Spring 5-1-2020

INTEGRATION OF MACHINE LEARNING FOR REACTIVE POWER CONTROL FOR A DISTRIBUTION FEEDER SIMULATION WITH EXTERNAL LOAD

Ivonne D. Acosta Molina
University of New Mexico - Main Campus

Ivonne Denisse Acosta Molina

Candidate

Electrical and Computer Engineering

Department

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Manel Martinez-Ramon , Chairperson

Marios Pattichis

Ali Bidram

INTEGRATION OF EXTERNAL LOAD AND MACHINE LEARNING FOR REACTIVE POWER CONTROL FOR A DISTRIBUTION FEEDER SIMULATION

BY

IVONNE DENISSE ACOSTA MOLINA

B.S., Electronics Engineering, Hermosillo Technological Institute, 2012

THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

Electrical Engineering

The University of New Mexico

Albuquerque, New Mexico

May 2020

Dedication

To all my grandparents that passed away while I was working on this achievement. And to Miguel Angel Gaytan Ordaz.

May you all rest in peace, I will love you forever.

Acknowledgments:

I would like to thank Professor Manel for always support me since the beginning until the end, I wouldn't have finished this manuscript if it weren't for him. Also, to professor Marios Pattichis for helping and supporting me when I needed it the most, it was such a pleasure to work with you.

Thanks to Professor Mammoli for giving me a chance of being part of his team. Thanks as well to Jee Won Choi and Mathew Robinson for all her help and kindness. Thanks to Matt Reno and Adam Summers at Sandia Labs also to Takayama-San and Guillermo Terren for sharing his research and work for this project.

Thanks to CONACYT in Mexico for their scholarship, I wouldn't have made them without them also.

Another person I would like to acknowledge its myself, to being able to complete this work at the same time I was suffering from major clinical depression, and all the people to help me coped through this tough time. And most importantly my family. My mom for being my backbone and my hero, my nieces for warming my heart and give me a reason to keep going. To Raissa and to all my friends that video called me to cheer me up and tell me "You got this!", because I got it thanks to y'all!

INTEGRATION OF MACHINE LEARNING FOR REACTIVE POWER CONTROL FOR A DISTRIBUTION FEEDER SIMULATION WITH EXTERNAL LOAD

by

Ivonne Denisse Acosta Molina

B.S., Electronics Engineering, Hermosillo Technological Institute, 2012

M.S., Electrical Engineering, University of New Mexico, 2020

ABSTRACT

It has been 32 years since the Brundtland Report was published. That was the first time that the term Sustainable Development (SD) was coined. In this context, renewable sources of energy play an important role on not to deplete our natural resources in order to meet our need without compromising future generations. Smart Grids are on the pathway to achieve the SD goals. This Thesis focuses on the integration of renewables, specifically Solar PV panels and inverters and its interactions with the distribution grid. Since the power injection caused by the PV inverter can alter the voltage range, Reinforcement Learning (RL) is applied as method for voltage regulation. This research aims to integrate all these elements in a Co-simulation Real-Time system. To achieve complexity and reality to this co-simulation frame, an external load is aggregated from an external source. Methodology and results are described, and conclusions and future work suggested.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

1.1 Microgrid background

Fossil Fuel for electric power generation it's fundamental in our current society for subsisting. Nevertheless, this dependence on Fossil Fuels have caused a detriment in our environment, since they're the main cause of environmental pollution and thus, global warming [1].

Therefore, it has become an ultimate priority to find energy alternatives that are sustainable, which means it needs to be a renewable resource and also safe for the environment and the people [2].

As a result of this, the Smart Grid emerges as one of the solutions for this predicament [3]. Since it's such a new term there's different definitions, such as the one from EPRI [4] that states:

"A Smart Grid is one that incorporates information and communications technology into every aspect of electricity generation, delivery and consumption in order to minimize environmental impact, enhance markets, improve reliability and service, and reduce costs and improve efficiency" [5]. In the same manner, according to the CERTS (Consortium for Electric Reliability Technology Solutions) criteria, a basic Microgrid is an interconnection of the following elements [6] [7]**,** which are best illustrated in figure 1. Distributed generating (DG) units such as photovoltaic energy, wind power, fuel cells, micro turbines, amongst others. Secondly, Energy Storage Systems (ESS) or devices, such as batteries, capacitors and flywheels for integration purpose. Additionally, groups of feeders for distribution and an Energy Management System (EMS) for power management and set point setting purposes for DG unit controllers.

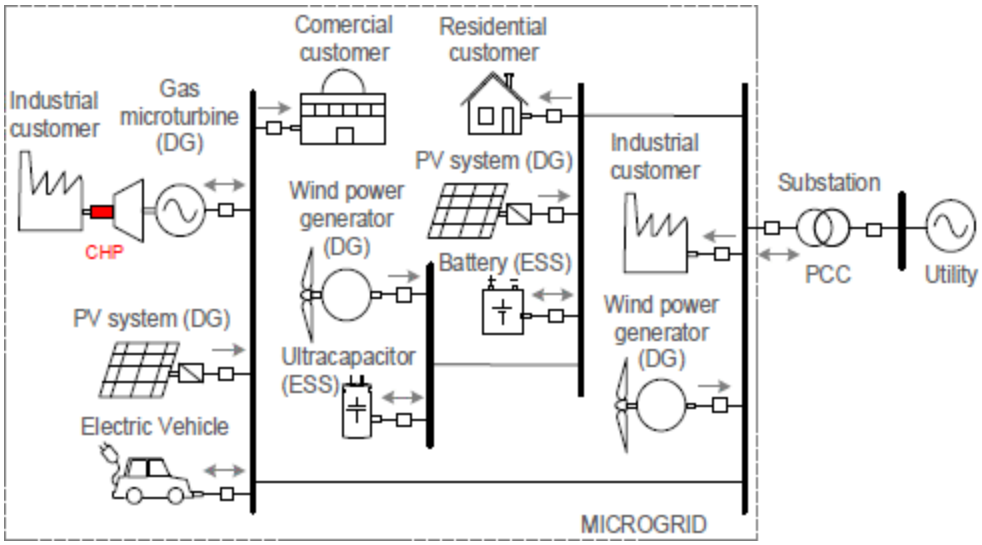

Figure 1. Microgrid Architecture [8].

1.2 Control of a Microgrid

Control of a Microgrid is essential for stable and economically efficient operation. The most fundamental functions of the control structure of a microgrid are voltage and frequency regulation for both islanded and grid connected modes, proper load sharing and Distributed Energy Resource coordination, all of this to achieve an optimal resynchronization with the main grid in grid mode [9], [10] . Also, there must be a Power Flow control between the microgrid and the main grid thus, resulting in an optimized microgrid operating cost [11].

One important feature to take in account in order to control a Microgrid is the PV integration impact since it can cause Voltage unbalance, specifically, voltage rise on the feeder and potential break of protection synchronization system [12][13].

For this reason, the literature shows various methods for controlling the voltage on the feeder, for instance, control of On-Load Tap Changing (OLTC) transformers, fixed or switched capacitors, Battery Storage (BS) systems, Power curtailment, Reactive Power Control, amongst others [14].

1.3 Thesis Scope and organization

This thesis is organized in the following chapters: This first chapter being the introduction, giving a brief background on Smart Grids and the importance of control of these. In chapter 2 the part of co-simulation and background of Reinforcement Learning are described.

Third chapter is the methodology of this study case. It starts with the model to be analyzed, then it explains the different parts that are going to be added to the model: the aggregated external load on a Raspberry Pi, the Reinforcement Learning block in Simulink and Irradiation data set. Then the integration part is explained. On the fourth chapter the results of the simulation are showed and discussed to be finished by a conclusion.

CHAPTER 2

BACKGROUND

2.2 Machine learning background

Machine learning, as the word contains it is about the process of learning, specifically speaking what's been learned is data or signals, convert it into information, and the purpose is the extraction of knowledge, which can be new data or information on the learning machine itself. This outcome can help us make decisions and/or predictions [15] [16]. Within the Machine Learning (ML) sphere, there's three main categories: unsupervised learning, supervised learning and reinforcement learning (RL). In this study, Reinforcement Learning is the case that's of interest.

2.2.1 Reinforcement Learning

Reinforcement Learning (RL) are iterative algorithms which learning process it through exploration of an unknown system, this provides the algorithms with experience for them to learn the optimal output or way to solve a problem [17].

The main of the elements on RL are the *agent* which interacts with its environment, on each time step, the agent receives the environment *state* S_t, in which $S_t \in S$, means the set of all possible states, based on that the agent "acts" selecting a specific *Action* $A_t \in A(S_t)$. As a consequence of this action, the environment sends a reward to the agent, and the main goal for the agent it's to maximize total reward overall [18]. This overall maximization is called Value function and the "rule" that the agent follows to achieve more rewards is called policy [19] which makes a mapping of probabilities distributions of each and one of the possibilities of actions that can be executed.

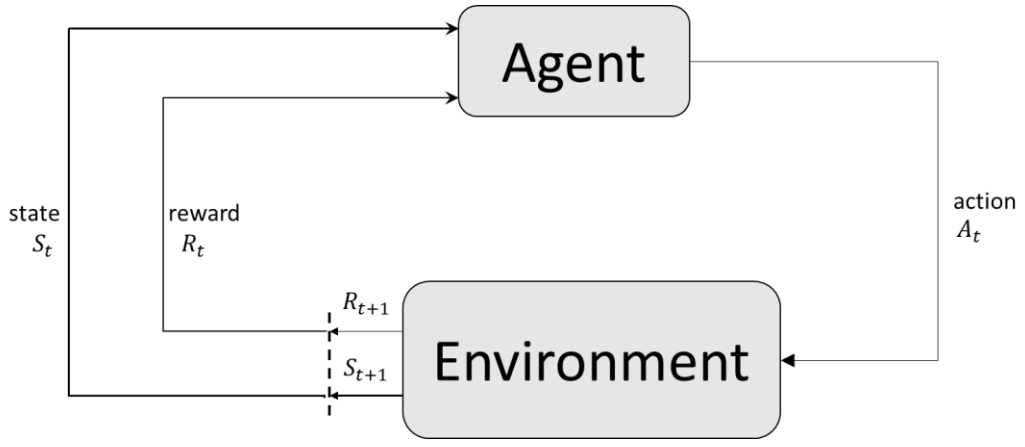

Figure 2 Basic Framework of Reinforcement Learning [18].

