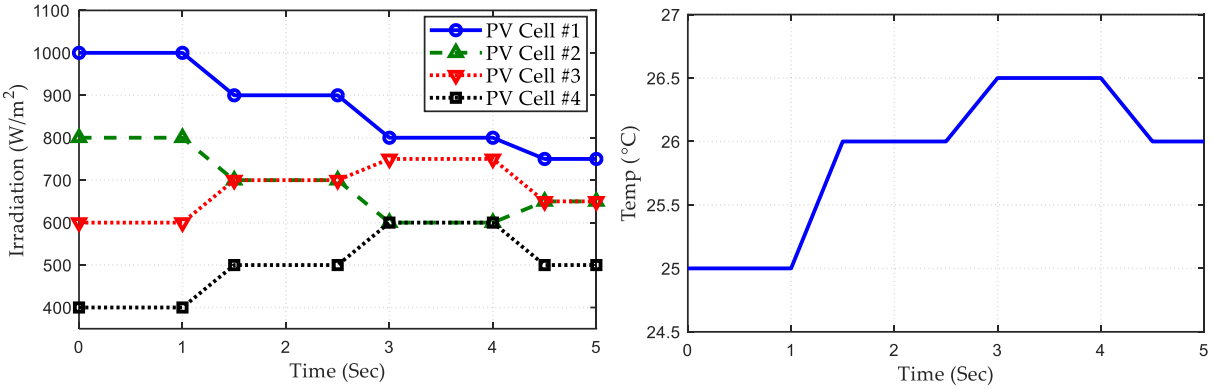
(a) (b)

Figure 9. PV system responses of seven methods obtained on the step change in solar irradiation with constant temperature. (a) Power and (b) Energy.

5.3 Ramp change in both solar irradiation and temperature

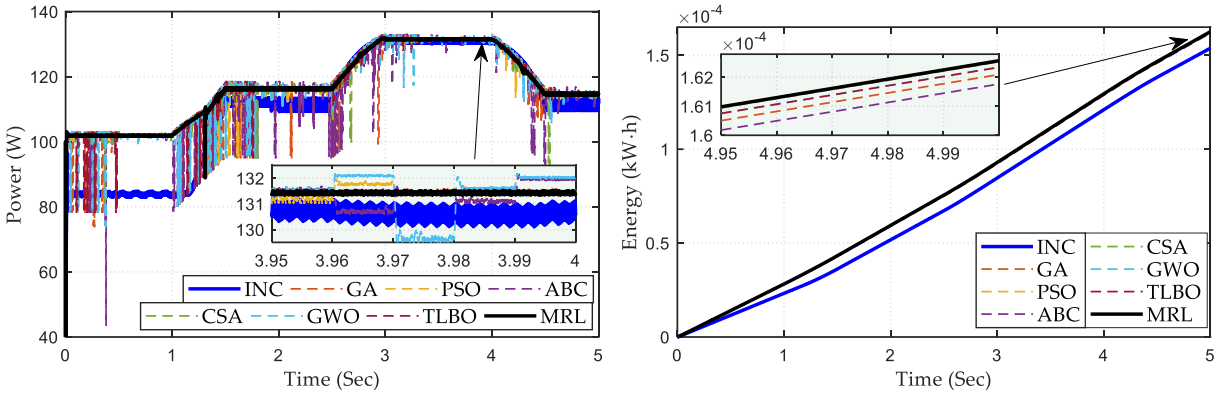
In a typical sunny day, both the solar irradiation and temperature increase when approaching the midday and thereafter decrease towards the evening. To investigate the MRL based MPPT performance under such scenario, a ramp change in both solar irradiation and temperature is emulated over a period of 5 *s*, in which the solar irradiation also ranges from 400 W/m^2 to 1000 W/m^2, as depicted in Fig. 10.

Figure 11 demonstrates the obtained performance of all algorithms under ramp change in both solar irradiation and temperature, which shows that the meta-heuristic algorithms except MRL still easily produce severer power fluctuations even in the presence of the relatively slow ramp gradual change of weather condition. In comparison, under the memetic computing framework, MRL can guarantee the highest convergence stability via a global information exchange between different local memeplexes. Besides, MRL can produce the highest energy of the PV system among all the algorithms, which is in excess of 5.71% of that obtained by INC.



(a) (b)

Figure 10. Gradual change in both solar irradiation and temperature. (a) Irradiation and (b) Temperature.



