Microgrid Expansion Planning Using Simulation-based Optimization and Reinforcement Learning

By

Stamatis Tsianikas

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Industrial and Systems Engineering

Written under the direction of

Professor David W. Coit

And approved by

_____

_____

_____

_____

_____

New Brunswick, New Jersey

October, 2020

ABSTRACT OF THE DISSERTATION

Microgrid Expansion Planning Using Simulation-based Optimization and Reinforcement Learning

by STAMATIS TSIANIKAS

Dissertation Director:

Professor David W. Coit

This dissertation provides an analytical framework for tackling the long-term microgrid expansion planning problem. In the wake of the highly electrified future that is ahead of us, the need for reliable and economical power supply will become more urgent than ever. The role of microgrids in fulfilling this need is expected to be highly crucial. While there is a lot of active research going on related to developing optimization models for such systems, the current work innovates by considering both economic and reliability aspects, as well as the stochastic nature of various components in the energy industry. Furthermore, the fact that the microgrids will be placed at the core of the future energy systems will naturally give birth to another important problem from the planning perspective; this problem concerns the derivation of optimal strategies when expanding the microgrids, both in storage and power capacity. The criticality of formulating systematic, analytical and novel methodologies to tackle this problem can be easily justified by considering the steady growth of load demand, the technological advancements continuously being made, and the high operating costs incurred in these processes. The research work that can

be found currently in the literature lacks in considering several peculiarities of microgrids. Moreover, many proposed approaches fail to provide realistic and complex-enough formulations due to the incompetence of traditional solution techniques in handling problems of that scale. The current research work serves as the first attempt to formulate a highly detailed long-term expansion planning problem in microgrid setting and solve it using advanced artificial intelligence techniques. Towards this direction, a simulation-based approach is developed to determine cost-optimal battery sizing under preset reliability constraints, and a unified dynamic optimization framework is built and used to derive holistic optimal expansion strategies. Overall, the goal of the present research work is to provide novel baseline models that give a well-shaped structure to the stochastic problem of long-term expansion planning, while utilizing advanced machine learning tools and techniques.

Acknowledgements

Firstly, I would like to thank my family for their immense love and for providing me whatever I needed (and more) to grow as a human being and as a student. In the present dissertation, I use a lot the concept of "delayed rewards". Well, I think that the people who introduced me first at this one are my parents: work hard now, to enjoy later. My father, Michalis, has always been a role model for me on how morality and ethics should guide us in life. He taught me how to organize my mind, my work and my goals and how to get up quickly after a failure, be "resilient". I wish I had inherited a bit more of his artistic talent, but I am sure he is proud that I am on my way to becoming a great engineer. My mother, Anastasia, has sacrificed so much time and energy into providing for me that I am afraid I will never be able to pay it back. I am also afraid that she convinced me that if I work hard "I can achieve anything that I want". There are some cases of course that this is impossible (unfortunately, I will never become a professional soccer player), but it seems that the word "impossible" does not exist for a mother. My brother, Vangelis, is an impressive example of maturity. I would get advice from him with no hesitation although he is 8 years younger than me. I am more than sure that he will have an incredible amount of growth and success in his life and academic journey. I wish they could all be present for my upcoming dissertation defense, but the current COVID-19 pandemic had other plans. In any case, I am sure we will all celebrate together when time comes.

Words are not enough to express my gratitude for my girlfriend, Maria, for sticking always by my side. Being such an amazing engineer herself, she had already set a pretty high standard for me, before I even start my journey at Rutgers. In terms of personality traits, what I truly admire most is her integrity, both in personal and academic context. I wish she starts believing in herself more, and then I am sure she

will create myriad success stories. I have to acknowledge that hearing about microgrids, optimization and crazy reinforcement learning jargon at 1 AM is not an easy task anyway. Thank you, my love, and I am sure you will get revenge with “medical devices”, “electroporation” and “gene delivery”.

I am also thankful for all these incredible friends of mine that have always my back; friends that I knew from Greece or I met at the US. For instance, Dionysios has been an example of hard work for me and a trustworthy counselor for my US life day and night. Vangelis has always been my buddy, in life and academia (same schools/universities for 20 years). Michalis, Antonella and Giannis were, apart from great friends, my study partners in undergrad. Jian and Nooshin started as just two lab mates for me but now they are my friends and I will always remember them, no matter in which corner of the world we are in the future. I would also like to thank Ayca for these studying all-nighters we pulled together for our PhD qualifying exams.

I am truly honored that I conducted my doctoral studies in the United States and at such an accredited and famous institution, such as Rutgers University. The exemplary academics that Rutgers hosts and the fierce competition in all disciplines are only two of the things that made my experience unforgettable. Although thousand miles from home, I truly had a wonderful time here and I will always be a proud Scarlet Knight.

They say that the choice of an academic advisor is one of the most important decisions someone has to make for his PhD journey. In my case, and although I did not know that at the time, I made one of the best decisions in my life. Professor Coit’s academic rigorousness, depth of knowledge and patience are three of the most crucial keys to my success. The fact that my research skills evolved so much during my PhD cannot be attributed to anyone else than him. Although I am sure we will always have

personal communication, from now on whenever I will hear the phrase “respected academic”, my mind will naturally go to him.

It is undoubtable that I received unwavering support from many professors during my PhD journey. Of course, the members of my dissertation committee belong to that group. Professor Rodgers could be called my co-advisor, given the fact that I bugged him continuously with minor and major details of my dissertation. Professor Birnie is the one responsible for the genesis of my dissertation topic via our collaboration during the first two years of my PhD. I should also thank Professor Xi who provided his valuable support and expertise in the area of energy storage. Special thanks to Professor Jafari, who is an excellent chairperson and helped me personally in numerous aspects during my PhD studies. Professor Felder shared his relevant expertise whenever needed. Moreover, I thank Professor Parlikad, who gave me the opportunity to visit UK and conduct research at a world-class institution, such as Cambridge University. Last but not least, I would like to thank Professor Gursoy because she is not only an amazing educator, but also the person who introduced me first to the world of stochasticity, random processes and the major of Industrial Engineering.

As a final note, I would like to thank my home country, Greece, and my undergraduate university, NTUA, for all the precious things I acquired from them and which I realized better when I left.

Table of Contents

List of Tables

List of Illustrations

Nomenclature

Section 2

SU_g	Start-up cost of unit g
SD_g	Shut-down cost of unit g
$F_g(\cdot)$	Fuel cost function for unit g
p_{gt}	Power generation/dispatch amount of unit g at time t
v_{gt}	Start-up action of unit g at time t, binary
w_{gt}	Shut-down action of unit g at time t, binary
$VOLL$	Value of lost load, \$/kWh
d_{it}	Load lost at bus i for time t
G	Set of generating units
T	Set of time intervals in the planning horizon
N	Set of buses
u_{gt}	Commitment decision, binary
ON_g	Minimum ON-duration
OFF_g	Minimum OFF-duration
$\|T\|$	Duration of a planning horizon
P_g^{min}	Minimum generation limit in an hour for generator g
P_g^{max}	Maximum generation limit in an hour for generator g
RD_g	Ramp-down rate for unit g
RU_g	Ramp-up rate for unit g
$F_g^e(\cdot)$	Emission function of unit g
SU_g^e	Start-up emission of unit g
SD_g^e	Shut-down emission of unit g
E^{max}	System emission limit
$E(\cdot)$	Expectation of load loss in a power system
ε_t	Loss allowance for time t
$L(\cdot)$	Loss function of the neural network
w	Neural network weights
α	Learning rate
W_{hx}, W_{hh}, W_{yh}	Input-hidden, hidden-hidden, hidden-output weight matrices
b_h, b_y	Biases for the hidden and output layer

f, g, φ	Activation functions
W_{fx}, W_{fh}	Input-hidden and hidden-hidden weights for forget gate
W_{ix}, W_{ih}	Input-hidden and hidden-hidden weights for input gate
W_{sx}, W_{sh}	Input-hidden and hidden-hidden weights for memory cell gate
W_{ox}, W_{oh}	Input-hidden and hidden-hidden weights for output gate
b_f, b_i, b_s, b_o	Biases for forget, input, memory cell and output gate
W_{zx}, W_{zh}	Input-hidden and hidden-hidden weights for update gate
W_{rx}, W_{rh}	Input-hidden and hidden-hidden weights for reset gate
b_z, b_r	Biases for update and reset gate
L^{CNN}	Number of hidden layers in the convolutional neural network
M^{CNN}	Number of filters in the convolutional neural network
w_p^l	Filter p for hidden convolutional layer l
N_{in}	Number of time steps in the input window
N_{out}	Number of time steps in the output window
n_f	Number of features in the dataset
n_h	Number of different facilities in the dataset
f^m	Single model estimator
f^j	Model estimator for the output time step j
S	Set of states
A	Set of actions
R	Set of rewards
S_t	Instance of a state at time t
A_t	Instance of an action at time t
R_t	Instance of a reward at time t
π	Policy of an agent
p	Probability transition matrix
G_t	Discounted sum of rewards up to time t
γ	Discount rate
v_π	State-value function
q_π	Action-value function
v_*	Optimal state-value function
q_*	Optimal action-value function

N	Number of possible states
Q, Q^A, Q^B	Look-up two-dimensional array
ε	Exploration/exploitation tradeoff parameter
d	Uniformly distributed random variable
θ	Set of parameters for a neural network

Section 3

B_r	Battery capacity, kWh
b	Battery price, $/kWh
B_{min}	Minimum battery energy value, kWh
DoD	Maximum depth of discharge of battery
$Q_B(t)$	Energy stored in battery at time t, kWh
$P(t)$	Power generated by PV array at time t, kW
$D(t)$	Load demand at time t, kW
$AEL(t)$	Amount of energy lost at time t, kWh
e	Round-trip efficiency of battery
η	Energy conversion efficiency of PV array
I_t	Solar irradiation at time t, W/m^2
A	PV array area, m^2
T	Simulated grid outage duration, hrs
K	Planning horizon, yrs
CAIDI	Customer Average Interruption Duration Index, (hrs/interruption)
SAIFI	System Average Interruption Frequency Index, (interruptions/yr)
$LOLP_{jk}$	Loss of load probability for j^{th} grid outage for year k
CCP	Chance constraint probability
TSC	Total system cost, $
N_k	Set of outages for year k
O_{jk}	Set of time intervals for j^{th} grid outage for year k
S	Total number of simulations
I_c	Initial charging state of battery when grid outage occurs
C_p	Proportion of critical load demand during an outage
t_{ijk}	The i^{th} time interval for the j^{th} outage in year k

T_{jk}	Duration of j^{th} grid outage of year k, hrs
α	Constraint limiting $E[LOLP]$
C_s	Cost of s^{th} simulated outage, $
β	Constraint limiting $LOLP_{jk}$
ζ	Constraint limiting CCP
$\delta(t)$	An indicator function for whether the load demand at time t is lost or not, binary
G_s	Set of time intervals for the s^{th} simulated grid outage
t_{is}	The i^{th} time interval of the s^{th} simulated grid outage
C_s	Cost of lost load corresponding to the s^{th} simulated grid outage
$LOLP_s$	Loss of load probability of the s^{th} simulated grid outage
λ_s	Indicator for the achieved $LOLP$ of the s^{th} simulated outage

Section 4

$VOLL^g$	Value of lost load for a facility g, $/kWh
C_p^g	Critical load factor for a facility g
G	Set of existing facilities in microgrid
p_c^i	Charging proportion of storage unit i
p_d^i	Discharging proportion of storage unit i
$P_{annuity}$	Annual payment amount of investment, $
$P_{principal}$	Principal payment amount of investment, $
IR	Annual interest rate
L^i	Lifetime of i^{th} storage device
$P_{solar}(t)$	Power production of solar array at time t, kWh
$P_{wind}(t)$	Power production of wind turbine at time t, kWh
n_{cpp}	Number of solar cells per panel
n_{pan}	Number of solar panels
ρ	Air density, kg/m^3
$W(t)$	Wind speed at time t, m/s
W_{in}	Wind cut-in speed, m/s
W_{out}	Wind cut-out speed, m/s
S^{tf}	Timing feature of state space
S^{ef}	External feature of state space

S^{if}	Internal feature of state space
SU	Set of storage units in the system
SL	Set of available expansion levels
f^{tf}	Timing component of state transition function
f^{ef}	External component of state transition function
f^{if}	Internal component of state transition function
p^{ef}	Transition matrix of the $\mathbf{s}^{ef}$ DTMC
$r_k(\cdot)$	Reward function for the k^{th} decision period of the problem
f^{RF}	Random forest function for the outage cost component
$N(t)$	Single Poisson process
μ	Parameter of $N(t)$ Poisson process
$N_1(t)$	Poisson process for regular outage events
$N_2(t)$	Poisson process for severe outage events
μ_1	Parameter of $N_1(t)$ Poisson process
μ_2	Parameter of $N_2(t)$ Poisson process
$N'(t)$	Superposed Poisson process for regular and severe outage events
Z_n	Type of the n^{th} outage event
T	Duration time for an outage event, hrs
T_1	Duration time for a regular outage event, hrs
T_2	Duration time for a severe outage event, hrs
κ	Parameter of the T Poisson distributed random variable
κ_1	Parameter of the T_1 Poisson distributed random variable
κ_2	Parameter of the T_2 Poisson distributed random variable

Section 5

PP	Set of power plants in the system
CL_i	Available expansion levels for microgrid entity i
EC_i	External characteristics of microgrid entity i
IC_i	Internal characteristics of microgrid entity i
γ	Discount rate
LT	Loan term, yrs
K	Number of decision periods

y	Number of years in a decision period
ME	Set of existing microgrid entities in the system
C_k^{inv}	Investment cost component for decision period k, \$
VC_p	Variable cost of power plant p, \$/kWh
$P_p(t)$	Power production of power plant p at time t, kW
$E(t)$	Electricity price at time t, \$/kWh
$eVC(t)$	Effective variable cost of all power plants at time t, \$/kWh
C_k^{opr}	Operational cost component for decision period k, \$
C_k^{los}	Outage cost component for decision period k, \$
OM^i	O&M rate for microgrid entity i
CC_k^i	Capacity of entity i installed at the system at decision period k
C_k^{om}	O&M cost component for decision period k, \$
RT^i	Retirement rate for microgrid entity i
C_k^{ret}	Retirement cost component for decision period k, \$
β^g	Constraint limiting $LOLP_{jk}^g$
ζ^g	Constraint limiting CCP_k^g
FP	Set of feasible policies of the constrained problem
λ_g	KKT multiplier for facility g
Q^A	Main Q network
Q^B	Target Q network
b	Experience replay buffer
$\delta_c(t,g)$	Counter for hours of unmet demand up to time t for facility g
z	Multiplier for controlling exponential $VOLL$

1. Introduction

The main topic of the dissertation is to present novel optimization frameworks for traditional long-term energy planning problems. Although the focus is given to microgrids, this research work intends to become the first comprehensive approach to tackle long-term planning problems with modern algorithmic techniques and to provide holistic approaches that could be generalized to other areas as well. Microgrids are local energy grids, close to distribution, which can operate connected or disconnected from the grid and are mostly comprised by distributed generators, batteries, and/or renewable resources (DoE 2014). In the forthcoming years, the optimization of planning, operation, and expansion, either in storage or in power capabilities, of the microgrids will be crucial and will affect millions of customers who currently, or will in the future, have their load demand served by these microgrids. The need for a reliable and resilient power grid will become even more urgent in the highly technological future that lies ahead and the microgrids could be the path to this kind of grid.

It is undoubtable that there are plenty of ongoing research studies and industry reports discussing the tradeoffs between conventional power plants and renewable energy resources, like solar or wind (Zhou, Tsianikas et al. 2019). The question of whether the world will stop using fossil fuels in the near future remains unanswered and depends heavily on the oil, gas and coal reserves around the globe (Helm 2016) (Covert, Greenstone et al. 2016). Oil peaking is an academic debate that has not yet concluded, and some optimists believe that the oil production rate will continue to rise until 2100 (Abas, Kalair et al. 2015). However, what is for sure known is that fossil fuels resources are finite and environmentally hostile compared to renewable energy resources (Atilgan and Azapagic 2015). Renewable energy is a physically replenished

source of energy which is coming either directly (thermal, photo-chemical and photo-electric) or indirectly (wind, hydropower) from the sun, or from other environmental processes (geothermal) (Ellabban, Abu-Rub et al. 2014).

1.1 Problem motivation

One of the most critical reasons that has contributed to the increasing attention that is given to microgrids is indubitably the cost-effectiveness of certain renewable energy sources. This is one aspect of the problem that has changed drastically during the last years and it can be verified by calculating and comparing the Levelized Cost of Energy (LCOE) for different energy generation technologies:

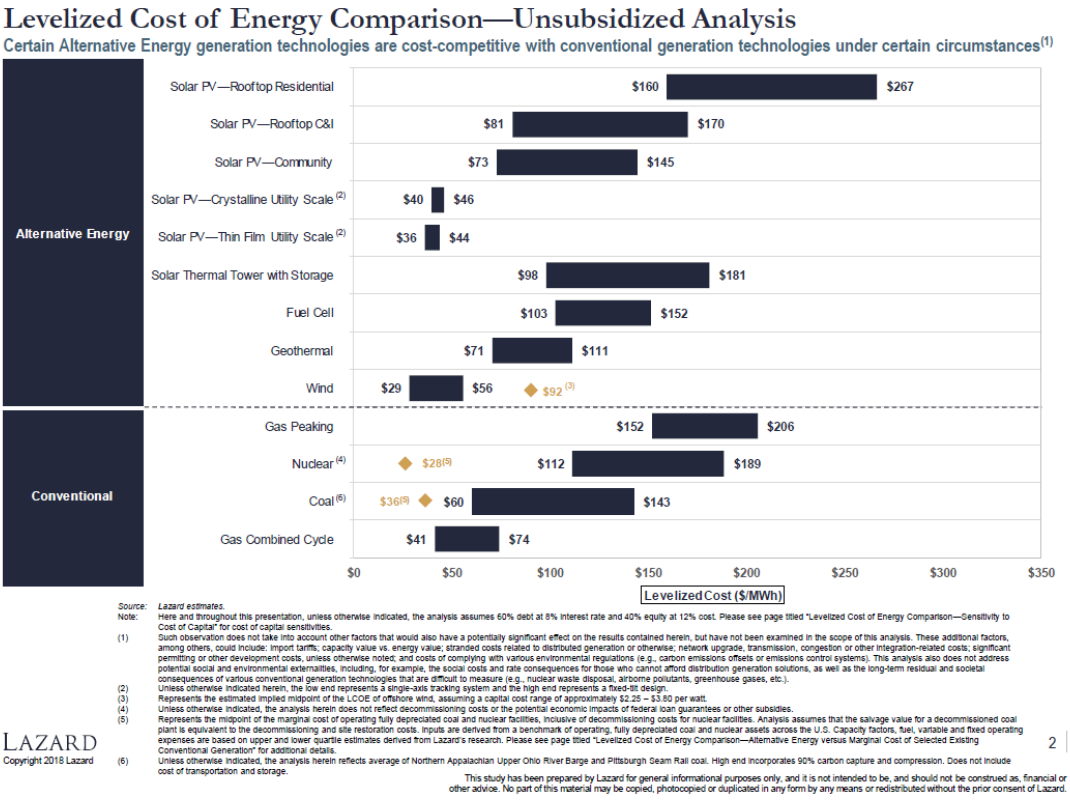

Fig. 1.1 LCOE comparison for various energy generation technologies (LAZARD 2018)

It can be seen in Fig. 1.1 that certain alternative energy generation technologies are now cost-competitive with some conventional generation technologies that were traditionally considered as more economical. Moreover, recent studies have shown that the impact of renewable energy consumption to the economic

growth of developed and developing countries may be much more statistically significant than what was initially thought (Inglesi-Lotz 2016). The enthusiasts around renewable energy sources are exponentially growing and the reasons around this fact are various and well beyond the economic reasons mentioned above. Inside this set of renewable technologies, solar energy is the one that has attracted the most interest, although its contribution to the worldwide energy supply is still negligible. However, its penetration in the energy market is constantly increasing, as it can be shown in Fig. 1.2:

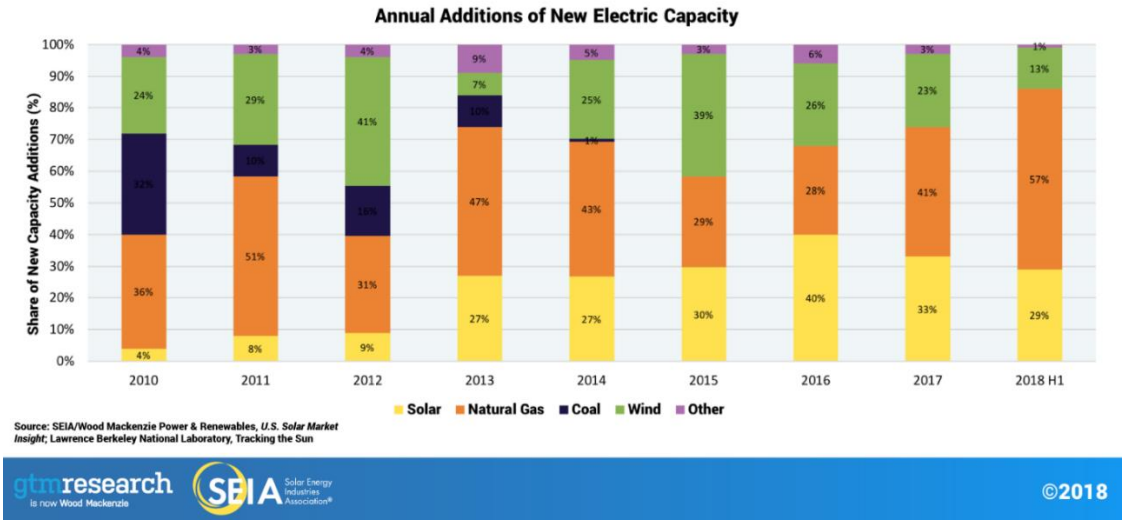

Fig. 1.2 Annual additions of new electric capacity in the U.S. (SEIA, 2018)

Except for solar energy's increased contribution to installed capacity, one thing that is indubitable by observing Fig. 1.2 is that coal's new annual additions are almost diminished, with the last significant ones dating back to 2014. This fact is indeed very optimistic for the future of renewable technologies in general.

However, there is a major problem occurring almost always when renewable energy resources are extensively deployed: intermittency. Naturally, renewable energy sources, such as solar or wind energy, are not dispatchable and cannot be predictable with an adequate assurance (Hakimi and Moghaddas-Tafreshi 2014) (Su,

Wang et al. 2014). The operation of solar PV and other distributed resources without energy storage still requires dependence on the grid, which impairs their ability to supply reliable power to customers during grid outages and to maintain system stability (Zhou, Tsianikas et al. 2019) (Chauhan and Saini 2014). Nevertheless, the addition of batteries or in general energy storage units to a PV system can balance the mismatch between the energy generation and the load demand (Caruana, Sattar et al. 2015).

