01 Aug 2007

Optimal Control of a Photovoltaic Solar Energy System with Adaptive Critics

Richard L. Welch

Ganesh K. Venayagamoorthy
Missouri University of Science and Technology

Recommended Citation

R. L. Welch and G. K. Venayagamoorthy, "Optimal Control of a Photovoltaic Solar Energy System with Adaptive Critics," *Proceedings of International Joint Conference on Neural Networks, 2007*, Institute of Electrical and Electronics Engineers (IEEE), Aug 2007.
The definitive version is available at https://doi.org/10.1109/IJCNN.2007.4371092

Proceedings of International Joint Conference on Neural Networks, Orlando, Florida, USA, August 12-17, 2007

Optimal Control of a Photovoltaic Solar Energy System with Adaptive Critics

Richard L. Welch, *Student Member, IEEE* and Ganesh K. Venayagamoorthy, *Senior Member, IEEE*

Abstract – **This paper presents an optimal energy control scheme for a grid independent photovoltaic (PV) solar system consisting of a PV array, battery energy storage, and time varying loads (a small critical load and a larger variable non-critical load). The optimal controller design is based on a class of Adaptive Critic Designs (ACDs) called the Action Dependant Heuristic Dynamic Programming (ADHDP). The ADHDP class of ACDs uses two neural networks, an "Action" network (which actually dispenses the control signals) and a "Critic" network (which critics the Action network performance). An optimal control policy is evolved by the action network over a period of time using the feedback signals provided by the critic network. The objectives of the optimal controller in order of decreasing importance is to first fully dispatch the required energy to the critical loads at all times; secondly to dispatch energy to the battery whenever necessary so as to be able to dispatch energy to the critical loads in any absence of energy from the PV array; and lastly to dispatch energy to the non-critical loads while not interfering with the first two objectives. Results on three different US cities show that the ADHDP based optimal control scheme outperforms the conventional PV-priority control scheme in maintaining the stated objectives almost all the time.**

I. Introduction

As the cost of fossil fuels rise and their availability falls, it is becoming important to look for alternate forms of energy. Currently, there are many alternative energy sources, including wind, solar, hydroelectric and geothermal. Of these, solar energy is perhaps the most well suited to employ on a wide scale, both supplying energy and possibly lowering stresses on the power grid through distributed generation. Additionally, photovoltaic (PV) arrays have no moving parts and therefore require very little maintenance and generally perform reliably while the sun is shining.

As the price of solar energy falls [1] through higher production volumes and technology improvements, its adoption rate has increased. However, even in light of rising utility prices, solar energy is still relatively expensive. The payback time (the time it takes for a PV installation to pay for itself) can be as high as 30 years (or more). Fortunately, the life span of many PV arrays usually matches this time.

In order to make the system cheaper, and hence shorten the payback period, optimal control can be employed to better dispatch the energy from the PV array to the system loads and battery storage. This optimal control can lead to a system with a smaller, less costly solar array while still powering all of the critical loads, such as critical refrigeration or communications equipment.

Traditionally, the energy control that is employed for PV systems is called the "PV-Priority" control scheme [2] and simply uses all available energy from the PV array to power the loads, and if there is any excess energy then it is stored in the battery, and if there is not enough energy coming from the PV array to power the loads then energy from the battery is used. Other types of energy controllers have been reported, such as a controller using Q-learning [2] and another using fuzzy logic [3].

In this paper, the proposed optimal energy dispatch controller is based on an adaptive critic design (ACD) approach called action dependant heuristic dynamic programming, or ADHDP [4, 5, 6]. Adaptive critic designs are based on a combination of reinforcement learning and dynamic programming. The ADHDP topology is the simplest form of ACD and the computationally simplest, using only 2 neural networks, one called the action (or actor) and the other called the critic. The objectives of the optimal controller in order of decreasing importance is to first fully dispatch the required energy to the critical loads at all times; secondly to dispatch energy to the battery whenever necessary so as to be able to dispatch energy to the critical loads in any absence of energy from the PV array; and lastly to dispatch energy to the non-critical loads while not interfering with the first two objectives.