Temporal Difference (TD) Learning

TD learning is the type of reinforcement that has both the quality of Monte Carlo methods, because it can learn from raw experience, without the dynamics of the environment's model and like Dynamic Programming methods, can update estimates without having to wait for the final outcome [20].

TD charactertic is the set of method that evaluate the value function, amongst them SARSA and Q learning. TD feature is that learns it's value function $V_{(s)}$ through a TD error, which is learned directly from experience, the way it does this is through bootstrapping, model free, online and fully incremental. This makes TD a prediction problem. To determine the value function update, the method is the equation 2.1:

$$V(S_t) \leftarrow V(S_t) + \alpha[R_{t+1} + \gamma V(S_{t+1}) - V(S_t)] \quad (2.1)$$

Being α the learning rate, γ a discount rate and the TD error update is defined by δt 2.2

$$\delta_t = r_{t+1} + \gamma V(s_{t+1}) - V(s_t) \quad (2.2)$$

Figure 3 ilustrates the algorithm for TD(0) which is tabular since "0" illustrates that is based on one-step return [15].

Input: the policy π to be evaluated
Initialize $V(s)$ arbitrarily (e.g., $V(s) = 0, \forall s \in \mathcal{S}^+$)
Repeat (for each episode):
  Initialize S
  Repeat (for each step of episode):
    $A \leftarrow$ action given by π for S
    Take action A; observe reward, R, and next state, S'
    $V(S) \leftarrow V(S) + \alpha\big[R + \gamma V(S') - V(S)\big]$
    $S \leftarrow S'$
  until S is terminal

Figure 3. Algorithm for TD(0) [21].

An off policy TD learning method is Q learning, which is defined by 2.3

$$Q(S_t, A_t) \leftarrow Q(S_t, A_t) + \alpha\left[R_{t+1} + \gamma \max_a Q(S_{t+1}, a) - Q(S_t, A_t)\right] \quad (2.3)$$

Being Q is the array storing the current action-value function estimate, and $Q(S_t, A_t)$ being an estimate of $Q^*(S_t, A_t)$for each state-action.

The problem with Q learning it that does not permit updates based on the rewards for mor than one iteration. The TD(λ) method described in the next section is the solution to this problem.

Eligibility traces

In addition to the TD(0) and Q method, there's the TD(λ) method, which now includes the concept of eligibility trace, which is a short term memory trace that lasts one episode and helps the learning by affecting the weight vector, which is a long-term memory, that lasts the entire duration of the system[21]. The way the elgibility trace works is based on the degree in which a state S_t has been visited in the past, once a reinforcement is collected, it updates all the states that have been recently visited given their eligibility [22]. TD(λ) is a way to unify the one step TD(0) and Monte Carlo methods, through this eligibility traces and the decay parameter λ, for predicting algorithms [15].

Function Approximation

TD is considered a tabular case, since the value function or the policy are stored in a tabular form, but it's limited by the number of states and actions, this becomes a problem when we take in account the memory and storage for large tables. The solution to this problem is make a generalization from previous states, this is what is called Function Approximation, which is an instance of supervised learning [20].

The method represented for function approximation is a parameterized functional form with parameter vector called weight vector $\boldsymbol{w} \epsilon \mathfrak{R}^n$, the formula used for the approximated value of state *S* given the weight vector is $\hat{v}(s, \boldsymbol{w}) \approx v_\pi(s)$ [18].

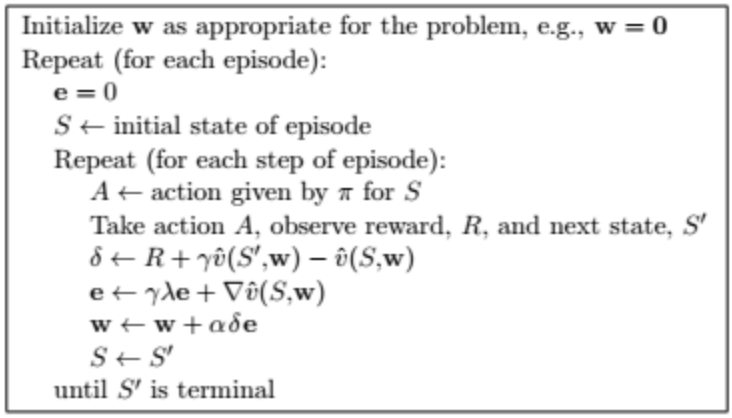

Figure 4. On-line gradient-descent TD(λ) for estimating vπ from [18].

Figure 4 presents the pseudo code for TD(λ) with function approximation, where $\nabla\hat{v}(s, \boldsymbol{w})$ is the gradient of the value approximation function with respect to *w*, this latter updated by the rule $\boldsymbol{w} \leftarrow \boldsymbol{w} + \alpha[r + \gamma\hat{v}(s', \boldsymbol{w}) - \hat{v}(s, \boldsymbol{w})]\nabla\hat{v}(s, \boldsymbol{w})$ [15].

Linear methods for function approximation

Above the gradient descent method for function approximation was defined, with this in mind, there's a special case of gradient-descent methods in which the approximate value function $\hat{v}$ is a linear function of vector $\boldsymbol{w}$, then 2.4 is now the approximate value function with x(s) as vector of features corresponding to every state S_t.

$$\hat{v}(s, \boldsymbol{w}) = w^T x(s) = \sum_{i=0}^{n} w_i x_i(s). \quad (2.5)$$

.Where $\nabla \hat{v}(s, \boldsymbol{w}) = \boldsymbol{x}(s)$, is the gradient of the approximate value function [21].

Actor-Critic Methods

Actor-Critic methods are RL methods that even though they compute action and state values, they don't use them directly for the action selection. Conversely, the policy with its own weights it's independently of any value function, is represented directly.

Usually, the critic is a state-value function that evaluate the new state and decides after each action selection. The way the critic evaluates the state is by using the TD error.

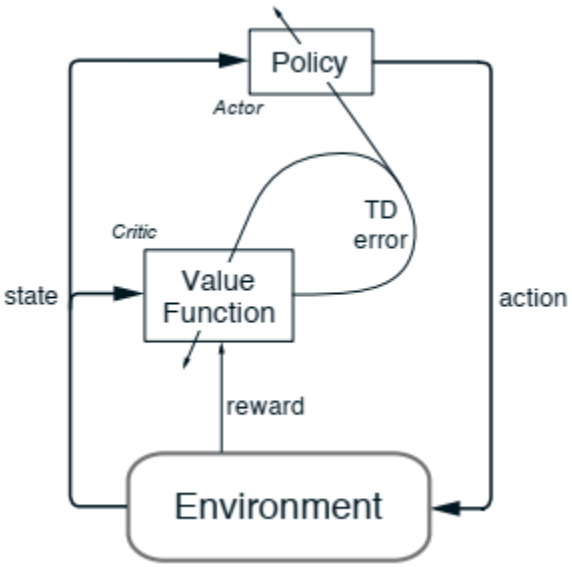

Figure 5. Actor-Critic Architecture

The advantages of using actor-critic is that they require minimal computation for action selection, since the policy is explicitly stored. The policy also might be stochastic [20].

2.3 Real-Time Simulation environment and Co-Simulation

As technology develops, simulation frameworks become more complex, since nowadays they incorporate physical, software and network aspects combined [23]. The development of these systems has been divided between different teams, or external supplier, each with its own expertise and its own tools. Each contributor plays a part developing its element that adds the system solving. All of this needs to be merged with all the other partial solution elements, this global system it's what is called Co-Simulation[24]. This concept can be seen as a holistic development process [25] where cutting-edge, multi-disciplinary and optimal solution can be accomplished. Challenges are the integration of this partial models and/or solutions, since they were developed with different software/hardware and it's possible they were developed by a specialized tool with intellectual property rights and can't be easily accessed.

Co-simulations as referred to in [24]:

"It consists of the theory and techniques to enable global simulation of a coupled system via the composition of simulators. Each simulator is a black box mock-up

of a constituent system, developed and provided by the team that is responsible for that system."

In this present work a Real Time Co-simulation is uses according to the diagram from [25], since it's a Power System with hybrid representations executed in individual runtime environments. As they describe, one challenge it's to synchronize the models and their solvers.

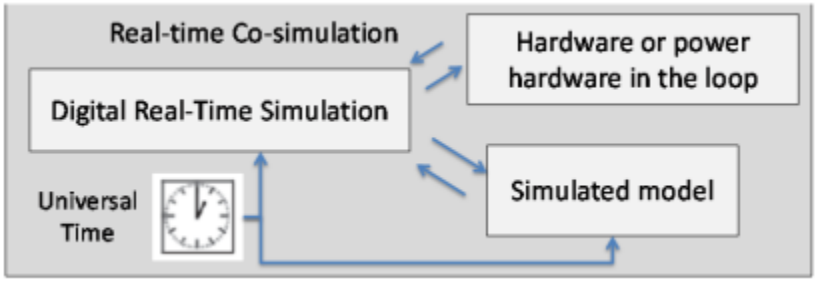

Figure 6 *Real-Time Co-simulation* [25].

Chapter 3:

Methodology

In this chapter the methodology used in this thesis is explained. First, we describe the study case which is a feeder model provided by Sandia National Labs in collaboration with the Center of Emerging Energies and Technology (CEET) center at the University of New Mexico. Subsequently, the aggregated external load part of a previous work also made by a master student at CEET and the real irradiance data gathering are described. After that, the Reinforcement Learning part of the project and the proposed model of unifying the three parts described above in a co-simulation environment is being detailed.