Studying the problem of the generation and capacity expansion planning of the grid was always a highly important task from both economic and reliability aspects. However, tailoring the methodologies to storage sizing and renewable energy plants additions in a microgrid, and more specifically creating dynamic programming models to solve these problems, are nowadays becoming even more crucial research topics. This fact can be supported based on two main reasons which are explained in the following paragraphs.

The transition from traditional fossil fuels-based plants to renewable energy plants and accordingly from a centralized grid to distributed energy resources seems inevitable for the foreseeable future. The serious concerns for avoiding, or at least mitigating the effects of, climate change coincides with an ever-increasing electrification of the grid, which may be even sharper if electric vehicles are massively adopted by the public (McKinsey 2018). These two facts combined will result in a compulsory need for a record-breaking reliable and resilient grid, which will simultaneously need to incorporate a high penetration of renewable energy sources. This need will bring microgrids and their optimal design in the first place of attention among other energy-related research topics.

Furthermore, the technological advancements that are currently being made and that will facilitate this whole process need to be highlighted. More specifically, it is already reported that solar panel investment costs, as well as battery costs, have significantly declined in the past years and are forecasted to continue in this decreasing trend (IRENA 2017). These reported results can be seen in Figs. 1.3-1.4:

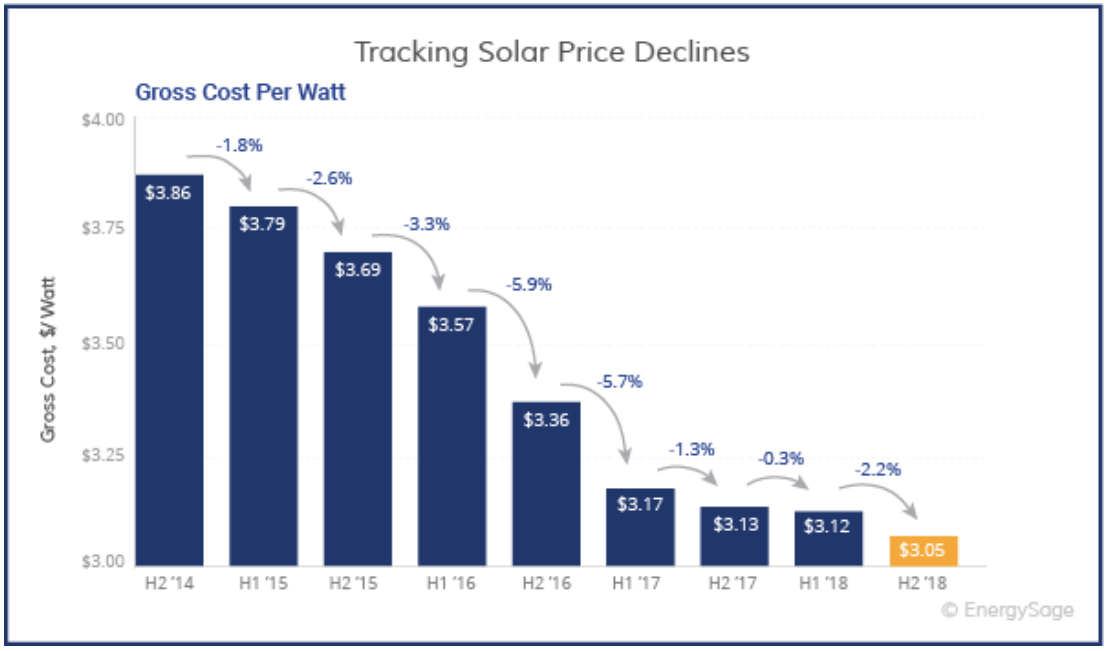

Fig. 1.3 Gross cost per watt for solar price (EnergySage 2019)

In Fig. 1.3, it is shown that the gross cost of the solar price decreased from 3.86 $/Watt in the second half of 2014 to 3.05 $/Watt in the second half of 2018. That means approximately a 21% cost reduction in only 4 years.

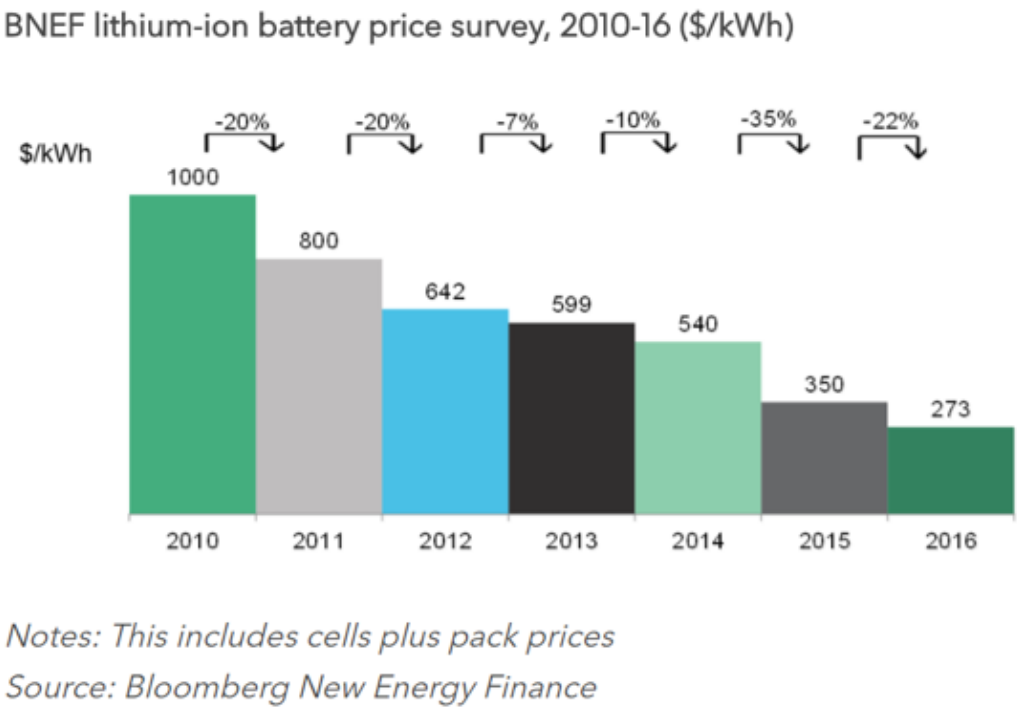

Fig. 1.4 Lithium-ion battery price, 2010-2016 (Colonnese 2017)

In Fig. 1.4, a lithium-ion (Li-ion) battery price survey is presented for the years of 2010-2016. The Li-ion battery is the dominant battery type at the moment for the majority of applications in the energy, as well as the automotive, industry. Herein, it is seen that the price of this specific Li-ion battery has fallen from 1000 $/kWh in 2010 to 273 $/kWh in 2016, indicating a 72.7% decrease in 6 years.

These findings are indeed very promising for the energy systems that will need to have reliability and resilience as a top priority in the future. When examined jointly with the electrification of the grid mentioned above, it is rational to conclude that in the future power grids, it will be increasingly unaffordable to experience outage interruptions and simultaneously cost-efficient to install larger storage capacities and to increase renewables penetration.

1.2 Research overview and objectives

In the present research work, all the necessary frameworks for the long-term microgrid expansion planning problem are provided. Encouraged not only by the

technological advancements which are expected to happen in the renewable energy industry, but also by the emerging need for reliable and cost-efficient decentralized power supply in several applications, a specific focus on microgrid studies is required. It is necessary therefore at this point to emphasize on the deficiencies that the current literature possess. First of all, most of these works seldom consider simultaneously the economic and reliability objectives of the system planners, as well as the stochastic nature that inherently exists in the problem. It is one of the main research objectives to fulfill this gap in the literature. But besides that, most of the current research studies fail to present a holistic optimization framework for determining optimal strategies when expanding the microgrids, both in storage and power capacity. More specifically, these studies do not tailor their applicability in the particular case of microgrids and even when they do so, most of them fail to formulate problems which can present reasonable and realistic levels of details. The reason for that is merely due to the unfitness of traditional solution techniques in solving problems of that scale. The current research work plans to serve as the first attempt to formulate a highly detailed long-term expansion planning problem in microgrid setting and to solve it using advanced artificial intelligence techniques.

The general optimization framework for solving the long-term microgrid optimization problem encompasses several sub-components, all of them serving its general goal. These sub-components are briefly described here before presented thoroughly in the subsequent sections of the dissertation.

Initially, a model for determining the optimal battery size to be attached in a given photovoltaic array is presented, along with its corresponding case study and results. In this work, a simulation-based optimization method is deployed to investigate the effects of battery size on meeting load demand of facilities at different

reliability levels during grid outages with minimal cost. The efficacy of the proposed simulation method is demonstrated by numerical examples using actual data sets of solar irradiation and example facilities' load profiles hourly throughout a whole year. Afterwards, specific focus is given in selected problem parameters, in order to examine how they affect the optimal solutions and how they interact with each other. The underlying relationship between the changes and uncertainty of unmet load penalty costs and battery price is explored by comparing the optimal total system cost under different scenarios. The results can provide guidance and insights regarding the impact of cost-related factors on photovoltaic + battery system design to make them grid-outage resilient and economically viable.

Secondly, a novel framework for dynamically expanding the microgrid's storage capacity using Markov Decision Processes (MDP) and reinforcement learning algorithms are shown and explained. The detailed mathematical formulation is given, as well as the necessary information for the algorithms used. In this section, highlighting and showcasing the importance of answering properly four questions is the priority: whether you actually need to add storage in your energy system, when to install this storage, how much capacity you should add and which storage technology you should choose. Along with the answers to these questions, supplemental research findings of this model are presented and discussed.

Afterwards, the focus is given on expanding the previous models to include the investment of power plants in the microgrid, as well as potential retirement of existing units. Therefore, the resulting model encompasses both power and storage expansion capabilities. Moreover, deep neural networks are utilized in conjunction with the reinforcement learning algorithms in order to pass the barrier of the increasing computational complexity of a realistic enough problem. The incorporation

of reliability in the present dynamic model is now accomplished via the introduction of KKT multipliers in the reward function. Finally, a sensitivity analysis is conducted through exploration of alternative scenarios for the future, such as the high electrification of the grid via the mass adoption of the electric vehicles.

In the last section of the dissertation before the conclusion, the problem of forecasting short-term residential electricity load is studied. The motivation behind this research part lies in the fact that in most real cases, perfect information is not available for the microgrid planner when decisions concerning energy scheduling need to be made. Herein, perfect information refers mostly to the actual energy demand, but it could also refer to meteorological data. Therefore, there is a crucial need to develop and test analytical methodologies to forecast the aforementioned time series. These methodologies need to be specifically tailored to short-term residential load; the reason is that residential load is significantly more unpredictable than other types of loads in the industrial or commercial sector. Consequently, in this section various deep learning-based methodologies are developed and tested against each other in three aspects: their general performance in forecasting the energy demand, their ability to capture the “peaks” and “valleys” which are inherently taking place in residential load data and finally their computational efficiency.

1.3 Research contributions

The major contribution of this research is to provide a general and comprehensive approach in tackling long-term capacity and power expansion planning problems, specifically tailored for microgrid applications. Despite the fact that the generalization of this framework and the reproducibility of the results are one of the author’s main priorities, the models are tested with real and location-specific environmental, load demand and grid outage data for the state of NY. Although the

most important novelty of the current research work is to derive optimal microgrid expansion plans through analytical methodologies and tested methods, there are various contributions which can collectively serve towards the desired direction. These research initiatives are presented as follows:

1. Provide a novel framework for the optimal storage sizing problem specifically tailored to renewable-based microgrid systems. In the wake of technological advancements which will consecutively bring lower storage investment costs, the significance of this research contribution becomes even higher.
2. Formulate the first-ever unified dynamic optimization problem which is able to derive optimal expansion policies for a finite time horizon, either these actions concern storage or power plants additions. It is safe to assume that the analytical consideration and incorporation of stochastic modeling for several aspects of the problem are able to further strengthen this research contribution.
3. Propose new modeling approaches and methodologies that can prove to be sufficient in their capability of solving large-scale expansion planning problems and can replace existing simplified formulations and techniques. Towards this direction, it is mandatory to utilize and expand advanced machine learning tools and techniques that have not been used in long-term energy planning problems before.

Overall, the current research work attempts to study the crucial topic of optimal expansion planning in microgrid setting, derive analytical methodologies to do so and provide a baseline for even more detailed frameworks in the future, which are going to be increasingly complex and realistic.

1.3.1 Theoretical contributions

In order to be more specific, presenting the research contributions should start by the two main theoretical ones. These contributions are specific to the topics of renewable energy systems combined with storage, power systems planning and dynamic programming. They are presented as follows:

1. Provide fresh optimization methods to system planners which allow them to simultaneously minimize their incurred costs and achieve their desired reliability performance. By exploiting the power of the created techniques, the whole process of selecting power plants and storage units for microgrid applications can be automated without the need of making unnecessary simplistic assumptions.
2. Put together a rigorous and novel long-term microgrid expansion planning problem, by studying and incorporating all the existing stochastic behaviors that can be found in the area. The inherent dynamic nature of the problem is accommodated with the use of Markov Decision Processes, which have seldom been used in this type of problems before.

Elaborating more on the first theoretical research contribution, it should be mentioned that there are several studies whose subject is to examine thoroughly the economics and the reliability of microgrids in general or more specifically the storage sizing problem in energy systems. However, these studies seldom consider the outage costs explicitly in the objective function, on top of any reliability constraints in the mathematical formulation. Moreover, a sensitivity analysis is conducted which links directly the desired reliability improvements in an energy system with the associated cost that it is incurred. Finally, another novelty of this work is the accurate

probabilistic outage modeling developed, which is based on real and location-specific outage data.

Concerning the second contribution, it should be noted that expansion planning problems in energy systems, either referring to generation, transmission or capacity expansion, are being solved for a long time by many researchers in the field. Moreover, reinforcement learning algorithms, like Q-learning, have also been used in the field but mostly in short-term planning problems, such as unit commitment or battery scheduling. However, the special novelty of this work is that it attempts for the first time to use Markov Decision Processes and reinforcement learning techniques, either established or modern ones, to solve long-term energy planning problems which are becoming more and more important in the wake of technological advancements and governmental attempts to pursue ambitious goals for the future of renewables in the energy sector.

1.3.2 Applied contributions

Herein, it is considered appropriate to continue with some other very important research contributions of the current work, which however relate to more practical issues and could be adopted in the near future by practitioners in the field of energy systems. These applied contributions include but are not limited to:

1. Obtain more realistic and better engineering solutions while studying the problem of sizing battery capacity in existing photovoltaic arrays. Original simulation-based optimization frameworks and real-time and location-specific datasets help to build towards this research contribution.
2. Introduce advanced reinforcement learning techniques to tackle large-scale versions of the proposed stochastic dynamic optimization

problem. That way, the solutions can be adapted to various scenarios in the future and can facilitate long-term energy planning via developing a priori awareness of what the outcomes of specific actions will be.

For the first out of these two contributions, the findings of the current research indicate that there can be situations when a decision to invest in a larger battery is accompanied by a reduction in total costs incurred. The reason for this lies in the fact that the savings from lowering unmet demand are greater than the investment cost for a larger battery. This is a highly significant result, which proves that the proposed methodology can be used successfully by many practitioners in the field whose tasks are to optimally design backup energy systems.

Concerning the second applied contribution, it is observed in the literature that many existing long-term planning models are seldom incorporating simultaneously a stochastic and dynamic behavior in their defined problems, a fact that makes the results of limited practical use for the future, no matter how scientifically accurate the proposed approach could be. Nevertheless, even in the cases where stochastic dynamic models have been proposed, the solution approaches chosen are classical dynamic programming approaches. These are algorithms of great theoretical importance, but they assume a perfect model for the environment and also, they are computationally expensive. Therefore, the proposed approach serves as a great baseline for more detailed and advanced frameworks which would be able to be used for extracting precise and flexible long-term expansion plans.

To sum up, a visual representation of the main research tasks and subtasks that the present work accomplishes is given in Fig. 1.5:

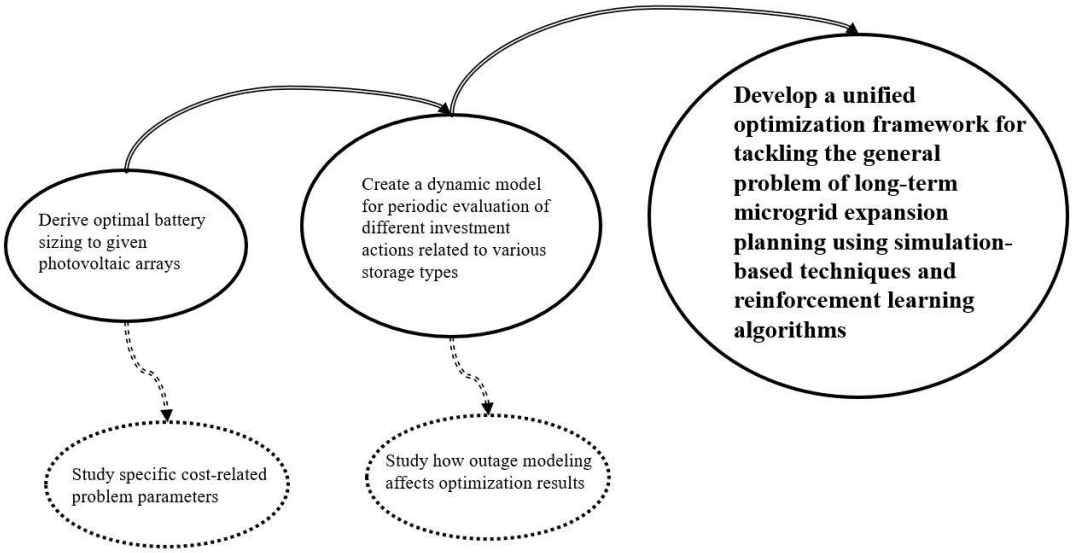

Fig. 1.5 Main research tasks and subtasks

2. Background and Literature Review

The literature review of this dissertation needs to cover a wide range of topics related to this research and therefore is broken into five main parts. Firstly, information about generating units in the electrical power systems is provided, as well as some important planning problems that arise from their study, such as the generation expansion planning (GEP) problem. In the second part, specific emphasis is given in the renewable energy sources and their crucial role in designing modern distributed energy systems is highlighted. In the third part of this literature review, the concentration is given on the importance of energy storage in microgrids, as well as on some existing and potential storage solutions in the market. Afterwards, it is explained why and how neural networks can prove a valuable tool for solving the problem of energy load forecasting. Lastly, Markov decision processes (MDP) are presented and the notions of dynamic programming and reinforcement learning are introduced; their clear understanding is proven very important in Sections 4 and 5 of the present dissertation.

2.1 Electrical power systems planning

Traditional electric power systems are designed with the main purpose being the power production in central generating stations and its delivery to the points of end use via transmission and distribution systems (NREL 2008). Herein, the role of generating units is pretty straightforward; they convert energy from other sources to electric energy, ready for industrial, residential or commercial consumption. However, the roles of transmission and distribution systems are more complicated and interrelated than the independent purpose of the generation system. Transmission system helps to transfer the power over sufficiently long distances and consequently make the operation of generating stations feasible, optimal and under the desired and

preset reliability levels. On the other hand, the distribution system is responsible for the delivery infrastructure in order for the power to be able to “meet” the load in the final destination. It should be mentioned that most distribution systems require bidirectional power flow between power substations and the end-user loads and also dictate the existence of sufficient power supply from the transmission systems. In Fig. 2.1, most of the aforementioned processes can be seen schematically:

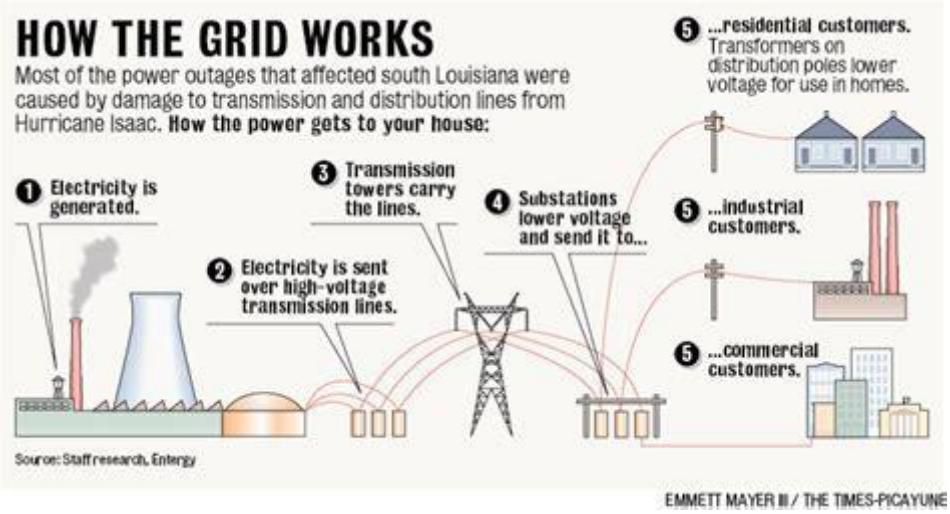

Fig. 2.1 Overview of how a typical electric grid works (Staff research, Entergy)

2.1.1 Overview of generating technologies

The backbones of electric power systems are all the various ways for actually producing electricity. Of course, there are multiple types of generating units that are used to satisfy the load demand, and there are various taxonomies which can be used for organizing the different technologies. Before giving a categorization of generating units, what is an electric generator should be clearly explained. It was the scientist Michael Faraday who discovered in 1831 that when a magnet is moved inside a coil of wire, an electric current flow in the wire (eia 2018). Therefore, an electric generator is generally described as a device that transforms a specific form of energy into electricity. Generators operate because of this exact feature of magnetism and electricity. A schematic representation of an electric generator can be seen in Fig. 2.2:

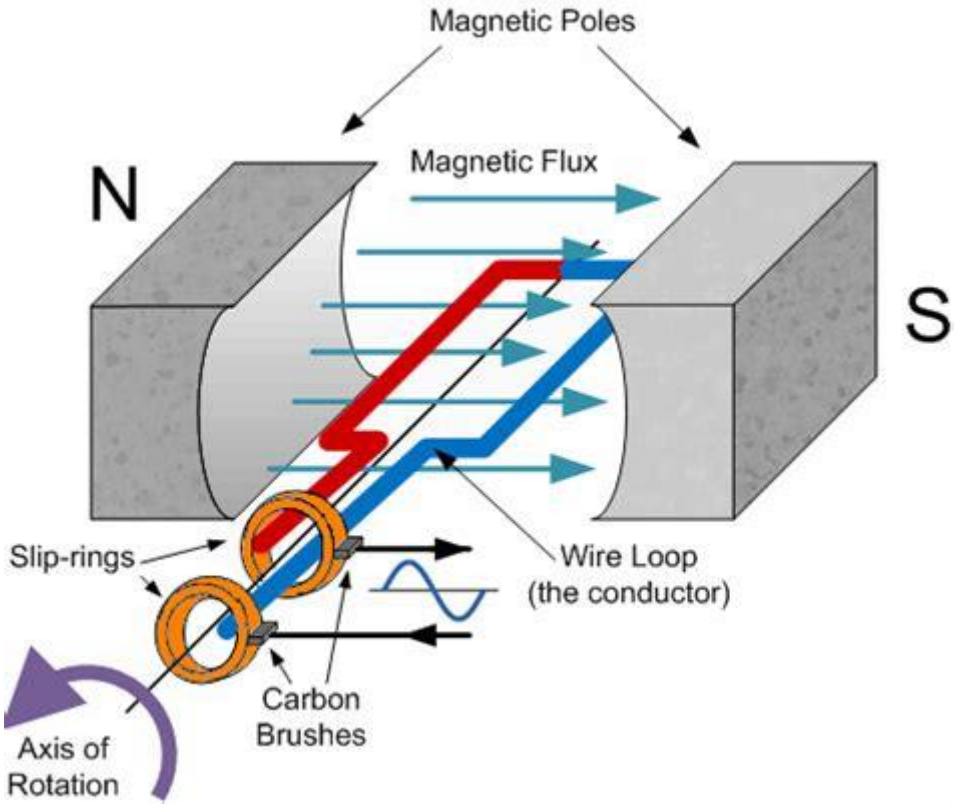

Fig. 2.2 Schematic representation of an electric generator (Source: ElectronicsTutorials)

Though it is true that the majority of the electricity generation in the United States is from power plants that use a turbine or a similar machine to drive electric generators, there are electricity generators which do not require a turbine in order to produce electricity. Therefore, the existence of a turbine in the electricity generation process could be a distinguishing factor among different generating technologies.