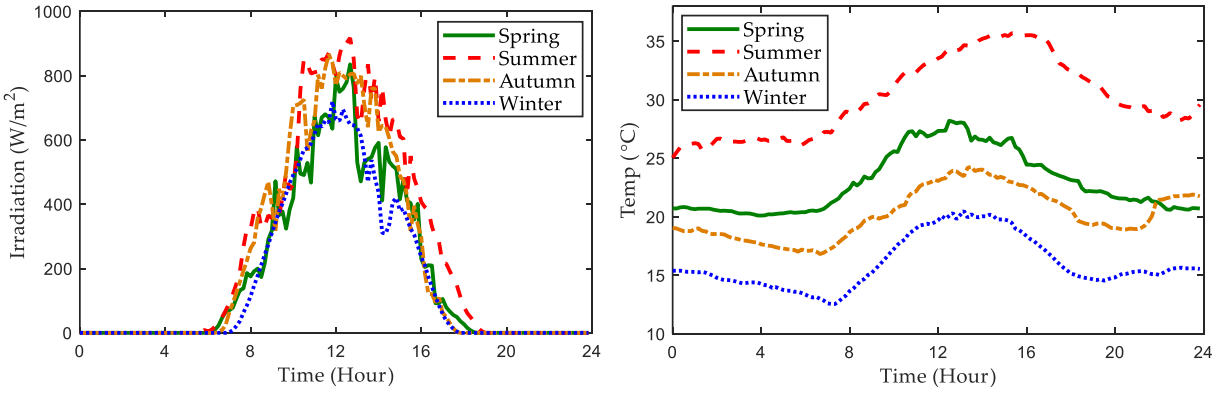
(a) (b)

Figure 11. PV system responses of seven methods obtained on the gradual change in both solar irradiation and temperature. (a) Power and (b) Energy.

5.4 Field atmospheric data in Hong Kong

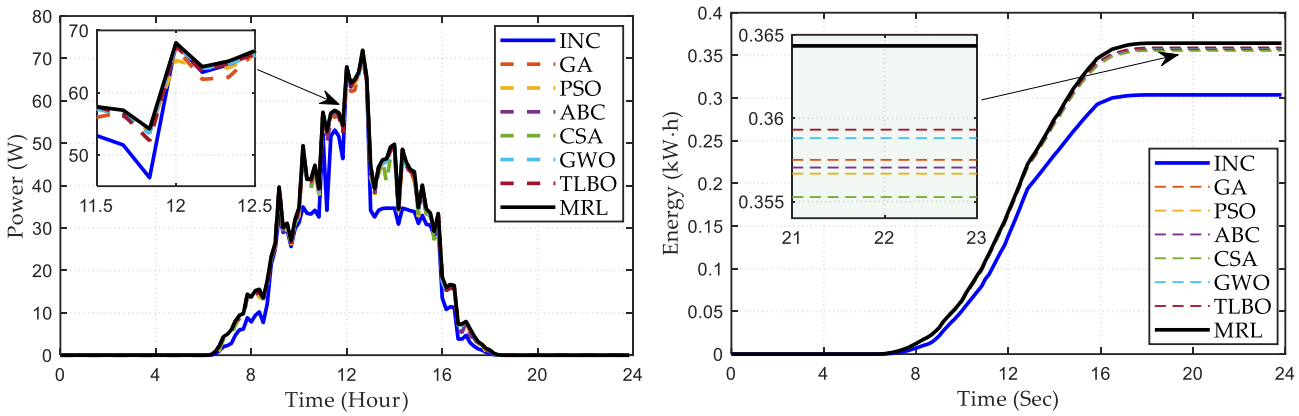
To further study the MRL based MPPT performance, it is tested by the field atmospheric data in Hong Kong (See Fig. 12) from four typical days of four seasons in 2012, where the solar irradiation ranges from 0 to 1000 W/m^2; and the data interval is 10 *min*.

Fig. 13 provides the obtained output power and energy of different algorithms for MPPT of spring in Hong Kong. It shows that all the meta-heuristic algorithms can generate much more energy than that of INC, in which the output energy obtained by MRL is up to 119.95% of that of INC. From the whole simulation period, although the MPPT performance difference of each meta-heuristic algorithm is relatively small, MRL still achieves superior performance to that of other meta-heuristic algorithms with the highest output energy.



(a) (b)

Figure 12. Daily field atmospheric data in Hong Kong. (a) Solar irradiation and (b) Temperature.



(a) (b)

Figure 13. PV system responses of seven methods obtained on the typical day of spring. (a) Power and (b) Energy.

5.5 Statistical studies

In order to quantitatively evaluate the power fluctuation of PV system, two indices are introduced as follows[47-49]:

$$\Delta v^{\text{avg}} = \frac{1}{T-1}\sum_{t=2}^{T}\frac{|P_{\text{out}}(t)-P_{\text{out}}(t-1)|}{P_{\text{out}}^{\text{avg}}} \tag{11}$$

$$\Delta v^{\text{max}} = \max_{t=2,3,\ldots,T}\frac{|P_{\text{out}}(t)-P_{\text{out}}(t-1)|}{P_{\text{out}}^{\text{avg}}} \tag{12}$$

where Δv^{avg} and Δv^{max} denote the average variability and the maximum variability of power output of PV system, respectively; t is the time period; T represents the total operation period; and P_{out}^{avg} means the average power output of PV system over the total operation period.