3.2 Feeder Model

3.2.1 Study Case

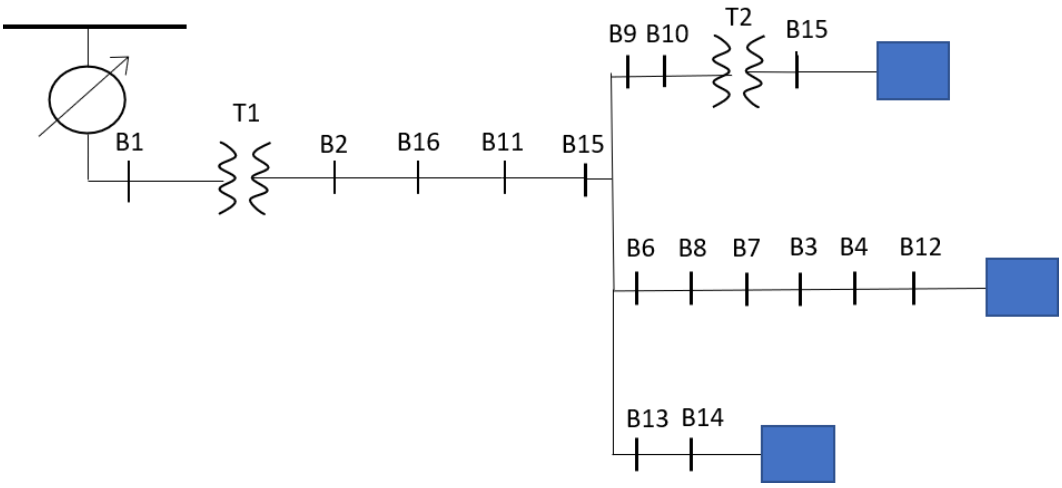

Figure 7. 15 Bus model provided by Sandia National Labs.

In this section, a 15 Bus model proportioned by Sandia National Labs was used to evaluate the proposed Reinforcement Learning method. The inverter that has the RL method is on Bus 12 and is a High Voltage (12.47kV) Distribution Network. Added to the model, there's an external aggregated load on Bus 12 supplied by Raspberry-Pi and altogether simulated in a MATLAB Simulink® environment.

3.2.1 Opal RT environment

There's various software packages that simulate distribution systems such as GridLAB-D used in [26], OpenDSS, PSCAD, Simulink, etc. In this research to add the Real-Time part into the simulation of the distribution Feeder, the

software/hardware used for the merging of the co-simulations is RT-Lab which works using a MATLAB Simulink environment.

Figure 7 shows the architecture of RT-Lab, it communicates via TCP/IP protocol and the basic setup is host computer (Windows OS) and the target computer by Opal RT, the model used for this studio is the OP 5600, with the following characteristics: Powerful real-time target with up to 12 INTEL processor cores 3.3 GHz, Real-time operating system Linux REDHAT and a FPGA based card for analog and digital inputs and outputs.

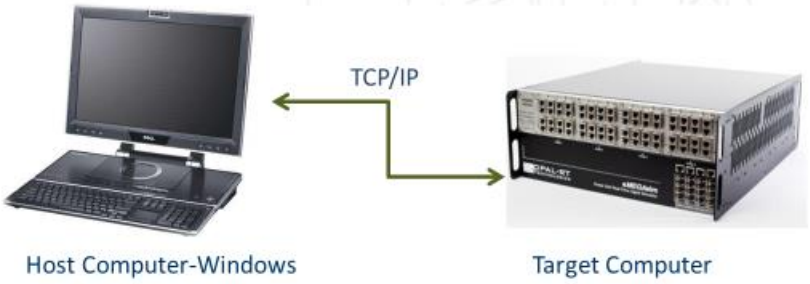

Figure 8. RT lab architecture[27]

3.2.2 Model Description

The distribution Feeder used in this simulation it's a 15 Bus Model, with a generator (Wye connection) three phase source 115 kV, 60 Hz, 2 transformers, one 115//12.47 kV, 30MW Nominal Power. Wye-Wye connection, and another at 12.47//480 V at Bus 10, one capacitor bank at bus 7, and a switched capacitor bank at bus 3. It also includes 4 unbalanced triphasic loads, 2 balanced triphasic loads. Table 1 shows the

details of the nominal characteristics of the static loads in the model.

Load Name	Characteristics
Load Bus 4	Triphasic Balanced 12.47 kV, 1.885 MW, 1.292 Mvar
Load Bus 5	Triphasic unbalanced Load A inductive-resistive 7.199kV, 28 W, 8.9 Var Load B, inductive-resistive 7.199 kV, 2288 W, 413 Var. Load C, inductive-resistive 7.199 kV, 393.99 W, 54.4 Var.
Load Bus 9	Triphasic unbalanced Load A, inductive-resistive 7.199kV, 97.18 kW, 22.6 kVar. Load B, inductive-resistive 7.199 kV, 100.579 kW, 22.407 kVar. Load C, inductive-resistive 7.199 kV, 116.317 kW, 25.615 kVar.
Load Bus 10	Triphasic unbalanced Load A, inductive-resistive 7.199kV, 107.384 kW, 11.558 kVar. Load B, inductive-resistive 7.199 kV, 108.561 kW, 11.911 kVar. Load C, inductive-resistive 7.199 kV, 118.93 kW, 12.605 kVar.
Load Bus 11	Unbalanced triphasic Load A, pure inductive, 7.199kV, 0 W, 3.43Var Load B, inductive-resistive 7.199kV, 386 kW, 489 Var Load C, inductive-resistive 7.199 kV, 670 W, 302 Var.
Load Bus 15	Triphasic Balanced, 480 V, 25.5kW, 19.2 kVar

Table 1. Description of the static loads on Bus15 Model.

Also, it includes 3 PLL (Phase Locked Loop Systems) that are used as inverters for the purpose of this project, the characteristics are shown below in table 2. In the case study the PLL used is on Bus12.

Name of PLL	Location	Nominal Characteristics
PLL	Bus 15	258 kVA, 480V Triphasic, 12.47 kV, 1.8Mvar
PLL1	Bus 12	10 MVA ,12.47 kV base Triphasic,
PLL2	Bus 14	10 MVA, 12.47kV base Triphasic

Table 2. Nominal Characteristics of PLL

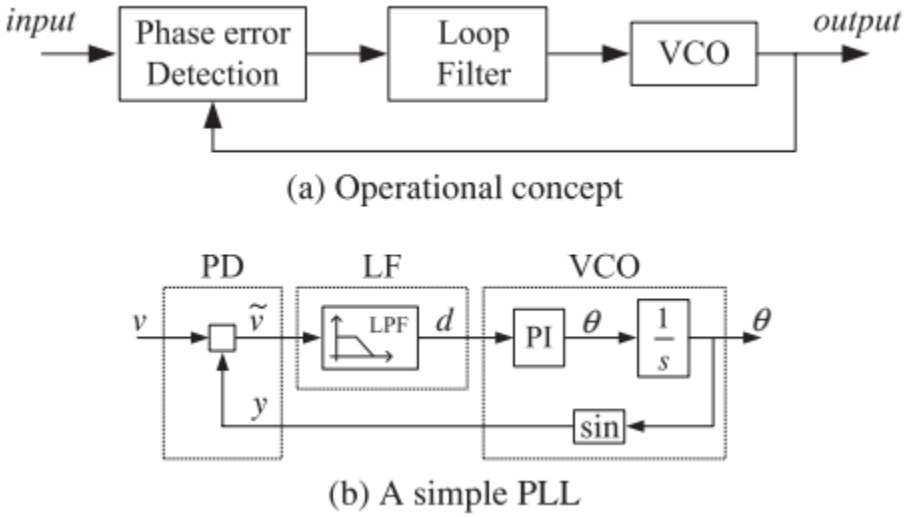

Figure 9. Basic PLL structure[28]

Phase Locked-Loop architecture (PLL)

The Diagram in Figure 7 shows the operation of a PLL, it has three main basic parts. An error phase detection (PD) that measures the phase difference between the input and output signals, then goes through the loop filter, which is generally a low pass filter that extracts the DC component from the phase error. Following that, the amplified component of the DC signal goes to the VCO, which can be a PI controller, as illustrated on Figure 7 (b), then the controller generates the output signal frequency and it's integrated to the phase of the signal output.

The output of the PD unit is defined by

$$\tilde{v} = vy = V_m sin\theta cos\theta_g = \frac{V_m}{2}\sin(\theta - \theta_g) + \frac{V_m}{2}\sin(\theta + \theta_g) \quad (3.1)$$

And the output of the loop filter is 3.2:

$$d = \frac{V_m}{2} sin[(\omega - \omega_g)t + (\phi - \phi_g)]. \quad (3.2)$$

This output is carried to the PI controller that will generate the frequency $\omega = \theta$, this until $d = 0$. Subsequently, this frequency gets integrated to the output signal, which is sent back to the error Phase Detection unit, making it a closed locked-loop.

The PLL controls that are used on the feeder as inverters with DQ frame using the technique three-phase applications is a PLL in the synchronously rotating reference frame (SRF-PLL) [26] to produce reactive and active power independently from the grid.

3.3 Aggregated external load

This sections describes the simulation framework of an aggregated load, from a previous developed work by [29] as a part of this co-simulation. For the purposes of this study, the part it would focus on would be on the aggregated load part. This load generator was used in the previous work to simulate a residential load in Mesa del Sol, Albuquerque. Figure 9 describes the entire feeder for Mesa del Sol, but for the purposes of this case, only the residential loads in 2 and 3 are to be considered.

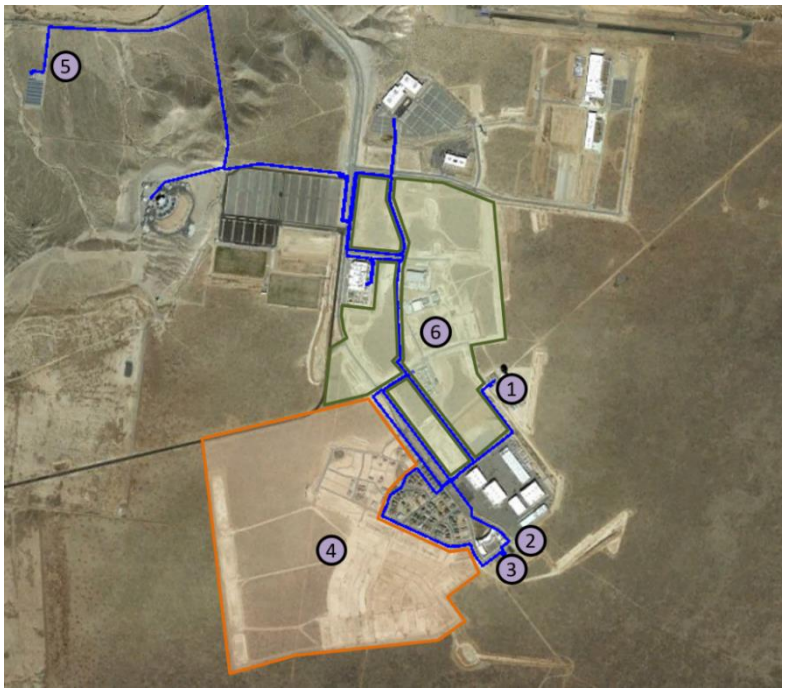

Figure 10. Mesa del sol layout, residential load is pointed out in 2 and 3.