The most common type of turbines is steam turbines. In steam turbines, hot water and steam are produced by burning a fuel in a boiler or by using a heat exchanger to capture heat from a fluid heated with other types of energy, like solar or geothermal (eia 2018). The steam is the one that gives motion to a turbine, which sequentially powers a generator. The fuels or in general energy sources that are typically seen in steam turbines include but are not limited to biomass, coal, geothermal energy, petroleum fuels, natural gas, nuclear energy, and solar thermal energy. A steam turbine can be seen in Fig. 2.3:

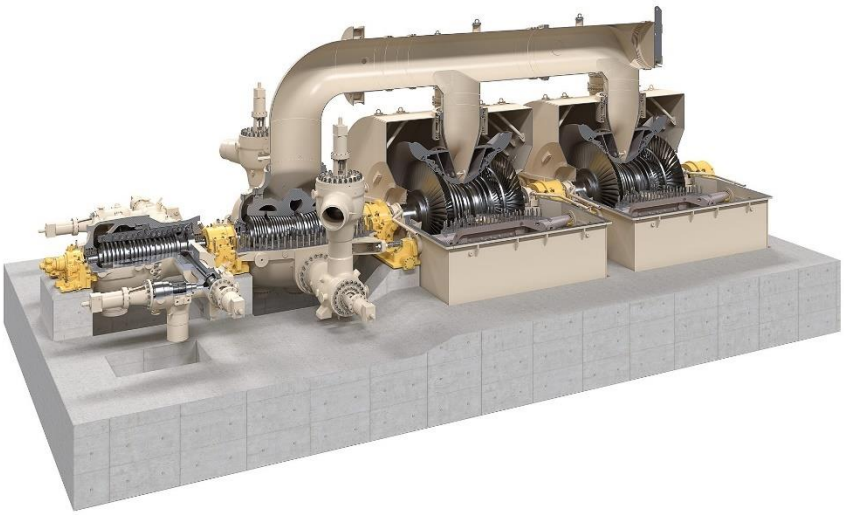

Fig. 2.3 Steam turbine (Source: Siemens)

Other types of turbines are the combustion gas turbines and the internal combustion turbines. Moreover, it should be mentioned that there are combined-heat-and-power (CHP) plants, which use the heat for other purposes, such as space heating or industrial process heat. The heat used by CHP plants is the one that cannot be immediately converted into electricity using the other types of turbines. Lastly, it should be reported that there are also hydroelectric turbines, which use as their driving force water to spin turbine blades and also wind turbines, which of course use the wind.

On the other hand, there are also generating technologies that do not require the use of a turbine to produce electricity. It is appropriate to say that the most typical examples in this category are the solar photovoltaic cells and the fuel cells. The former converts the solar irradiation directly to electricity, while the latter convert fuels into electricity through a chemical process. The most common type of fuel used in fuel cells is hydrogen. An example of a fuel cell is shown in Fig. 2.4:

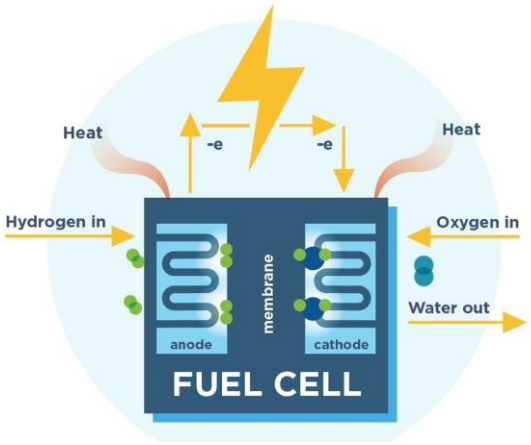

Fig. 2.4 Fuel cell (Source: FCHEA)

2.1.2 Short-term operation planning

The first subclass of problems that the energy planners are required to solve have a short-term nature and are usually solved on an hour-by-hour basis. This subclass includes but is not limited to problems, such as unit commitment and determining scheduling regime of conventional units, i.e. their cyclical operating modes in response to dispatch requirements: on/off operation, low-load cycling operations and load following (Sadeghi, Rashidinejad et al. 2017). These processes are highly important in the optimal operation of an energy system from both economic and reliability aspects.

In this section, the mathematical formulation of a specific problem in this subclass is presented, which is called unit commitment dispatch. But before that, two other problems should be explained; economic dispatch, and unit commitment. These two problems combined build up to the unit commitment dispatch (Rodgers 2016). Firstly, the economic dispatch problem refers to the problem of determining the optimal combination of generators in order to meet the load demand at a given time interval. On the other hand, the unit commitment problem relates to the optimal scheduling of units, such as start-up and shut-down times, in order again to become feasible for a specific forecasted load to be met.

As mentioned above, the unit commitment dispatch model is a combination of the economic dispatch model and the unit commitment model. This model is a mixed integer nonlinear optimization problem that minimizes start-up costs, shutdown costs, and variable costs from energy generation. The decision variables in this problem include dispatching decisions as well as start-up and shutdown decisions (Rodgers 2016). As an example, a formulation for a unit commitment problem is presented (Huang, Pardalos et al. 2017).

$$\min \quad \sum_{g\in G}\sum_{t\in T}\left(SU_g v_{gt} + SD_g w_{gt}\right) + \sum_{g\in G}\sum_{t\in T} F_g\left(p_{gt}\right) + VOLL\sum_{i\in N}\sum_{t\in T} d_{it} \tag{2.1}$$

$$\text{s.t. } u_{gt} - u_{g(t-1)} \le u_{g\tau} \qquad \forall g \in G, t \in T, \tau = t,...,\min\left\{t + ON_g - 1, |T|\right\} \tag{2.2}$$

$$u_{g(t-1)} - u_{gt} \le 1 - u_{g\tau} \qquad \forall g \in G, t \in T, \tau = t,...,\min\left\{t + OFF_g - 1, |T|\right\} \tag{2.3}$$

$$v_{gt} \ge u_{gt} - u_{g(t-1)} \qquad \forall g \in G, t \in T \tag{2.4}$$

$$w_{gt} \ge -u_{gt} + u_{g(t-1)} \qquad \forall g \in G, t \in T \tag{2.5}$$

$$P_g^{\min} u_{gt} \le p_{gt} \le P_g^{\max} u_{gt} \qquad \forall g \in G, t \in T \tag{2.6}$$

$$-RD_g \le p_{gt} - p_{g(t-1)} \le RU_g \qquad \forall g \in G, t \in T \tag{2.7}$$

$$\sum_{g\in G}\sum_{t\in T}\left(F_g^{\,e}\left(p_{gt}\right)u_{gt} + SU_g^{\,e}\, v_{gt} + SD_g^{\,e}\, w_{gt}\right) \le E^{\max} \qquad \forall g \in G, t \in T \tag{2.8}$$

$$\mathrm{E}\left(\sum_{i\in N} d_{it}\right) \le \varepsilon_t \qquad \forall t \in T \tag{2.9}$$

$$u_{gt}, v_{gt}, w_{gt} \in \{0,1\} \qquad \forall g \in G, t \in T \tag{2.10}$$

$$p_{gt} \ge 0 \qquad \forall g \in G, t \in T \tag{2.11}$$

The objective function (2.1) is composed of three main terms. The first cost component is affected directly by the day-ahead decisions of the control unit, such as startup and/or shutdown decision for each generating unit. The second cost

component includes the fuel costs and finally, the third cost component denotes the possible unserved energy penalty. Constraint (2.2) denotes the minimum ON-time of generators, while constraint (2.3) denotes the minimum OFF-time of generators. Moreover, constraints (2.4) and (2.5) are the start-up and shut-down action constraints respectively. (2.6) is the thermal generation constraint and (2.7) is a basic constraint to address generation ramping. An emission constraint is presented in (2.8) and (2.9) is the unserved energy constraint. Finally, (2.10) and (2.11) are binary and nonnegativity constraints.

2.1.3 Long-term expansion planning

Besides short-term planning problems, there is a second subclass of problems in energy planning which is also more tightly related to the current research. This subclass consists of the well-known and studied expansion planning problems. Generally, electric power system expansion can be carried out in generation, transmission and distribution sectors (Hemmati, Hooshmand et al. 2013). However, since the investment on generation expansion planning (GEP) and transmission expansion planning (TEP) is much more than the relevant investment in the distribution expansion, these two parts become significantly more important from the energy planning perspective.

More specifically, GEP refers to the monetary and unit investment needed for energy production. This kind of planning should take into account the sizing, timing, technology of new generation units, etc. In GEP, the objective is to expand the existing power system to serve the growing demand in the future, achieving the lowest possible cost and satisfying the preset reliability criteria (Hemmati, Hooshmand et al. 2013). The planning horizon in GEP problems is typically 10-30 years. It is clearly a challenging problem due to the ever-increasing demand for

uninterrupted electricity supply in almost every aspect of the modern everyday life as well as in every part of most production and service systems in society (Sadeghi, Rashidinejad et al. 2017).

Known as one of the most complicated types of power system planning problems, the GEP problem has been broadly investigated through numerous studies. These studies present a wide range of objectives ranging from cost minimizing in monopoly markets to profit maximization in the competitive environment. Therefore, in this section, it is tried to present several existing research studies on GEP problems and use a specific criterion for breaking down the different approaches. This criterion is closely related to the topic of the current research work and is the environmental scope under which the author is trying to examine the GEP problem. This environmental scope may refer to the extensive usage of renewable technologies, to deploying carbon capture mechanisms or to ways of mitigating health implications caused by power generation plants. Consequently, in the first part of this section, existing general research studies on GEP are reviewed, while in the second part the emphasis is given on those research studies that examined the GEP problem through the environmental lens, as it is described above.

System dynamics are investigated very frequently in joint with the optimal solution of a GEP problem. Valinejad *et al.* presented a new framework to study the generation capacity expansion in a multi-stage horizon in the presence of strategic generation companies (GENCOs) (Valinejad, Marzband et al. 2017). In this context, they proposed a three-level model as a pool-based network-constrained electricity market that is presented under uncertainty in the predicted load demand modeled by a discrete Markov model. In the same direction, Pereira *et al.* created a model to solve the GEP problem in competitive electricity markets (Pereira and Saraiva 2011). In this

research work, they developed an approach which recognizes that the objective of the various generating companies in the problem is to maximize their monetary earnings. Besides that, this study is able to incorporate the several uncertainties in this problem, which include but are not limited to demand, fuel prices, investment and maintenance costs, as well as the electricity price. Furthermore, Park *et al.* have mentioned long ago the need for building sophisticated solutions for the GEP problem in the competitive environment, due to the high complexity of the problem, which arises from the conflicts among generation companies (Park, Kim et al. 2002). The novelty of this work is that in order for someone to find the optimal solution of a specific GENCO competing in the environment, they have to successfully solve the conflicting and correlated subproblems among GENCOs.

From a different perspective, the GEP problem has also been solved with primary criterion the most accurate modeling of the problem's stochasticity. For example, Coit *et al.* identified the high level of uncertainty in the GEP problem, due to its high level of space and timing complexity (Coit, Selcuklu et al. 2015). In order to facilitate the process of the decision-making process under uncertainty, they introduced a systematic way of dealing with uncertainty. This approach involves the selection of non-dominated solutions for the Pareto optimal set. The proposed methodology of this research work is very significant, because it successfully preserved the non-deterministic information of solutions in the objective space, allowing the decision makers to consider simultaneously the multiple objectives and the stochastic aspects of their problem. A somewhat more risk-averse approach was adopted in (Tekiner-Mogulkoc, Coit et al. 2015), in which the authors incorporated demand uncertainties considering conditional-value-at-risk and maximum regret as risk measures. Furthermore, Tekiner *et al.* used Monte-Carlo simulation to generate

numerous scenarios related to system sub-component availabilities and anticipated electricity load demand, in order to find a Pareto front for the multi-objective GEP problem (Tekiner, Coit et al. 2010). In this study, they solved the GEP problem as a two-stage stochastic programming problem. In that way, they managed to tackle the problem of simultaneously incorporating reliability and expansion planning optimization.

Nevertheless, it would be interesting to examine the previous research studies on GEP problems from another scope; the optimization techniques. Several different solvers have been used in the past for solving the stochastic and dynamic, by nature, GEP problem. One of the most common approaches used to solve the GEP problem is without doubt the genetic algorithms. For example, Firmo *et al.* used a special type of chromosome, christened pointer-based chromosome (PBC), in order to develop an iterative genetic algorithm for solving the investment subproblems in GEP (Firmo and Legey 2002). Moreover, genetic algorithms were the chosen solver approach used in (Pereira and Saraiva 2011) and (Park, Kim et al. 2002), too. On the other hand, a mixed-integer linear programming (MILP) framework is considered appropriate in several other studies. As an example, Bakirtzis *et al.* presented a MILP model for the solution of the centralized GEP problem (Bakirtzis, Biskas et al. 2012). In their research work, the authors aimed to minimize the total present value of the investment, operation and unmet penalty energy costs net the salvage value of the new units at the end of the planning horizon. Moreover, Zhang *et al.* took into account line losses and energy storage systems in their deterministic single-stage MILP model (Zhang, Hu et al. 2013). Finally, metaheuristic techniques, in general, have been used widely to solve the GEP problem, such as in (Kannan, Slochanal et al. 2005). Therein, the authors applied and compared several metaheuristic techniques in their capability

of solving the GEP problem; differential evolution, evolutionary programming, ant colony optimization, and tabu search are only some of these techniques. It should be mentioned here that these studies seldom considered dynamic programming approaches and if so, they did with relaxed assumptions or reduced problem dimensionality. Of course, this fact serves as one of the main and ultimate goals of the current research.

Last but not least, it is useful to mention that the GEP problem has often been studied in accordance with the TEP problem. Nemati *et al.* proposed a static model for coordinated generation and transmission expansion planning (CGTEP) (Nemati, Latify et al. 2018). While the main objective is still to minimize investment costs, operation costs and energy not served within the system, their model also targets in making the power more robust against physical deliberate attacks in the planning horizon. Finally, Zhang *et al.*, although considered solely the transmission expansion planning problem, they successfully incorporated energy storage systems planning in their formulation (Zhang, Hu et al. 2013).

At this point, there should be a clear attempt to pay specific attention in several GEP studies which incorporated elements that made them distinct from others, in the sense that these studies had transparent environmental considerations in their formulations and results. If needed to create a broad taxonomy of these research works, it could have been said that they mainly belong in two categories, based on the scope they examine the environmental impacts in their formulation. In the first category, researchers are trying to give an emphasis on minimizing emissions as one of their objective functions or measure health externalities occurring from environmentally-hostile GEP solutions. Studies incorporating climate change scenarios fall into this category, too. On the other hand, there are several studies

which are mainly concentrated on the higher penetration of renewables, either combined with storage or not. It is safe to assume that these studies, although serving the overall goal of providing "green" solutions of the GEP problem, are different from the studies of the first category in the way they approach and solve their formulated problems.

Sirikum *et al.* presented an application of genetic algorithms for solving a mixed integer nonlinear programming (MNILP) version of the GEP problem, under consideration of power generation limits and load demand levels, loss of load probability (LOLP) levels and environmental limitations (Sirikum and Techanitisawad 2006). This study is mainly motivated by the widespread social awareness about environmental concerns and also by the continuously increasing attention that expansion planning policies pay to environmental costs. Therefore, the authors added the environmental factor into the GEP problem, by trying to control emissions. In another study, Tekiner *et al.* tried to include the minimization of air emissions as a separate objective in their multiple-objective formulation (Tekiner-Mogulkoc, Coit et al. 2012). In order to do so, they considered the impact of several smart grid technologies by testing and comparing different scenarios, based on the level of deployment of these technologies. At this point, the focus should be given to the studies that tried to quantify the health or other externalities caused by large penetration of fossil fuels in the fuel mix of the resulting solution in the GEP problem. The purpose of these research works is mainly to highlight the importance of proceeding with "green" technologies, even though their economic impacts are not accurately measurable in most of the cases. Rodgers *et al.* solved exactly this problem, by enabling decision makers to directly assess the health implications of power grid expansion decisions by explicitly estimating the total societal costs and by

quantifying externalities of any proposed planning strategy (Rodgers, Coit et al. 2019). One of the key research findings in this work is that by enforcing emission limits and by deploying renewable portfolio standards it is possible to reduce health and other damages, and consequently minimize the total societal costs. In another study, Rodgers *et al.* were able to provide significant algorithmic research advancements by developing an analytical metamodeling framework for direct estimation of health damages incurred in the process of GEP (Rodgers, Coit et al. 2019). Khan *et al.* attempted to internalize environmental externalities on the GEP problem, with the ultimate goal of making a large scale integration of renewable energy sources economically feasible (Khan, Sun et al. 2014). Indeed, the results of this study showed that the planning interest inclined in favor of RES confirming that including externalities in the planning models can be used as an extremely efficient policy mechanism for the mass explosion and deployment of RES. Lastly, there are existing studies which considered exclusively climate change and made it their driver for the optimal solution of GEP problem. Li *et al.* supported that climate change is and will be affecting power plant investment decisions in new and more uncertain ways (Li, Coit et al. 2016). Therefore, the authors formulated a robust electric power GEP optimization model minimizing the expected total cost under different climate change scenarios. According to their methodology, the authors successfully proved that the stochasticity in the climate change affects directly not only the demand for electricity but also supply, reliability and other related factors. Therefore, it is considered necessary to be able to derive systematic methodologies incorporating stochastic modeling and also come up with new and efficient solution methods.

Herein, the examination relates to the studies that fall into the second category; these studies mention directly the higher penetration that renewable energy

sources should have in the future grids and they formulate their problems based on that. Distributed energy resources are placed in the core of this research work, due to the fact that microgrid power and storage expansion planning is one of the main topics. Therefore, this part of the literature review requires special attention. Firstly, Rajesh *et al.* investigated the impact of the introduction of solar power plants with a storage facility in the GEP problem (Rajesh, Karthikeyan et al. 2016). This study, while attempting to highlight the challenges of the general decision-making process in introducing solar plants into an existing system, provided a four-level hierarchy which helped energy planners in understanding several policy-related issues that might arise in GEP and therefore in tailoring the corresponding solutions to each case separately. On the other hand, Luz *et al.* presented a multi-objective model for expansion with high penetration of renewable energy (Luz, Moura et al. 2018). In this study, the authors claimed that it was possible to meet 90% of the annual load with renewable sources (with 23% being supplied by non-hydro) and that the capacity of solar power could be increased from 21 MW to 40,000 MW by 2030. Furthermore, wind farms is a major consideration for the GEP problem (Hemmati, Hooshmand et al. 2016). In this research work, the authors used probabilistic modeling and Monte-Carlo simulation to consider the wind power uncertainty and to conclude that GENCOs and transmission companies (TRANSCOs) are able to achieve their goal of profit maximization if they incorporate in great detail in their optimization models the wind farm stochasticity. Last but not least, Careri *et al.* investigated the GEP problem under the prism that humans live in the Age of Green Economy, as they called it in (Careri, Genesi et al. 2011). In this paper, the impact of some of the most popular incentive systems (namely feed-in tariffs, quota obligation, emission trade, and carbon tax) on generation planning is studied. The authors formulated and extended a comprehensive

GEP model with an appropriately modified objective function and additional policy-related constraints. The results of this research work showed that incentive systems for RES deployment and CO_2 mitigation could greatly affect the strategic generation planning for a GENCO operating in a liberalized environment. Renewable-based generation expansion planning is also studied in (Pereira, Ferreira et al. 2017). The special characteristic of this work can be found in the fact that the authors considered explicitly the variable output of renewables in their binary mixed integer non-linear mathematical formulation. Lastly, Sadeghi *et al.* used the gravitational search algorithm to solve their comprehensive GEP model, with the ultimate goal of enforcing the GENCOs in more renewable energy sources investments (Sadeghi, Mohammadian et al. 2014). After their tests and results, they were able to conclude that it is possible to limit significantly emitted contaminations through careful and optimally designed RES-based penetration. Overall, it is safe to assume that renewable energy sources are lying in the core of future grids and should be thoroughly studied and optimized until they finally become a standard.

2.2 Distributed generation in energy systems

In this section of the dissertation, after providing the necessary definitions and explanations, the study is tailored specifically to distributed generation, renewable energy sources and microgrids. In the first subsection, a definition of a microgrid is given, the operation of a typical microgrid is explained and also several types of microgrids are presented. In the second subsection, specific emphasis is given to the various types of renewable energy sources used widely in microgrids, especially when it comes to solar cells or wind turbines. Finally, it is considered necessary to further highlight the importance of optimal planning in backup systems in general, and in microgrids specifically.

The main factors driving microgrid development and deployment in locations with existing electrical grid infrastructure fall into three broad categories: energy security, economic benefits and clean energy integration (Hirsch, Parag et al. 2018), which are simultaneously the backbone of the current research motivation. To be more specific, the main reason behind the mass interest given in microgrid applications in the United States has been their potential to increase the resilience (the ability to bounce back from a problem quickly) and reliability (the fraction of time an acceptable level of service is available) of "critical operations" in sectors such as transportation, communications, health care and emergency response infrastructure (Hirsch, Parag et al. 2018).

It should be mentioned that the penetration of distributed generation (DG) at medium and low voltages, both in utility networks as well as downstream of the meter, is increasing in developed and developing countries worldwide (Hatziargyriou, Asano et al. 2007). It is considered true that a large-scale deployment of distributed energy resources (DERs) can potentially offset the requirement for traditional system expansion. However, managing a huge number of DERs is itself an operation that creates challenges. One of these major challenges is operating and controlling the network under preset safety and efficiency standards and luckily this can be addressed by microgrids at a satisfactory level.