Table 2 tabulates the statistical results obtained by eight methods under four cases. One can directly find that MRL owns the highest output energy and the smallest power variations (highlighted in bold) in all cases among all algorithms. In particular, the average variability of MRL under step change in solar irradiation is only 35.29%, 84.30%, 88.70%, 93.58%, 75.56%, 87.18%, 88.70% to that of INC, GA, PSO, ABC, CSA, GWO, and TLBO, respectively. Besides, the output energy generated by MRL in winter is 120.86%, 101.84%, 101.10%, 102.19%, 101.61%, 101.43%, and 101.27% to that of INC, GA, PSO, ABC, CSA, GWO, and TLBO, respectively. As a consequence, MRL can achieve the most satisfactory MPPT performance with the highest stability of generated power/energy.

Table 2 Statistical results obtained by eight methods under four cases

Normal test									
Test	Index	INC	GA	PSO	ABC	CSA	GWO	TLBO	MRL
Start-up	Energy(10^{-6} kW ħ)	6.9219	8.2740	8.4132	8.4007	8.4060	8.3887	8.3320	**8.4190**
	$\Delta v^{\max}$(%)	0.6173	1.9572	0.0742	0.1047	0.0573	1.9154	1.3484	**0.0298**
	Δv^{avg}(%)	0.0403	0.0148	0.0066	0.0074	0.0080	0.0086	0.0123	**0.0064**
Step change	Energy(10^{-6} kW ħ)	89.1534	130.5378	131.1098	131.033	130.719	130.843	130.9785	**131.0347**
	$\Delta v^{\max}$(%)	56.5161	38.5988	38.4304	38.4529	38.5453	38.5087	38.4689	**38.4524**
	Δv^{avg}(%)	0.0289	0.0121	0.0115	0.0109	0.0135	0.0117	0.0115	**0.0102**
Ramp change	Energy(10^{-6} kW ħ)	153.7881	162.0980	162.5478	161.7704	162.5377	162.355	162.3459	**162.5748**
	$\Delta v^{\max}$(%)	32.7632	31.0837	30.9976	31.1466	30.9996	31.0345	31.0362	**30.9925**
	Δv^{avg}(%)	0.0267	0.0102	0.0089	0.0118	0.0085	0.0098	0.0092	**0.0081**
Field atmospheric data in Hong Kong									
Season	Index	INC	GA	PSO	ABC	CSA	GWO	TLBO	MRL
Spring	Energy (kW ħ)	0.3037	0.3575	0.3567	0.3570	0.3553	0.3588	0.3593	**0.3643**
	$\Delta v^{\max}$(%)	245.09	124.20	142.43	139.17	159.57	138.28	139.23	**138.29**
	Δv^{avg}(%)	14.0297	14.0203	13.7675	13.9927	15.0213	13.9135	14.0876	**13.7189**
Summer	Energy (kW ħ)	0.4538	0.4880	0.4854	0.4883	0.4902	0.4893	0.4902	**0.4959**
	$\Delta v^{\max}$(%)	161.84	87.93	80.91	87.18	79.68	88.87	87.19	**83.99**
	Δv^{avg}(%)	9.9617	9.9847	11.0842	10.5370	10.2591	10.3998	10.0304	**10.0936**
Autumn	Energy (kW ħ)	0.3843	0.4277	0.4310	0.4265	0.4305	0.4295	0.4303	**0.4354**
	$\Delta v^{\max}$(%)	209.19	82.43	78.22	86.91	78.46	80.41	79.28	**79.71**
	Δv^{avg}(%)	13.6754	11.3672	10.6816	11.3706	10.5944	10.5189	10.5390	**10.5438**
Winter	Energy (kW ħ)	0.2708	0.3214	0.3238	0.3203	0.3221	0.3227	0.3232	**0.3273**
	$\Delta v^{\max}$(%)	87.69	86.04	80.85	89.53	87.55	86.80	82.75	**86.68**
	Δv^{avg}(%)	9.46	9.08	8.73	8.98	8.82	8.54	8.42	**8.4474**

5.6 Sensitivity analysis

This section is performed to investigate the influence of the input parameters on the system outputs obtained

by different algorithms. In this study, the input parameters include the solar irradiation S of each PV array and the working temperature T_c, and the system outputs include the average power output P_{out}^{avg}, the maximum variability Δv^{max}, and the average variability Δv^{avg}.