Section II presents the grid independent PV solar energy system studied in this paper. Section III describes the standard PV-priority controller. Section IV describes the ADHDP optimal controller design. Section V presents the evaluation and comparison of the ADHDP optimal PV controller and the standard PV-priority controller performances on Typical Meteorological Year (TMY) data of three cities in the United States of America. Finally, the conclusion is given in Section VI.

II. Grid Independent PV Solar Energy System

The complete photovoltaic system model is composed of the PV array, maximum power point tracker, controller, battery charge controller, batteries, inverter, critical loads and

---

The support from the National Science Foundation under the CAREER grant ECCS # 0348221 is gratefully acknowledged by the authors.

Richard L. Welch and Ganesh K. Venayagamoorthy are with the Real-Time Power and Intelligent Systems Laboratory (www.ece.umr.edu/RTPIS), Department of Electrical and Computer Engineering, University of Missouri-Rolla, MO 65409 USA (e-mails: rwelch@ieee.org and gkumar@ieee.org).

non-critical loads. The critical load consists of loads that should not be dropped (such as refrigeration, emergency radio communication), while the non-critical load contains items which are non-essential (television, etc).

In order to simplify the simulation and focus on the controller aspect of this system, all of the supporting system components (such as the inverter, maximum power point tracker, wiring, etc), are assumed to operate at 100% efficiency. Also, the efficiency of the PV array model is taken as 11% to account for various non-optimal conditions (such as array misalignment, dust on the arrays, etc). This value is representative of the current commercially available range of efficiencies for PV arrays. Generally, PV panels vary in efficiency from 6% to up to 30%; although the high efficiency panels are generally reserved for spacecraft usage because of their high radiation tolerances and higher power-to-weight ratio. A rough equivalent to the PV arrays being simulated in this paper would be an array of eight Kyocera KC200GT panels. These panels are over 16% efficient and will output 200W during optimal conditions [7]. The minimum charge for the battery of 30% is required to supply energy to the loads (this is consistent with standard deep cycle lead-acid batteries).

Due to insufficient and no PV energy during winter months and nights respectively, a control system is required to decide the amount of energy to be dispatched the different loads including the charging of the battery. The complete system in schematic diagram form is shown below in Fig. 1 (energy flows in the direction of the arrows).

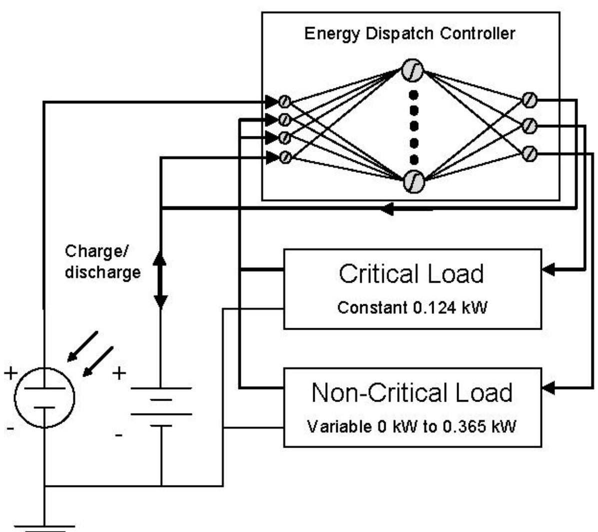

Fig. 1. Schematic diagram of the PV system model (this control is applicable only when there is insufficient PV collector energy to supply the critical loads, non-critical loads and the charge the battery).

III. PV-Priority Controller

The standard controller called the "PV-Priority" controller is a very simple controller which always tries to meet the loads (the critical and then the non-critical) before charging the battery. At any one time, if there is not enough energy from the PV array to supply the loads then the balance is drawn from the battery. If instead there is an excess, then whatever is left over after supplying the loads is dispatched to the battery. In this way, the controller will attempt to power all loads and charge the battery as best it can, without any considerations given to the time varying states of the system.

This controller works well when there is sufficient PV energy. However, when there is not sufficient PV energy, then the battery will not be fully recharged and the loads will be dropped. The weather and user loads are stochastic in nature; therefore there is no one definitive model at all times. Thus, it makes sense to look at intelligent model-free learning methods of controlling such a system.