2.2.1 Definition and types of microgrids

Although someone can find multiple definitions of microgrids in the literature, a formal one is given by the U.S. Department of Energy (Ton and Reilly 2017); a microgrid is a group of interconnected loads and distributed energy resources (DERs) with clearly defined electrical boundaries that acts as a single controllable entity with respect to the grid. It can connect and disconnect from the grid to enable it to operate

in both grid-connected or island modes. The characteristics mentioned in the definition above, present microgrids as relatively small-sized power systems with certain capabilities. These capabilities include but are not limited to self-supply of load demand, operating in islanding mode and generating and distributing simultaneously electricity to local customers (Parhizi, Lotfi et al. 2015). It may seem that microgrids play solely the role of backup generation. However, microgrids are more than just that. Their main difference is that backup generation units' main purpose is to provide a temporary supply of electricity to local loads when there are operating problems and disconnections from the main grid. On the other hand, microgrids are entitled to a broader spectrum of abilities and are significantly more flexible than just a backup generation. In Fig. 2.5, the reader can see a schematic overview of a microgrid:

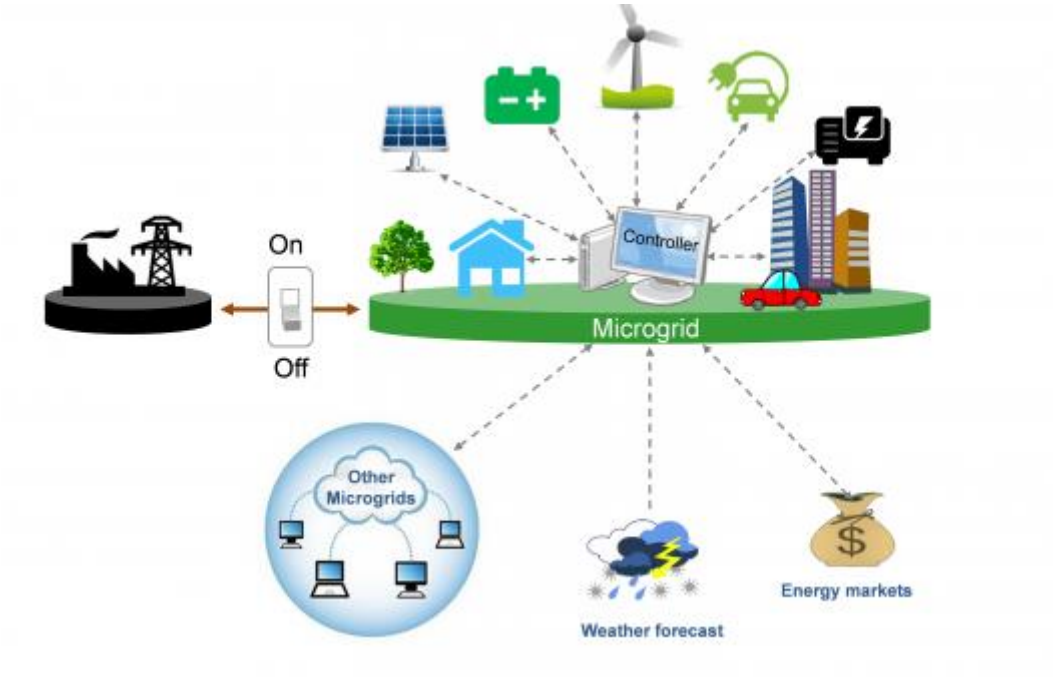

Fig. 2.5 A schematic overview of a microgrid (Source: Microgrids at Berkeley Lab)

Before proceeding with how a microgrid works, it should be useful to understand how the main grid works. The main grid is able to connect the power generation from the main power plants to residential, commercial and industrial lines

via the transmission network. Though it is known for its local character, a microgrid actually has similar functionality. However, despite the conventional energy sources that the main grid occasionally utilizes, a microgrid is highly dependent on renewable energy sources, such as solar cells, wind turbines and energy storage systems (Mbuwir, Ruelens et al. 2017). Besides the above-mentioned distributed energy sources, as well as the storage options, that a microgrid incorporates, it is crucially important to mention that almost every microgrid has a point of common coupling (PCC) - ON/OFF "switch" in Fig. 2.5 -, which allows the controller to operate the microgrid in a connected or disconnected mode from the main grid. The PCC may seem a negligible feature of a microgrid, but actually, it is one of the major advantages in contrast with the main grid. Microgrid's capability of breaking off the tie with the grid and operating in an isolated mode is the characteristic that energy system planners desire for a reliable and resilient power generation and distribution. Last but not least, the PCC also maintains the voltage at the same level as the main grid, under normal operating conditions (Lantero 2014). Except the energy stability the microgrid has to offer, there are also potentially huge cost savings via leveling the cost of energy, due to the decentralization driven by renewables (Telegraph 2019).

While there is a vast amount of great theoretical research around microgrids and their benefits, it could be really useful to track some real-life microgrid applications. According to Navigant Research, which has recorded microgrid applications since 2011, the United States has been the historical leader in deployed capacity; today, though, the U.S. and Asia have roughly the same capacity of live or ongoing microgrid projects, each with 42% of the market. Europe trails with 11%, Latin America with 4%, and the Middle East and Africa currently have just a 1% share (Hirsch, Parag et al. 2018).

At this point, it is considered appropriate to present the most common types of microgrids that someone can find operating nowadays around the globe (Hirsch, Parag et al. 2018). Firstly, campus microgrids use mostly combined cooling, heat, and power (CCHP, also known as “trigeneration”) technologies with the various loads collocated on a campus facility owned by the same entity. In another microgrid type, on-grid or off-grid military microgrids provide a quintessential characteristic; cost-effective energy security. An illustrative case study of military microgrid research can be found at the Smart Power Infrastructure Demonstration for Energy Reliability and Security (SPIDERS) Joint Capability Technology Demonstration (JCTD), a three-phase program (DoD 2015). Probably the most common type of microgrids available right now, are the residential microgrids. Abundant around the world, they are able to supply cost-efficient and reliable power supply at the level of individual homes, by using microgrid technologies. As a natural extension of the residential microgrids, someone can find also many community microgrids. These are again small-scale microgrids, however, now comprised by different facilities which require uninterrupted electricity supply at different reliability levels but are served by the same set of distributed energy resources and storage systems. Finally, an emerging type of microgrid lately, is the remote, or so-called “off-grid” microgrid, which is mainly located in developing and underdeveloped countries which currently lack access to uninterrupted power supply. These specific microgrids, which combine clean generation, storage and are sometimes supported by novel mobile payment platforms, can provide life-saving features to many people living in rural areas, allowing children to study at night or hospitals and medical centers to operate continuously (Hirsch, Parag et al. 2018). At this point, it should be made clear that the case studies that the current research attempts to explore concern mainly community

microgrids which possess islanding capabilities during grid outages via using a PCC. However, the theoretical methodologies developed and applied here could be easily expanded and modified in order to encompass the optimization of other types of microgrids.

2.2.2 Generation options in microgrid

According to the issue of power generation in a microgrid, the energy system planners have a wide variety of options. There are multiple studies in the literature, which cover exactly the topic of elaborating on the different distributed energy resources that can be incorporated in a microgrid structure (Akorede, Hizam et al. 2010) (El-Khattam and Salama 2004) (Mariam, Basu et al. 2016). However, in this context, it is considered appropriate to mention only some of them and then elaborate more on specific renewable generation units, which are lying in the core of this research.

Firstly, one of the options available in the market are the diesel and spark ignition reciprocating internal combustion engines. This option has several advantages and disadvantages (Hirsch, Parag et al. 2018); for example, these engines are dispatchable and have a quick startup, but on the other side they generate a large amount of noise and more importantly they contribute to a significant amount of greenhouse gas emissions. The next option for power generation is the microturbines; this technology is mechanically simple, provides multiple fuel options but has a reasonable amount of greenhouse gas emissions, too. Furthermore, there is a recent research development in the area of fuel cells, which have zero on-site pollution and high efficiency rates, but they are still more expensive than the rest of the options. Last but not least, the most common type of technology used in modern microgrids are renewable energy sources. In this broad class of technologies, someone can find

solar photovoltaic cells, wind turbines and mini-hydro plants. The apparent advantages following these options are the zero emissions and the zero variable/fuel operating costs. On the other hand, renewable-based generation is still not dispatchable without storage and it is also considered highly variable.

Towards expanding the topics which are of main interest in this research work, it is considered necessary to provide detailed information about two of these renewable sources: solar cells and wind turbines.

Only solar energy itself, it could be able to satisfy the whole global energy demand, if the necessary technology advancements are made (Kabir, Kumar et al. 2018). Almost four million exajoules of solar energy reach the earth annually and a significant percentage is believed to be harvestable (iea 2018). Of course, it should be noted here that the solar power distribution and intensity are highly dependable on the exact location under examination.

A very interesting matter that arises from the increasing penetration of solar energy in the market, is its implications on the reliability and resilience of the power supply. In recent years, it is observed that power outages on several systems around the globe occur more frequently. For instance, the large blackout caused by Superstorm Sandy affected 8.2 million people in 17 states (Hines, Apt et al. 2008). According to (Klinger, Landeg et al. 2014), only in the first quarter of 2013, there had been 14 power outages in the US, affecting more than half a million people. But even very recently, a massive blackout left millions of people without power in South America (KCBD 2019). It was considered as an unprecedented catastrophe and several hours after the outage occurred, half of the Argentina's population was still without power. It is self-proven at this point that solar energy could play a crucial role in addressing this type of problems because distributed PV generation enables

customers to access electricity even during grid failure. As a result, several studies have explored the supplemental resilience value of adding battery capacity to PV systems (Birnie 2014) (Zhou, Tsianikas et al. 2019).

On the opposite side, the main drawback of solar energy systems that holds back their full exploitation is their cost-efficiency. Although the costs of solar panels have declined rapidly, there are still technological gaps that need to be filled in order to achieve highly scalable deployment (Lewis 2016). For instance, improved thermal storage fluids would provide longer-term storage to compensate for cloudy days in areas of high direct insolation. In another area, some recently proposed ideas include the use of perovskite in the manufacturing of solar cells, which could theoretically increase their efficiency tremendously (EnergySage 2018). Besides the movements that need to be made from the technical scope, there are also policy incentives which can help in a great degree the solar growth. The solar Investment Tax Credit (ITC) is one of the most important federal policy mechanisms to support the deployment of solar energy in the United States (SEIA 2019). Solar panel installations peaked in 2016 ahead of potential drop down of the ITC, but an extension in late 2015 has ensured federal policy stability through 2021. The yearly US installations, along with the ITC landmark dates, can be seen in Fig. 2.6:

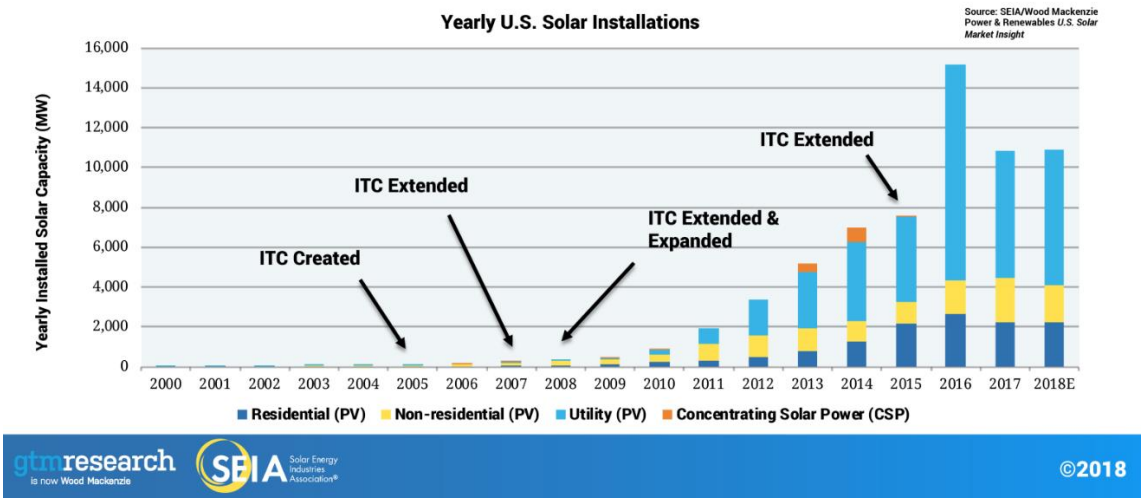

Fig. 2.6 Yearly U.S. Solar Installations by sector, with ITC landmarks

Besides the clear increasing trend of solar installations that can be noticed in Fig. 2.6, another thing worth mentioning is that the same trend exists in every sector. In particular, for residential customers, the amount of solar capacity installed is almost tripled up from the year 2013 to the year 2016. To sum up, it is required to mention that innovative applications and usages of solar energy are currently being studied in the literature. For example, Birnie researched the capabilities of vehicle solar roofs on energy capture (Birnie 2016). Birnie analyzed the operation of vehicle-installed solar panels in the roof from the perspective of a commuter's travel and one of his major findings was that the energy capture may be blocked by battery capacity saturation.

Another renewable energy source that is continuously increasing its share in the renewables market is undoubtably wind power. The wind is actually a form of solar energy caused by a combination of three concurrent events: the sun heating unevenly the atmosphere, irregularities of the earth's surface and the rotation of the earth (Office 2014). The wind turbines are responsible for converting the kinetic energy of the wind to mechanical energy and then a generator finally converts into electricity. Wind turbines are mainly categorized based on their rotation axis; vertical-axis turbines or horizontal-axis turbines. The size of utility-scale wind turbines can vary from a few kilowatts to many megawatts (Office 2014). In 2017, cumulative installed wind power capacity in the United States increased by 8.3 percent to reach around 89 gigawatts, according to (statista 2019). A schematic representation of the U.S. cumulative installed wind power capacity for the years 2000-2017 is presented in Fig. 2.7:

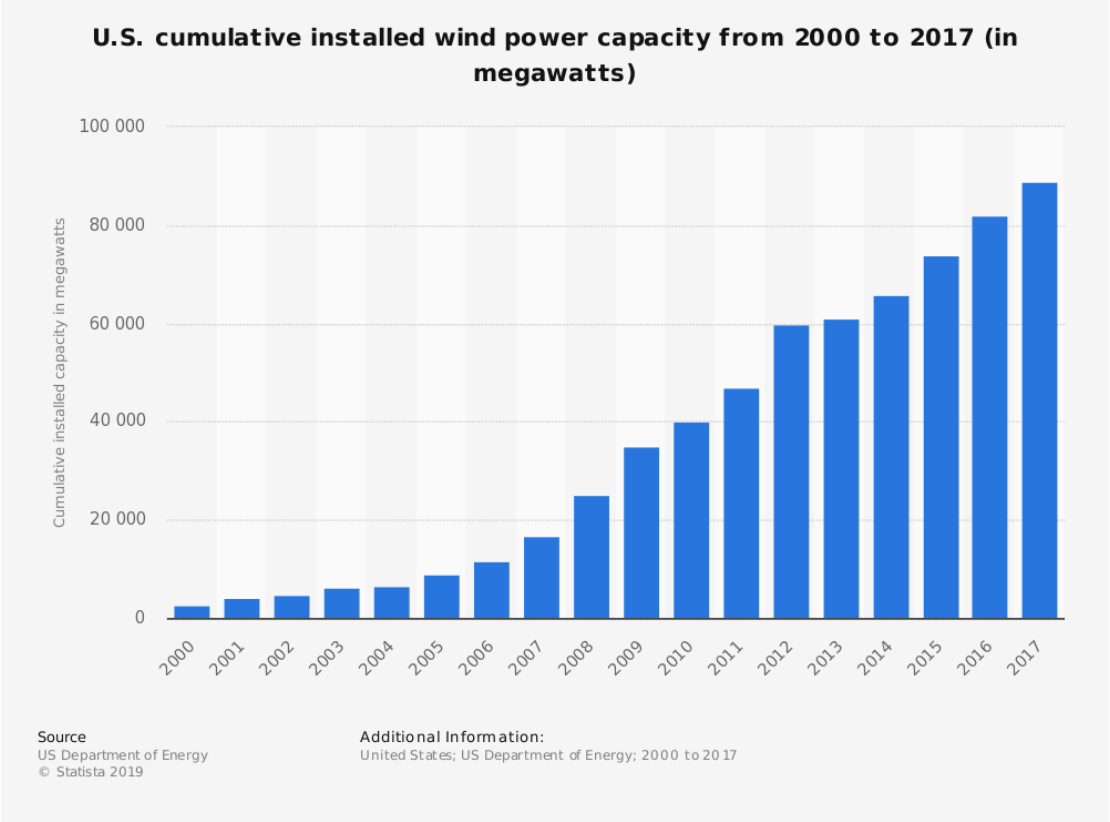

Fig. 2.7 Yearly U.S. Wind Power Capacity, 2000-2017

While the increasing penetration of wind power in the energy markets is clearly a positive signal for the energy planners working with renewable sources, its economic feasibility as long as its capability for reliable power supply still remains debatable. Towards the direction of settling wind turbines as a standard for renewable energy plants' design, there are several studies ongoing concerning the optimization of onshore or offshore wind farms. These studies break up into several subcomponents of the optimization process; turbine layout design, wind uncertainty modeling, and wake effect models are only some of them. As an illustration, Song *et al.* developed a two-stage optimization model for an offshore wind farm that combined optimal layout design and turbine maintenance strategies (Song, Li et al. 2018). In this research work, the authors used probabilistic models to tackle the stochasticity of wind speed and direction and they applied their proposed algorithm in a case study for a wind farm along the New Jersey coast. Overall, their results are

reasonably optimistic for the future of this specific technology, since they are able to maximize the wind energy profit by integrating the decisions about layout design and maintenance scheduling.

2.2.3 Importance of optimal planning on backup systems

While it is well-established up to now that reliable and cost-efficient power supply is a crucial element of operating energy systems of any size or any type, it is considered mandatory at this point to pay more attention to backup systems and microgrids in general. This decision can be sufficiently justified by two main reasons: the various uncertainties incorporated in the operation of the local grids which mainly utilize distributed energy resources, and the uninterrupted power supply dictated by the type of facilities they have promised to serve. While the uncertainties are abundant in almost every microgrid around the globe, the second reason applies mainly to this category of microgrids which encompass hospital facilities, governmental buildings and even education centers. In these facilities, the unmet penalty costs incurred by potential loss demand are much more devasting, if properly quantified. However, besides the economic damages resulting from power losses, there are also other types of costs incurred, even more important in some cases; health implications and societal costs are only some of them. Consequently, there is an emerging need for further research whose purpose would be to optimally design these systems from both economic and reliability aspects.

Intermittency is not the only reason behind the multiple uncertainties existing in the planning process of microgrids. Moreover, there is a high degree of variability in forecasted topical load demand as well as in market prices. Consequently, many researchers have studied the expansion planning of microgrids while carefully modeling and closely examining the various uncertainties. Ceseña *et al.* adopted

methods of finance, and more specifically real options, to solve a stochastic mixed integer linear program for the optimization of distributed generation system expansion planning subject to relevant uncertainties (Cesena, Capuder et al. 2016). Continuing in the real options approaches, Farzan *et al.* modeled the inherent uncertainties of the microgrid investment problem by utilizing closed form contingent analysis (Farzan, Mahani et al. 2015). The authors provided an analytical framework in assessing the value of the option to invest instead of directly making this decision. This work is particularly important in this research context, due to the stochastic decision-making problem that is formulated and solved in Section 4, too. The problem's stochasticity is also a primary consideration in (Khodaei, Bahramirad et al. 2015). The authors decomposed the microgrid planning into two phases; an investment master problem and an operation subproblem. The specific uncertainties considered in this study include but are not limited to load forecasts, renewable energy generation, and market prices. However, one novelty of this research is the incorporation of another type of uncertainty, specifically related to when and how long disturbances occur in the main grid, or so called "grid outages". Grid outage events, which are rarely studied in great detail in the literature, is one of the primary research considerations in this dissertation. An accurate modeling of grid outages is essential, due to the fact that it significantly affects the optimization results in microgrid planning. The issue of when, where and how long natural disasters will happen and last is another major concern and source of uncertainty. This happened to be exactly the motivation behind the research presented in (Chen and XI 2018). The authors used a joint power flow analysis and Monte Carlo simulation approach to optimally design a microgrid under natural disasters. By testing different scenarios of increasing severity, they were able to provide a reliability-based two-stage design selection process for the distributed

generation resources used in the microgrid. Nevertheless, it is worth mentioning that uncertainties are considered in various other studies in the literature such as in (Khayatian, Barati et al. 2018) and in (Wagar, Wang et al. 2015).

Wang *et al.* concentrated on isolated microgrids and their optimal expansion plans (Wang, Chen et al. 2017). Isolated microgrids are the responsible energy suppliers for remote and rural areas and therefore they certainly fall into the category described before. In their research work, Wang *et al.* formulated a tri-level expansion planning framework, composed of a demand expansion layer, a capacity optimization layer and an operation optimization layer. They are also able to test a case study based on Weizhou Island in Guangxi Province, China and proved that the yearly profit could increase by 25% by utilizing properly controllable loads. Solar-powered microgrids are the core topic of research in (Mahani, Liang et al. 2019). In their research work, the authors posed the goal of optimizing operation and maintenance (O&M) policies in such types of microgrids. Swartz *et al.* developed a methodology for the proper sizing of renewable energy systems in microgrids (Swartz, Ghofrani et al. 2017). By adopting relevant physics equations and by integrating the design suggestions into a hybrid energy system, the authors are able to test their methodologies in the software TRNSYS and conclude to the optimal system configurations. In a similar topic, Hajipour *et al.* took into account reliability issues and cost factors in their study about stochastic expansion planning of remote microgrids (Hajipour, Bozorg et al. 2015). Recognizing that this type of microgrids is generally supported by diesel generators, the authors highlighted the high costs incurred by this fact. Therefore, they proposed a stochastic programming approach along with scenario reduction techniques, in order to successfully deploy renewable energy sources, such as wind farms, and storage systems in isolated microgrids. A critical issue when dealing with this type of

problems is without any doubt the proper inclusion of unmet demand in the optimization models. While reliability is used as a model constraint in many optimization models existing in the literature, the attempt to properly quantify the penalty costs of unmet load demand is less. Khodaei and Shahidehpour presented an algorithm whose objective is to minimize the total system planning costs jointly with the expected cost of unserved energy (Khodaei and Shahidehpour 2013). The problem that the authors solved in their study is a microgrid-based co-optimization of generation and transmission planning and they managed to do so by explicitly defining the expected energy not served by the microgrid and by running a case study in a modified IEEE 118-bus system.

In addition, it should not be omitted to pay special attention to the recurring challenges faced when dealing with generation planning in microgrids; intermittency and manufacturing barriers, which keep efficiency down and simultaneously costs up. Therefore, it is obvious that the introduction of an emerging technology is needed, which also happens to present a declining trend in its incurred costs: energy storage (World 2018). By using recent advancements in software and management tools, it could become possible to utilize this technology at maximum; this fact could not only tackle the problem of renewable power output variability, but could also drive the total system costs of a microgrid down and therefore make it obviously more attractive. Overall, it would be useful to elaborate more on this exact research field in the next section of the dissertation.

