

A Combined Reinforcement Learning and Sliding Mode Control Scheme for Grid Integration of a PV System

Aurobinda Bag, Bidyadhar Subudhi, *Senior Member, IEEE*, and Pravat Kumar Ray, *Senior Member, IEEE*

Abstract—The paper presents development of a reinforcement learning (RL) and sliding mode control (SMC) algorithm for a 3-phase PV system integrated to a grid. The PV system is integrated to grid through a voltage source inverter (VSI), in which PV-VSI combination supplies active power and compensates reactive power of the local non-linear load connected to the point of common coupling (PCC). For extraction of maximum power from the PV panel, we develop a RL based maximum power point tracking (MPPT) algorithm. The instantaneous power theory (IPT) is adopted for generation reference inverter current (RIC). An SMC algorithm has been developed for injecting current to the local non-linear load at a reference value. The RL-SMC scheme is implemented in both simulation using MATLAB/SIMULINK software and on a prototype PV experimental. The performance of the proposed RL-SMC scheme is compared with that of fuzzy logic-sliding mode control (FL-SMC) and incremental conductance-sliding mode control (IC-SMC) algorithms. From the obtained results, it is observed that the proposed RL-SMC scheme provides better maximum power extraction and active power control than the FL-SMC and IC-SMC schemes.

Index Terms—Fuzzy logic, incremental conductance, instantaneous power theory, reinforcement learning, sliding mode controller.

NOMENCLATURE

S	Set of states.
A	Set of actions.
R	Reward.
V_{pv}	PV output voltage.
P_{pv}	PV output power.
I_{pv}	PV output current.
i	Number of exploration round.
m	Multiplier.
α	Learning rate.
γ	Discount factor.
N_{max}	Maximum number of exploration round.
R_c	Inverter interface resistor.
L_c	Inverter interface inductor.

R_g	Grid interface resistor.
L_g	Grid interface inductor.
V_g	Grid voltage.
I_g	Grid current.
R_L	Load Resistor.
L_L	Load Inductor.
I_c	Inverter output current.
V_{dc}	DC capacitor voltage.
V_{dcref}	Reference dc capacitor voltage.
D	Duty cycle.
C_{dc}	DC capacitor.
P_g	Grid active power.
Q_g	Grid reactive power.
P_L	Load active power.
Q_L	Load reactive power.
V_p	Voltage at point of common coupling.

I. INTRODUCTION

HARVESTING maximum possible power from a solar PV system and then connecting it to the utility grid is a challenging task owing to intermittent nature of solar PV system, like variation of solar insolation and temperature. Many MPPT techniques have been implemented for extraction of maximum power from the PV system, which can be categorized in to indirect method and direct method [1]. Indirect methods use the database like parameters, PV array characteristic curves at different insolation and temperature for extraction of maximum power. The indirect method include curve fitting [2], look up table [2], open circuit voltage [2], short circuit current [3]. Indirect methods are not suitable in case of changing atmospheric conditions. Direct methods include the methods, which use the voltage current measurement of PV panel, such as hill climbing [2], perturb and observe (P&O) [2], incremental conductance [4], fuzzy logic (FL), artificial intelligence (AI), SMC. However, hill climbing and P&O MPPT algorithm encounter serious problem such as oscillations at the MPP due to continuous perturbations and reduced efficiency [5]. Further, another demerit of P&O is that, it is sensitive to changing atmospheric condition in [6]. IC MPPT algorithm has a serious recursion near the MPP [7]. FL control based MPPT offers fast tracking and expert knowledge is required to design fuzzy parameters [8]. On the other hand, AI based MPPT algorithms have been reported in [9]. SMC MPPT algorithm for PV system has been discussed in [2].

Manuscript received September 29, 2016; revised June 7, 2018; accepted July 23, 2018. Date of publication December 30, 2019; date of current version August 13, 2019.

A. Bag and P. K. Ray are with Department of Electrical Engineering, National Institute of Technology, Rourkela-769008.

B. Subudhi (corresponding author, email: bidyadhar@iitgoa.ac.in) is with School of Electrical Sciences, Indian Institute of Technology Goa, Ponda-403401.

DOI: 10.17775/CSEEJPES.2017.01000

These are off-line MPPT algorithms and applicable for fixed problems.

In order to overcome the above mentioned problems, we exploited RL concept to develop an effective MPPT algorithm that yields reaching to the MPP faster for achieving better MPP tracking under change in solar insolation and temperature. In [10], the reinforcement learning MPPT algorithm is proposed for wind energy conversion system. Also, the RL MPPT algorithm is reported for a standalone PV system in [11], [12]. A RL MPPT algorithm using new state definition is also introduced in [13].

However, to the best of author's knowledge no work is reported to date on RL MPPT for a grid connected PV system (GCPVS). We employed the concept of RL for development of an MPPT algorithm for a grid integrated PV system in this paper. In [14], a number of grid synchronization schemes for renewable energy system has been discussed. IPT is one of the popular methods used for generation of reference inverter currents proposed in [15], which is used in this paper. For obtaining improved steady state and transient performance, non-linear control algorithms are suitable grid synchronization of a PV system. Some non-linear control algorithms, optimal control, artificial neural network have been discussed in [16]. Application of SMC strategy for grid synchronization of PV system is discussed in [17]. In order to achieve robust operation under change in solar irradiance, temperature or load variation for integration of PV system to grid, sliding mode control algorithm is developed in this paper to generate switching signals for controlling the VSI.

The contribution of this paper is to exploit the learning capability of an agent based RL algorithm to achieve maximum power extraction from a GCPVS under varied solar irradiance and temperature. SMC algorithm is employed for generation of switching signals of the VSI. The performance of the proposed RL-SMC scheme is compared with that of FL-SMC and IC-SMC schemes, in simulation and on a prototype PV experimental setup.

The paper is organized as follows. The proposed RL based MPPT scheme is presented in Section II. In Section III, the sliding mode controller is described. Section IV describes reference current generation of VSI using IPT algorithm and in Section V the IC MPPT algorithm is described. In Section VI, simulation results are provided with a comparative assessment of performances of the proposed RL-SMC, FL-SMC and IC-SMC. Section VII presents the experimental results with discussions. Finally Section VIII provides the conclusion of the paper.