IV. ADHDP Optimal Controller

One such intelligent system can involve the use of adaptive critic designs. ACDs utilize neural networks and are capable of optimization over time in conditions of noise and uncertainty. A family of ACDs was proposed by Werbos [4] as a new optimization technique combining the concepts of approximate dynamic programming and reinforcement learning. With ACDs, for a given series of control actions that must be taken sequentially (and not knowing the effect of these actions until the end of the sequence), it is possible to design an optimal controller using the traditional supervised learning based neural network.

The adaptive critic method determines an optimal control for a system by adapting two neural networks: an *Action* network and a *Critic* network. The Action network is responsible for driving the system to the desired states, while the Critic network is responsible for providing the Action network with performance feedback with respect to reaching the desired states over time. With this feedback, the Action network is able to adapt its parameters continuously to maximize its objective. The Critic network learns to optimize the Action network by approximating the Hamilton-Jacobi-Bellman equation associated with optimal control theory.

This Actor-Critic adaptation process starts with a non-optimal or suboptimal policy by the action network; the Critic network then guides the Action network toward an optimal solution at each successive adaptation. During the adaptations, neither of the networks needs any "information" of an optimal trajectory, only the desired cost needs to be known. Furthermore, this method determines optimal control policy for the entire range of initial conditions. Additionally, it needs no external training, unlike other neural-controllers [5].

The design ladder of ACDs includes three basic implementations: Heuristic Dynamic Programming (HDP), Dual Heuristic Programming (DHP) and Globalized Dual Heuristic Programming (GDHP), in the order of increasing power and complexity. The interrelationships between members of the ACD family have been generalized and explained in [6]. In this paper, an Action dependent HDP (ADHDP) approach is adopted for the design of an optimal PV controller. Action dependent adaptive critic designs do not need system models to develop the optimal control policy (action network output).

As mentioned, the objective of the optimal PV control is threefold - to maximize or fully dispatch the required energy to the critical loads at all times, dispatch energy to charge the battery whenever necessary so as to dispatch energy to the critical loads in the absence of energy from the collector and the last objective is to dispatch energy to the non-critical loads not comprising on the first two objectives. The optimal controller is not used for instances where there is sufficient solar energy to power all loads as well as completely charge the battery. When this occurs, all loads are satisfied and the battery is completely charged.

This optimal controller uses two networks (the Action and Critic networks) as previously mentioned. The inputs to the Action network correspond to the states of the system while the outputs correspond to the amount of energy to be dispatched to the critical loads, battery and non-critical loads. The inputs to the Critic consist of the inputs to the Action network at time *t*, *t-1*, and *t-2*, as well as the outputs of the Action network at time *t*, *t-1*, and *t-2*. The Critic then uses the information from the current states and actions in the current time step (as well as from the recent past) to derive the Action network over time to evolve an optimal control policy. Fig. 2 shows the connection between the Action network, Critic network and the PV system.

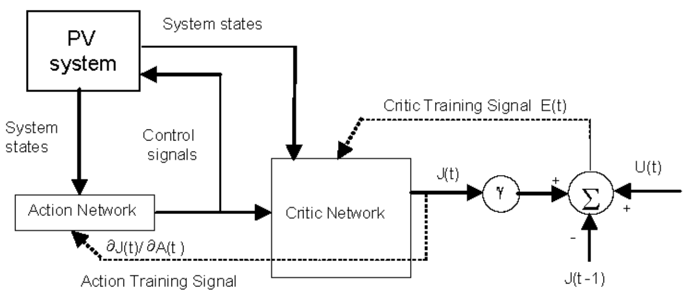

Fig. 2. Structure of the ADHDP based optimal PV controller design.

A. Critic Neural Network

The Critic network is a multilayer feedforward network trained using the standard backpropagation (BP) training algorithm. The input, hidden and output layers consists of twenty-two linear neurons, twenty sigmoidal neurons and one linear neuron respectively. As previously mentioned, the inputs to the Critic network are the outputs and inputs of the action network, at times *t*, *t-1* and *t-2*. A diagram of the Critic network is shown in Fig. 3.