2.3 Storage solutions for microgrids

The inclusion of energy storage systems (ESS) in almost all types of microgrids becomes increasingly important (Mbuwir, Ruelens et al. 2017) and the reason behind that is not only the problem of intermittency as mentioned before. ESS

can benefit the microgrid in multiple other ways, such as peak shaving, smoothing power flow, volt-ampere reactive (Casares, Lopez-Luque et al.) support and other behind-the-meter benefits (Mahani, Nazemi et al. 2020). This fact along with the forecasted decline in their investment costs can definitely make ESS an integral part of most medium or large-scale microgrids. In order to prove that argument, it should be mentioned that according to Navigant Research, ESS are present in more than 40% of the installed new project capacity after their update in 4Q of 2016 (Knowledge 2016). More specifically, the chart below in Fig. 2.8 shows the percentage of ESS utilization by microgrid segment for both the 4Q 2015 and the 2Q 2016:

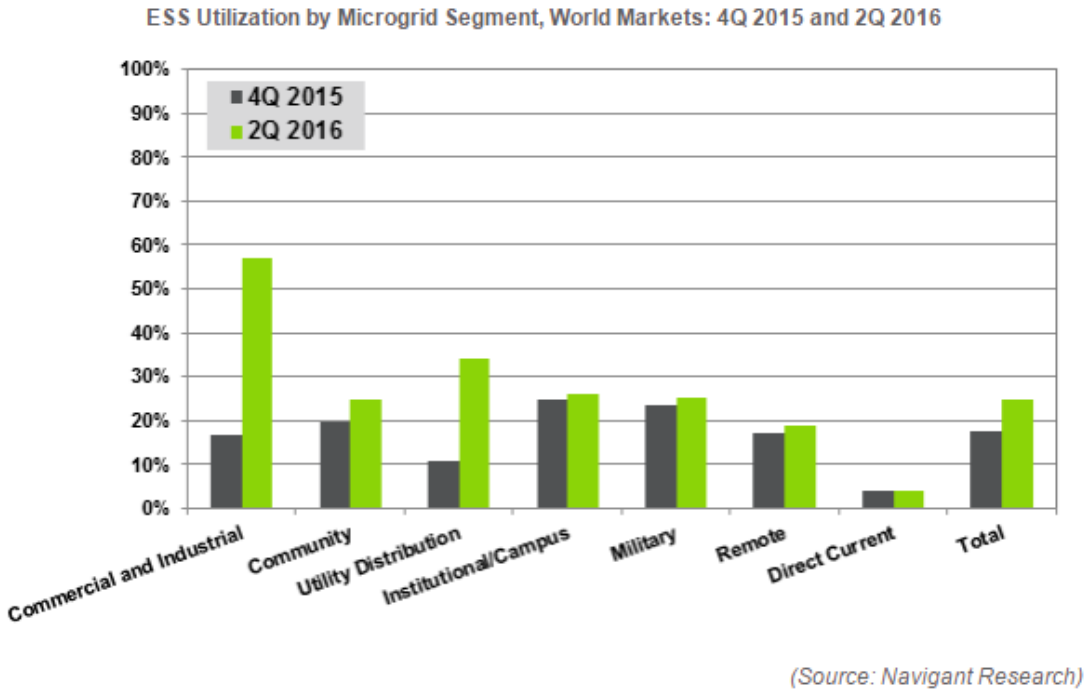

Fig. 2.8 ESS Utilization by Microgrid Segment

While it is clear that ESS utilization grew in almost every sector, it is worth observing that the biggest increase occurred in the commercial and industrial sectors. The findings of this report would become even more useful when combined with another finding of Navigant Research: the solar PV capacity grew by almost 840 MW since the company's last update (Knowledge 2016). It is already known that renewable energy sources, such as solar or wind energy, have a higher need for ESS

due to their inherent characteristics. Consequently, it would be very reasonable to assume that a more explosive deployment of ESS could coincide with an equally significant deployment of solar PV capacity.

Towards this direction, many recent research studies have been concentrated on exploring photovoltaic and battery systems as a model for strengthening the resilience of distributed power generation to reduce power interruptions of critical facilities (O'Brien and Hope 2010) (Prehoda, Schelly et al. 2017). The more frequent power outages have adverse impacts on industrial operations and personal lives (Amin 2008) (Zhou, Huang et al. 2018). However, adoption of ESS allows customers to be electric grid independent in some areas where the grid is distant or is not very reliable (Bhattacharyya 2012). Additional benefits of these systems are their independence and energy-saving potential as well as their environmentally friendly character (Lucio, Valdes et al. 2012). PV + battery systems were not economically viable until recently as the battery prices have come down. Batteries though remain relatively expensive especially in contrast with other distributed backup energy resources, e.g., diesel generators. Therefore, there is a pressing need to understand battery size and cost tradeoffs for adding battery storage capacity to any existing PV systems, especially when evaluating the economic consequences to health-care, manufacturing, or other sectors under circumstances of outages, while making careful assessment of the reliability improvements for meeting load demand (Hontoria, Aguilera et al. 2005) (Mellit, Kalogirou et al. 2009) (Jakhrani, Othman et al. 2012).

2.3.1 Review of energy storage system technologies

Given that the ESS importance in microgrid applications is already illustrated up to now, it is considered appropriate to provide a comprehensive review of the different types of ESS at the market at this point.

Mechanical energy storage systems are beneficial because of their flexible operation and their ability to deliver the stored power when required for mechanical work (Göğüş 2009). Three major subclasses that fall into the greater family of mechanical energy storage systems are flywheel energy storage systems, compressed-air energy storage systems and gravity energy storage systems (Faisal, Hannan et al. 2018). The flywheel is a massive rotating cylinder that is supported on a stator by magnetically levitated bearings (Chen, Cong et al. 2009) and is the main part of the flywheel energy storage systems (FESS). FESS can be divided into two main categories: high-speed and low-speed FESS. While it is not one of the major scopes of the current research work to elaborate in great detail about the mechanisms underlying the energy storage systems, it is considered appropriate here to mention some of the advantages of the FESS technology: high cycle life (hundreds of thousands), long calendar life (more than 20 years), fast response, high round trip efficiency, high charging/discharging rates, high power density and low environmental impacts (Amiryar and Pullen 2017). Therefore, it is easy to see why this technology has great potential to be used in many real-life applications in the future. However, it should be noted that it is still very expensive to install, and especially for small-scale microgrids.

Electrochemical storage systems (EcSS) are definitely the largest and most common group of energy storage devices and refer mostly to conventional rechargeable batteries and flow batteries (Faisal, Hannan et al. 2018). The biggest advantage of this storage category is that it is available in different sizes and possesses different characteristics depending on the battery type that someone chooses. Undoubtedly batteries play a major role in the current research work, so a deeper dive in some battery types commonly used in microgrid applications is appropriate.

Concerning mobile consumer electronics, the most important type of battery at the moment is without any doubt the Li-ion battery. Li-ion batteries exchange lithium ions between the anode and the cathode, which are made from lithium intercalation compounds (IRENA 2017). A schematic representation of a Li-ion cell composed by lithium metal oxide cathode and carbon-based anode can be seen in Fig. 2.9:

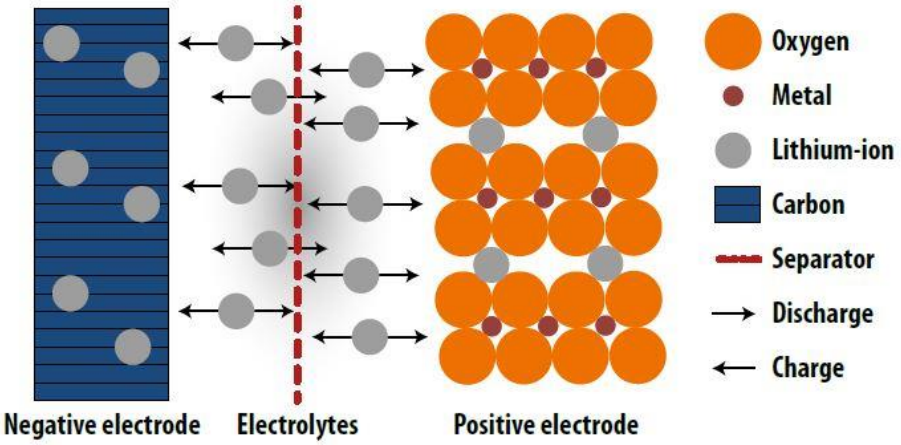

Fig. 2.9 A typical Li-ion cell

Overall, Li-ion batteries have the advantage of high energy and power density when compared to other battery technologies. It is also known that they have high power discharge capabilities, excellent round-trip efficiency, a relatively long lifetime and a low self-discharge rate. On the other hand, someone could mention that their costs are still high and also there are some concerns about their thermal stability and safety (IRENA 2017). Another important type of battery is the lead-acid battery, which is mainly famous for its cost-effectiveness. Most of the lead-acid batteries use liquid sulfuric acid as an electrolyte. The advantages and disadvantages of a lead-acid battery are well-known and studied; it is probably the less costly type of storage available in the market and it can be easily implemented in large-scale storage applications, but on the other hand, has a comparably low round trip efficiency and also has a low cycle life. The last type of battery that is worth a special reference, is

the family of redox flow batteries. Flow batteries (FBs) operate in charged or discharged mode by a reversible chemical reaction (Faisal, Hannan et al. 2018). This chemical reaction is happening between the battery electrolytes, which are contained in separate tanks. The vanadium redox flow battery (VRFB) specifically is based on redox reactions in the cell that are fed by active ionic vanadium materials from the tanks, resulting in electron transference in the circuit (IRENA 2017). The mechanism of a VRFB is shown illustratively in Fig. 2.10:

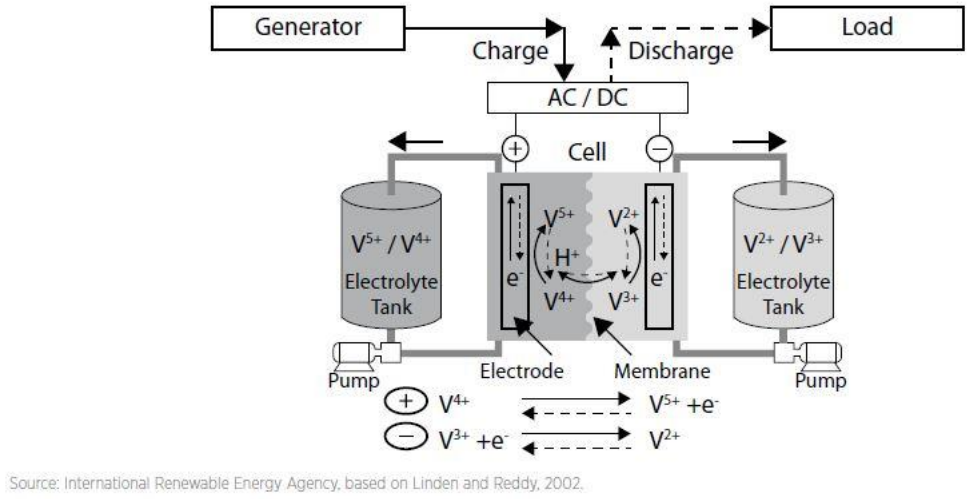

Fig. 2.10 Operation mechanism of a vanadium redox flow battery system

There are really plenty of advantages accompanying this specific battery type; long cycle life, long duration discharge feature and fast response times are only some of them (IRENA 2017). Its low efficiency, compared to Li-ion, and its high repair and maintenance costs are the only drawbacks of VRFB that are worth-mentioning besides its cost. This last factor, which seems to be rapidly declining, has the potential to become a tie-breaking factor on the competition between VRFB and the rest of the ESS.

Although the main energy storage systems used in microgrids is a topic already covered, it is now necessary to proceed with some other options that the energy planners have or may have in the near future. For example, there are chemical

storage systems and the most famous system in this family is undoubtedly the hydrogen fuel cell (HFC) storage system. It is a topic that has recently attracted a lot of research attention and this of course happened for specific reasons; it can be used for zero-emission electricity generation and also can be applied in distributed generation to the electric automobile industry (Faisal, Hannan et al. 2018). The main process involved in this storage system is that HFC burns at a quick rate and releases water vapor only into the environment. As long as the environmental footprints of HFC continue to be negligible and its associated raw materials costs are decreasing, it should not be surprising to see this technology attracting more and more enthusiasts. Last but not least, it is useful to note here that there are also even more storage options available like supercapacitors, which belong to the broad family of electrical storage systems, or like sensible heat systems, which belong to the category of thermal storage systems.

Herein, it should be mentioned that no matter the type of energy storage system chosen for an application, there are always inherent uncertainties concerning the state-of-charge or the general management of the system. Xi *et al.* have covered exclusively the topic of battery diagnostics, prognostics and uncertainty management in (Xi, R. et al. 2016). Furthermore, Dahmardeh *et al.* studied the topic of state-of-charge uncertainty in (Dahmardeh and Xi 2019). In their research work, the authors provided a unified framework for state-of-charge estimation given the variability that exists from cell-to-cell. Based on the number of cells that exist in a battery pack, they propagated the model parameters in order to provide an estimate for the overall state-of-charge.

It is mandatory to conclude the taxonomy of different energy storage systems by providing some information about hybrid energy storage systems (HESS). HESS

can be considered as a separate category of storage systems, but it is actually the combination of two or more aforementioned ESS. Given the fact that it is widely recognized that there is no single storage type that can provide all the desirable features from the planner's perspective, integrating more ESS in the same application could actually be a great idea (Faisal, Hannan et al. 2018). Extensive literature review in this area has proved that efficient configurations of HESS can be achieved by optimizing the power management strategies involved (Jing, Lai et al. 2017) or by combining storage systems with supplemental characteristics (Bocklisch 2015); an illustrative example for the latter is the integration of conventional batteries (high energy) with flywheel energy systems (high power) or fuel cells (high energy) with batteries (high power).

2.3.2 Previous research on energy storage systems in microgrids

After demonstrating the urgent need for utilizing ESS in the modern grids, it is time to turn the attention to the existing research works that studied the optimal ESS sizing and integration in microgrids. Although the problem is roughly the same in all cases, there are some distinct traits in these studies depending on whether the authors are considered isolated ("off-grid") or connected ("on-grid") systems. Before presenting the relevant works, it should be noted that these studies seldom considered the dynamic nature of storage investments in a microgrid and/or whether hybrid energy storage systems would provide benefits to the local grids. It should be highlighted that the close examination that is attempted at these topics is another research novelty of the current work.

Alsaidan *et al.* formulated a comprehensive model with the goal of optimal battery storage sizing in microgrid applications (Alsaidan, Khodaei et al. 2018). The real novelty of their work though lies in the fact that they incorporated accurate

modeling for several problem-related factors that are usually ignored; distributed deployment, the impact of depth-of-discharge (*DoD*) and the number of charging/discharging cycles in battery life are only some of them. Bahramirad *et al.* put specific emphasis on the reliability constraints of the ESS optimization sizing problem (Alsaidan, Khodaei et al. 2018). In another study, Nandi and Ghosh ran a feasibility test for a hybrid power system in Bangladesh, composed by wind-PV-battery (Nandi and Ghosh 2010). Their study showed that the optimized version of this hybrid system was more cost-efficient than all the other possible configurations that they tried. Mahani *et al.* presented a mathematical model to assess distributed energy resources, combined with energy storage, from both economic and operational perspective (Mahani, Jamali et al. 2020). Alsaidan *et al.* dealt again with the problem of optimal battery sizing in microgrids, however from a different perspective; through their mathematical formulation, they were able to compare two distinct ESS configurations: the aggregated and the distributed one (Alsaidan, Khodaei et al. 2016).

Dufo-López and Bernal-Agustín presented a novel triple multi-objective design of isolated hybrid systems (Dufo-Lopez and Bernal-Agustin 2008). Their three objectives consisted of minimizing the total cost, pollutant emissions, and unmet load. The authors used an evolutionary algorithm, supported by a secondary genetic algorithm, to solve their problem and they concluded that by using their methodology, the system designers are able to select the most adequate solution from the non-dominated set of solutions obtained, based on their application-specific criteria. Mandelli *et al.* examined closely the effect of load profile uncertainty on the optimization process of an isolated PV + battery system used for rural electrification (Mandelli, Brivio et al. 2016), while Diaf *et al.* made a technical and economic evaluation of hybrid PV/wind/battery system in Corsica island (Diaf, Belhamel et al.

2008). Kaldellis *et al.* tailored his focus on autonomous small islands (Kaldellis, Zafirakis et al. 2010); the authors proposed a comprehensive methodology in order to determine the sizing of an autonomous system at the lowest possible cost. Another really useful study in this research area can be found in (Maleki and Pourfayaz 2015). The authors of this research work compared different evolutionary algorithms in their performance when solving the optimization problem of sizing a hybrid wind/PV/battery system. Another research contribution of this work lies in the fact that they have used a constraint for the maximum allowable loss of power supply probability ($LPSP_{max}$), which resulted in more reliable final system configurations. Arun *et al.* presented a methodology for the optimum sizing of a PV + battery system with different reliability levels considering solar insolation uncertainty (Arun, Banerjee et al. 2009). Cabral *et al.* developed a stochastic model for optimal sizing of stand-alone PV + battery systems based on given load, which is shown to be more reliable and realistic (Cabral, Oliveira et al. 2010). Birnie focused on using known seasonal/daytime insolation variability to predict likely steady through-the-night emergency power delivery levels for PV + battery systems for arbitrary PV array size and battery storage capacity, assuming a system that can shift to island mode in response to a storm-damaged grid outage (Birnie 2014). In another study focused on a remote island, Ma *et al.* used HOMER software to run simulations and evaluate an "off-grid" hybrid PV/wind/battery system from technical and economic aspects (Ma, Yang et al. 2014). In (Yang, Lu et al. 2007), Loss of Power Supply Probability (LPSP) model and Levelized Cost of Energy (LCOE) are used in joint with a model for a hybrid PV/wind/battery system, in order to construct an analytical and detailed Hybrid Solar-Wind System Optimization Sizing (HSWSO) model. Lastly, Zhou et al. presented a very intriguing research work in (Zhou, Lou et al. 2010). The authors'

main research purpose is to present a detailed overview of the research progress happened so far in the area of optimum sizing of stand-alone hybrid energy systems. Besides that, the authors are able to motivate and inform the readers for several areas in which more research and development efforts need to be done.

Finally, there are research attempts in an area whose main scope is slightly deviated from what it is seen up to now. For example, Hemmati *et al.* orchestrated a dual optimization problem for their microgrid optimization: short-term scheduling and long-term expansion planning (Hemmati, Saboori et al. 2017). Although the presence of renewable energy sources in joint with energy storage is present in this research paper too, the fact that they attempted to coordinate two problems of different time scale dissociates their work from others. In another work by Hemmati *et al.* the optimization focus had been given in large scale energy storage systems and how their optimal sizing can affect generation expansion planning costs and environmental footprints (Hemmati 2016). Finally, Mahani *et al.* formulated two stochastic optimization problems concerning the energy storage allocation and scheduling in systems with high penetration of renewables (Mahani, Farzan et al. 2017). In their illustrated example, they proposed an approximation model which managed to bring a 45% reduction in computational complexity.

Overall, it is safe to assume that although the relevant studies in the field are abundant, research attempts to provide unified and holistic dynamic frameworks for tackling the problem of optimally expanding the power and the storage of microgrids in a finite time horizon are rarely seen. Also, the integration of different ESS in the so-called hybrid energy storage systems (HESS) is clearly missing from the vast majority of those studies. Herein, it should be mentioned that it is the author’s

responsibility for filling these gaps in the literature with the research that is presented in this framework.

2.4 Energy load forecasting using neural networks

In this section of the dissertation, the attention is given to neural networks and how these can be used as a forecasting tool for energy load. Artificial neural networks have proven to be a reliable methodology on various forecasting tasks. This of course includes applications closely related to microgrids' environments (Hernandez, Baladron et al. 2014). The techniques used when it comes to energy load forecasting include traditional discrete-time neural network-based approaches, but also time horizon independent methods (Xie, Parlikad et al. 2019). In the first subsection, three classical neural network approaches are explained in great detail. In the second subsection, the attention is turned into techniques that can be utilized in order to enable the development of neural networks with multiple parallel input and multi-step output. Multiple parallel input refers to multiple synchronous demand datasets coming from various sources and facilities. In other words, someone could be able to produce forecasts for many different load profiles using a single neural network architecture. Moreover, by using these techniques, someone could also boost the prediction performance of the neural networks with the addition of meteorological data. This type of data has extensively been used when energy demand (or renewable energy output) forecasting is the problem (Chen, Duan et al. 2011). On the other hand, multi-step output refers to the fact that many times the forecast needs to be made for not just one, but multiple steps ahead.

2.4.1 Classical neural network approaches

The most used neural network-based frameworks for short-term energy load forecasting are illustrated in this section. Neural networks can be broadly defined as

powerful learning models which have already proved their efficacy in a wide range of supervised and unsupervised learning tasks (Lipton, Berkowitz et al. 2015). Their success is mainly based on their ability to learn hierarchical representations and complex underlying features in various datasets by utilizing the ever more increasing computational power and storage capability that the new digital era has brought to the table. Despite their superiority over traditional techniques in various tasks, they still have limitations and therefore room for improvement. After defining the vanilla neural network, more advanced configurations are explained, such as the recurrent and the convolutional neural networks.

Artificial neural network (ANN) is a network whose inspiration is heavily based on the neural structure of the human brain (Muralitharan, Sakthivel et al. 2018). The neurons in an ANN are, like in human brain, trained using given data points and with the sole purpose to minimize the prediction errors for this specific dataset. The overall goal is therefore to find the optimal weights that connect the existing neurons in a specified ANN architecture.

The next question that arises here is how these neural nodes of computation could be organized so they become capable of recognizing complex patterns and representing highly nonlinear functions. Feedforward networks are a restricted class of networks which form a directed and acyclic graph of nodes. Using this configuration, the neural nodes are arranged in the following layers: input, hidden and output. Of course, depending on the desired depth of the ANN someone could use any number of hidden layers in this configuration. In order to introduce nonlinearity in the ANN, activation layers are succeeding some or all the hidden layers of the network. These activation layers transform their given input based on rules defined by

activation functions. Common activation functions nowadays are the sigmoid or the rectified linear unit (ReLU).

Using the sequential organization defined above, the feedforward ANN is able to map the input $\mathbf{X}$ to the desired output $\hat{\mathbf{y}}$. The input is fed to the input layer of the network, flows to the hidden and activation layers of the network and finally output is generated in the output layer. The most common and successful algorithm used to train feedforward ANNs is called backpropagation (Lipton, Berkowitz et al. 2015). The goal is to minimize the loss function $L(\mathbf{y},\hat{\mathbf{y}})$, whose goal is to penalize the distance between the predicted output $\hat{\mathbf{y}}$ and the actual output $\mathbf{y}$. In order to do so, the backpropagation algorithm uses the chain rule to calculate derivatives of the loss function with respect to each weight and then update these weights using gradient descent. Nowadays, most neural networks are trained using several variants of stochastic gradient descent with mini batches, such as the popular Adam algorithm (Kingma and Ba 2014). For example, stochastic gradient descent with mini-batch size equal to 1, updates the weights based on:

$$w \leftarrow w - \alpha \nabla_w L_i \tag{2.12}$$

Where w is the network weights, α is the learning rate and $\nabla_w L_i$ is the gradient of the loss function with respect to the weights as calculated by a single training example (x_i, y_i).