5.6.1 Influence of working temperature

The working temperature is varied from 8 ℃ to 36 ℃, while the temperature interval between two adjacent points is set to be 2 ℃; and the solar irradiations of four PV arrays are set to be the same as that of start-up test. As illustrated in Fig. 14(a), the average power output obtained by MRL decreases from 102.98 W to 101.25 W (a decrease of 1.68%) as working temperature increases. Furthermore, INC has a more significant decrease (11.40%) of average power output, while the obtained average power output is much lower than that of MRL. This also demonstrates that INC may easily trap at a low-quality LMPP for the PV system. Compared with other meta-heuristic algorithms, MRL also has a higher average power output of each scenario, as shown in Fig. 14(b). Moreover, the average power outputs obtained by some meta-heuristic algorithms (e.g., GA) change randomly instead of monotonically decreasing due to their high convergence randomness. Similarly, it can be found from Fig. 14(c) that the maximum variability obtained by each algorithm also changes randomly, in which that of MRL is the lowest among all the algorithms for each working point. On the other hand, the average variability obtained by INC monotonically increases from 0.0383% to 0.0414% (an increase of 8.12%) as the working temperature increases. In contrast, the average variability obtained by MRL changes much more significantly from 0.0050% to 0.0085% (an increase of 70.00%), while that obtained by some meta-heuristic algorithms (e.g., GWO) also change randomly. In summary, the working temperature can result in a major effect on the power outputs by INC, as well as on the average variability obtained by MRL.