The output of the critic network is the estimated *cost-to-go* function *J* of Bellman's equation of dynamic programming, which is given by (1).

$$J(t)=\sum_{i=0}^{\infty}\gamma^{i}U(t+i) \tag{1}$$

Where γ is the discount factor for finite horizon problems with the range of [0, 1] and is chosen to be 0.8 in this study. *U(t)* is known as the utility function or the local cost function. This utility function guides the Critic in critiquing the Actor's performance. In this study, *U(t)* in (2) is chosen to be a function of critical load (CL), state of battery charge (BC) and non-critical load (NCL).

$$\begin{aligned} U(t) = {} & (30/23)*abs(1-(ECL/(CL+M*MCL)))+\\ & (15/23)*abs(1-(EB/((MBC-CBC)+M*MBC)))+\\ & (13/23)*abs(1-(ENCL/(NCL+M*MNCL))) \end{aligned} \tag{2}$$

Where:

ECL = Energy Dispatched to the Critical Load
CL = Critical Load
MCL = Maximum Critical Load
EB = Energy Dispatched to the Battery
MBC = Maximum Battery Charge
CBC = Current Battery Charge
ENCL = Energy Dispatched to the Non Critical Load
NCL = Non Critical Load
MNCL = Maximum Non Critical Load
M = Multiplier (used to ensure divisor is non-zero; for this experiment, a value of 0.1 was used).

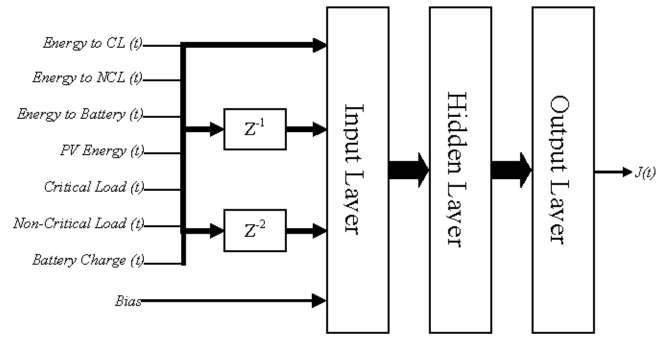

Fig. 3. Critic neural network.

In the *U(t)* function given in (2), a higher priority is given to meeting the critical load at all times over the batteries being charged or the non-critical load being supplied by assigning different weightings - 30/23 to the CL term, 15/23 to the BC term, and 13/23 to the NCL term. This U(t) meets the threefold objective for the optimal PV controller design.

In the training of the Critic network, the objective is to minimize (3) given below.

$$\sum_{t=0}^{\infty} E^2(t) \tag{3}$$

where

$$E(t)=U(t)+\gamma J(t)-J(t-1) \tag{4}$$

The weight change and update equations for the Critic network using the BP algorithm is given by (5) and (6) respectively.

$$\Delta W_c(t)=\eta_c . E(t).\frac{\partial J(t)}{\partial W_c} \tag{5}$$

$$W_c(t+1)=W_c(t)+\Delta W_c(t) \tag{6}$$

Where η_c and W_c are the learning rate and the weights of the Critic neural network respectively.

B. Action Neural Network

The Action network is a multilayer feedforward network trained using the BP algorithm. The input, hidden and output layers of the Action network consists of five linear neurons, thirty sigmoidal neurons and three linear neurons respectively, as is shown in Fig. 4. The Action network inputs consist of the following:

- Solar energy from the PV array (as a fraction of total possible energy from the PV array)
- Critical load (as a fraction of total load)
- Non-critical load (as a fraction of total load)
- Current battery charge (as a fraction of total charge)

- Bias term.

The Action network outputs consist of the following:

- Energy dispatched to the critical load (ECL)
- Energy dispatched to the non-critical load (ENCL)
- Energy dispatched to the battery (EB); this can be positive or negative, depending on whether the battery is being charged or being used as a source.

Additionally, the Action network's outputs are checked to ensure the sum of energy dispatched is no more than is available at the inputs. This is accomplished by performing the following series of steps immediately after calculating the outputs from the action network:

i). Verify that the energy dispatched to each of the loads does not exceed the load demand, and is not negative. Also ensure that the energy to the battery is not higher than the energy collected by the PV arrays.
ii). Verify that the battery is not being overcharged, or over depleted.
iii). The outputs (including the energy dispatched to the battery if it is being charged) are scaled by the ratio of energy inputs to outputs.
iv). After scaling (step iii), another round of checks is made on the Action network outputs in order to be certain that they are not greater than the load or less than zero.