II. RL BASED MPPT CONTROL ALGORITHM

RL algorithm is an adaptive algorithm, where an agent learns from its personal experience by interacting continuously with the environment through states, actions and rewards. An agent must be capable to sense the state of the environment and take actions to change the state. RL algorithm aims to map from states to actions for maximizing the reward. The agent always receives a reward and takes actions to transit from one state $s \in S$ to other state $s' \in S$ towards the optimal

value. Fig. 1 shows the structure of the RL algorithm for a PV system. Most of the reinforcement learning algorithms such as temporal difference learning, Monte Carlo method, Q learning deal with finite Markov's decision process (MDP). A finite MDP is defined by a set of states S , a set of actions A , and the transition probability function T that defines the transition probability from state $s \in S$ to s' when the action $a \in A$ is selected and reward $r \in R$ that receives after each state transition.

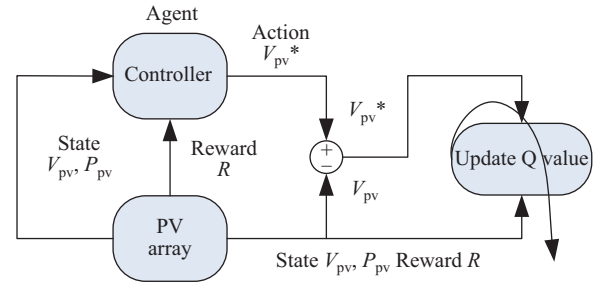


Fig. 1. Structure for RL algorithm of PV system.

In RL algorithm, an agent receives a state signal, which represents states $s \in S$ of the environment. According to the state action mapping, it takes an action $a \in A$. The action is taken according to the ϵ -greedy action selection policy. The ϵ -greedy policy is such that, selecting random actions with uniform distribution from available set of actions. Using the ϵ -greedy policy either we can select random action with ϵ probability and we can select an action with $(1-\epsilon)$ probability that gives maximum reward in given state. After one time step, action $a \in A$, the state transition occurs from s to s' corresponding to the action selection strategy. Then the agent receives reward $r \in R$, according to the instant effect of action 'a' of the change of state. The role of an agent in reinforcement learning algorithm is to achieve action selection strategy for maximizing future reward.

A. Development of RL Based MPPT for PV System

In order to apply the reinforcement learning, for MPPT, three main constituents such as state space, action space and reward are necessary to define. Fig. 1 shows the structure of the RL algorithm for a PV system. The structure of RL algorithm for PV system describes that, the agent detects the state of the environment $s \in S$, the operating point for the PV system and accordingly pick out a discrete control action $a \in A$ using ϵ -greedy strategy. In the mean time, the Q value $Q(s,a)$ is recalled. After that, the PV system enters into another operating point. In the mean time, an immediate reward $r \in R$ is received by the agent and maximum Q value is picked out using ϵ -greedy strategy. The previous Q value will be updated using (4). For obtaining maximum power from a PV panel, each time a higher action value should be picked out with a higher possibility.

1) *State*

The state, which describes the PV system's condition, is crucial for the performance of RL algorithm. The state should be descriptive enough and include all required information to

describe system condition. Also, excess information can lead to a very large state space and longer learning period, where as insufficient information may lead to decreased learning accuracy. The state space is described for application of RL MPPT is divided into number of states $s_1, s_2, s_3, s_4 \dots s_n$ as defined in (1).

$$S = [s_1, s_2, s_3, s_4 \dots s_n] \quad (1)$$

2) Action

The action space A can be defined as finite number of desired perturbation, that can be applied on the PV system to bring the change of system's operation. The action space can be defined as

$$A = \{-2\Delta v, -\Delta v, 0, \Delta v, 2\Delta v\} \quad (2)$$

where $2\Delta v$ is a large change, Δv is a small change and 0 represents no change in PV's output voltage. In each state, an agent has five actions for changing the system's operation. Every time, the agent selects action on the basis of state measurement and action selection policy from the action space for extracting of the MPP of a PV system.

3) Reward

After selecting an action, the agent experiences the reward for evaluation of the action selected. The reward function can be defined as

$$R = \begin{cases} +1, & \text{if } P_{pv,t+1} - P_{pv,t} > \delta \\ 0, & \text{if } P_{pv,t+1} - P_{pv,t} < \delta \\ -1, & \text{if } P_{pv,t+1} - P_{pv,t} < -\delta \end{cases} \quad (3)$$

where $P_{pv,t+1}$ and $P_{pv,t}$ denotes the PV's output power at two consecutive time steps $t+1$ and t respectively. δ represents a positive number of small value. The reward +1 means that, the action to be taken for increment of P_{pv} for the agent, where the reward -1 is for decrement of P_{pv} .

B. Reinforcement Learning Algorithm for MPPT of PV System

The RL algorithm learns from its experience by interacting continuously through the environment and performs the best action after completion of learning period according to its return. The PV MPPT control problem is a deterministic problem since the PV transitions (i.e. change of state from s to s' according to action) will be the same for every state-action combination under the same environmental conditions. So a deterministic policy can be learned and the applied exploration strategy does not have to be constant to obtain an up to date model. Therefore, random exploration is necessary for state-action space before policy exploitation. The algorithm explores for a certain number of rounds according to the action space. The number of exploration rounds can be defined as $i = a * m$, where a is the number of action defined and m is the multiplier. The value of m should be large enough for more state action visits. When all the state actions are explored then the algorithm may be mature enough to converge to an optimal policy. The flowchart for the RL MPPT algorithm of PV system is shown in Fig. 2. The flowchart is described as follows.