For the purposes of this study, it simulates a fictional residential aggregated load on the bus 12 of the 15 Bus Feeder. The residential community it's conformed by 200 homes with typical electric appliances, these are clothes dryer, air conditioner, domestic hot water, refrigerator, cooking range and lights. As well as the previous work, to simulate physical separation and add a more real component to the RT simulation, the load aggregator is executed on a Raspberry Pi 3.

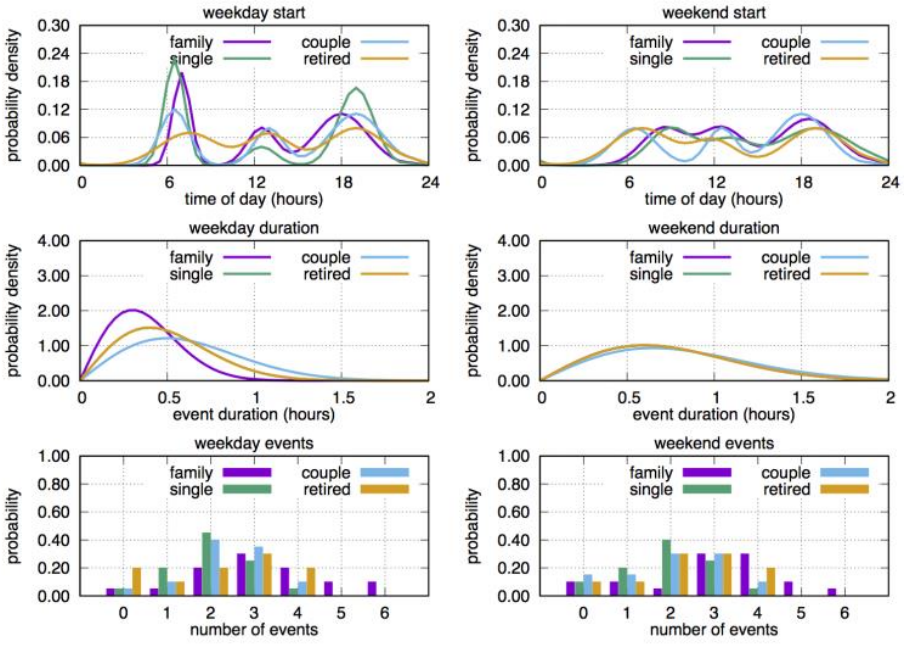

Figure 11. Probabilistic characterization of the use of an electric range, in terms of probability density function for start-time, probability density function for duration, and discrete probability for number of use events. [29]

In the previous developed work , a statistical approach was used to represent the effects of human factor on the distribution feeder, in order to represent the interaction of the community within the house on their home appliances. A bottom up statistical approach is used to represent the interactions of humans-in-the-loop. Figure 3 shows the probability density used to describe the use of each appliance, modelled after customer usage patterns. Likewise, each use of the appliance it's being statistically characterized in regard to specific social groups (working singles, couples, retirees, families. The integral combination of stochastic processes to generate the aforementioned load, illustrated on figure 4, there's an example of a

1000 house residential complex bases on demographics and real-time demand response in distribution systems produces realistic scenarios that can be used for several applications. For this study is used to analyze and control the effect of voltage level on the bus. Figure 5 shows the contrast between the aggregated load in the previous work, now applied to the model in this study, 200 houses residential load, in a period of two hours outputting every second.

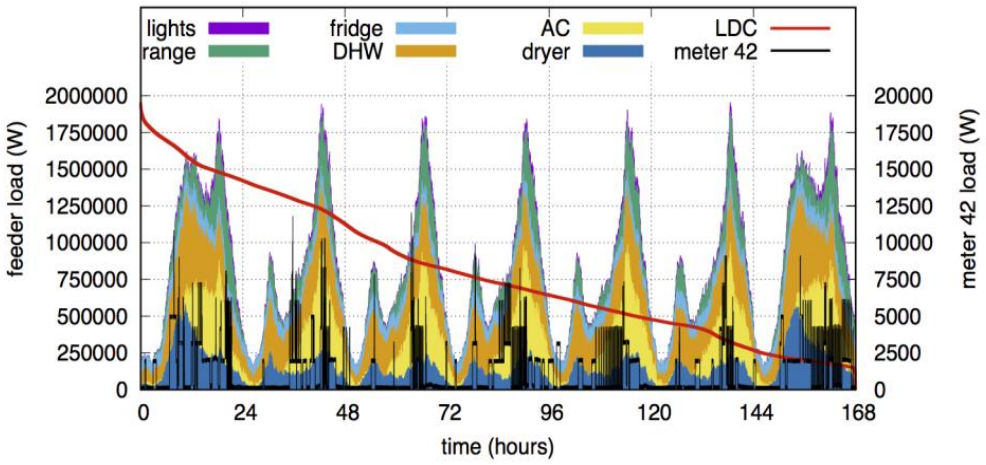

Figure 12.Total feeder load resulting from 1000 houses, also indicating the aggregated contribution from each appliance. The total load for an individual meter (42) is also shown for comparison. The load duration curve (shown in red) indicates significant opportunity for shifting and deferring loads.

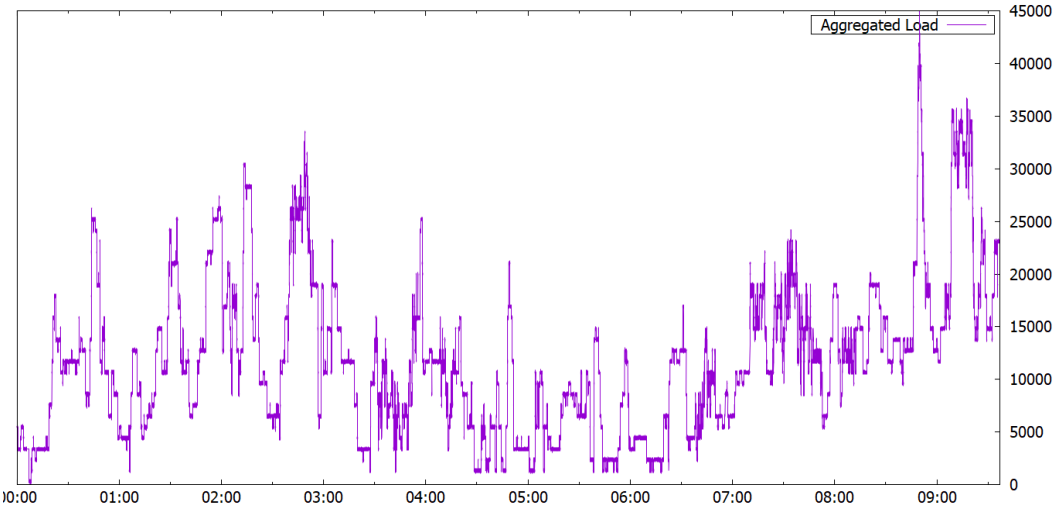

Figure 13. External aggregated load from Bus 12 in the 15 Bus Feeder.

3.4 Proposed Method

3.4.1 Raspberry pi co-simulation environment

As previously stated, the model of the load is a Real-time each second outputted load. In order to have a more realistic model, the external load is not simulated within the Simulink model, but using OPAL-RT as an interface between them. Figure 13 shows the diagram of the simulation environment. The step size of the model is 80μs, to ensure a real-time simulation approach. The first step is load the model into the RT-lab software, this generates an executable in C and sends it to the OP5600, which programs the FPGA to receive the PWM signal from the Raspberry-Pi and executes the Simulink GUI in the host computer.

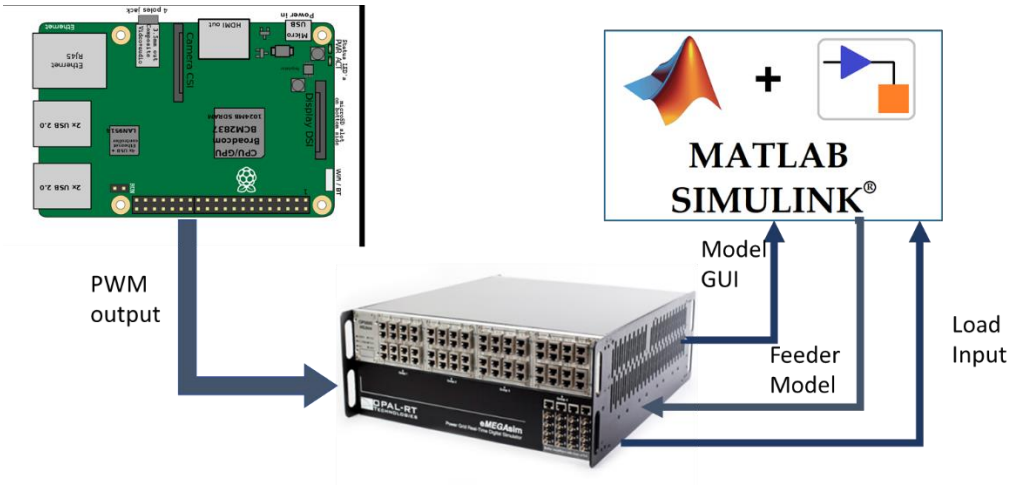

Figure 14.Simulation environment

3.4.2 Temporal Difference Learning

In this section, the outline of the proposed method is discussed, based on [19] work. As figure 6 states the diagram of TD(λ) method, reviewed in chapter 2.

Figure 8 illustrates how the proposed method is applied. The agent is the PV inverter and the distribution network it's the environment. Moreover, the state S_t is the interconnecting point between the PV and the network, the action $a_{t,}$ being the reactive power output Q at the inverter, and the reward r_{t+1} is the operation rule of distribution network. Summarizing, the actor (inverter) observes the state (voltage) at the interconnecting point and determines an action translated on reactive power at the output of itself.