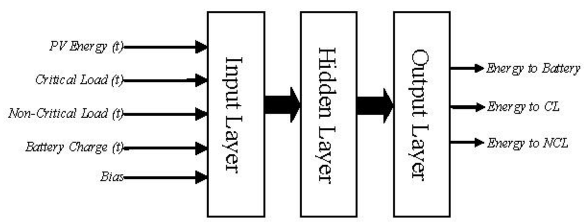

Fig. 4. Action neural network.

The change in the Action network weights ΔW_A are calculated by backpropagating the current $J(t)$ value back through the trained Critic network as shown in Fig. 2 to obtain $\partial J/\partial A$. The error in the Action network output is given by (7).

$$E_A(t) = \partial J(t) / \partial A(t) \tag{7}$$

The change in the Action network's weights ΔW_A obtained using the BP algorithm and update weight equations are given by (8) and (9) respectively.

$$\Delta W_A(t) = \eta_A . E_A(t) . \frac{\partial A(t)}{\partial W_A} \tag{8}$$

$$W_A(t+1) = W_A(t) + \Delta W_A(t) \tag{9}$$

Here η_A and W_A are the learning rate and the weights of the Action neural network respectively.

C. Actor/Critic Training

The flowchart in Fig. 5 outlines the training steps for both the Critic and Action networks. During the iterative training phase, several metrics can be used to determine if the Actor's performance is increasing. For this study, the simple sum of the utility function for each cycle of training the Action network is used. This means that when the sum of the utility function is decreasing, the performance of the Action network is improving. The simulation is run for a fixed number of iterations, but if the sum of the utility function increases during training then the new Action network weights are discarded and replaced with the previous best weights. When this happens, a very small perturbation (a random number between -0.01 and 0.01) is added to the Action network weights such that the network avoids getting stuck in a local minimum.

After the best Action network weights are found, these weights are then used to optimally dispatch energy to the critical loads, the non-critical loads and the battery.

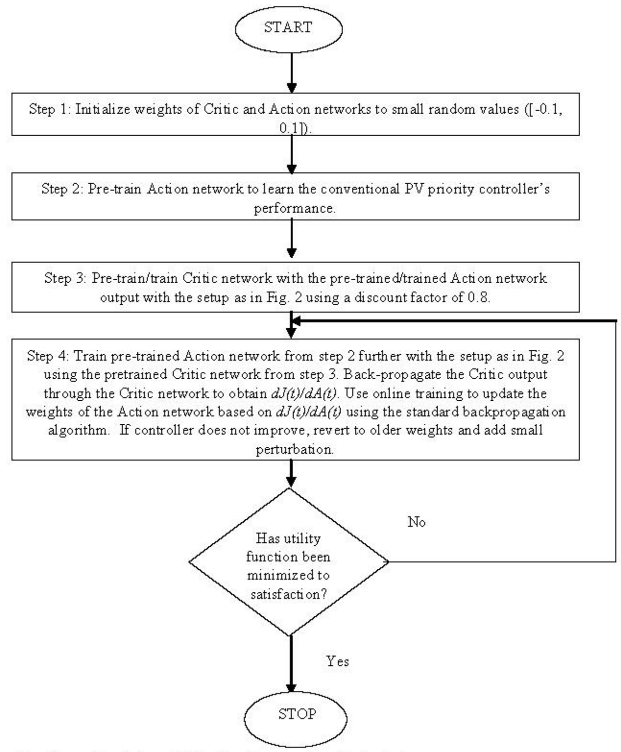

Fig. 5. Flowchart for Critic/Action network training.

V. Results

A one year simulation of the PV system is carried out for the following areas: Phoenix, AZ, Springfield, MO and Caribou, ME. These simulations used data from the TMY2 database [8].

Phoenix receives more solar radiation than Springfield, and that Springfield receives more solar radiation than Caribou.

The PV energy captured by the solar array is optimally dispatched to power a time varying load (as shown in Fig. 6). This figure (displaying the first 200 hours of the year) shows that for the Caribou trained controller using the Caribou data set, the optimal PV controller is able to continue to power the critical loads (at the expense of the non-critical loads) while the PV priority controller is not. For each city, an action network is separately trained and then this network is tested against the data for each city. The results of these tests are shown in Tables I, II and III. Additionally, a row labeled "Total Score" shows the weighted sum of the results from each test – the higher the score, the better the result. The weightings used in this calculation are the same weights used in the utility function (2). To find this score, each metric in the column was multiplied by the appropriate weight and all values in a column are summed.