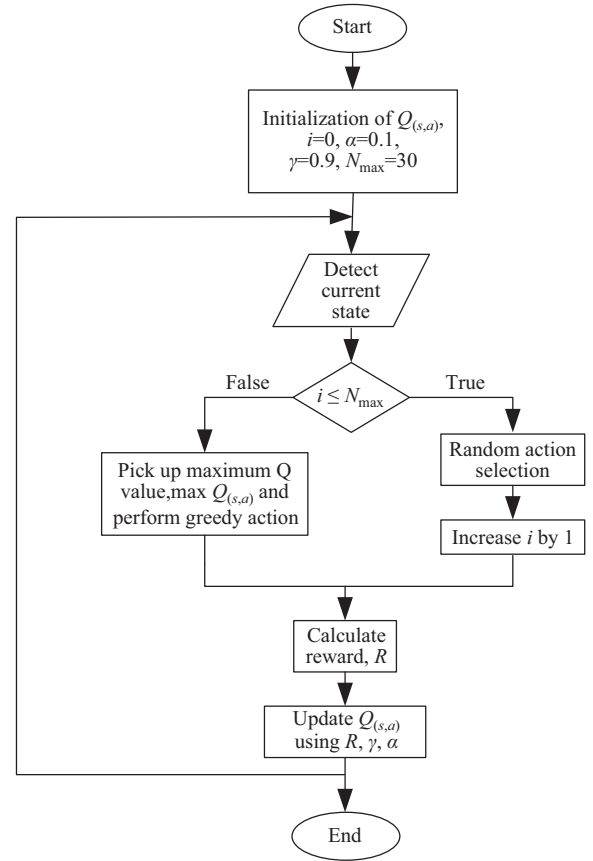


Fig. 2. Flow chart for proposed RL MPPT for PV system.

The Q-table can be initialized by simple setting of all initial values to zero. Then in each cycle, of learning V_{pv} , P_{pv} are sensed for determining the indexes of state in the Q-table. At that time action 'a' is taken randomly from the set of action. In the mean time, the parameter is increased by 1. The RL MPPT controller starts by exploring the different state action until m number of exploration rounds per state is completed. At the beginning of each sampling interval, a new state is found. The reward R as in (3) can be manipulated by taking the difference between two successive PV output powers. The maximum value in the next state is picked out, which is expressed as $\max_{a'} Q(s', a')$. Q-learning update rule is used to calculate the state action values as given below.

$$Q'(s,a) = Q(s,a) + \alpha[R + \gamma \max_{a'} Q'(s', a') - Q(s,a)] \quad (4)$$

where α is the learning rate and γ is the discount factor. Large learning rates will allow faster convergence, while however might cause oscillations between non-optimal values. Therefore, a small α is chosen, that will allow convergence. The discount factor γ will indicate the significance of future reward and therefore a close to 1 value should be chosen. At the end of exploration, an optimal policy is chosen for the best state action combination.

III. DESIGN OF SMC CONTROLLER FOR GCPVS

The circuit configuration for grid integration of PV system is shown in [17]. The SMC controller is used for tracking of actual inverter current to RIC as in [17].

IV. REFERENCE CURRENT GENERATION FOR VSI

IPT is implemented to generate RIC for VSI as in [17].

V. INCREMENTAL CONDUCTANCE MPPT ALGORITHM

The ICMPPT algorithm is usually used for its good tracking properties [18]. The flowchart for IC MPPT is shown in Fig. 2 in [4].

VI. SIMULATION RESULTS

The parameters of the grid connected PV system and controller are used for simulation which are given in Table 1 in [17]. The simulation is performed using power system tool box of MATLAB/SIMULINK software. The performance of RL-MPPT is compared with IC-MPPT and FL-MPPT [19], [20] and summarized in Section VII-C.

A. Performance of RL-SMC with Change in Irradiance

Figure 3(a) and Fig. 3(b) respectively show the output voltage and current of PV panel with change in irradiance at

0.2 s (from 700 W/m² to 300 W/m²), 0.4 s (from 300 W/m² to 700 W/m²), 0.6 s (from 700 W/m² to 300 W/m²) and 0.8 s (from 300 W/m² to 700 W/m²). Fig. 3(a) and Fig. 3(b) show that, when irradiance reduces, voltage reduces slightly but current reduces considerably. It is observed that, when irradiance increases from 300 W/m² to 700 W/m², the grid current reduces to its original value and inverter current increases to normal values, which is observed from in Fig. 3(c) and Fig. 3(d). For this reason, the load current is maintained constant, which is observed from Fig. 3(e). From Fig. 3(f) it is observed that, the grid voltage and current waveform of phase 'a' are in the same phase.

B. Comparison Between IC-SMC, FL-SMC and RL-SMC

From Fig. 4 it is seen that the THD of the grid current in case of IC-SMC is 3.27% and in FL-SMC 3.12% which is reduced to 2.95% in case of RL-SMC. Initially, the time taken for achieving the steady state in RL-SMC is 0.09 s, but in IC-SMC, 0.07 s and in FL-SMC, 0.08 s.