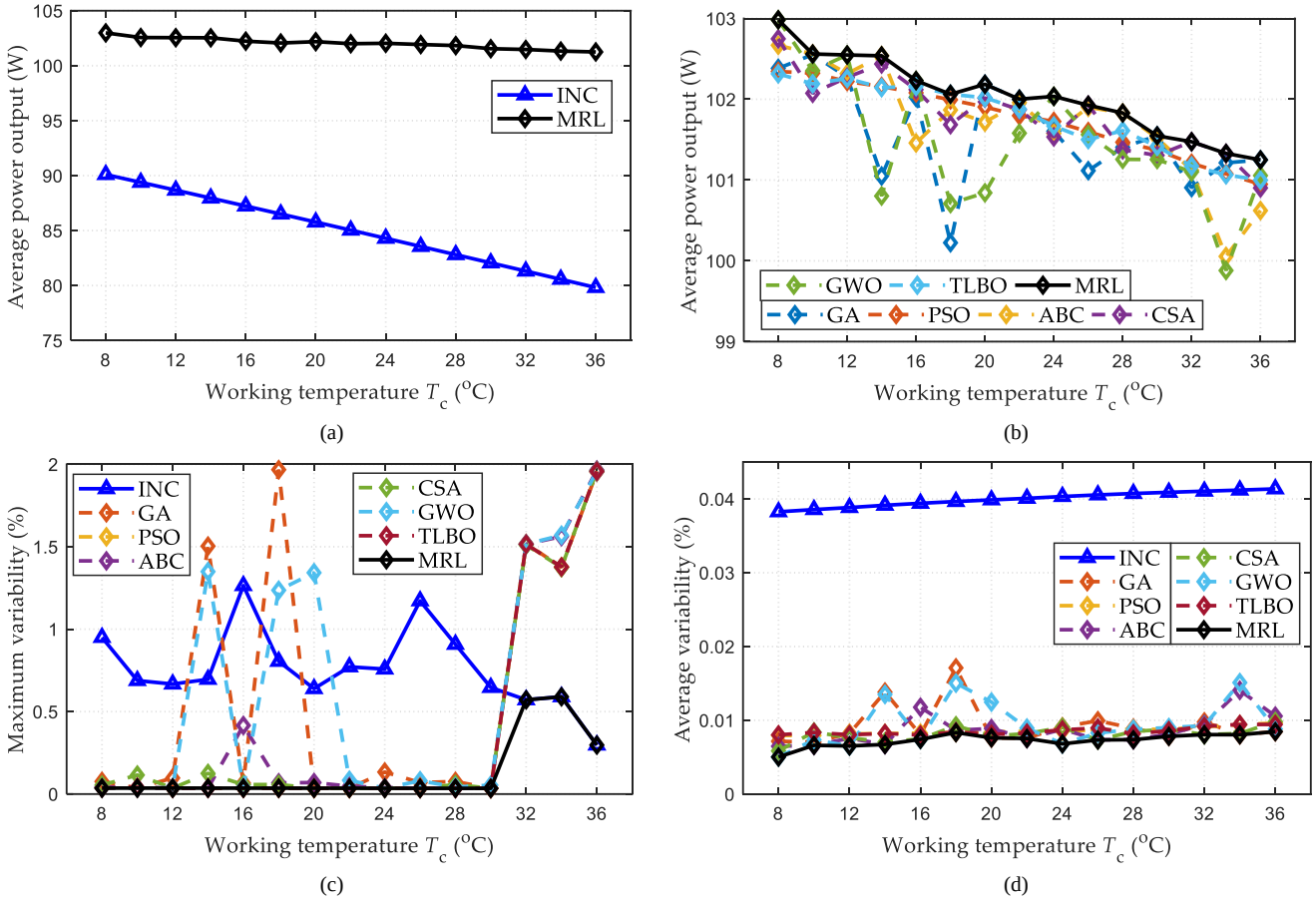


Figure 14. Effect of the working temperature variation on the PV system outputs. (a) Average power outputs obtained by MRL and INC, (b) Average power outputs obtained by MRL and meta-heuristic algorithms, (c) Maximum variability, and (d) Average variability.

5.6.2 Influence of solar irradiation

The solar irradiations of the start-up test are regarded as the benchmark irradiations in this sensitivity analysis, while the irradiation intensity ratio is varied from 5% to 100%; and the ratio interval between two adjacent points is set to be 5%. Besides, the working temperature is set to be 25 ℃. As given in Fig. 15(a), the average power output obtained by each algorithm is monotonically increasing with the increase of irradiation intensity ratio, and the variation gradient is close to 1. This reveals that the solar irradiation results in a much more effect on the average

power output than that influenced by the working temperature, while their influences are inverse. Among all algorithms, MRL and INC also obtain the highest and the lowest average power outputs of each working point, respectively. Similarly, the maximum variability (See Fig. 15(b)) obtained by each algorithm also varies randomly as the irradiation intensity ratio increases, as well as for the average variability (See Fig. 15(c)). When the solar intensity ratio exceeds 20%, the average variability obtained by MRL dramatically decreases from 0.0396% to 0.0076% (a decrease of 80.73%). This also demonstrates that the influences of the working temperature and solar irradiation are inverse on the average variability when the solar intensity ratio exceeds 20%.