Interestingly, the optimal PV controller trained using the Caribou, ME data set seems to always do better than any other controller. Also, except for 1 case in Tables I and III, it seems that the performance of the PV-priority controller is always lowest. Fig. 7 shows the state of battery charge for Caribou, ME for the period of late fall and early winter using both the PV-priority controller as well as the optimal PV controller trained using Caribou data. This result was chosen because it is the most demanding situation.

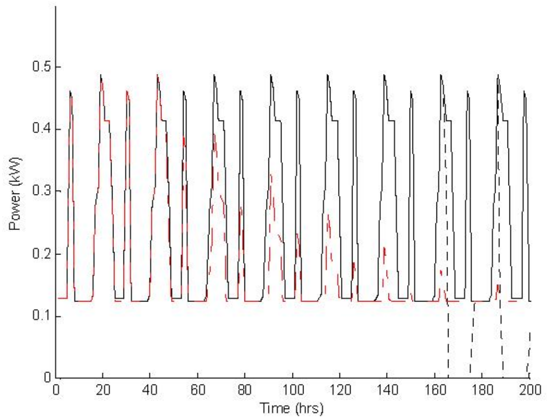

Fig. 6 – Sum of both critical and non-critical loads (solid black line) being satisfied by the PV priority controller (small dashed black lines) and the optimal controller (long dashed red lines).

Table I – Table of all controllers using data from Phoenix, AZ

City:	Phoenix, AZ			
Controller:	PV Priority	Optimal PV controller trained with data from:		
		Phoenix, AZ	Springfield, MO	Caribou, ME
Critical Load Satisfied:	98.68% [1071.9 kWh]	98.62% [1071.2 kWh]	100% [1086.2 kWh]	100% [1086.2 kWh]
Non-Critical Load Satisfied:	98.19% [989.9 kWh]	98.26% [990.6 kWh]	89.06% [897.8 kWh]	93.28% [940.4 kWh]
Average Battery Charge:	84.37% [29.2 kWh]	84.37% [29.2 kWh]	90.75% [31.4 kWh]	91.13% [31.5 kWh]
Total Score*:	2.392	2.392	2.400	2.426

*Computed using (2)

Table II – Table of all controllers using data from Springfield, MO

City:	Springfield, MO			
Controller:	PV Priority	Optimal PV controller trained with data from:		
		Phoenix, AZ	Springfield, MO	Caribou, ME
Critical Load Satisfied:	91.64% [995.4 kWh]	91.29% [991.6 kWh]	100% [1086.2 kWh]	100% [1086.2 kWh]
Non-Critical Load Satisfied:	89.05% [897.7 kWh]	89.42% [901.5 kWh]	73.56% [741.6 kWh]	76.89% [775.2 kWh]
Average Battery Charge:	72.66% [25.1 kWh]	72.66% [25.1 kWh]	82.62% [28.6 kWh]	84.37% [29.2 kWh]
Total Score*:	2.173	2.235	2.259	2.289

*Computed using (2)

Table III – Table of all controllers using data from Caribou, ME

City:	Caribou, ME			
Controller:	PV Priority	Optimal PV controller trained with data from:		
		Phoenix, AZ	Springfield, MO	Caribou, ME
Critical Load Satisfied:	84.22% [914.8 kWh]	83.44% [906.4 kWh]	96.25% [1045.5 kWh]	96.54% [1048.7 kWh]
Non-Critical Load Satisfied:	77.21% [778.4 kWh]	78.05% [786.8 kWh]	59.22% [597.0 kWh]	61.87% [623.7 kWh]
Average Battery Charge:	63.87% [22.1 kWh]	63.87% [22.1 kWh]	72.96% [25.2 kWh]	74.35% [25.7 kWh]
Total Score*:	1.951	1.946	2.066	2.094

*Computed using (2)

As can be seen from Fig. 7, the optimal controller keeps the charge of the battery higher than the conventional PV-priority controller. Additionally, it meets much more of the critical load (dropping it only during the winter months, when the battery is completely depleted), but a little less of the non-critical load. It is also interesting to note that this controller changed behavior as the battery charge increased; when it was lower it would power less of the non-critical load and focus on the critical load, and when it was more fully charged it would attempt to power both loads. This is in sharp contrast to the PV priority controller which always tried to power both loads, leading to a much lower average state of charge for the battery (and less met critical load demand). This behavior results from the coefficients used in the utility function (2).