After learning, RL-SMC tracks MPP faster than IC-SMC

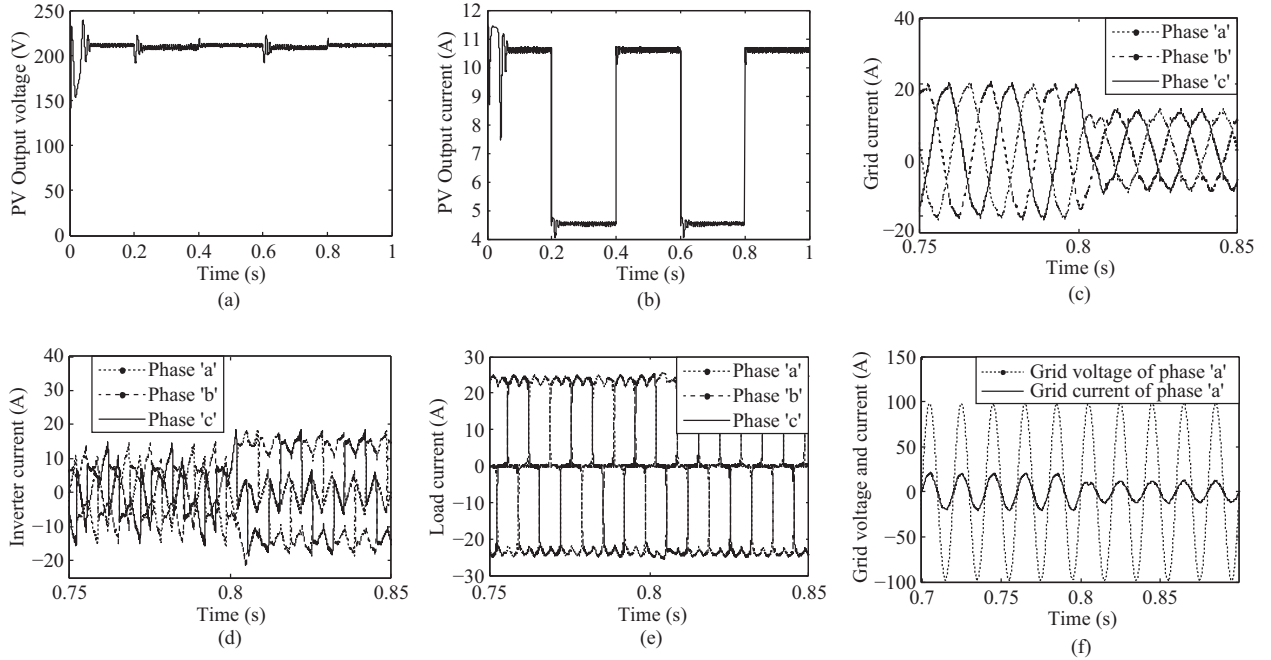


Fig. 3. Steady state performance of RL-SMC: (a) Output voltage (b) Output current (c) Grid current (d) Inverter current (e) Inverter voltage (f) Grid voltage and current.

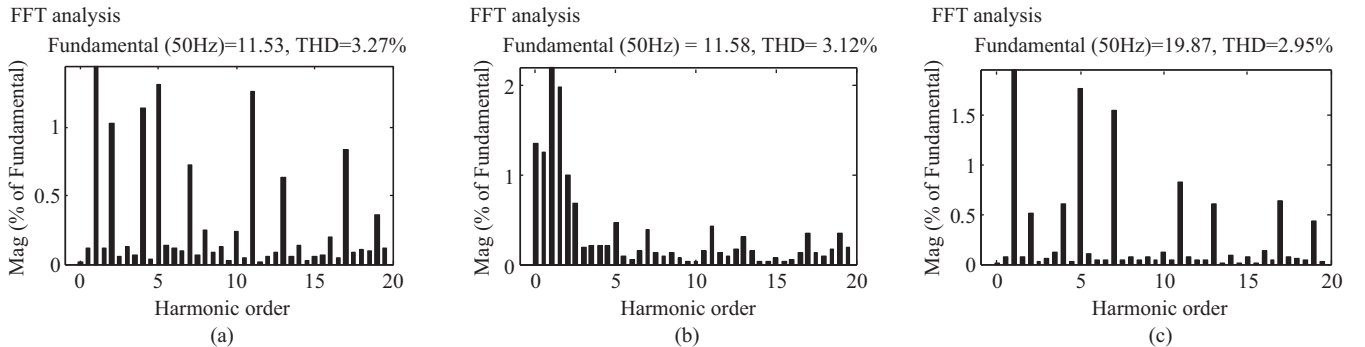


Fig. 4. THD performance of grid current: (a) THD in IC-SMC (b) THD in FL-SMC (c) THD in RL-SMC.

and FL-SMC at changing irradiance, which is shown in Fig. 5. Less ripple is resulted in case of RL-SMC as compared to IC-SMC and FL-SMC when the irradiance is changed from 700 W/m² to 300 W/m². From Fig. 6, it is observed that in

RL-SMC the time taken for reaching the steady state is 0.09 s where, it is faster (takes 0.07 s) in case of IC-SMC and (takes 0.08 s) in case of FL-SMC. But after learning, faster tracking is achieved in case of RL-SMC than the IC-SMC and FL-SMC,

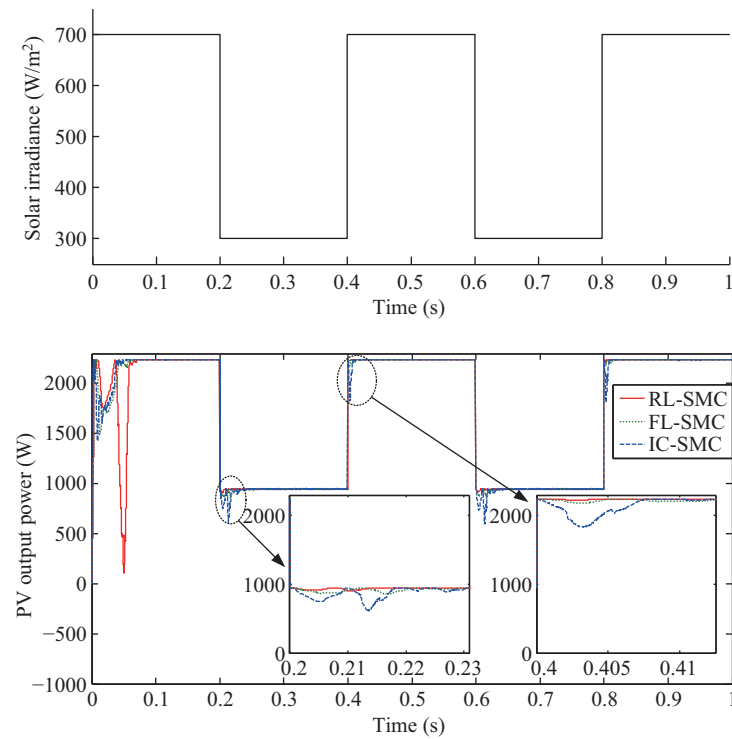


Fig. 5. PV output power at variable irradiance.