At this point, it is well proven that many feedforward neural networks have achieved state-of-the-art performance in various supervised or unsupervised learning tasks. However, their high performance depends on the assumption for independence between training and testing data (Zheng, Xu et al. 2017). This assumption clearly does not hold when dealing with time series datasets. In this scenario, there is

correlation between present and past data and therefore between training and testing datasets.

On the contrary, Recurrent Neural Networks (RNNs) have been specifically designed to work with this kind of datasets. RNNs allow a bidirectional information flow, by incorporating recurrent edges that connect adjacent time steps (Lipton, Berkowitz et al. 2015). By doing so, they are able to capture the temporal relationships inherently existing in the time series dataset. At each time point, the nodes in the network receive information from both the current data points but also from the hidden node values that stored information from previous time steps. Given an input time series $\mathbf{x}=\{x_1, x_2, ..., x_T\}$, the RNN computes a hidden sequence $\mathbf{h}=\{h_1, h_2, ..., h_T\}$ and an output sequence $\mathbf{y}=\{y_1, y_2, ..., y_T\}$ by using the following equations:

$$h_t = f\left(W_{hx}x_t + W_{hh}h_{t-1} + b_h\right) \tag{2.13}$$

$$y_t = g\left(W_{yh}h_t + b_y\right) \tag{2.14}$$

where W_{hx}, W_{hh} and W_{yh} are the input-hidden, hidden-hidden and hidden-output weight matrices. The biases for the hidden and the output layer are represented by b_h and b_y respectively and finally f and g are the activation functions.

In this work, the emphasis is given on two specific RNN architectures. The first one is the most popular architecture up to now, called Long Short-Term Memory (LSTM) and initially introduced in (Hochreiter and Schmidhuber 1997). The second one is a slight variation of LSTM, called Gated Recurrent Unit (GRU) that is now considered equally powerful but computationally more efficient than its predecessor.

The motivation behind the development of the LSTM networks is mainly the vanishing gradients problem of the classic RNN architecture when dealing with long-

term dependencies. Therefore, it is considered necessary to introduce the notion of memory in the RNNs. In order to do so, the LSTM architecture utilizes four integral parts: the input gate, the forget gate, the output gate and the memory cell state. A schematic representation of an LSTM can be seen in Fig. 2.11:

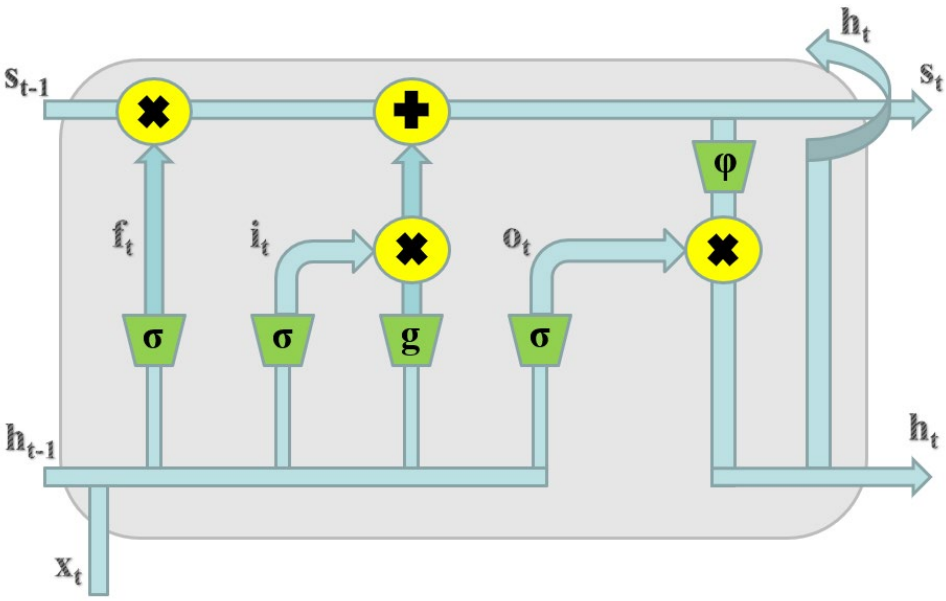

Fig. 2.11 Schematic representation of an LSTM module

First, the forget gate is responsible for helping the network to forget useless past information and update accordingly the memory cell. The input gate controls the new memory content to be added. Sigmoid functions σ are used to compute their activations:

$$f_t = \sigma\left(W_{fx}x_t + W_{fh}h_{t-1} + b_f\right) \tag{2.15}$$

$$i_t = \sigma\left(W_{ix}x_t + W_{ih}h_{t-1} + b_i\right) \tag{2.16}$$

In order to compute the new memory cell state, two element-wise multiplications are needed, between the output of the input gate and the new values to be added and between the output of the forget gate and the previous memory cell state:

$$s_t = g\left(W_{sx}x_t + W_{sh}h_{t-1} + b_s\right) \odot i_t + s_{t-1} \odot f_t \tag{2.17}$$

Finally, using the information from the forget and input gate, and from the updated cell state, the network is able to compute the output gate o_t and the final output h_t :

$$o_t = \sigma\left(W_{ox}x_t + W_{oh}h_{t-1} + b_o\right) \tag{2.18}$$

$$h_t = \varphi\left(s_t\right) \odot o_t \tag{2.19}$$

where φ is the activation function of the memory cell. Of course, it should be noted again that $W_{fx}, W_{fh}, W_{ix}, W_{ih}, W_{sx}, W_{sh}, W_{ox}$ and W_{oh} represent input-hidden and hidden-hidden weight matrices for the forget gate, input gate, memory cell and output gate, respectively. Moreover, b_f, b_i, b_s, and b_o denote the bias units. It can be seen here that, unlike the traditional recurrent unit, the LSTM is able to make decisions on whether to keep or forget incoming sequential information (Chung, Gulcehre et al. 2014).

In a similar fashion, GRUs were first introduced in 2014 in (Cho, Van Merriënboer et al. 2014) to make the recurrent unit capable of capturing temporal information from different time scales. An illustration of a GRU unit can be seen in Fig. 2.12:

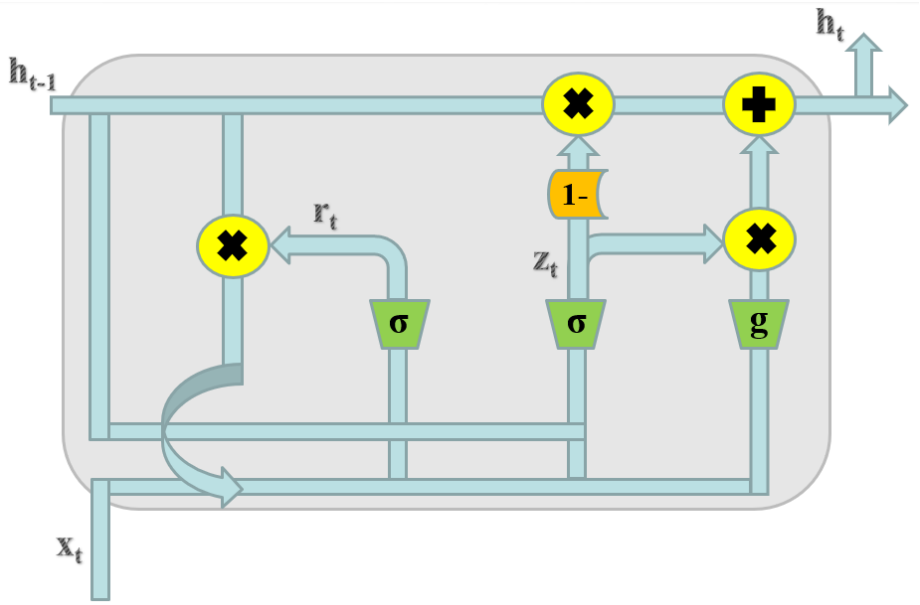

Fig. 2.12 Schematic representation of gated recurrent unit (GRU)

GRU is able to combine both input and forget gates to one single, update gate. In addition to this gate, it uses another gate called reset gate. The equations for these newly introduced update z_t and reset gate r_t are presented here:

$$z_t = \sigma\left(W_{zx}x_t + W_{zh}h_{t-1} + b_z\right) \tag{2.20}$$

$$r_t = \sigma\left(W_{rx}x_t + W_{rh}h_{t-1} + b_r\right) \tag{2.21}$$

The resulting network is a simplified version of the LSTM, with fewer weights to compute and therefore it is more efficient in terms of computational cost.

Convolutional neural networks (CNNs) were first developed for handling two-dimensional image data, with the idea of local connectivity. In order to achieve this local connectivity, the weighted sums of the previous neural networks are now replaced with convolution operations. In each layer of the CNN, the input is convolved with the weight matrix (called filter in this context) in order to create a feature map. The intuition behind this approach is to be able to create a network that is extracting all the useful spatial information from the input and results into more efficient training.

The input to a traditional convolutional layer is usually three-dimensional: the height, weight and number of channels. In each layer of the network, the input is convolved with a set of three-dimensional M^{CNN} filters, in order to create the feature output map. The third dimension of these filters is always the number of channels in the input (Borovykh, Bohte et al. 2017). However, a one-dimensional CNN is a convolution network which is designed to operate over a 1D sequence $\mathbf{x} = \{x_1, x_2, ..., x_T\}$. As in the traditional CNN hidden layer, the 1D convolutional layers are followed by pooling layers, whose responsibility is to preserve only the most salient features. In each of these hidden convolutional layers $l = 1, 2, ..., L^{CNN}$, the input feature map x^{l-1} is convolved with a set of M_l^{CNN} filters w_p^l, where $p = 1, 2, ..., M_l^{CNN}$. It should be noted here that in the case of time series data, the input shape is of dimension $N_{in} \times n_f$, where N_{in} is the number of time steps in the input window and n_f is the number of features in the dataset. The output of the convolution operation is then passed through the nonlinearity activation layer g to finally produce the output feature map x^l. This whole process is described in Eq. (2.22):

$$x^l(i,h) = g\left(\left(w_p^l * x^{l-1}\right)_i\right) = g\left(\sum_{j=-\infty}^{+\infty}\sum_{m=1}^{M_l^{CNN}} w_p^l(j,m)\, x^{l-1}(i-j,m)\right) \tag{2.22}$$

The output of L^{CNN} convolutional layers in total is $x^{L^{CNN}}$. In the 1D CNN, this output is then usually passed through a flatten layer to reduce the feature maps to 1D vectors and finally by one or more densely connected layers in order to produce the final predicted output $\hat{\mathbf{y}}$. Of course, the purpose of the optimization algorithm is again to penalize the distance between the predicted output $\hat{\mathbf{y}}$ and the actual target $\mathbf{y}$.

2.4.2 Pre-processing techniques for multiple parallel input and multi-step output

For the context of this section, let us consider a matrix $\mathbf{X}^{in}$, consisting of n_f stacked columns, where each of these columns represents an input feature. Herein, by input feature it means either a real time series vector of electricity load demand for a specific facility or any other feature available to use, i.e. weather predictors such as temperature or precipitation. Therefore, it is possible to define $\mathbf{X}^{in} = \left\{ \mathbf{x}_1^{in}, \mathbf{x}_2^{in}, ..., \mathbf{x}_{n_f}^{in} \right\}$. Each column in this matrix contains N_{in} observations, where N_{in} is the number of time steps of the input window, so $\mathbf{x}_i^{in} = \left\{ \mathbf{x}_{i,1}^{in}, \mathbf{x}_{i,2}^{in}, ..., \mathbf{x}_{i,N_{in}}^{in} \right\}$ where $i \in \left(1, 2, ..., n_f\right)$. Now, define N_{out} as the number of time steps of the output window and n_h the number of different facilities that their demand should be predicted. Therefore, the task of the multi-step ahead forecasting is using $\mathbf{X}^{in}$ to predict $\mathbf{X}^{out}$, which can be defined as $\mathbf{X}^{out} = \left\{ \mathbf{x}_1^{out}, \mathbf{x}_2^{out}, ..., \mathbf{x}_{n_h}^{out} \right\}$ where $\mathbf{x}_i^{out} = \left\{ \mathbf{x}_{i,1}^{out}, \mathbf{x}_{i,2}^{out}, ..., \mathbf{x}_{i,N_{out}}^{out} \right\}$ and $i \in \left(1, 2, ..., n_h\right)$.

The main strategies that are used at this task of multi-step ahead forecasting are three-folded (Zheng, Xu et al. 2017): the recursive strategy, the direct strategy and the multiple-input multiple-output (MIMO) strategy. It would be useful here to see what each of these strategies represent. Firstly, the recursive strategy is probably the most intuitive of all these. It refers to generating one-step ahead predictions and incorporate these in the input dataset one-step at a time in a recursive fashion. Using this strategy and the notation defined above, it is possible to produce a matrix $\mathbf{X}^{out}$ in a total of N_{out} times, where $\mathbf{X}^{j,out} = f^m\left(\mathbf{X}^{j,in}\right)$ is the prediction for the output time step j and $\mathbf{x}_i^{j,in} = \left\{ \mathbf{x}_{i,j}^{j,in}, \mathbf{x}_{i,j+1}^{j,in}, ..., \mathbf{x}_{i,j+N_{out}}^{j,in} \right\}$ where $i \in \left(1, 2, ..., n_f\right)$ and f^m is the model estimator. On the other hand, the direct strategy refers to constructing N_{out} different models, each for every time step of the output window. Therefore, in that case it is

true that $\mathbf{X}^{j,out} = f^{j}\left(\mathbf{X}^{j,in}\right)$ where $j \in \left(1,2,...,N_{out}\right)$ and f^{j} is the estimator for the output time step *j*. One advantage of the direct strategy against the recursive is that it does not utilize any forecasted data as part of the input data and therefore it is less prone to accumulated errors. However, given the fact that all models are trained separately in the direct strategy, there is a risk of degrading performance due to the statistical independence among the N_{out} forecasted values. Both aforementioned techniques are considered single output techniques as they are able to map multiple inputs (vector) to a single output (scalar). Lastly, the third strategy mentioned above (MIMO) should be presented. This strategy utilizes multiple inputs to produce directly, and by constructing only one model, multiple outputs. This is the reason why it is referred to as vector-to-vector technique. Using MIMO, it is possible to produce $\mathbf{X}^{out} = f^{m}\left(\mathbf{X}^{in}\right)$, where $\mathbf{X}^{out}$ is the whole output matrix (for all output time steps) and f^{m} is a single model estimator. It is intuitive now that the advantage of this strategy is to preserve the statistical dependency among the forecasted time series.

In Section 6 of the present dissertation, the MIMO strategy is adopted. In order to utilize a vector-to-vector approach, a specific data preprocessing procedure is needed though. First of all, given the matrices $\mathbf{X}^{in}$ and $\mathbf{X}^{out}$, the task is to produce a predicted $\hat{\mathbf{X}}^{out}$. In order to do so, the data sequences should be split in such a way that they can be used from neural networks. Each observation consists of different matrices $\mathbf{X}^{in}$ and $\mathbf{X}^{out}$, depending of course on the time period and the given input and output time windows N_{in} and N_{out} respectively. Afterwards, it is feasible to use a vector-output model. Using this model, it is possible to produce directly the output in the desired output time window without using any intermediate models. To do so, the output layer of the neural networks should be a dense “flatten” layer with size equal to

the total size of the desired predictions. At this point, it is worth mentioning another model that someone could use: the encoder-decoder model. This one is specifically developed to handle situations that it is required to predict an output sequence given an input sequence. As its name suggests, this model is comprised of two sub-models: the encoder and the decoder. Firstly, the encoder's responsibility is to interpret the input sequence and output a fixed-length vector representing this interpretation. On the other hand, the decoder should receive this interpretation, once for each output time step, and produce the desired output sequence. Of course, as in all types of neural networks, multiple layers can be stacked together to produce deeper versions of the aforementioned architectures.

2.5 Markov decision processes and reinforcement learning

In the last section of the literature review, it is considered appropriate to dig deeper into the areas of machine learning and more specifically, reinforcement learning. The exploration of these notions starts by defining and explaining the concept of Markov Decision Processes (MDPs) during the first subsection of this chapter. After that, the topic of dynamic programming is explained and brief details are given about classical approaches on this type of problems. In the third and last subsection, emphasis is given on reinforcement learning and the core algorithmic approaches involved in this area are provided. It should be mentioned here that the proper understanding of these topics could be proven very useful in Section 4 and Section 5 of this dissertation.

2.5.1 Markov decision processes

Learning from interaction and achieving a goal is the main and sole purpose of reinforcement learning. In order to understand how this process works, the reader should be familiar with a specific class of stochastic processes which is called Markov

Decision Processes (MDPs). Key definitions in MDPs are the notions of the agent and the environment (Sutton and Barto 2015). The agent is basically the decision-maker in the problem and the one who is responsible for learning. The environment is all the things that the agent should interact with, in order to get information. The agent and the environment are interacting continuously and the process which describes these communications is pretty straightforward: the agent takes actions and the environment, based on these actions, gives feedback to the agent called reward. Overall, the purpose of the agent is to maximize the sum of earned rewards over a finite (or infinite) time horizon. This process is illustrated in Fig. 2.13:

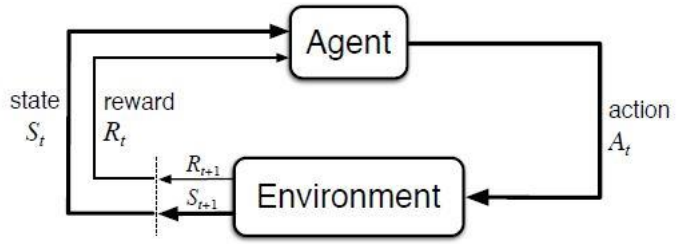

Fig. 2.13 Agent-environment interactions in reinforcement learning setting (Sutton and Barto 2015)

To be more specific, the agent and the environment interact at specific discrete time steps, $t = 0, 1, 2, 3...$ At each time step *t,* the agent receives a representative description of the environment's state $S_t \in S$, where *S* is the set of possible states of the environment and selects an action $A_t \in A(S_t)$, where $A(S_t)$ is the set of possible actions in state S_t. Consecutively, the environment sends back a numerical signal to the agent, which is usually affected by the agent's chosen action. This signal is called a reward in this context and it is denoted $R_{t+1} \in R$. The agent then is responsible to do a mapping at each time step from states to actions. This mapping is called the agent's policy, is denoted by π_t and basically $\pi_t(\alpha|s)$ refers to the probability that $A_t = \alpha$,

given that $S_t = s$. Finally, the system transits to a new state S_{t+1} and this procedure should continue iteratively until convergence is reached (Sutton and Barto 2015).

Continuing with the terminology used above, the necessary equations are presented now in order to fully define an MDP (Sutton and Barto 2015):

$$p\left(s', r \middle| s, a\right) = \Pr\left\{S_{t+1} = s', R_{t+1} = r \middle| S_t = s, A_t = a\right\} \tag{2.23}$$

$$r(s,a) = E\left[R_{t+1} \middle| S_t = s, A_t = a\right] = \sum_{r \in R} \sum_{s' \in S} p\left(s', r \middle| s, a\right) \tag{2.24}$$

$$p\left(s' \middle| s, a\right) = \Pr\left\{S_{t+1} = s' \middle| S_t = s, A_t = a\right\} = \sum_{r \in R} p\left(s', r \middle| s, a\right) \tag{2.25}$$

$$r(s,a,s') = E\left[R_{t+1} \middle| S_t = s, A_t = a, S_{t+1} = s'\right] = \frac{\sum_{r \in R} r\, p\left(s', r \middle| s, a\right)}{p\left(s' \middle| s, a\right)} \tag{2.26}$$

Eq. (2.23) gives the probability of each possible pair of next state and reward, s', r, given the current state s and action a. Eq. (2.24) defines the expected reward that is received if an action a is chosen at state s. Eq. (2.25) provides the state transition probabilities from state s to state s', given the fact that action a was chosen and finally, Eq. (2.26) calculates the expected rewards for a triple of state-action-next state s, a, s'.

The issue that naturally arises here is that, besides the notation and equations previously mentioned, the metrics and the rules that actually determine how good the agent is doing in the overall goal of maximizing its rewards are required. In order to do so, the notion of value functions in reinforcement learning should be properly defined. By formally defining the total discounted returns up to time *t* as:

$$G_t = \sum_{k=0}^{T-t-1} \gamma^k R_{t+k+1} \tag{2.27}$$

where γ is the discount factor of the rewards and by recalling that π is the agent's policy, it is possible, therefore, to define the value of a state *s* given policy π as:

$$v_\pi(s) = E_\pi\left[G_t \middle| S_t = s\right] = E_\pi\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \middle| S_t = s\right] \tag{2.28}$$

It should be noted here that the value of the terminal state should always be zero. Therefore, Eq. (2.28) gives the expected value of the discounted sum of rewards, if being in state *s* and following policy π. v_π is officially called the state-value function of policy π.

If Eq. (2.28) is doublechecked, it is clear that there is no action involved in the state-value function definition. However, in a similar way the value of taking action a while in state *s* and following policy π can be defined:

$$q_\pi(s,a) = E_\pi\left[G_t \middle| S_t = s, A_t = a\right] = E_\pi\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \middle| S_t = s, A_t = a\right] \tag{2.29}$$

in which q_π is intuitively called the action-value function of policy π.

At this point, it should be seen that Eq. (2.28) has actually a recursive nature and correlates the value of a state *s* to the value of its successor states. Mathematically, this feature can be given by the following derivations:

$$\begin{aligned} v_\pi(s) &= E_\pi\left[G_t \middle| S_t = s\right] \\ &= E_\pi\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \middle| S_t = s\right] \\ &= E_\pi\left[R_{t+1} + \gamma \sum_{k=0}^{\infty} \gamma^k R_{t+k+2} \middle| S_t = s\right] \\ &= \sum_a \pi(a|s) \sum_{s'} \sum_r p(s', r|s, a)\left[r + \gamma E_\pi\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+2} \middle| S_{t+1} = s'\right]\right] \\ &= \sum_a \pi(a|s) \sum_{s',r} p(s', r|s, a)\left[r + \gamma v_\pi(s')\right] \end{aligned} \tag{2.30}$$

The final version of Eq. (2.30) is called Bellman equation. It basically defines the relationship between the value of state *s* and the value of all its successor states and it is a fundamental equation in dynamic programming. The computation of the value function of a specific state depends highly on the next states that are going to be

encountered, so a mathematical relationship such as the Bellman equation provides the baseline for looking ahead at what is going to happen on those states.