Afterwards, the critic obtains the reward and voltage value after having controlled the reactive power output. This critic evaluates the result of the actor and updates

the value function, in turn the actor updates the probability policy based on the result of the critic's evaluation and thus, learning reactive power optimal output.

In this case, the policy utilized to update the value function in equation 2.1 is the following normal distribution function equation 3.1

$$\pi(a_t|\mu_p\sigma_p) = \frac{1}{\sigma_p\sqrt{2\pi}} exp\left(-\frac{(a_t-\mu_p)^2}{2\sigma_p}\right). \quad (3.3)$$

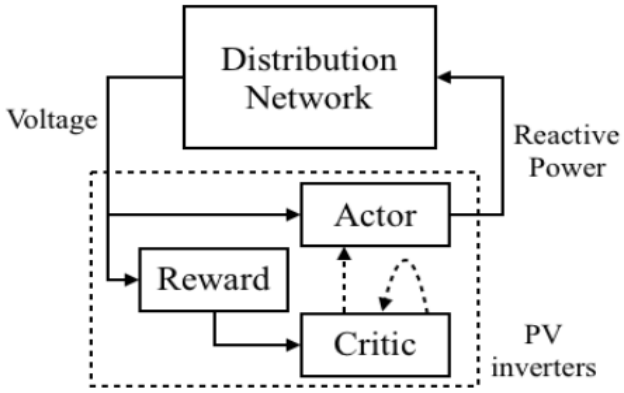

Figure 15. Diagram of Reinforcement learning applied to the proposed method [19].

On this actor critic method, π is the policy which is defined by the probability distribution of action at *t* using the average μ_p and the standard deviation σ_p. Thus, an optimum reactive power output is a function of the voltage as a state S. The average μ is calculated using 3.4

$$\mu_p \begin{cases} a(s_t - V_d) & (s_t \geq 1.0(p.u.)) \\ 0 & (otherwise) \end{cases} \quad (3.4)$$

Where Vd is the transmission voltage at the pole transformer and *a* is the gradient of the reactive power at *S.* In this case the reactive power is controlled by the inverter when the voltage raises, generating a reverse power flow when St is greater than the transmission voltage, whereas when the voltage at state St is lower than transmission voltage it does nothing. This is the reason that for the purposes of this study we only focus on the upper limit of 1.019 p.u. and not the lower limit.

So having this in mind the work in [19] had two targets optimizing the distribution network voltage through reactive power control. One is dissolving the deviation in voltage range, the second one, is minimizing the amount of reactive power. The proposed reward signal for the project is the following:

$$r_{t+1} = \begin{cases} -1 + (a_{t-1} - a_t) & (s_{t+1} \geq 1.019(p.u.)) \\ a_{t-1} - a_t & (otherwise) \end{cases} \quad (3.5)$$

This reward signal compares the reactive power output *a* at *t-1* and *t*, if the voltage doesn't go above the limit, that means if $a_{t-1} - a_t$ is positive, it receives the reward signal. On the other hand, if the reactive power a_t is above 1.019 p.u. it sends a penalty.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Simulation Results

For this study, various scenarios were simulated, the details are explained in the following section. There were two phases of the project, one using a random uniform distribution of the power inverter. This to corroborate the method was working. The original parameters of RL are shown in table 2. This parameter was established by one of the authors of the original work.

Parameter	Value
α	0.1
γ	0.9
σ_p	5
σ_i	5
c_i	5 intervals from 0.995 to 1.035 (p.u.)

Table 3. RL parameters

4.1.2 Test Run

First scenario was a test run that serves to corroborate that the algorithm works. Since it was only a test run, it was only for one-hour duration. Figure 14 shows the Bus 12 voltage of the Feeder, and it shows some regulation on the RL model, it has a lot of fluctuations in voltage and is clearly unstable, even though is keeping the voltage within range.

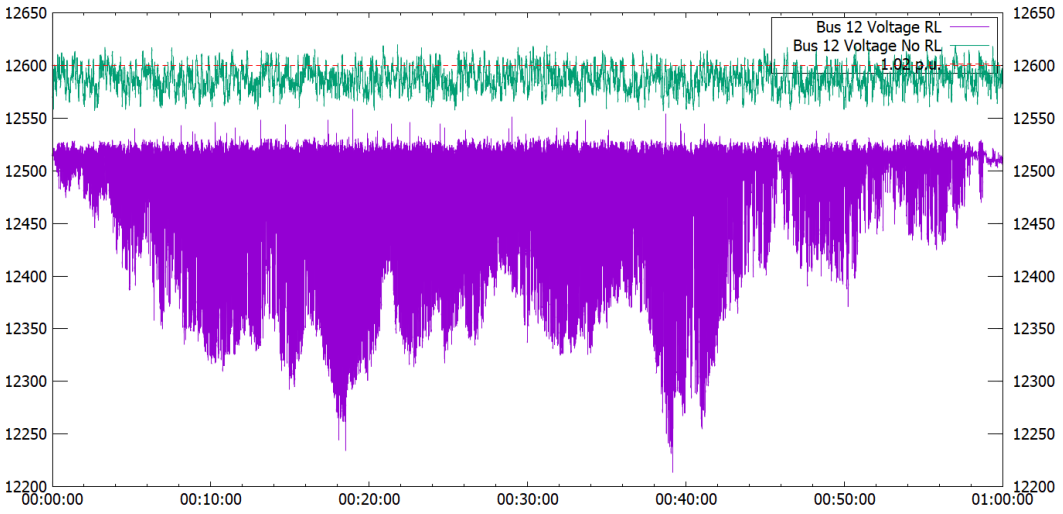

Figure 16 Bus Voltage of feeder.

In figure 16, Q output of inverter is shown, it can be obvious that the RL algorithm is not learning on a good pace, making it unstable and outputting more reactive power than it's capable of. To avoid this, the parameters of RL will be changed to see if stability can be reached.

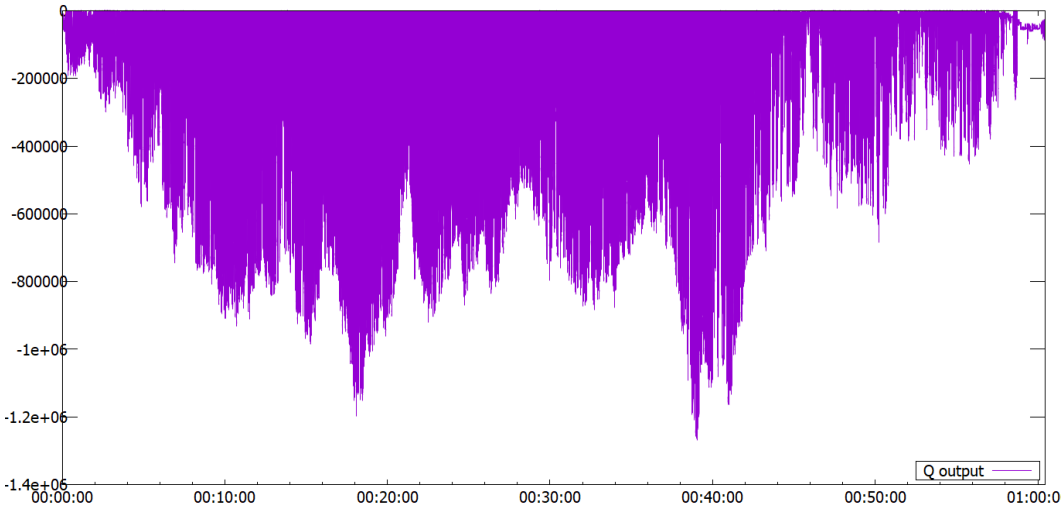

Figure 17. Q output of inverter.

In contrast we have in Figure 17 a graph that shows the external load Output and the Bus Voltage for a period of 30 minutes. The purpose of this graph is to analyze whether the external load was relevant on the bus voltage. In the next section we corroborate it has no effect, since the bus is next to the inverter, it's really affected by its output and the load has almost no impact on the bus voltage.

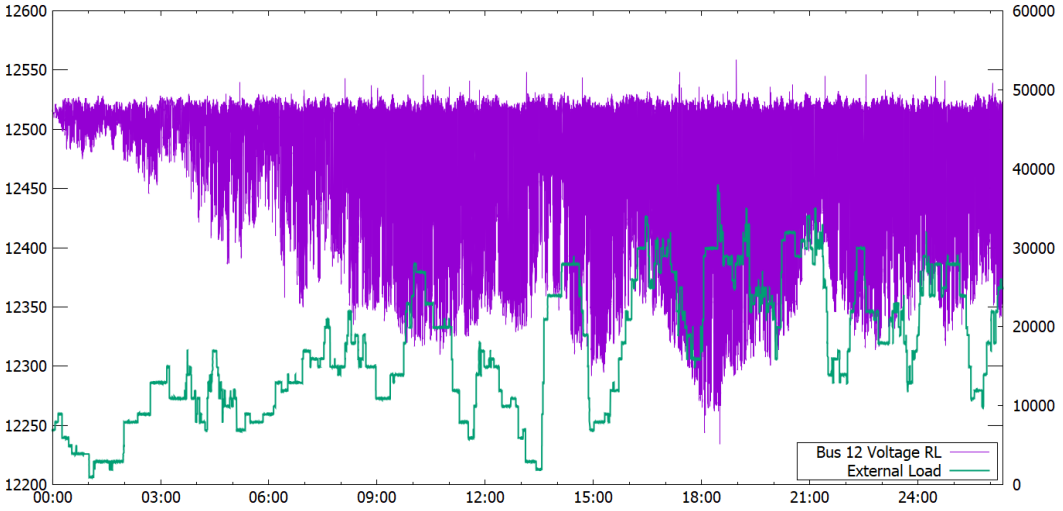

Figure 18. Bus Voltage and external load.

On the second test run, the original parameters of RL where changed, changing α from 0.1 to 0.001 and γ from 0.9 to 0.99 to improve learning and to see if the model would gain stability. Table 4 shows the new parameters, only α and γ were changed, the rest of the parameter remained unchanged.

Parameter	Value
$\boldsymbol{\alpha}$	0.001
$\boldsymbol{\gamma}$	0.99
$\boldsymbol{\sigma_p}$	5
$\boldsymbol{\sigma_i}$	5
$\boldsymbol{c_i}$	5 intervals from 0.995 to 1.035 (p.u.)

Table 4. New parameters of RL.