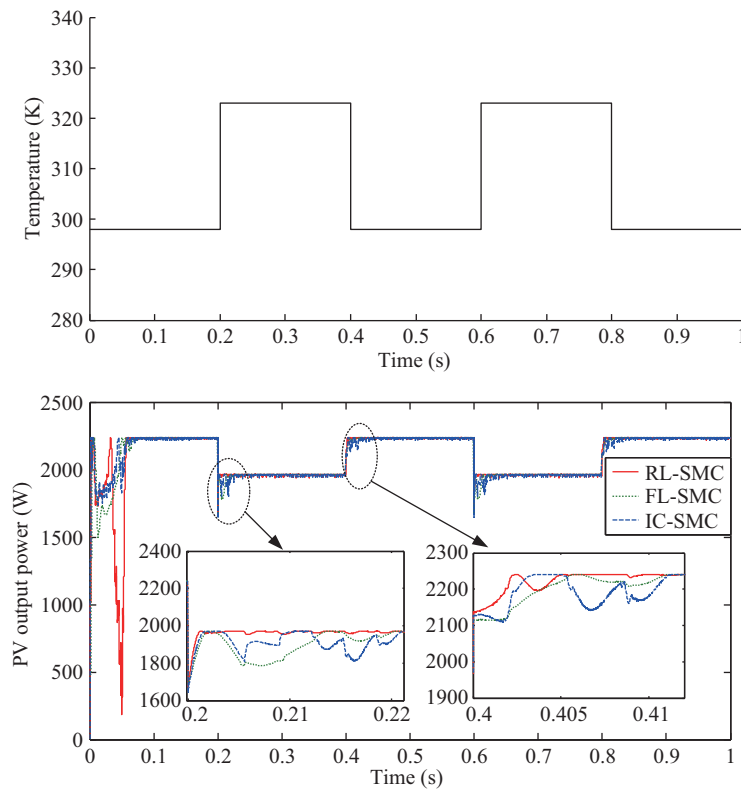


Fig. 6. PV output power at variable temperature.

when temperature changes from 298 K to 323 K. In Fig. 6, at the time 0.2 s, temperature of solar panel increases from 298 K to 323 K, in the mean time the undershoot in RL-SMC is less than IC-SMC and FL-SMC and yields less fluctuation after reaching the steady state. At 0.4 s, when temperature reduces from 323 K to 298 K, less fluctuation is seen in the PV output power in case of RL-SMC than IC-SMC and FL-SMC.

RL-SMC takes more time for learning from its own experience by taking random actions from the action space till all the state actions are explored. Once learned, RL-SMC outperforms in terms of faster tracking with taking the required perturbation. But in case of IC-SMC and FL-SMC, the perturbation is fixed. From the above observations, it is concluded that the RL-SMC algorithm exhibits better performance under changing atmospheric condition than IC-SMC and FL-SMC. RL-SMC exhibits better performance in terms of reduction of distortion in PV output power, reduction in THD and reaching the steady state faster than that of IC-SMC and FL-SMC schemes.

VII. EXPERIMENTAL RESULTS

An experimental setup has been developed in our Renewable Power Control Laboratory at NIT Rourkela. The details of parameter for experimental set up are given in [17]. Fig. 7 shows the block diagram for experimental setup for GCPVS. The detailed picture of the experimental setup is shown in Fig. 7 in [17].

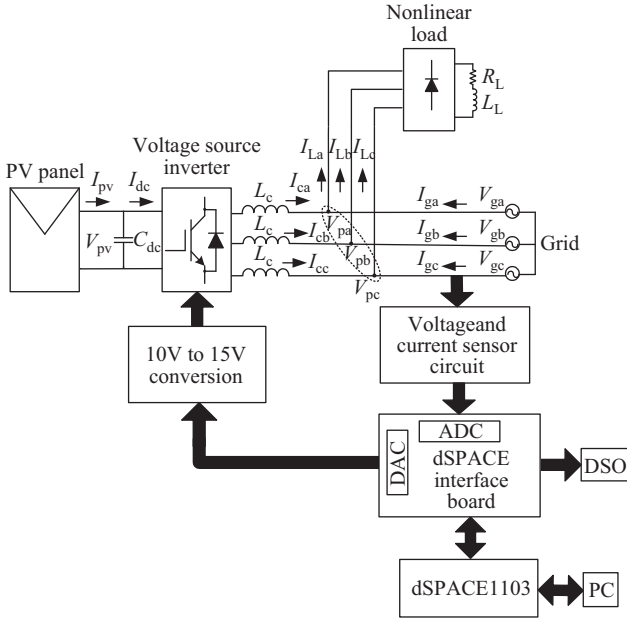
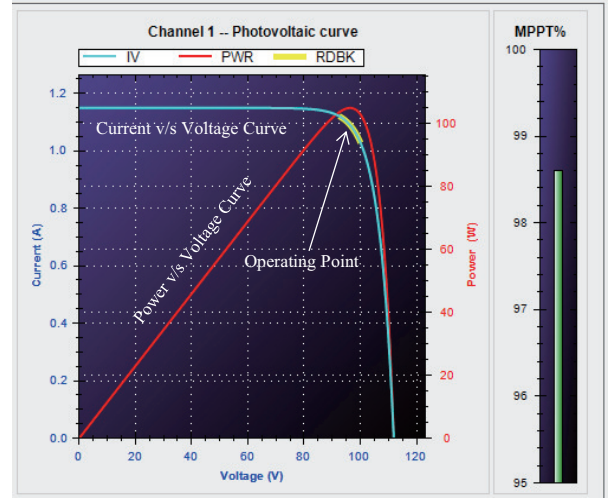


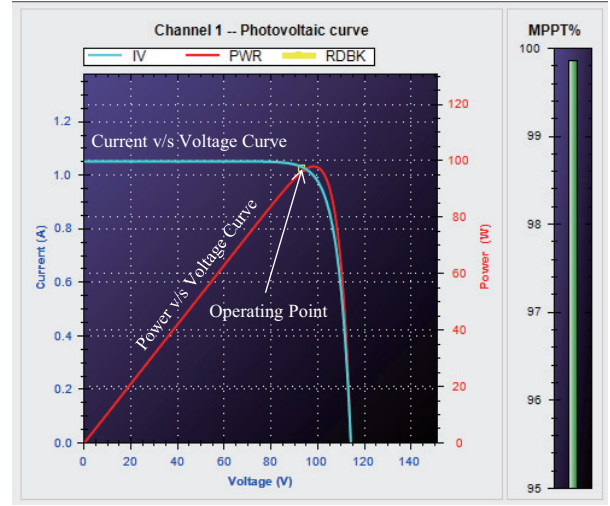
Fig. 7. Block diagram for experimental setup of GCPVS.

A. Steady State Performance of RL-SMC

The performance under steady state of RL-SMC is verified in the aforementioned experimental setup and compared with that of IC-SMC and FL-SMC. Fig. 8(a) shows the P-V and I-V curves for solar array simulator with 98.6% of MPPT achieved by IC-SMC, where as Fig. 12(a) shows the P-V and



(a) Performance of MPPT tracking in RLMPPT algorithm



(b) Performance of MPPT tracking in ICMPT algorithm

Fig. 8. MPP tracking performance: (a) Performance of MPPT for IC-SMC (b) Performance of MPPT for RL-SMC.