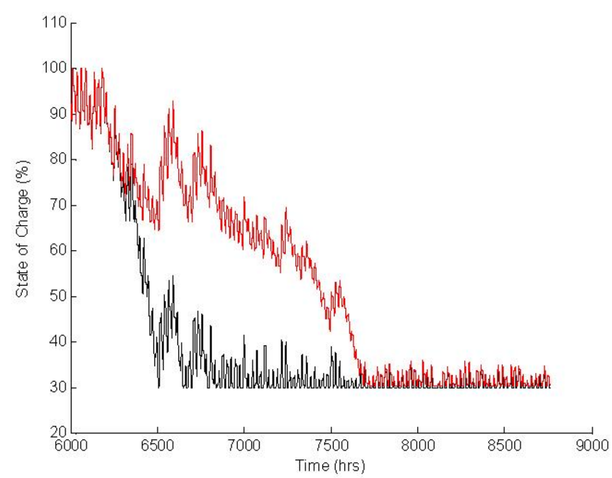

Fig. 7 – State of battery charge using the PV-priority controller (black line) and the optimal PV controller trained with Caribou data (red line) for late fall and early winter.

Another interesting result was that the optimal controller trained using the Phoenix, AZ data set usually did not perform any better than the priority controller, and always not as well as the other optimal controllers. This is most likely because the region received so much more sunlight than the other regions that it had less 'opportunity' to learn a more optimal technique, since there were more periods of excess sunlight where all loads could be satisfied. This most likely lead to a different operating characteristic that did not lend itself well to other locations that received less sunlight. Further verifying this point is the Caribou-trained controller, which received the least amount of sunlight of all 3 locations but seemed to perform better than any of the controllers.

The average charge of the batteries for the entire year for each of the three cities using the Caribou trained controller and PV-priority controller follow in Figs. 8, 9, and 10. In these figures it is evident that the optimal controller is far more capable of sustaining a higher average battery charge than the PV-priority controller.

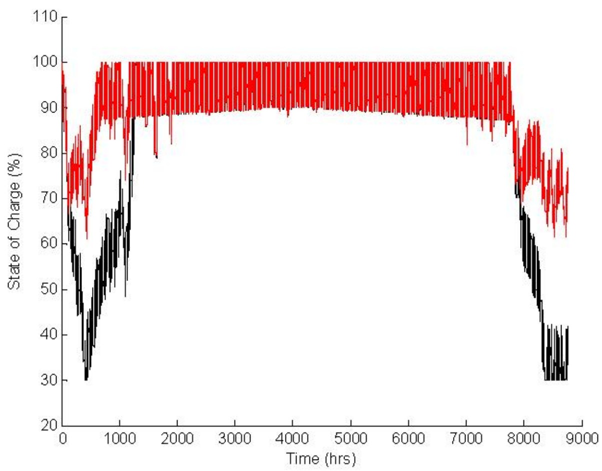

Fig. 8 – State of charge of the battery using the PV-priority controller (black line) and the optimal PV controller trained with Caribou data (red line) for the entire year in Phoenix, AZ.

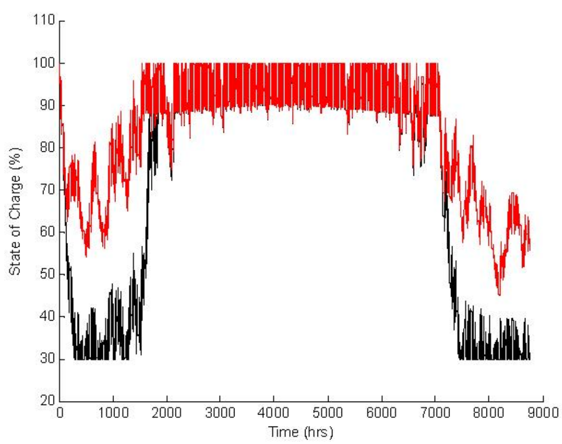

Fig. 9 – State of charge of the battery using the PV-priority controller (black line) and the optimal PV controller trained with Caribou data (red line) for the entire year in Springfield, MO.