For the next test run, we wanted to test the robustness of the RL with the new parameters, so the capacitor bank at bus 3 is activated and the results are shown below in figure 18.

As it can be appreciated, the voltage without RL raised excessively and so the voltage with RL, but this remained still within the limits, but still with a difference of 200 volts higher. Thus, for the next simulations, the capacitor bank was switched off.

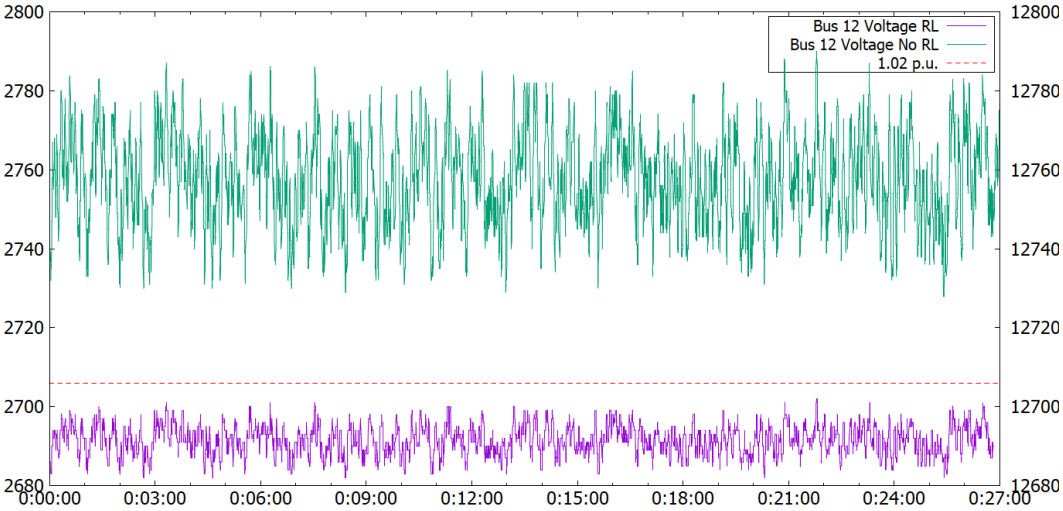

Figure 19. Test drive capacitor bank switched on.

4.2 Sunny day case

After having adjusted the RL parameters, and run some test runs we used real data for the inverter. From Albuquerque, NM on the day of June 22, 2018, being considered an average for a day with good amount of irradiance. This data was provided by [30], which captured the global irradiance for the entire year of 2018 in the mentioned city.

The global irradiance of that day is shown in figure 19. The time where the simulation was run, was between 7:30 am and 6:22pm, being the values above and below these so low, were not taken in account for this study case.

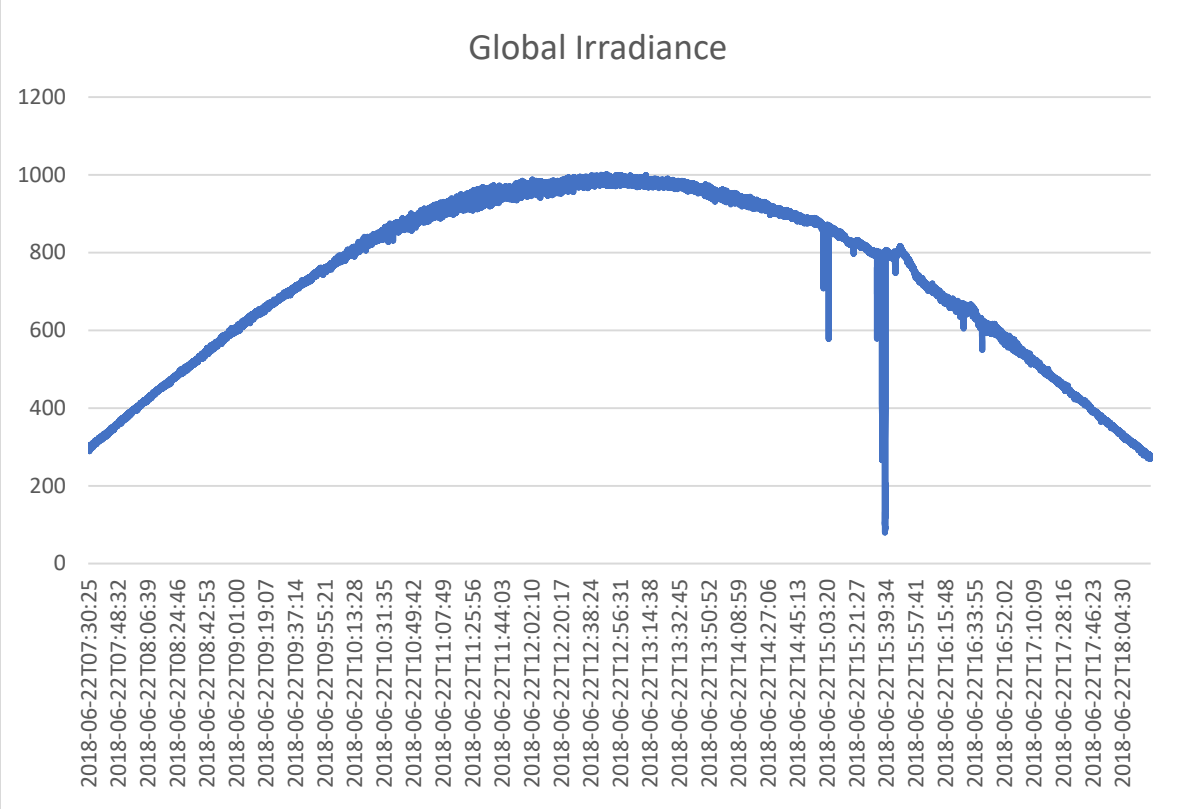

Figure 20. Global Irradiance in W/m2

With the parameters changed and now taking in account Real-Time values of Irradiance (green) in the PV inverter, the Voltage on Bus 12 can be shown in Figure 20 (purple). It can be appreciated that the RL maintains voltage well-regulated and with a variance of +-10 volts and well within range from 1.002 to 1.004 p.u.

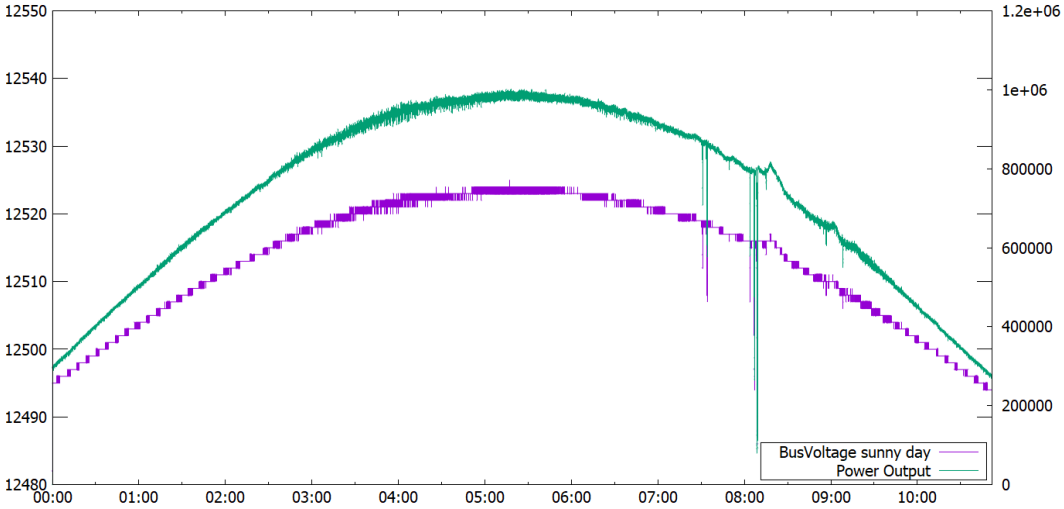

Figure 21. Bus Voltage and Power Output of Inverter.

On the other hand we have figure 21 that shows the output of both simulations with and without RL, and it's clear that the voltage with RL is better regulated, and the one without RL is above the 1.01 p.u. limit with a rise of 100 volts more, which is really significant for sensitive components.

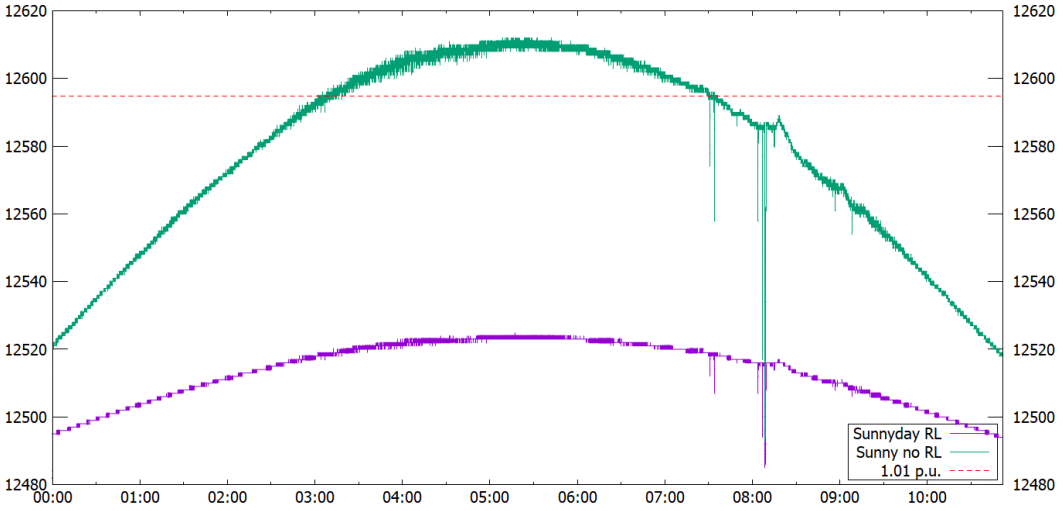

Figure 22. High irradiance on Bus Voltage 12 with and without RL.

In the same manner, we corroborate that the external load that it's been shown in figure 22 has no effect on the RL voltage regulation, since it basically mimics the behavior of the inverter Output.