I-V characteristic curves for solar array simulator with 99.04% of MPPT achieved by FL-SMC and Fig. 8(b) shows the P-V and I-V curves for solar array simulator with 99.8% of MPPT achieved by RL-SMC. Fig. 9(a)–(c) shows the grid current, inverter current and load current of phase ‘a’ respectively. The total current taken by load is 2.697 A. Out of the load current 2.693 A, the PV-VSI combination supplies 1.191 A and 1.524 A supplies by grid. The total consumption of load active power is 203 W. Out of 203 W, PV-VSI combination supplies 95 W and grid supplies 101 W, which is shown in Fig. 9(d)–(f). The load reactive power is compensated by PV-VSI combination, so that the grid power factor is unity. The THD of grid current is 3.3%, which was before compensation same as the load current 25.3% and the inverter current THD is 23.4% as in Fig. 9(g)–(i). Fig. 10(a) represents for phase ‘a’ grid voltage and current, which are in the same phase. The dc capacitor voltage is shown in Fig. 10(b).

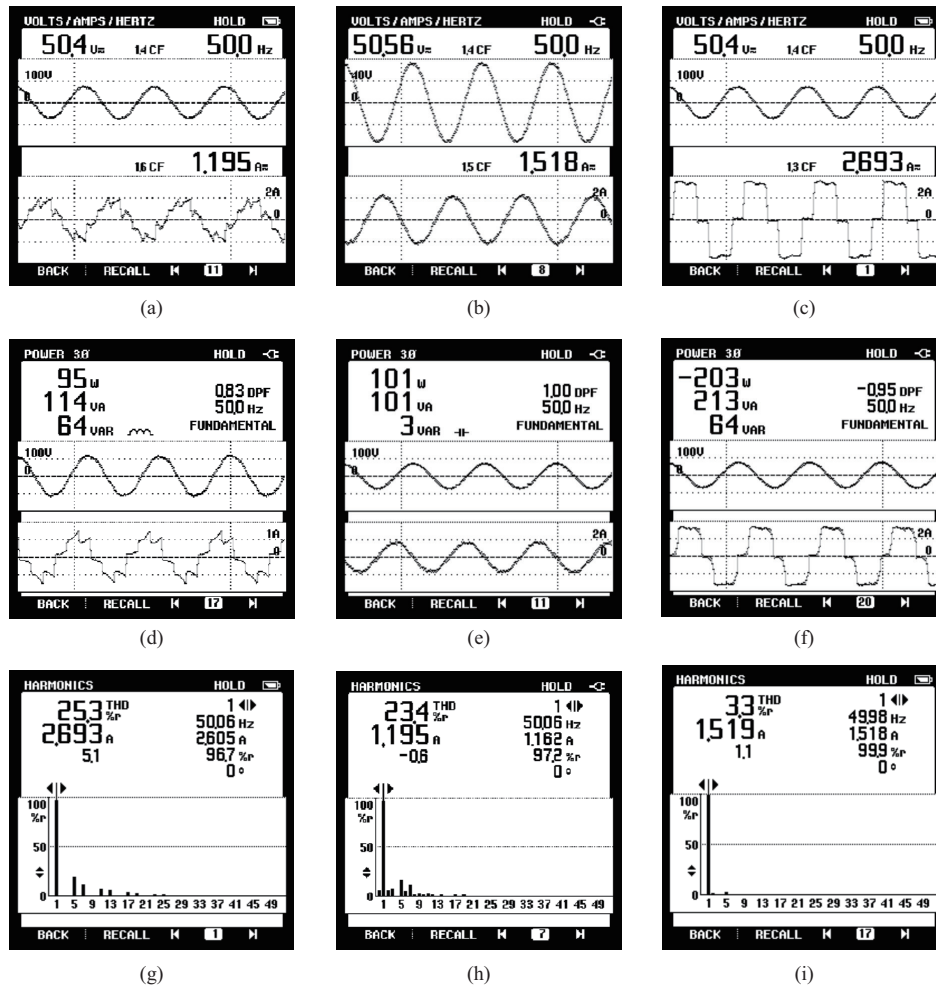


Fig. 9. Steady state performance of RL-SMC: (a)–(c) Grid output voltage (L-L) and grid output current, Inverter output current, Load current; (d)–(f) Active power, Reactive power and power factor of grid, VSI and load; (g)–(i) THD for load current, inverter current and grid current.

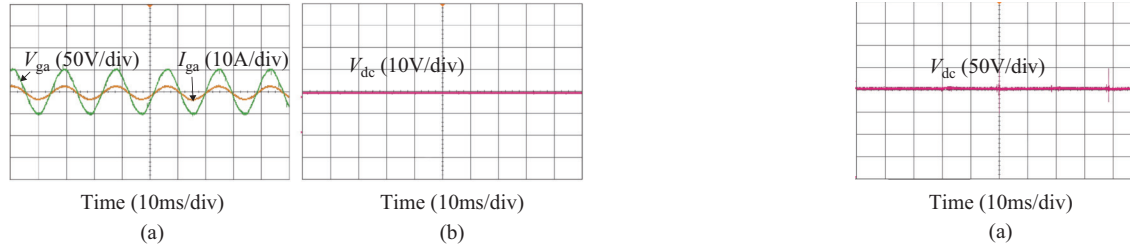


Fig. 10. Steady state performance of RL-SMC: (a) Voltage and Current of grid (b) DC link capacitor voltage in RL-SMC.