VI. Conclusion

An optimal energy control scheme based on adaptive critic designs for the photovoltaic solar energy system has been developed and compared with the conventional PV-priority control scheme used today. The ACD method optimizes the control policy over time to ensure that the critical load demand is met primarily all the time and then the non-critical load demand. The state of the battery charge is also maintained as high as possible to ensure energy supply to the critical loads during nights and the winter months. This in turn provides the benefit of extended battery life. Results have been presented for three different US cities with different radiations and the optimal PV controller has exhibited better control in almost all cases. The comparison between the two control schemes shows that for the most part, the optimal PV controller satisfies the critical load and some of the non-critical loads demand better than the PV-priority control scheme, while keeping a higher battery charge. The hardware of implementation of such an ACD controller is feasible and cheap compared to savings as a result of proper energy management. This scheme is adaptive and therefore can fine tune to different locations and weather profiles within a short period of time. Thus, it is a promising and a potential inexpensive technique for practical deployment on growing solar energy systems.

Future work involves real-time investigations to try to further optimize the energy controller to follow more closely the load profiles and provide even better performance. In addition to improving performance, integration with other energy sources and hybrid forms of storage, such as hydrogen fuel cells with battery will be investigated.

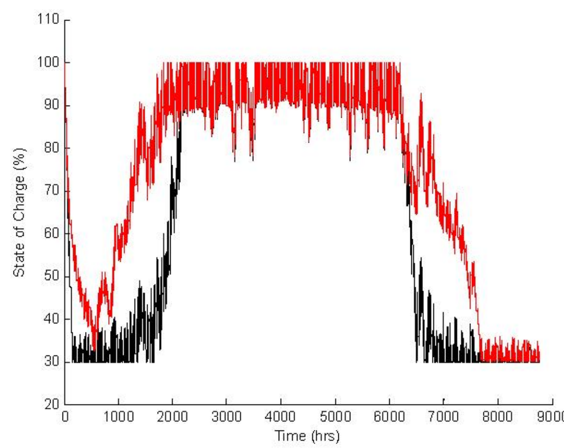

Fig. 10 – State of charge of the battery using the PV-priority controller (black line) and the optimal PV controller trained with Caribou data (red line) for the entire year in Caribou, ME.

VII. References

[1] R. A. Messenger, J. Ventre, *Photovoltaic System Engineering*, CRC Press, 2004.

[2] G. P. Henze, R. H. Dodier, "Adaptive Optimal Control of a Grid-Independent PV-System", *Trans. ASME Journal of Solar Energy Engineering*, vol. 125, Feb. 2003, pp. 34 – 42.

[3] A.U. Chuku, B. Oni, F. Kuate, and E. Overton, "Making Solar Energy a Viable Stand-alone Alternative Through Efficient Control Algorithm", in *International Conference on Power Systems Operation and Planning – VI (ICPSOP) 2005*, pp. 71-79

[4] P. J. Werbos, "Approximate Dynamic Programming for Real-time Control and Neural Modeling", in *Handbook of intelligent control*, White and Sofge, Eds. New York: Van Nostrand Reinhold, 1992, pp 493-525.

[5] G. K. Venayagamoorthy, R.G. Harley and D.C. Wunsch, "Comparison of Heuristic Dynamic Programming and Dual Heuristic Programming Adaptive Critics for Neurocontrol of a Turbogenerator," *IEEE Trans. Neural Networks*, vol. 13, pp. 764 -773, May 2002.

[6] D. Prokhorov, D.C. Wunsch, "Adaptive Critic Designs", *IEEE Trans. Neural Networks*, vol. 8, pp. 997-1007, Nov. 1997.

[7] Kyocera KC200GT data sheet, http://www.kyocerasolar.com/pdf/specsheets/KC200GT.pdf, accessed April 30, 2006.

[8] "TMY2 User's Manual", Jun. 1995. National Renewable Energy Laboratory, Golden, Colorado. [Online] Available: http://rredc.nrel.gov/solar/old_data/nsrdb/tmy2/.