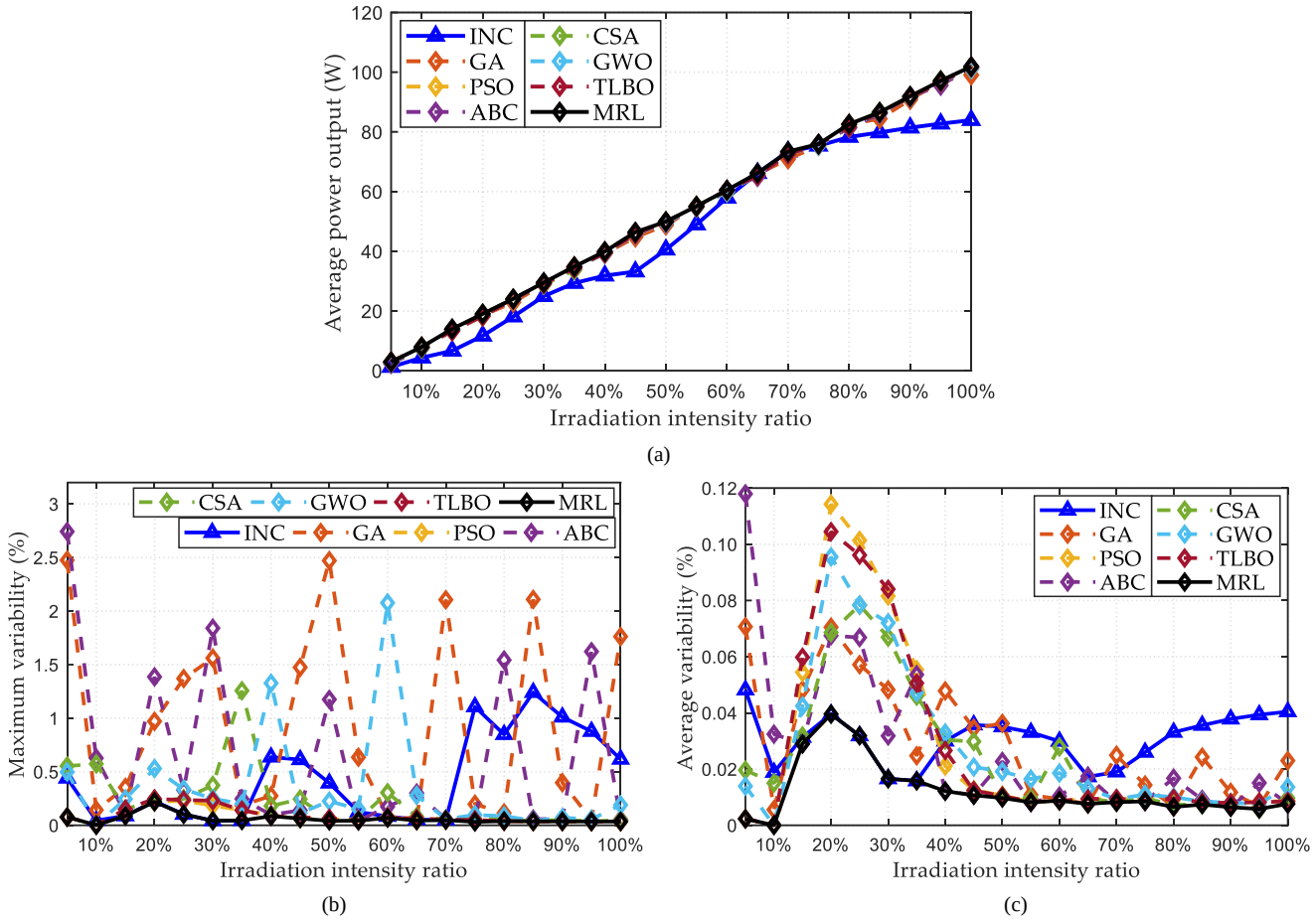


Figure 15. Effect of the irradiation intensity variation on the PV system outputs. (a) Average power outputs, (b) Maximum variability, and (c) Average variability.

6. Conclusions

This paper proposes a novel MRL based MPPT technique for PV system under PSC, which contributions/novelties are summarized as follows:

(1) Through incorporating the memetic computing framework into RL, the searching ability of MRL can be significantly enhanced based on an effective coordination between the local search and the global information exchange.
(2) MRL can obtain a higher quality optimum for MPPT under PSC compared with that of other traditional meta-heuristic algorithms, e.g., it can approximate the GMPP more closely. Thus, the PV system can generate higher energy under different weather conditions in different seasons;
(3) Due to the high convergence stability of MRL, it produces the smallest power fluctuations of the PV system than that of other traditional meta-heuristic algorithms under the same weather conditions;

Acknowledgments

The authors gratefully acknowledge the support of Research and Development Start-Up Foundation of Shantou University (NTF19001), National Natural Science Foundation of China (51477055, 51777078), and China Southern Power Grid Key Research Program (YNKJQQ00000279)

References

[1] Yang, B.; Yu, T.; Shu, H.C.; Dong, J.; Jiang, L. Robust sliding-mode control of wind energy conversion systems for optimal power extraction via nonlinear perturbation observers, *Applied Energy* **2018**, 210: 711-723.

[2] Liu, J.; Wen, J.Y.; Yao, W.; Long, Y. Solution to short-term frequency response of wind farms by using energy storage systems. *IET Renewable Power Generation* **2016**, 10(5): 669-678.

[3] Yang, B.; Jiang, L.; Wang, L.; Yao, W.; Wu, Q.H. Nonlinear maximum power point tracking control and modal analysis of DFIG based wind turbine. *International Journal of Electrical Power & Energy Systems* **2016**, 74: 429-436.

[4] Yang, B.; Zhang, X.S.; Yu, T.; Shu, H.C.; Fang, Z.H. Grouped grey wolf optimizer for maximum power point tracking of doubly-fed induction generator based wind turbine. *Energy Conversion and Management* **2017**, 133: 427-443.

[5] Liao, S.W.; Yao, W.; Han, X.N.; Wen, J.Y.; Cheng, S.J. Chronological operation simulation framework for regional power system under high penetration of renewable energy using meteorological data. *Applied Energy* **2017**, 203: 816-828.

[6] Yang, B.; Yu, T.; Shu, H.C.; Zhu, D.N.; An, N.; Sang, Y.Y.; Jiang, L. Energy reshaping based passive fractional-order PID control design and implementation of a grid-connected PV inverter for MPPT using grouped grey wolf optimizer. *Solar Energy* **2018**, 170: 31-46.

[7] Kandemir, E.; Cetin, N.S.; Borekci, S. A comprehensive overview of maximum power extraction methods for PV systems. *Renewable and Sustainable Energy Reviews* **2017**, 78: 93-112.

[8] Belhachat, F.; Larbes, C. A review of global maximum power point tracking techniques of photovoltaic system under partial shading conditions. *Renewable and Sustainable Energy Reviews* **2018**, 92: 513-553.

[9] Ashouri-Zadeh, A.; Toulabi, M.; Dobakhshari, A.S.; Taghipour-Broujeni, S.; Ranjbar, A.M. A novel technique to extract the maximum power of photovoltaic array in partial shading conditions. *International Journal of Electrical Power & Energy Systems* **2018**, 101: 500-512.

[10] Bingöl, O.; Özkaya, B. Analysis and comparison of different PV array configurations under partial shading conditions. *Solar Energy* **2018**, 160: 336-343.