B. Comparison Between IC-SMC, FL-SMC and RL-SMC

Figure 11(a) and Fig. 11(b) respectively show the grid voltage (L-L), grid current for phase 'a' and grid current THD for phase 'a' at steady state of IC-SMC. The dc capacitor voltage for IC-SMC is shown in Fig. 11(c). The dc capacitor voltage in case of IC-SMC contains spikes, which can be observed from Fig. 11(c). Fig. 12(b) and Fig. 12(c) respectively show the dc capacitor voltage and grid current THD for phase 'a' at steady state of FL-SMC. The dc capacitor voltage in case of FL-SMC contains spikes, which can be observed from Fig. 12(c). The current waveform obtained by applying RL-SMC is found to

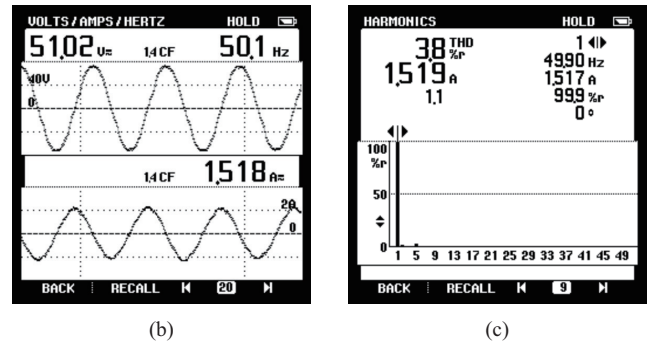


Fig. 11. Steady state performance of IC-SMC: (a)–(c) dc link capacitor voltage, grid voltage (L-L) and grid current, THD of grid current.

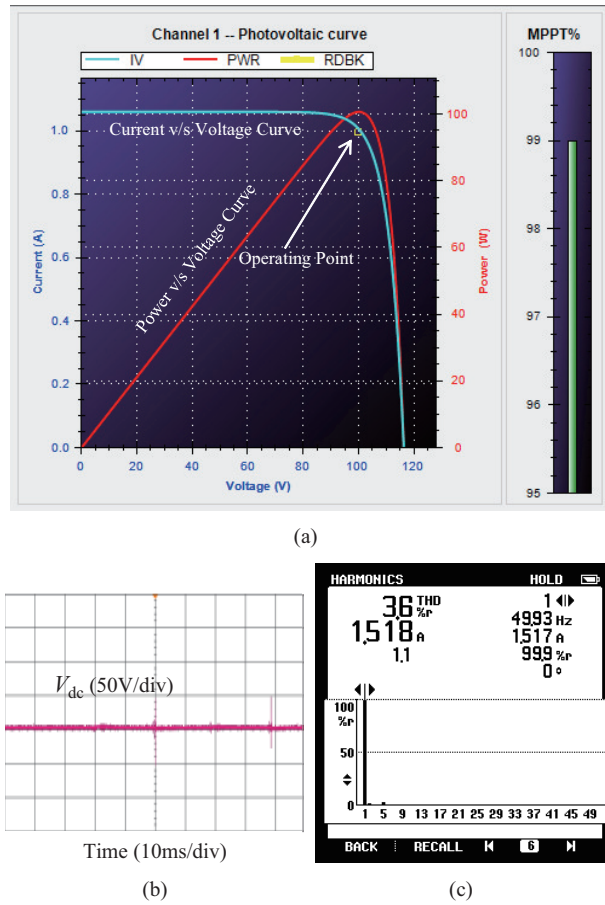


Fig. 12. Steady state performance of FL-SMC: (a) Performance of MPPT (b) DC link capacitor voltage (c) THD of grid current.

be sinusoidal, but in the case of IC-SMC and FL-SMC, it is a distorted sinusoidal waveform. The THD of grid current of RL-SMC is 3.3%, which is found to be less than the THD obtained in IC-SMC i.e. 3.8% and FL-SMC i.e. 3.6%.

RL-SMC is compared with IC-SMC and FL-SMC as shown in Section VII subsection-C satisfies the IEEE 519 standards. From subsection.C, it is observed that RL-SMC outperforms once it is learned than the IC-SMC and FL-SMC in terms of more power extraction, reduced THD and reduced ripple.

C. Comparison Between IC-SMC, FL-SMC and RL-SMC

IC-SMC: online learning ability: no; MPPT Performance (%): 98.60; Nature of grid current waveform: distorted sinusoidal waveform; Duration (s) for sustain of ripple in PV output power with application of change in irradiance: 0.023; THD (in%) under steady state (simulation): 3.27; THD (in%) under steady state (experimental): 3.8; Duration(s) for sustain of ripple content in PV output power with application of change in temperature: 0.02; Ripple present in dc link voltage: present. FL-SMC: online learning ability: no; MPPT Performance (%): 99.04; Nature of grid current waveform: better waveform than IC-SMC; Duration (s) for sustain of ripple in PV output power with application of change in Irradiance: 0.022; THD (in%) under steady state (simulation): 3.12; THD (in%) under steady state (experimental): 3.6; Duration(s) for sustain of ripple content in PV output power with application

of change in Temperature: 0.015; Ripple present in dc link voltage: present. RL-SMC: online learning ability: no; MPPT Performance (%): 99.80; Nature of grid current waveform: sinusoidal waveform; Duration (s) for sustain of ripple in PV output power with application of change in Irradiance: 0.005; THD (in%) under steady state (simulation): 2.95; THD (in%) under steady state (experimental): 3.3; Duration(s) for sustain of ripple content in PV output power with application of change in Temperature: 0.005; Ripple present in dc link voltage: no ripples.

VIII. CONCLUSION

A new RL-SMC scheme is proposed for integration of a PV system to grid, in which the MPPT is achieved by Reinforcement Learning algorithm for obtaining maximum power form the PV panel under changing solar irradiances and temperature. Further, a sliding mode controller is designed that connects the VSI, for robust performance satisfying the IEEE-929 standard. RL-SMC scheme outperforms under changing atmospheric condition as compared to IC-SMC and FL-SMC schemes. RL-SMC provides better MPPT tracking performance after properly learned as compared to IC-SMC and FL-SMC, which can be observed from the simulation and experimental result. The RL-SMC yields more output power as compared to IC-SMC and FL-SMC. The THD of grid current in RL-SMC is lower than that yielded in case of IC-SMC and FL-SMC, which is within IEEE-519 standard.