In order to conclude the introduction to MDPs, the element of optimality should be added. It is appropriate to find formal ways to compare policies among each other and be able to extract the policies that are to best interest; by best interest here, it always means the overall goal of maximizing the discounted sum of rewards in a finite (or infinite) time horizon. A formal definition of a better policy can be given as: a policy π is better than π' if and only if $v_{\pi}(s) \geq v_{\pi'}(s) \quad \forall s \in S$, that is if its expected rewards are greater than the ones of π', for all states. Towards the goal of optimality, firstly two more functions are provided:

$$v_*(s) = \max_{\pi} v_{\pi}(s) \tag{2.31}$$

$$q_*(s,a) = E\left[R_{t+1} + \gamma v_*(S_{t+1}) \middle| S_t = s, A_t = a\right] = \max_{\pi} q_{\pi}(s,a) \tag{2.32}$$

Eq. (2.31) is called the optimal state-value function and Eq. (2.32) is called the optimal action-value function. Both of them provide necessary tools for extracting the overall optimal policy and are used in multiple ways in the next sections. Of course, both of these equations should satisfy the properties of the Bellman equation and can be rewritten in that context. Next, the so-called Bellman optimality equation is presented for the optimal value function (Sutton and Barto 2015):

$$\begin{aligned} v_*(s) &= \max_{a \in A(s)} q_*(s,a) \\ &= \max_a E_*\left[G_t \middle| S_t = s, A_t = a\right] \\ &= \max_a E_*\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \middle| S_t = s, A_t = a\right] \\ &= \max_a E_*\left[R_{t+1} + \gamma \sum_{k=0}^{\infty} \gamma^k R_{t+k+2} \middle| S_t = s, A_t = a\right] \\ &= \max_a E_*\left[R_{t+1} + \gamma v_*(S_{t+1}) \middle| S_t = s, A_t = a\right] \\ &= \max_a \sum_{s',r} p(s', r | s, a)\left[r + \gamma v_*(s')\right] \end{aligned} \tag{2.33}$$

Eq. (2.33) illustrates the recursive nature of the optimal state-value function by formally correlating the optimal state-value of state *s* with the optimal state-value of its successor states. In a similar fashion, the Bellman optimality equation is defined for the action-value function:

$$\begin{aligned} q_*(s,a) &= E\left[R_{t+1} + \gamma \max_{a'} q_*(S_{t+1}, a') \middle| S_t = s, A_t = a\right] \\ &= \sum_{s',r} p(s', r|s, a)\left[r + \gamma \max_{a'} q_*(s', a')\right] \end{aligned} \tag{2.34}$$

It should be mentioned that Eqs. (2.33) and (2.34) have unique solutions for finite-horizon MDPs. These two equations actually are consisted of two subsets of equations, where each set consists of *N* equations, where *N* is the number of possible states of the problem. The problem of solving these equations is the actual optimization framework and is going to be discussed in the next sections.

2.5.2 Dynamic programming

Dynamic programming refers to a set of algorithmic techniques used to find optimal policies, given a perfect model for the existing environment. The classical dynamic programming approaches that are going to be presented in this section are of great theoretical significance, but they have two serious limitations: they are computationally inefficient and they always assume a perfect model for the environment. This section of the dissertation is organized as follows: firstly, the necessary tools for evaluating policies are provided, followed by ways for improving policies and lastly, two classical dynamic programming approaches are presented: policy iteration and value iteration.

Examining again Eq. (2.28) and (2.30), it can be observed that the state-value function can be expressed in closed form and computed accurately if the environment's dynamics are completely known (Sutton and Barto 2015). Eq. (2.30) becomes then a set of *N* linear equations with *N* unknowns, where *N* is the number of

possible states, |*S*|. However, this method would prove computationally expensive and an iterative solution would be preferred. By using the Bellman equation for v_π, someone can compute:

$$\begin{aligned} v_{k+1}(s) &= E_\pi\left[R_{t+1} + \gamma v_k(S_{t+1}) \middle| S_t = s\right] \\ &= \sum_a \pi(a|s) \sum_{s',r} p(s', r|s, a)\left[r + \gamma v_k(s')\right] \end{aligned} \tag{2.35}$$

It should be mentioned here that the sequence $\{v_k\} \to v_\pi$, as $k \to \infty$. By using Eq. (2.35), the state-value function can be computed for all possible states. This operation is called a full backup because each computation is based on all possible next states than just a sample (Sutton and Barto 2015). Consequently, in order to use the aforementioned way for policy evaluation, it is required to examine the entire state space, which can be proven extremely inefficient, even for modern computational machines.

Now that the necessary tools for determining "how good" a policy is are defined, the next tool that is required is a way to improve these policies. If it is thought as a one-step problem, it could be restated as follows: if in state *s* following policy π, would it be better to choose α and then continue following policy π or should the previous plans hold? The generalization of this question is called policy improvement theorem. If, for given deterministic policies π, π' (by deterministic, the policies themselves define the state transitions in a deterministic fashion), it is true that for all $s \in S$:

$$q_\pi(s, \pi'(s)) \geq v_\pi(s) \tag{2.36}$$

It means that policy π' should be better than policy π. In order to construct the final version of the policy improvement theorem, it should be observed that, given the fact that there is a way to improve a policy at a given state, what is needed is

actually to expand this methodology to all possible states and actions. Therefore, a greedy policy can be computed for all states $s \in S$ by:

$$\begin{aligned}\pi'(s) &= \arg\max_a q_\pi(s,a) \\ &= \arg\max_a E\left[R_{t+1} + \gamma v_\pi\left(S_{t+1} \middle| S_t = s, A_t = a\right)\right] \\ &= \arg\max_a \sum_{s',r} p\left(s', r \middle| s, a\right)\left[r + \gamma v_\pi(s')\right]\end{aligned} \tag{2.37}$$

At the end, this greedy policy is able to give the optimal action α at every possible state *s*. Given the aforementioned ways for evaluating and improving policies, it is time to mention the two most important classical dynamic programming approaches: policy iteration and value iteration. The former approach is actually self-explanatory: by utilizing the fact that a finite MDP has always a finite number of policies, it iteratively uses policy evaluation and policy improvement until it finally converges to the optimal policy. The algorithm is shown in Fig. 2.14:

1. Initialization
   $V(s) \in \mathbb{R}$ and $\pi(s) \in \mathcal{A}(s)$ arbitrarily for all $s \in \mathcal{S}$

2. Policy Evaluation
   Repeat
      $\Delta \leftarrow 0$
      For each $s \in \mathcal{S}$:
         $v \leftarrow V(s)$
         $V(s) \leftarrow \sum_{s',r} p(s', r|s, \pi(s))[r + \gamma V(s')]$
         $\Delta \leftarrow \max(\Delta, |v - V(s)|)$
   until $\Delta < \theta$ (a small positive number)

3. Policy Improvement
   policy-stable $\leftarrow$ *true*
   For each $s \in \mathcal{S}$:
      $a \leftarrow \pi(s)$
      $\pi(s) \leftarrow \arg\max_a \sum_{s',r} p(s', r|s, a)[r + \gamma V(s')]$
      If $a \neq \pi(s)$, then *policy-stable* $\leftarrow$ *false*
   If *policy-stable*, then stop and return V and π; else go to 2

Fig. 2.14 Policy iteration algorithm (Sutton and Barto 2015)

As it can be seen, there are three main steps involved in the algorithm: firstly a random policy is chosen, afterwards it is evaluated for all the possible states and lastly it is improved for each one of these states. It should be noted that the last two steps of

the algorithm are iterative themselves. This iteration stops when it is possible to conclude in an optimal policy, based on preset convergence criteria.

The second algorithm presented, value iteration, attempts to solve a serious problem involved in the first approach; at each iteration, all policies are evaluated again and again. This fact, of course, hurts in terms of computational cost. It is proven that policy evaluation steps can be truncated without losing the convergence guarantees (Sutton and Barto 2015). Using this exact fact and by integrating it with policy improvement, the value iteration algorithm uses its own backup update:

$$\begin{aligned} v_{k+1}(s) &= \max_a E_\pi\left[R_{t+1} + \gamma v_k(S_{t+1}) \middle| S_t = s, A_t = a\right] \\ &= \max_a \sum_{s',r} p\left(s', r \middle| s, a\right)\left[r + \gamma v_k\left(s'\right)\right] \end{aligned} \tag{2.38}$$

Eq. (2.38) holds for all $s \in S$ and now with $\{v_k\} \to v_*$, as $k \to \infty$. In order to understand it in a simpler way, the reader can think the value iteration algorithm as converting the Bellman optimality equation for state-value function that it was seen in Eq. (2.33) to an update rule such as the one in Eq. (2.38). An illustrative representation of the value iteration algorithm can be shown in Fig. 2.15:

Initialize array V arbitrarily (e.g., $V(s) = 0$ for all $s \in \mathcal{S}^+$)

Repeat
  $\Delta \leftarrow 0$
  For each $s \in \mathcal{S}$:
    $v \leftarrow V(s)$
    $V(s) \leftarrow \max_a \sum_{s',r} p(s', r|s, a)\left[r + \gamma V(s')\right]$
    $\Delta \leftarrow \max(\Delta, |v - V(s)|)$
until $\Delta < \theta$ (a small positive number)

Output a deterministic policy, π, such that
  $\pi(s) = \text{argmax}_a \sum_{s',r} p(s', r|s, a)\left[r + \gamma V(s')\right]$

Fig. 2.15 Value iteration algorithm (Sutton and Barto 2015)

Concisely, value iteration joins successfully the procedure of policy evaluation and the procedure of policy improvement in each of its iterations. To sum up, it

should be mentioned that there is a whole class of truncated policy iteration algorithms that are although out of the scope of the current research.

2.5.3 Reinforcement learning

Reinforcement learning algorithms are actually only one branch of the broad family of machine learning algorithms. Before presenting the theory and definitions required, it is mandatory to understand the broad categorization of machine learning areas. In order to do that, a taxonomy of all learning approaches existing right now in the literature is presented. A widely used classification of machine learning is given in Fig. 2.16:

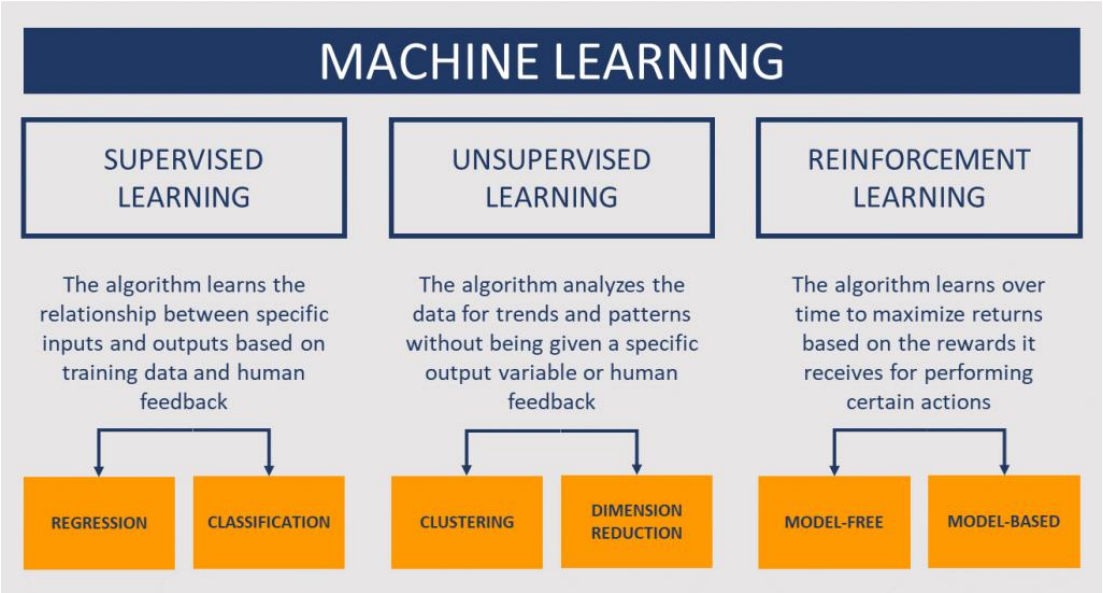

Fig. 2.16 Categorization of machine learning techniques (Source: Enaxis Consulting)

As it can be seen, the first category belonging to the family of machine learning is called supervise learning; it is probably the most well-known branch of machine learning and it refers to situations where the target variable is known. In this case, the target variable is present in the dataset and the model should learn based on the value of this target variable (Solutions 2017). The main subproblems belonging to this category are the famous regression and classification problems. The second type of machine learning is somewhat more complicated and is called unsupervised

learning; the main difference between supervised and unsupervised learning is that in the latter, the model should learn to optimize without having a specific target as a purpose. Examples of this category include but are not limited to clustering ("grouping" inputs based on their distinct attributes) and dimensionality reduction ("learning" what are the most important features in a dataset). The third and last type of machine learning is called reinforcement learning and is the one that is going to be used heavily in this research work; it could easily be said that reinforcement learning lies in a gray area between supervised and unsupervised learning, given the fact that there is still no target from which the model can extract information from, but however there is a specific form of interaction between an environment and an agent responsible to make the optimal decisions. Therefore, in reinforcement learning, there are reward signals that the agent should receive from an environment and therefore decide which actions are good or bad for the problem that it is trying to solve. The applications of reinforcement learning nowadays are abundant given the data-centric era that is approaching and the number of processes requiring accurate and optimal decision-making. In the last part of this section, the existing applications of reinforcement learning are covered in greater detail and specifically tailored for the energy systems research field, which are therefore related to a high degree to the current research work.

There are two main criteria that can be used to classify different reinforcement learning approaches. The first criterion refers to whether there is a perfect model of the environment (or if someone can build and use such a model) or not. Model-based methods know or "learn" how the environment works and so they can predict the next states that they are going to enter or the rewards that they are going to receive. However, most model-based approaches become impractical when dealing with very

large state-action spaces (Huang 2018). On the contrary, model-free approaches simply learn by trial-and-error. These approaches do not require knowledge about the environment and therefore can “learn” an optimal policy by iteratively experiencing trajectory paths and their corresponding sets of rewards and states. These are the kind of algorithms that have attracted a lot of interest recently and the reason behind that is straightforward: computational efficiency. The second criterion that can be used to classify reinforcement learning algorithms is whether the algorithms “learn” off-policy or on-policy; off-policy means that the agent estimates its returns given that a greedy policy is followed after, even though the agent itself may not follow this greedy action. On the other hand, on-policy calculates the expected rewards based on the policy that is currently following. The first algorithm that is described and presented in this subsection is probably the most famous reinforcement learning algorithm, called Q-learning. In the subsequent paragraphs after that though, some different approaches that could be taken are elaborated. These are motivated by some limitations that the Q-learning algorithm imposes.

Q-learning is a model-free, off-policy learning algorithm that depends heavily on the well-known Bellman equation that was presented in (2.34). By using (2.34), the Q-learning algorithm is able to update its Q-values by using the following update rule:

$$q(s,a) \leftarrow q(s,a) + \alpha\left[r + \gamma \max_a q(s',a) - q(s,a)\right] \tag{2.39}$$

where α here denoting the learning rate of the algorithm, i.e. how fast to approach the optimal solution. Herein, it can be observed that the Q-learning update rule resembles a lot to the value iteration algorithm. However, this approach has the advantage that it does not require the transition probability matrix of the classical

dynamic programming approach. In the following Table 2.1, the Q-learning algorithm is presented in steps:

Table 2.1 Q-learning algorithm

Algorithm: Q-learning

1: **initialization:** Q table
2: **for** *every training episode* **do:**
3: **initialization:** starting state s
4: **for** *every decision period* **do:**
5: select action α based on Q and ε-greedy policy
6: observe reward r and next state s'
7: $Q(s,a) \leftarrow Q(s,a) + \alpha\left[r + \gamma \max_a Q(s',a) - Q(s,a)\right]$
8: $s \leftarrow s'$
9: **end for**
10: **end for**

The algorithm is pretty straightforward: in line 1, a look-up table Q is initialized, which is actually a matrix, in which different states are in different rows and actions are in different columns. For every training episode of the algorithm, the state of the environment is initialized and then for every decision period (or time epoch), an action is chosen based on the look-up table that there is at hand and following an ε-greedy policy. The ε-greedy policy refers to the exploration/exploitation tradeoff; initially, the agent chooses actions almost randomly (so ε should be high) but as convergence is approaching, the agent is forced to choose actions mostly based on the maximum q-values of the look-up table, depending on the specific state (row) that it is found at any time (so ε should be low). More information about exploration/exploitation tradeoff can be found in (ADL 2018). After that, the agent observes the reward that it received and the next state that it goes to. Finally, it updates the corresponding element of the look-up table based on the Q-learning update rule and also updates its next state. The Q-learning has been proven to converge to the optimal solution, given Markov properties in the state-to-state transitions and an infinite number of visits to each state-action pair. A simple proof of

convergence can be found in (Melo 2001). The final result of the Q-learning algorithm is actually the "trained" look-up table with size $|S|\times|A|$, which can give the optimal action for every state by simply checking the column which has the maximum q-value. An example of the look-up table can be seen in Fig. 2.17:

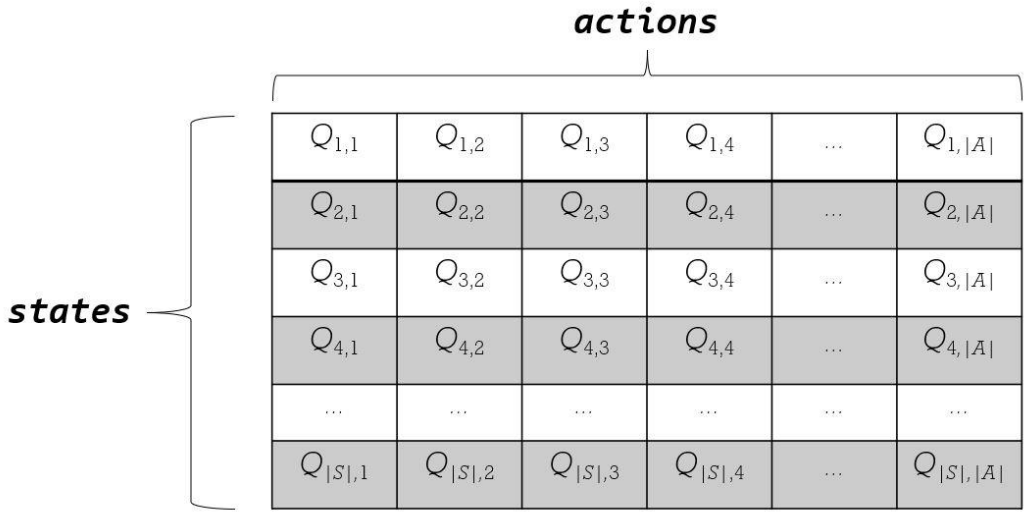

Fig. 2.17 Look-up table of Q-learning algorithm

Despite its simplicity and its high usage rate in many MDP settings, Q-learning has been proven that suffers from some serious underperformance issues. van Hasselt showed that the algorithm's performance can have really poor performance in stochastic MDPs due to a large overestimation of action values (van Hasselt 2010). This overestimation comes from the fact that positive bias is inherent to the Q-learning algorithm from using the maximum action value as an approximation of the maximum expected action value. Q-learning uses the single estimator approach for estimating the value of the next state; $\max_a q(s',a)$ is an estimate for $E\left[\max_a q(s',a)\right]$, but then, in turn, it is used as an estimator for $\max_a E\left[q(s',a)\right]$. van Hasselt proved that this estimator is biased in highly stochastic environments because instead of the expectation over the next state, only the average over all

possible results of the experiment is computed. It would be useful here to see an illustrative example of this problem in Fig. 2.18:

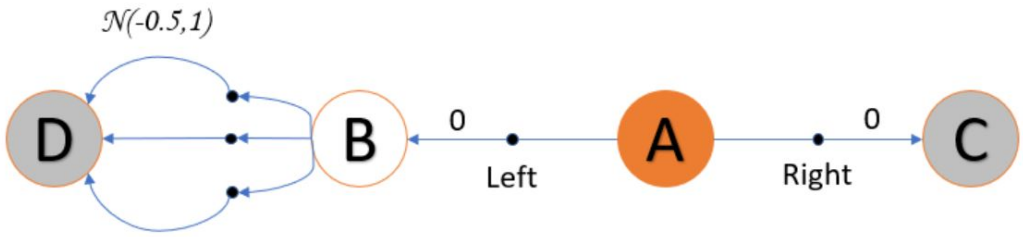

Fig. 2.18 An example of an MDP with four states (Salloum 2018)

Let Fig. 2.18 define an MDP with four states and where D, C represent the terminal states. All the rewards are 0, except for the reward associated with the transition of the agent from state B to state D: $r_{B\to D} \sim N(-0.5,1)$. It is intuitively clear from this example, that the optimal strategy for the agent would be to move right to the terminal state C directly. However, Q-learning may be tricked in this setting and think that the optimal strategy is to move left because some of the rewards incurred by these actions would be positive. Therefore, in order to solve the aforementioned problem, van Hasselt proposed a new approach, called Double Q-learning (van Hasselt 2010); the intuition behind this approach is that the selection of the best action should be decorrelated with the evaluation of this action. The algorithm is presented in Table 2.2:

Table 2.2 Double Q-learning algorithm

Algorithm: Double Q-learning

initialization: Q^A table and Q^B table
for *every training episode* **do:**
    initialization: starting state s
    for *every decision period* **do:**
        select action α based on Q^A, Q^B and ε-greedy policy
        observe reward r and next state s'
        generate $d \sim U(0,1)$

8: **if** $d \leq 0.5$ **do:**
9: $Q^A(s,a) \leftarrow Q^A(s,a) + \alpha\left[r + \gamma Q^B\left(s', \text{argmax}_a Q^A(s',a)\right) - Q^A(s,a)\right]$
`10: **else do:**
11: $Q^B(s,a) \leftarrow Q^B(s,a) + \alpha\left[r + \gamma Q^A\left(s', \text{argmax}_a Q^B(s',a)\right) - Q^B(s,a)\right]$
12: **end if**
13: $s \leftarrow s'$
14: **end for**
15: **end for**

The idea is simple and straightforward; instead of one Q-table, two are going to be used: Q^A and Q^B. At each iteration of the algorithm, only one of these is arbitrarily selected to be updated. The crucial point of the algorithm and its' characteristic difference with Q-learning is that the selection of the optimal action is based on one look-up table, while the evaluation of the action-value is based on the other look-up table. That way, it is possible to avoid the pitfall of overestimation bias that the Q-learning imposes.