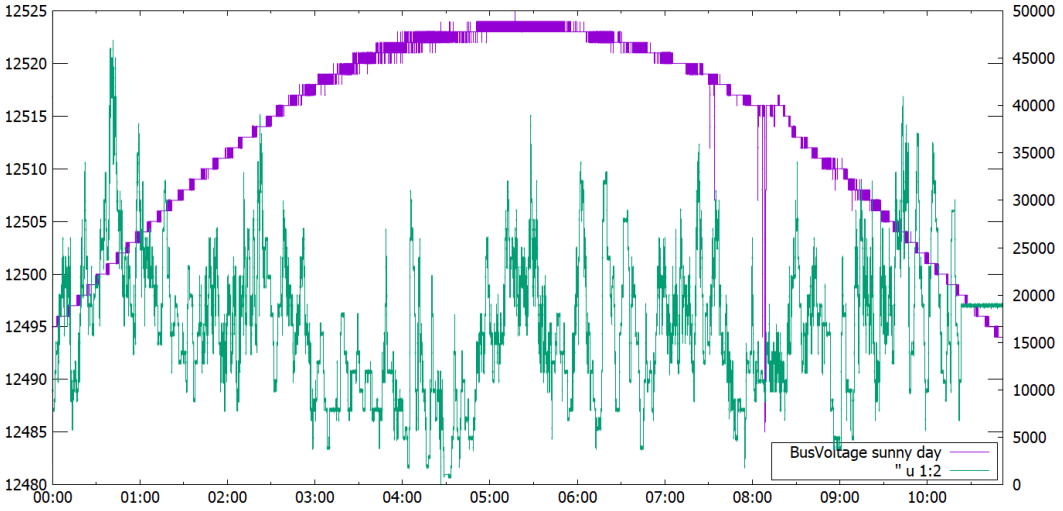

Figure 23. External load Power and Bus 12 Voltage.

Now, in this simulation the Q output in figure 23 (purple) has a good behavior, not oscillating on large values, with a more realistic approach.

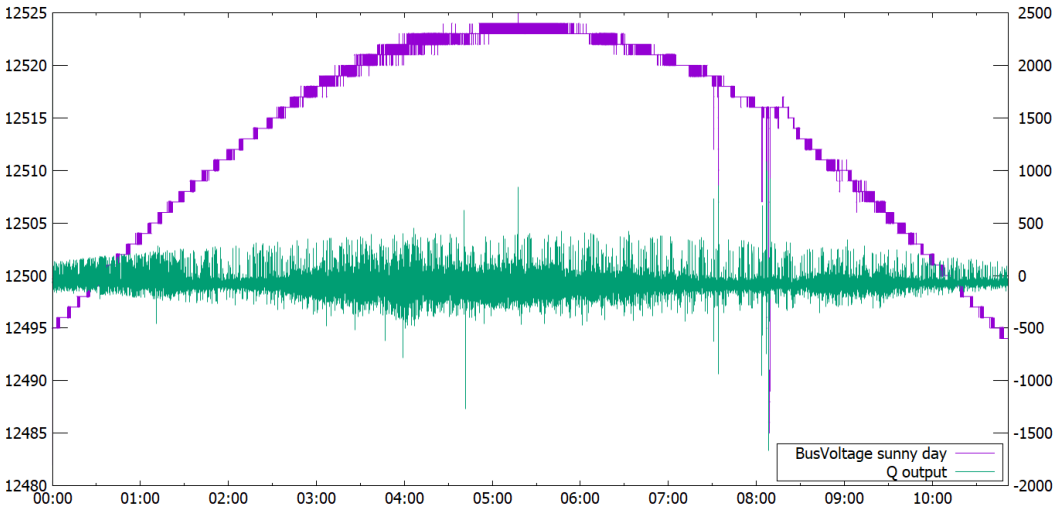

Figure 24. Bus Voltage and Q output of Inverter.

4.2 Cloudy day case

Next Case scenario was taking a day with irregular patterns of irradiance, a cloudy day. For this research the day of August 22 was selected, due to its irregularities, the irradiance of that day is shown in Figure 24.

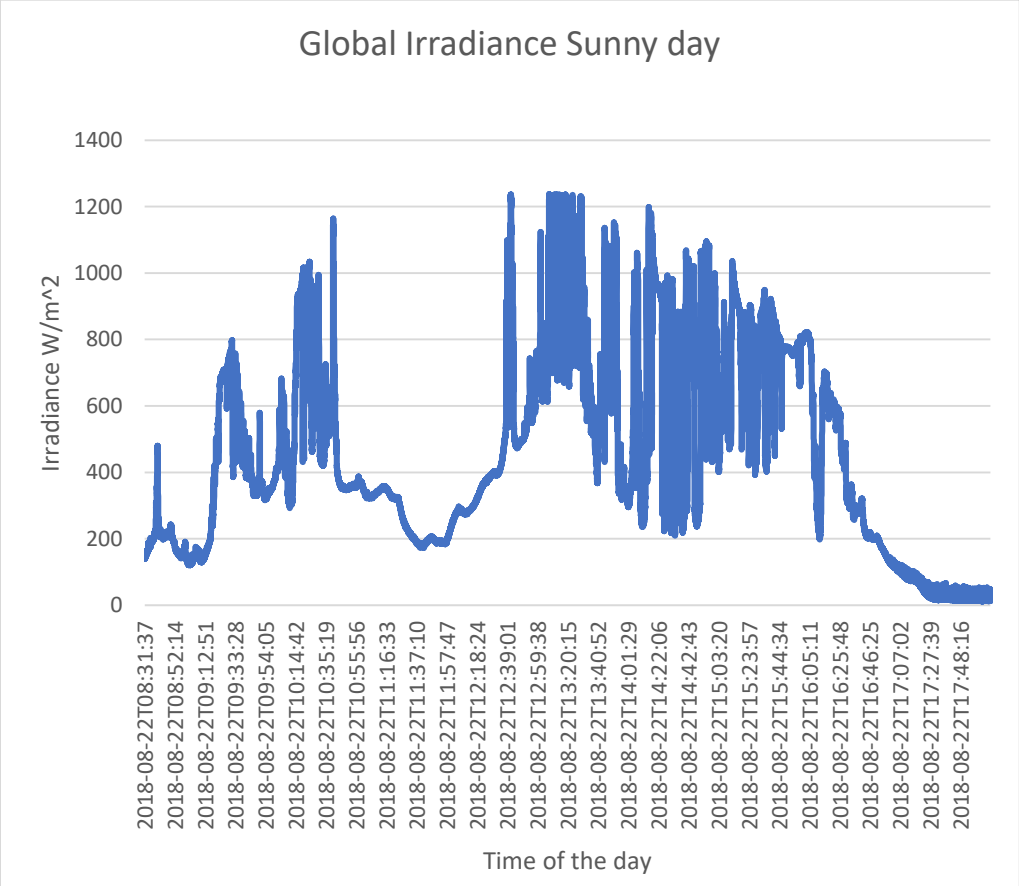

Figure 25. Irradiance on August 22 2018.

As expected from the previous case, Figure 25 shows the Voltage on the two models and they both mimic again the behavior of the inverter output shown in figure 26.

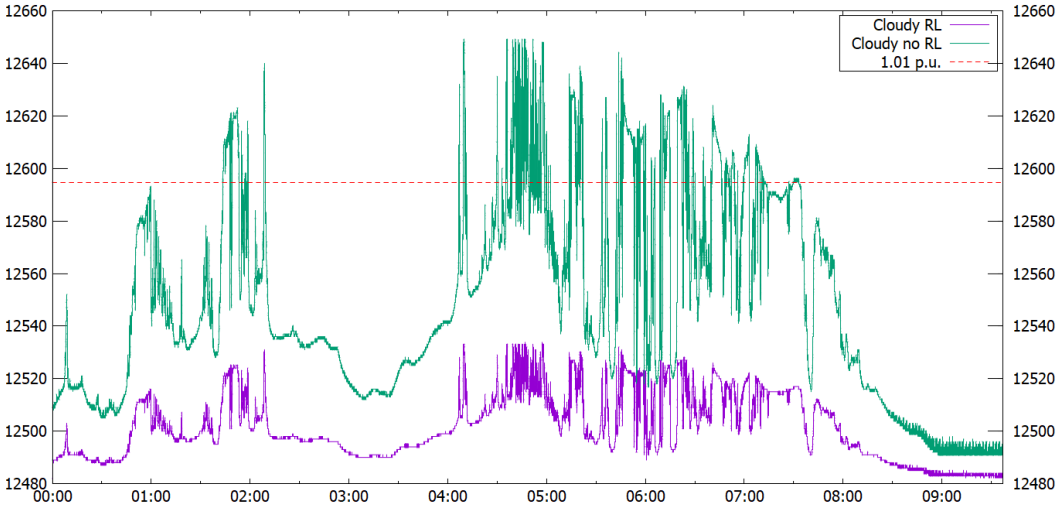

Figure 26 Bus 12 Voltage outputs on a cloudy day.

In this case we have the RL controlled voltage being from $\cong 1.00\, p.u.$ to 1.004 p.u., which coincides with the parameter c_i of a standard deviation of 1.035 p.u. and the voltage without regulation is 200 volts higher than the one being regulated.

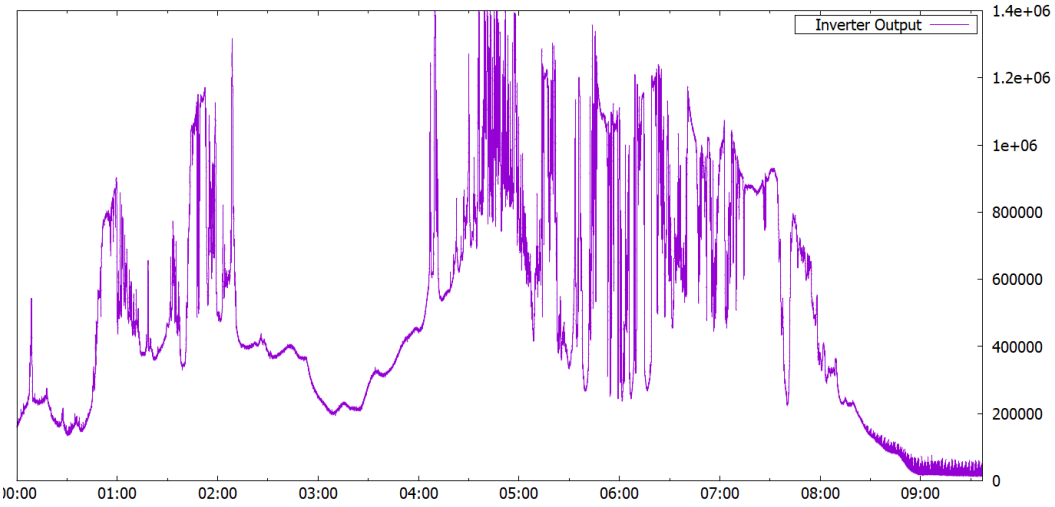

Figure 27 Inverter Output on a cloudy day.