[11] Zakzouk, N.E.; Elsaharty, M.A.; Abdelsalam, A.K; Helal, A.A. Improved performance low-cost incremental conductance PV MPPT technique. *IET Renewable Power Generation* **2016**, 10(4): 561-574.

[12] Tanaka, T.; Toumiya, T.; Suzuki, T. Output control by hill-climbing method for a small scale wind power generating system. *Renewable Energy* **2014**, 12(4): 387-400.

[13] Mohanty, S.; Subudhi, B.; Ray, P.K. A grey wolf-assisted Perturb & Observe MPPT algorithm for a PV system. *IEEE Transactions on Energy Conversion* **2017**, 32(1): 340-347.

[14] Amir, A.; Amir, A.; Selvaraj, J.; Rahim, N.A.; Abusorrah, A.M. Conventional and modified MPPT techniques with direct control and dual scaled adaptive step-size. *Solar Energy* **2017**, 157: 1017-1031.

[15] Alik, R.; Jusoh, A. An enhanced P&O checking algorithm MPPT for high tracking efficiency of partially shaded PV module. *Solar Energy* **2018**, 163: 570-580.

[16] Rong, F.; Gong, X.; Huang, S. A novel grid-connected PV system based on MMC to get the maximum power under partial shading conditions. *IEEE Transactions on Power Electronics* **2017**, 32(6): 4320-4333.

[17] Manickam, C.; Raman, G.P.; Raman, G.R.; Ganesan S.I.; Chilakapati, N. Fireworks enriched P&O algorithm for GMPPT and detection of partial shading in PV systems. *IEEE Transactions on Power Electronics* **2017**, 32(6): 4432-4443.

[18] Rezk, H.; Fathy, A.; Abdelaziz, A.Y. A comparison of different global MPPT techniques based on meta-heuristic algorithms for photovoltaic system subjected to partial shading conditions. *Renewable and Sustainable Energy Reviews* **2017**, 74: 377-386.

[19] Daraban, S.; Petreus, D.; Morel, C. A novel MPPT (maximum power point tracking) algorithm based on a modified genetic algorithm specialized on tracking the global maximum power point in photovoltaic systems affected by partial shading. *Energy* **2014**, 74(5): 374-388.

[20] Li, H.; Yang, D.; Su, W.; Lu, J.; Yu, X. An overall distribution particle swarm optimization MPPT algorithm for photovoltaic system under partial shading. *IEEE Transactions on Industrial Electronics* **2018**. DOI: 10.1109/TIE.2018.2829668.

[21] Benyoucef, A.S.; Chouder, A.; Kara, K.; Silvestre, S.; Sahed, O.A. Artificial bee colony based algorithm for maximum power point tracking (MPPT) for PV systems operating under partial shaded conditions. *Applied Soft Computing* **2015**, 32: 38-48.

[22] Ahmed, J.; Salam, Z. A maximum power point tracking (MPPT) for PV system using Cuckoo search with partial shading capability. *Applied Energy* **2014**, 119: 118-130.

[23] Diab, A.A.Z.; Rezk, H. Global MPPT based on flower pollination and differential evolution algorithms to mitigate partial shading in building integrated PV system. *Solar Energy* **2017**, 157: 171-186.

[24] Rezk, H.; Fathy, A. Simulation of global MPPT based on teaching-learning-based optimization technique for partially shaded PV system. *Electrical Engineering* **2017**, 99: 847-859.

[25] Kumar, N.; Hussain, I.; Singh, B.; Panigrahi, B.K. MPPT in dynamic condition of partially shaded PV system by using WODE technique. *IEEE Transactions on Sustainable Energy* **2017**, 8(3): 1204-1214.

[26] Mohanty, S.; Subudhi, B.; Ray, P. K. A new MPPT design using grey wolf optimization technique for photovoltaic system under partial shading conditions. *IEEE Transactions on Sustainable Energy* **2015**, 7(1): 181-188.

[27] Yang, B.; Yu, T.; Zhang, X.S.; Li, H.F.; Shu, H.C.; Sang, Y.Y.; Jiang, L. Dynamic leader based collective intelligence for maximum power point tracking of PV systems affected by partial shading condition. *Energy Conversion and Management* **2019**, 179: 286-303.

[28] Yang, B.; Zhong, L.E.; Yu, T.; Li, H.F.; Zhang, X.S.; Shu, H.C.; Sang, Y.Y.; Jiang, L. Novel bio-inspired memetic salp swarm algorithm and application to MPPT for PV systems considering partial shading condition. *Journal of Cleaner Production* **2019**, 215: 1203-1222.

[29] Kaelbling, L.P.; Littman, M.L.; Moore, A.W. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research* **1996**, 4(1): 237-285.

[30] Zhang, X.; Yu, T.; Yang, B.; Cheng, L. Accelerating bio-inspired optimizer with transfer reinforcement learning for reactive power optimization. *Knowledge-Based Systems* **2017**, 116: 26-38.