REFERENCES

- [1] V. Salas, E. Olías, A. Barrado, and A. Lazaro, "Review of the maximum power point tracking algorithms for stand-alone photovoltaic systems," *Solar Energy Materials and Solar Cells*, vol. 90, no. 11, pp. 1555–1578, Jan. 2006.
- [2] B. Subudhi and R. Pradhan, "A comparative study on maximum power point tracking techniques for photovoltaic power systems," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 1, pp. 89–98, Jan. 2013.
- [3] T. Esmar and P. L. Chapman, "Comparison of photovoltaic array maximum power point tracking techniques," *IEEE Transactions on Energy Conversion*, vol. 22, no. 2, pp. 439–449, Jun. 2007.
- [4] A. Bag, B. Subudhi, and P. K. Ray, "Grid integration of pv system with active power filtering," in *2016 2nd International Conference on Control, Instrumentation, Energy & Communication (CIEC)*. IEEE, 2016, pp. 372–376.
- [5] M. A. Elgendy, B. Zahawi, and D. J. Atkinson, "Assessment of perturb and observe mppt algorithm implementation techniques for pv pumping applications," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 1, pp. 21–33, Jan. 2012.
- [6] M. A. Elgendy, B. Zahawi, and D. J. Atkinson, "Operating characteristics of the perturb and observe algorithm at high perturbation frequencies for standalone pv systems," *IEEE Transactions on Energy Conversion*, vol. 30, no. 1, pp. 189–198, Mar. 2015.
- [7] M. A. Elgendy, B. Zahawi, and D. J. Atkinson, "Assessment of the incremental conductance maximum power point tracking algorithm," *IEEE Transactions on Sustainable Energy*, vol. 4, no. 1, pp. 108–117, Jan. 2013.
- [8] P. Singh, D. Palwalia, A. Gupta, and P. Kumar, "Comparison of photovoltaic array maximum power point tracking techniques," *Int. Adv. Res. J. Sci. Eng. Technol.*, vol. 2, no. 1, pp. 401–404, Jan. 2015.
- [9] A. Mellit and S. A. Kalogirou, "Mppt-based artificial intelligence techniques for photovoltaic systems and its implementation into field programmable gate array chips: Review of current status and future perspectives," *Energy*, vol. 70, pp. 1–21, Jun. 2014.
- [10] C. Wei, Z. Zhang, W. Qiao, and L. Qu, "Reinforcement-learning-based intelligent maximum power point tracking control for wind energy conversion systems," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 10, pp. 6360–6370, Oct. 2015.

- [11] R. C. Hsu, C.-T. Liu, W.-Y. Chen, H.-I. Hsieh, and H.-L. Wang, "A reinforcement learning-based maximum power point tracking method for photovoltaic array," *International Journal of Photoenergy*, vol. 2015, pp. 1–12, Mar. 2015.
- [12] A. Youssef, M. E. Telbany, and A. Zekry, "Reinforcement learning for online maximum power point tracking control," *Journal of Clean Energy Technologies*, vol. 4, no. 4, pp. 245–248, Jul. 2016.
- [13] P. Kofinas, S. Doltsinis, A. Dounis, and G. Vouros, "A reinforcement learning approach for mppt control method of photovoltaic sources," *Renewable Energy*, vol. 108, pp. 461–473, Aug. 2017.
- [14] F. Blaabjerg, R. Teodorescu, M. Liserre, and A. V. Timbus, "Overview of control and grid synchronization for distributed power generation systems," *IEEE Transactions on industrial electronics*, vol. 53, no. 5, pp. 1398–1409, Oct. 2006.
- [15] N. D. Tuyen and G. Fujita, "Pv-active power filter combination supplies power to nonlinear load and compensates utility current," *IEEE Power and Energy Technology Systems Journal*, vol. 2, no. 1, pp. 32–42, Mar. 2015.
- [16] B. Singh and S. R. Arya, "Back-propagation control algorithm for power quality improvement using dstatcom," *IEEE transactions on industrial electronics*, vol. 61, no. 3, pp. 1204–1212, Mar. 2014.
- [17] A. Bag, B. Subudhi, and P. K. Ray, "Comparative analysis of sliding mode controller and hysteresis controller for active power filtering in a grid connected pv system," *International Journal of Emerging Electric Power Systems*, vol. 19, no. 1, 2018.
- [18] K. Visweswara, "An investigation of incremental conductance based maximum power point tracking for photovoltaic system," *Energy Procedia*, vol. 54, pp. 11–20, Aug. 2014.
- [19] B. N. Alajmi, K. H. Ahmed, G. P. Adam, and B. W. Williams, "Single-phase single-stage transformer less grid-connected pv system," *IEEE Transactions on Power Electronics*, vol. 28, no. 6, pp. 2664–2676, 2013.
- [20] T. Yetayew and T. Jyothsna, "Evaluation of fuzzy logic based maximum power point tracking for photovoltaic power system," in *Power, Communication and Information Technology Conference (PCITC), 2015 IEEE*. IEEE, 2015, pp. 217–222.



Pravat Kumar Ray (M'13–SM'18) received the B.S. degree in Electrical Engineering from Indira Gandhi Institute of Technology Sarang, Odisha, India, in 2000, the M.E. degree in Electrical Engineering from Indian Institute of Engineering Science and Technology, Shibpur, Howrah, India, in 2003, and the Ph.D. degree in Electrical Engineering from NIT Rourkela, India, in 2011. He is currently an Associate Professor with the Department of Electrical Engineering, NIT Rourkela. He was also a postdoctoral fellow at Nanyang Technological University, Singapore during Jan 2016 to June 2017. His research interests include signal processing and soft computing applications to power system, power quality and grid integration of renewable energy systems.



Aurobinda Bag received B.E. degree in Electrical Engineering from Biju Patnaik University of Technology, Rourkela, Odisha, India, in 2005, M.E. degree in Power System Engineering from Veer Surendra Sai University of Technology (VSSUT), Burla, Odisha, India, in 2012 and Ph.D degree in Electrical Engineering from National Institute of Technology (NIT), Rourkela, Odisha, India. He is currently Senior Assistant Professor in Madanapalle Institute of Technology, Andhra Pradesh, India. His research interests include active power filtering and

control of grid integrated PV system.



Bidyadhar Subudhi (M'94–SM'08) received Bachelor Degree in Electrical Engineering from National Institute of Technology (NIT), Rourkela, India, Master of Technology in Control & Instrumentation from Indian Institute of Technology, Delhi, India in 1988 and 1994 respectively and Ph.D. degree in Control System Engineering from Univ. of Sheffield in 2003. He was a post doc in the Dept of Electrical and Computer Engineering, National University of Singapore in 2005. Currently he is a professor in the School of Electrical Science at IIT Goa. He is a

Senior Member of IEEE and Fellow IET. His research interests include system and control theory, control of PV power system and active power filtering.