Then again, figure 27 shows the Q output (green) being reasonably stable and helping out maintain the voltage within a regulated range. Although, the voltage here is less stable than the high irradiance case, but this is normal due to the irregular output of a cloudy day.

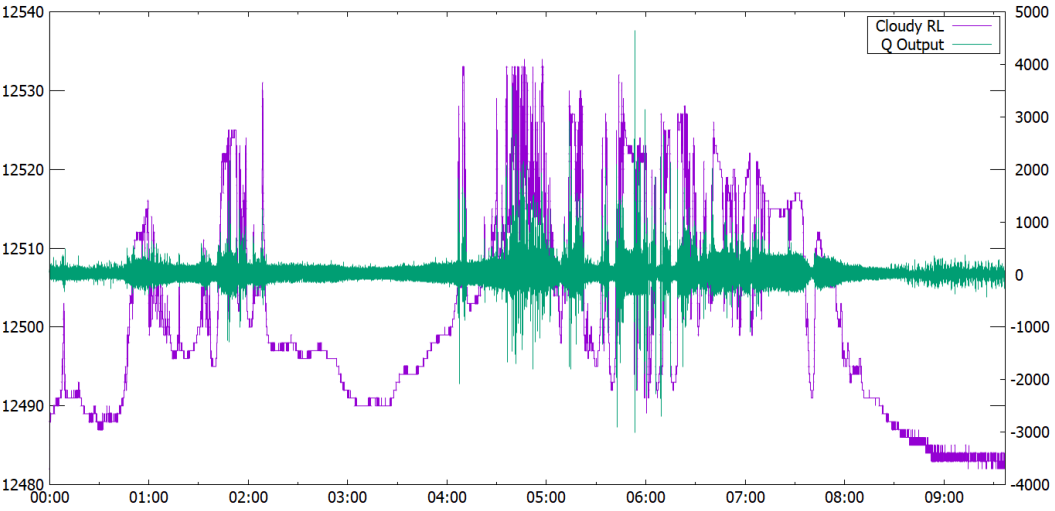

Figure 28. Bus Voltage and Q output on a cloudy day.

Chapter 5:

Conclusions and Future Work

5.1 Conclusions

One of the objectives of this work was to integrate an external load on real-time outputted every second physically to through Opal-RT and analyze a Distribution Feeder Model made in Simulink, which was achieved successfully. Although, for the robustness of the voltages being simulated it proved that what has most influence in the bus voltage variation was the inverter output, being this one a high large PV array (1MW).

The second objective was able to put real data on the Feeder Model outputted every second on the model for a true Real-Time simulation and see how it affected the bus voltage being regulated by Reinforcement Learning.

Lastly the Reinforcement Learning for reactive power control proved to be successful in lower considerably the bus voltage, 100 volts less for a high irradiance day, and 200 volts for a irregular irradiance case.

On the work published by [19], it's says that flexibility and robustness of the proposes method it's something that's needs to be evaluated in the future. In this work this robustness is evaluated since in the previous work, they work with low penetration PV and Low Voltage (LV) network, specifically 210 volts. While in the model presented on this work the target is a medium/high voltage network, working on the 12.47kV and a PV penetration of 1 MW, in contrast with the previous work that had a PV penetration of 5 kW. Therefore the robustness hypothesis of proposed method in [19] it's being evaluated successfully here.

5.2 Future work:

Even though the method worked in maintain the regulated voltage within range, it might need some adjustments and further analysis in the RL and Simulink parameters. Since the non-regulated voltage, even though it was 200 volts higher did not really exceed the 1.02 p.u., maybe this can be fixed using different RL parameters or equations or a different policy or reward. This because, when we switched on the bank capacitor the RL regulated voltage did not followed the RL policy.

It also can be evaluated with other voltage ranges that are lower than the one presented in this work. Since this work wanted to focus in high voltage with high PV penetration for Distribution Feeders. This way the RL would probably work better.

Another consideration for the future is implement more real and physical simulation by adding a real PV inverter simulator or real PV inverter data, for a more realistic behavior of inverter output, input and PV voltage and irradiance.

References

[1] J. N. McDougall, *Fuels and the National Policy*. Butterworths Toronto, ON, 1982.

[2] A. Keyhani and M. Marwali, *Smart power grids 2011*, vol. 53. 2012.

[3] IEC, "IEC Smart Grid Standardization Roadmap," *IEC Rep.*, no. June, pp. 1–136, 2010.

[4] M. E. El-Hawary, "The smart grid—state-of-the-art and future trends," *Electr. Power Components Syst.*, vol. 42, no. 3–4, pp. 239–250, 2014.

[5] A. Parisio, E. Rikos, and L. Glielmo, "A model predictive control approach to microgrid operation optimization," *IEEE Trans. Control Syst. Technol.*, vol. 22,o. 5, pp. 1813–1827, 2014.

[6] J. H. Eto *et al.*, "The CERTS Microgrid Concept, as Demonstrated at the CERTS/AEP Microgrid Test Bed," *US Dep. Energy, Berkeley*, vol. 53, 2018.

[7] Y. Nagaraja, T. Devaraju, and M. Vijaykumar, "Optimal distributed control of renewable energy-based microgrid–an energy management approach," *Int. J. Ambient Energy*, pp. 1–8, 2019.

[8] A. M. R. Lede, M. G. Molina, M. Martinez, and P. E. Mercado, "Microgrid architectures for distributed generation: A brief review," in *2017 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)*, 2017, pp. 1–6.

[9] K. L. Butler-Purry and M. Marotti, "Impact of distributed generators on protective devices in radial distribution systems," in *Transmission and Distribution Conference and Exhibition, 2005/2006 IEEE PES*, 2006, pp. 87–88.

[10] Y.-K. Wu, C.-S. Chen, Y.-S. Huang, and C.-Y. Lee, "Advanced analysis of clustered photovoltaic system's performance based on the battery-integrated voltage control algorithm," *Int. J. Emerg. Electr. Power Syst.*, vol. 10, no. 4, 2009.

[11] A. Bidram and A. Davoudi, “Hierarchical structure of microgrids control system,” *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1963–1976, 2012.

[12] R. Albarracín and H. Amarís, “Power quality in distribution power networks with photovoltaic energy sources,” in *Proceedings of International Conference on Environment and Electrical Engineering*, 2009, pp. 10–13.

[13] R. Tonkoski, L. A. C. Lopes, and T. H. M. El-Fouly, “Coordinated active power curtailment of grid connected PV inverters for overvoltage prevention,” *IEEE Trans. Sustain. Energy*, vol. 2, no. 2, pp. 139–147, 2010.

[14] M. Karimi, H. Mokhlis, K. Naidu, S. Uddin, and A. H. A. Bakar, “Photovoltaic penetration issues and impacts in distribution network – A review,” *Renew. Sustain. Energy Rev.*, vol. 53, pp. 594–605, Jan. 2016.

[15] Y. Li, “Deep reinforcement learning: An overview,” *arXiv Prepr. arXiv1701.07274*, 2017.

[16] J. L. Rojo-Álvarez, M. Martínez-Ramón, J. M. Marí, and G. Camps-Valls, *Digital signal processing with Kernel methods*. Wiley Online Library, 2018.

[17] J. G. Vlachogiannis and N. D. Hatziargyriou, “Reinforcement learning for reactive power control,” *IEEE Trans. power Syst.*, vol. 19, no. 3, pp. 1317–1325, 2004.

[18] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[19] S. Takayama and A. Ishigame, “Autonomous decentralized control of distribution network voltage using reinforcement learning,” *IFAC-PapersOnLine*, vol. 51, no. 28, pp. 209–214, 2018.

[20] R. S. Sutton and A. G. Barto, “Reinforcement learning. 1998,” *Kluer Acad. Pub*.

[21] R. S. Sutton, “Reinforcement learning: Past, present and future,” in *Asia-Pacific Conference on Simulated Evolution and Learning*, 1998, pp. 195–197.

[22] L. P. Kaelbling, M. L. Littman, and A. W. Moore, “Reinforcement learning: A survey,” *J. Artif. Intell. Res.*, vol. 4, pp. 237–285, 1996.

[23] C. B. Nielsen, P. G. Larsen, J. Fitzgerald, J. Woodcock, and J. Peleska, “Systems of systems engineering: basic concepts, model-based techniques, and research directions,” *ACM Comput. Surv.*, vol. 48, no. 2, p. 18, 2015.

[24] C. Gomes, C. Thule, D. Broman, P. G. Larsen, and H. Vangheluwe, “Co-simulation: State of the art,” *arXiv Prepr. arXiv1702.00686*, 2017.

[25] H. Van der Auweraer, J. Anthonis, S. De Bruyne, and J. Leuridan, “Virtual engineering at work: the challenges for designing mechatronic products,” *Eng. Comput.*, vol. 29, no. 3, pp. 389–408, 2013.

[26] V. Ayon, M. Robinson, A. Mammoli, A. Fisher, and J. Fuller, “Integration of bottom-up statistical models of loads on a residential feeder with the GridLAB-D distribution system simulator, and applications,” in *2017 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, 2017, pp. 1–6.

[27] O.-R. Techologies, “eMEGASIM with SimPowerSystems.” pp. 3–4, 2017.

[28] Q.-C. Zhong and T. Hornik, *Control of power inverters in renewable energy and smart grid integration*, vol. 97. John Wiley & Sons, 2012.

[29] V. H. Ayon, “Co-simulation Framework for Power Distribution System Analysis with Humans in the Loop,” 2017.

[30] O. Garcia-Hinde *et al.*, “Evaluation of dimensionality reduction methods applied to numerical weather models for solar radiation forecasting,” *Eng. Appl. Artif. Intell.*, vol. 69, pp. 157–167, 2018.