[31] Zhang, X.; Bao, T.; Yu, T.; Yang, B.; Han, C. Deep transfer Q-learning with virtual leader-follower for supply-demand Stackelberg game of smart grid. *Energy* **2017**, 133: 348-365.

[32] Neri, F; Cotta, C. Memetic algorithms and memetic computing optimization: A literature review. *Swarm and Evolutionary Computation* **2012**, 2: 1-14.

[33] Jately, V.; Arora, S. Development of a dual-tracking technique for extracting maximum power from PV systems under rapidly changing environmental conditions. *Energy* **2017**, 133: 557-571.
[34] Yang, B.; Yu, T.; Shu, H.C.; Zhu, D.N.; An, N.; Sang, Y.Y.; Jiang, L. Perturbation observer based fractional-order sliding-mode controller for MPPT of grid-connected PV inverters: design and real-time implementation. *Control Engineering Practice* **2018**, 79: 105-125.
[35] Yang, B.; Yu, T.; Shu, H.C.; Zhu, D.N.; Zeng, F.; Sang, Y.Y.; Jiang, L. Perturbation observer based fractional-order PID control of photovoltaics inverters for solar energy harvesting via Yin-Yang-Pair optimization. *Energy Conversion and Management* **2018**, 171: 170-187.
[36] Qi, J.; Zhang, Y.; Chen, Y. Modeling and maximum power point tracking (MPPT) method for PV array under partial shade conditions. *Renewable Energy* **2014**, 66: 337-345.
[37] Chen, K.; Tian, S.; Cheng, Y.; Bai, L. An improved MPPT controller for photovoltaic system under partial shading condition. *IEEE Transactions on Sustainable Energy* **2017**, 5(3): 978-985.
[38] Neri, F.; Cotta, C.; Moscato, P. Handbook of Memetic Algorithms, in: Studies in Computational Intelligence, vol. 379, Springer-Verlag, 2012.
[39] Eusuff, M.; Lansey, K.; Pasha, F. Shuffled frog-leaping algorithm: a memetic meta-heuristic for discrete optimization. *Engineering Optimization* **2006**, 38(2):129-154.
[40] Watkins, J.C.H.; Dayan, P. Q-learning. *Machine Learning* **1992**, 8(3-4): 279-292.
[41] Tan, Z.; Zhang, X.; Xie, B.; Wang, D.; Liu, B.; Yu, T. Fast learning optimiser for real-time optimal energy management of a grid-connected microgrid. *IET Generation, Transmission & Distribution* **2018**, 12(12): 2977-2987.
[42] Bianchi, R.A.C.; Celiberto, L.A.; Santos, P.E.; Matsuura, J.P.; Mantaras, R.L.D. Transferring knowledge as heuristics in reinforcement learning: a case-based approach. *Artificial Intelligence* **2015**, 226: 102-121
[43] Eusuff, M.M..; Lansey, K.E. Optimization of water distribution network design using the shuffled frog leaping algorithm. *Journal of Water Resources Planning & Management* **2015**, 129(3): 210-225.
[44] Krynicki, K.; Houle, M.E.; Jaen, J. An efficient ant colony optimization strategy for the resolution of multi-class queries. *Knowledge-Based Systems* **2016**, 105: 96-106.
[45] Yang, B.; Yu, T.; Shu, H.C.; Zhang, Y.M.; Chen, J.; Sang, Y.Y.; Jiang, L. Passivity-based sliding-mode control design for optimal power extraction of a PMSG based variable speed wind turbine. *Renewable Energy* **2018**, 119: 577-589.
[46] Li, X.; Wen, H.; Hu, Y.; Jiang, L.; Xiao, W. Modified beta algorithm for GMPPT and partial shading detection in photovoltaic systems. *IEEE Transactions on Power Electronics* **2018**, 33(3): 2172-2186.
[47] Yao, W.; Jiang, L.; Wen, J.Y.; Wu, Q.H.; Cheng, S.J. Wide-area damping controller for power system inter-area oscillations: a networked predictive control approach. *IEEE Transactions on Control Systems Technology* **2015**, 23(1): 27-36.
[48] Shen, Y.; Yao, W.; Wen, J.Y.; He, H.B.; Chen, W.B. Adaptive supplementary damping control of VSC-HVDC for interarea oscillation using GrHDP. *IEEE Transactions on Power Systems* **2018**, 33(2): 1777-1789.
[49] Chen, J.; Yao, W.; Zhang, C.K.; Ren, Y.; Jiang, L. Design of robust MPPT controller for grid-connected PMSG-Based wind turbine via perturbation observation based nonlinear adaptive control. *Renewable Energy* **2019**, 134: 478-495.