Although the algorithms presented up to now are powerful (and are used later on in Section 4), there are still some drawbacks concerning their ability to generalization. These algorithms use mainly two-dimensional arrays to store their Q-values, a procedure which resembles a lot to dynamic programming (Huang 2018). Therefore, it is clearly seen that in higher dimensions, this could potentially pose a danger; the agent has no knowledge for unseen states or at least less knowledge for less seen states. As the dimensionality of the problem increases, this impact becomes more serious, too. The solution to this problem can come from the utilization of deep Q-networks (DQN). Given a state *s* as an input, the DQN is able to output a vector of action-values $Q(s,:;\theta)$, where θ are the parameters of the network. The key notions behind deep Q-learning are two: experience replay and target network (Choudhary 2019). The former one refers to the systematic way that input data are selected for the training of the deep neural network. More specifically, the experiences of the agent

are not immediately used to train the network, but instead are stored in a buffer to be used later. At predetermined intervals, a random sample of the buffer's experiences are chosen and then the deep neural network can be trained based on this sample. This fact ensures that the correlations between training samples is low, while this would not be the case if there was not this specific experience buffer. It should be mentioned here that by experiences, it is actually meant a tuple of (state, action, reward, next state). Concerning the target network, it should be noted that if a single network was used for calculating both the predicted and the target value, there could be a lot of divergence between these two (Choudhary 2019). A good analogy that is often given, is that this process looks like someone is trying to hit a moving target. Therefore, a separate network is used to estimate the target. The target network has similar architecture with the main one, but its' parameters are "frozen". Again, at predetermined intervals, the parameters of the main network can be copied to the target network and this could lead to a more stable training. Overall, a visualization of how deep Q-learning relates to tabular Q-learning can be seen in Fig. 2.19:

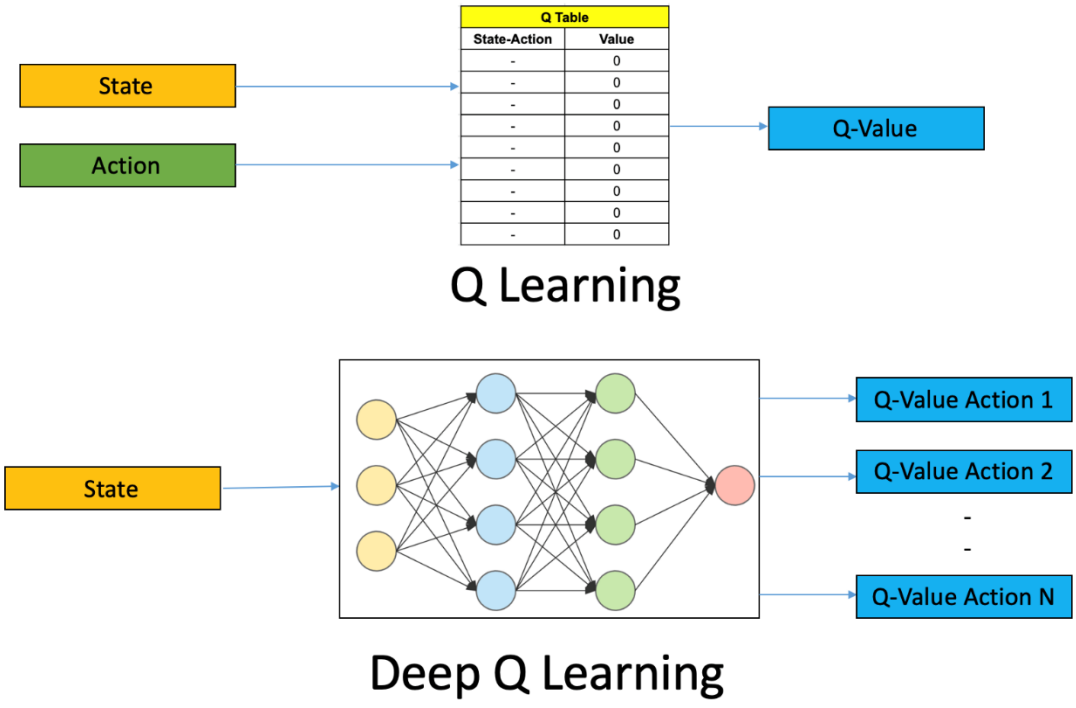

Fig. 2.19 Relationship between tabular and deep Q-learning (Choudhary 2019)

In a similar fashion presented above for the Double Q-Learning theory, van Hasselt presented that his framework can be extended to include neural networks, resulting in a network called Double Deep Q-Network (DQN) (van Hasselt, Guez et al. 2016). Double DQN has achieved exciting performance in many applications, with the most notable one being the DeepMind successful attempt to use it in playing Atari games (Mnih, Kavukcuoglu et al. 2013).

Finally, it is considered appropriate to mention a few other well-known reinforcement learning techniques that have already achieved significant results. State-Action-Reward-State-Action (SARSA) is a very similar methodology with Q-learning. The key difference is that SARSA is an on-policy algorithm. Therefore, it implies that SARSA updates the Q-values based on the policy currently followed and not a greedy policy (Huang 2018). Furthermore, a variant to the Deep Q-learning algorithm is the Deep Deterministic Policy Gradient (DDPG) framework. The motivation for the DDPG algorithm arises from situations where the action space is continuous (or close to continuous – largely discretized). DDPG relies on the actor-critic architecture (Huang 2018). Although further details on this architecture are out of the scope of the current research, it should be noted that DDPG utilizes temporal differences in a learning procedure called Temporal-Difference (TD) learning, which is actually a superset of the Q-learning presented in this section.

Reinforcement learning techniques have been previously used in the energy systems research field. The following paragraphs of this section are devoted to presenting such applications. It should be mentioned though that almost all of these works, if not all, utilized reinforcement learning frameworks in order to solve mostly short-term planning problems, such as battery scheduling or unit commitment.

Consecutively, another novelty of the current research work is that it is one of the first attempts to deploy reinforcement learning algorithms to expansion planning problems on a multi-year horizon. Although this fact poses additional limitations and dangers, mainly concerning the higher degree of stochasticity involved, it is the author's belief that by using detailed and analytical modeling and methodologies it is able to present a comprehensive framework for dealing with this kind of problems.

Mbuwir *et al.,* intrigued by recent advancements in batch reinforcement learning, presented an approach in which they utilized this method in battery energy management (Mbuwir, Ruelens et al. 2017). After modeling their agent and environment setting for the battery using MDP, they used an algorithm called fitted Q-iteration which is used by the agent to derive optimal control policies. The authors tested their methodology in Belgian residential customers and they showed that they can achieve a performance increase of 19%. Yousefi *et al.* used a Q-learning-based approach to develop a new dynamic maintenance methodology for systems with individually repairable components (Yousefi, Tsianikas et al. 2020). In another study, Raju et *al.* dealt with the battery scheduling optimization in a microgrid setting where photovoltaic cells and battery are present (Raju, Sankar et al. 2015). The algorithm used in that case is called Coordinated Q-learning (CQL) and it involves multiple agents taking actions in the same environment. Multi-agent reinforcement learning for deriving optimal control policies in microgrids is also used in other studies (Dimeas and Hatziargyriou 2010) (Li, Wu et al. 2012). Furthermore, a very intriguing study is presented in (François-Lavet, Taralla et al. 2016). The novelty of this work, which is also relevant to the current research topic, is that the authors considered microgrids which incorporate different types of storage systems. After presenting the deep reinforcement learning framework, they tailored the state, action and reward

definition to their microgrid case study and they proved that the deep network variation of reinforcement learning is able to generalize efficiently in situations where unseen configurations of the system are met. This is a result of great importance to this research because it can show that this specific framework can be used in environments with high level of stochasticity, such as the environment that exists in this research work. Finally, Kuznetsova *et al.* proposed an algorithm for battery scheduling in a microgrid scenario with a hybrid wind and battery system (Kuznetsova, Li et al. 2013). The authors defined their own sets of scenarios, actions, and rewards for the battery scheduling problem, which they solved it by using a 2 steps-ahead reinforcement learning algorithm. This example is indeed an authentic approach in the research area of multi-state and multi-criteria decision making for medium-term energy storage management.

3. Static Storage Expansion Planning in Microgrids

This section of the dissertation is concerned with the static problem of storage expansion planning in microgrids. This problem refers to determining the optimal battery size to be attached in a given PV array. By static, it is meant that any investment in energy storage in the microgrid should take place at the start of the time horizon under examination. In the first subsection, the general problem is presented, along with the optimization frameworks used and the results of the case study. In the second subsection, the emphasis is given in the trade-off between investment and unmet load penalty costs.

3.1 Benefits of adding battery capacity to solar/battery microgrids

Power systems with photovoltaic (PV) arrays combined with battery backup storage are frequently analyzed to assess their capability of working in power island mode during grid outages. The problem under examination is to determine the optimal battery sizes for PV + battery systems with given solar array sizes, from both reliability and economic perspectives. Concisely, the methodology used to optimize battery capacity for PV + battery systems is presented, which can operate in island mode to supply customers with reliable power economically during a grid outage. Sensitivity analysis of the impact of cost on reliable energy supply for facilities is also performed. The results can finally provide with insight into the trade-off between minimal system total cost and reliable power supply of PV + battery systems.

3.1.1 Simulation-based optimization as a preliminary model

In this section, the proposed optimization problem formulation for battery sizing for PV + battery systems, with the reliability level of power supply as a constraint, is presented. The main variables and parameters considered in the optimization problem are:

1) Total system cost, which is the objective of the problem that needs to be minimized. Total system cost includes two parts: the investment cost of added battery capacity and the loss of load cost which denotes the incurred cost of unsatisfied load demand of customers.

2) Solar irradiation, which represents the solar radiation input of the PV panel. Real solar insolation data of hourly GHI are considered (NREL 2016) (global horizontal irradiance, combining both direct and scattered light hitting a level local reference plane).

3) Load demand, which is the electricity usage demand of customers. Actual data of load demand for specific model facilities are applied in this work (NREL 2013).

4) Loss of load probability (*LOLP*), which is calculated as the proportion of time when load demand of facilities cannot be met by PV + battery system during a grid outage. It is adopted as one of the reliability metrics for the electricity supply of PV + battery system, which has already been used by some researchers to develop sizing algorithms to minimize system cost (Kazem and Khatib 2013) (Khatib 2012) (Yang 2008).

The corresponding mathematical model is shown as follows (Zhou, Tsianikas et al. 2019):

$$\min \quad B_r \times b + VOLL \times \sum_{k=1}^{K} \left((1+\gamma)^{-k} \sum_{t} AEL(t) \right)$$

$$\text{s.t. } AEL(t) = \begin{cases} D(t), & \text{if } Q_B(t) + \int_t^{t+\Delta t} \left(P(u) - D(u) \right) \times e \, du < B_{\min} \\ 0, & \text{otherwise} \end{cases} \tag{3.1}$$

$$Q_B(t+\Delta t)=\begin{cases}\min\left\{Q_B(t)+P(t)\times e, B_r\right\}, & \text{if } Q_B(t)+\int_t^{t+\Delta t}\left(P(u)-D(u)\right)\times e\,du < B_{\min}\\ Q_B(t)+\int_t^{t+\Delta t}\left(P(u)-D(u)\right)\times e\,du, & \text{if } B_{\min}\le Q_B(t)+\int_t^{t+\Delta t}\left(P(u)-D(u)\right)\times e\,du \le B_r\\ B_r, & \text{if } Q_B(t)+\int_t^{t+\Delta t}\left(P(u)-D(u)\right)\times e\,du > B_r\end{cases}$$

(3.2)

$$P(t)=\eta\times I(t)\times A,\quad \forall\, t \tag{3.3}$$

$$B_{\min}=B_r\times\left(1-DoD\right) \tag{3.4}$$

$$\Pr\left\{LOLP\le\beta\right\}\ge 1-\alpha \tag{3.5}$$

where

$$LOLP=\sum_{t\in DNS} t\Big/T,\quad DNS=\left\{t;\ Q_B(t)+\int_t^{t+\Delta t}\left(P(u)-D(u)\right)\times e\,du < B_{\min}\right\}$$

$$0\le\alpha,\quad \beta\le 1,\quad B_{\min}\ge 0,\quad B_r\ge 0$$

The objective function consists of two parts, i.e., $B_r\times b$ represents battery investment cost and $VOLL\times\sum_{k=1}^{K}\left((1+\gamma)^{-k}\sum_t AEL(t)\right)$ denotes the total cost of lost load during grid failure for all k years of the planning horizon K, discounted with a rate γ. B_r is battery capacity (kWh), and b denotes battery price ($/kWh). $VOLL$ is the penalty cost of unmet demand ($/kWh), which is a typical monetary expression for the costs caused by electricity interruptions (Leahy and Tol 2011) (Mandelli, Brivio et al. 2016) (Willis and Garrod 1997). Time is a discrete parameter in this formulation with increments of Δt, typically one hour in these analyses.

Eq. (3.1) presents the amount of load demand that is not satisfied by the PV + battery system at time t. B_{min} is minimum battery energy value (kWh), and $Q_B(t)$ is the energy stored in the battery (kWh) at time t. $P(t)$ is the power generated by PV array (kW) at time t and $D(t)$ is load demand (kW) at time t. e represents the efficiency of

discharging/charging process of the battery. The assumption made in this aspect is that no partial demand satisfaction is allowed. Eq. (3.2) presents the change of energy stored in the battery over a short time period Δt due to the power generation of the PV array and load demand of the facilities. It describes the energy balance of the PV + battery system, and several previous methods are proposed based on this topic (Bucciarelli 1986) (Klein and Beckman 1987). Eq. (3.3) denotes the energy produced by the PV array at time *t.* η is the energy conversion efficiency of the PV array. $I(t)$ is solar irradiation (W/m^2) at time *t*. *A* is PV array area (m^2). Eq. (3.4) denotes how B_{min} is calculated with *DoD*, which is the abbreviation of the maximum depth-of-discharge of the battery (%). Eq. (3.5) presents the definition of Chance Constraint Probability (*CCP*), and 1 – *a* denotes the minimum desired probability. *LOLP* is obtained as $\sum_{t \in DNS} t$ divided by *T*, which is limited to be not more than β. *T* is grid outage duration, and *DNS* denotes the set of time when demand is not satisfied by the power output of the PV + battery system during a grid outage. The smaller the β and *a* are, the more strict and limiting the constraints are.

Herein, it is considered appropriate to elaborate more on the simulation methodology that has been developed and applied to solve the problem, and also on several assumptions considered for the environment. Like other industrial systems, the life cycle of a PV + battery system should be taken into account when a cost-benefit problem is analyzed. Whereby, the investment of battery is designed for 20-year operation of a PV + battery system in the simulation model (Koutroulis, Kolokotsa et al. 2006), i.e., decisions for the PV + battery system are made considering 20 years of operation, and loss of load cost is accumulated for these 20 years. It should be noted that loss of load cost in each year is calculated as the present value with a discount rate of 4%.

Three metrics are used to evaluate the islanding capability of the PV + battery system from both economic and reliability aspects, which are total system cost, *CCP* and system achieved *LOLP*. Chance constraint probability is estimated as the proportion of grid outages in which the *LOLP* constraint are satisfied by the islanding operation of PV + battery system with specific battery sizing, i.e., an estimate of the probability that the *LOLP* constraint could be met by PV + battery system energy supply during outage duration. System achieved *LOLP* denotes the actual *LOLP* achieved by islanding mode of PV + battery system during 20 years of operation. These two metrics represent the reliability level of the PV + battery system power supply during a grid outage.

CAIDI and SAIFI are two standard customer-oriented reliability indices of electricity distribution systems (Association 2012). SAIFI is reported in average interruptions per year on the system. CAIDI is reported as the average length of time that a customer's outage lasts in minutes (Association 2014). In this work, grid outage duration T is assumed to follow a shifted or scale adjusted Poisson distribution in which $T > 0$ and the expected outage duration is approximately equal to CAIDI. The number of outages per year equals to SAIFI. It should be noted that CAIDI is applied with the unit of hours in this context. The starting time of a grid outage is assumed to be uniformly distributed throughout a year. It is also assumed that *VOLL* is uniformly distributed in a specific range, which is determined based on the criticality of the facility, as described below. *VOLL* is chosen as a random value within a range because the economic impact of each outage could vary depending on specific missions or unique characteristics of the facilities.

Overall, the simulation steps can be seen in Table 3.1:

Table 3.1 Simulation procedure

Simulation procedure

```
1: initialization: random outages for all simulation runs
2: for every array size do:
3:     for every simulation run do:
4:         calculate PV array production
5:         simulate charging/discharging of battery
6:         compute amount of energy lost and associated cost
7:     end for
8:     compute total system cost and reliability metrics for a specific array
9: end for
```

By iteratively following this procedure, it is possible to obtain the optimization results that are presented in Section 3.1.2 for the case study.

3.1.2 Case study results and analysis

A hospital facility, subject to grid outage and served only by the PV + battery system, is selected for demonstration purposes. Most of the previous methods which describe the energy balance of the PV + battery system assume a constant daily load demand and even rely on random models for generating solar radiation data (Casares, Lopez-Luque et al. 2014), which are inconsistent with real situations to some extent. In the present work, actual historical data for solar irradiation and load demand are applied to numerical examples as an improvement, which makes the results more realistic (NREL 2013) (NREL 2016). The considered facility is a hospital in Islip, Long Island, NY, USA.

Concerning the numerical assumptions considered this case study, it should be mentioned that the battery price equals 162 $/kWh (Curry 2017). *VOLL* is within the selected 90% confidence-level (*CL*) range of (5 $/kWh - 25 $/kWh), which is based on the expectation of *VOLL* level in 2030 (van der Welle and van der Zwaan 2007). The solar radiation profile and load profile of the facilities used are for Islip, Long Island, NY, USA (NREL 2013) (NREL 2016). It is commonly assumed that the generator efficiency remains constant (Ambrosone, Catalanotti et al. 1985); here the PV array conversion efficiency is set to be 16% (Aggarwal 2019). The efficiency of

the battery charging/discharging process is set to be 85%, and *DoD* is 70%. Values of CAIDI and SAIFI are 8.23 and 1.17 for the Long Island, NY area (Service 2018).

The selected, model hospital is a large (average load demand is 1,138 kW) and critical facility, so *VOLL* for this facility is selected between 22 $/kWh and 25 $/kWh. The *LOLP* constraint is 10%. Simulations (100,000 in total) are performed on three different solar array sizes, i.e., 200 m^2, 5200 m^2, 10200 m^2, with battery capacities ranging from 200 kWh to 20,000 kWh. The final model's goal is to determine the optimal battery size for the given PV array. It should be noted that each outage hour is considered to be independent.

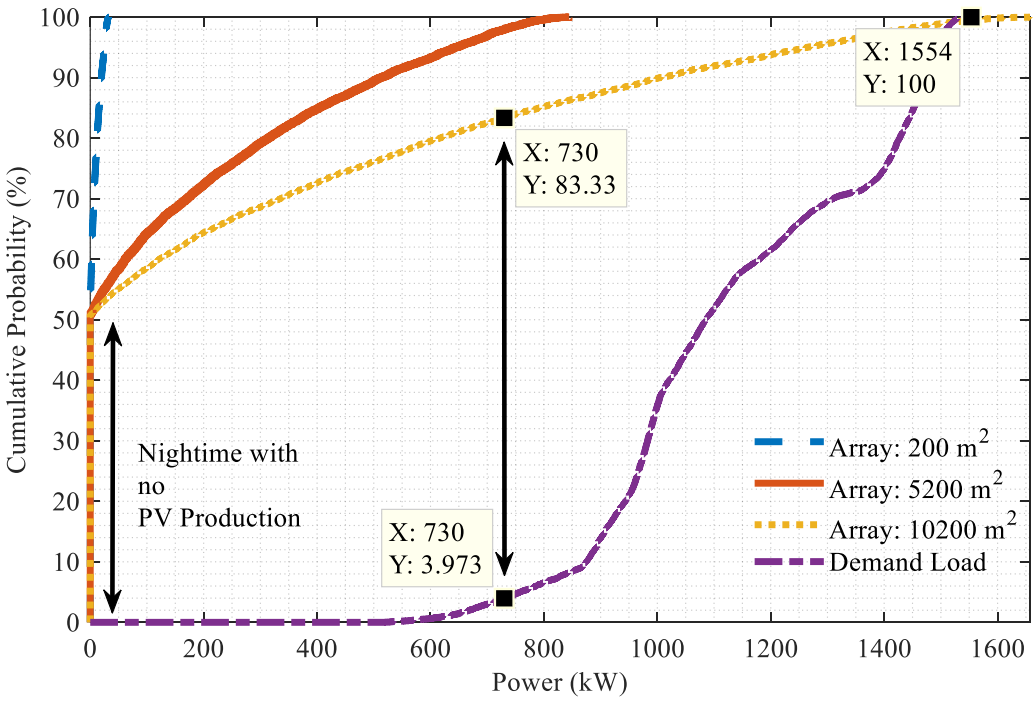

Fig. 3.1 Probability distribution of hourly PV generation and demand power

Fig. 3.1 shows the cumulative probability distributions for both the hourly energy generated by the PV array and the hourly load demand vs. the fraction of hours in a year that experience less PV energy generation or less load demand. As the labeled dots show, the probability that a randomly chosen hourly PV generation from the largest solar array size of 10200 m^2 is less than 730 kW is about 83%, while the

probability that a randomly chosen hourly demand power is less than 730 kW is around 4%. Thus, the hourly energy generated by the three considered solar array sizes can rarely satisfy the hourly load demand. Since there is no sunlight during the night, there is an approximately 50% probability that a randomly chosen PV hourly energy generation is 0 kWh.

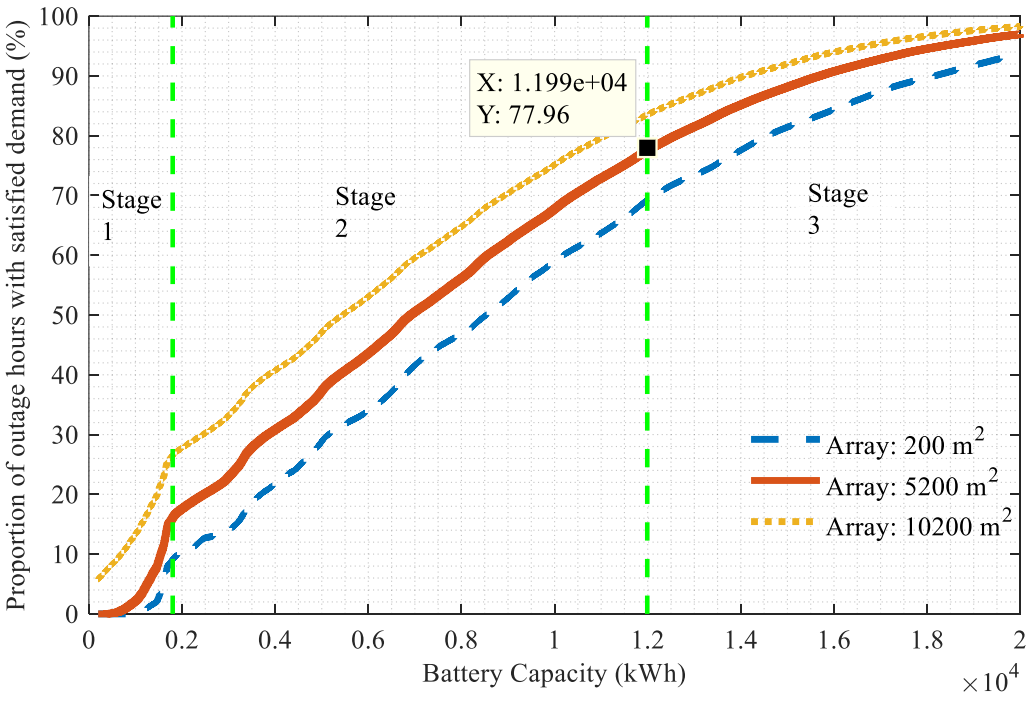

Fig. 3.2 Proportion of outage hours when load demand is satisfied

Fig. 3.2 shows the results of the simulation. For the three considered array sizes, battery capacity is the *x*-axis vs. the proportion of outage hours when load demand is satisfied by the PV + battery system as the *y*-axis. As it can be observed, the increasing trend of each curve can be divided into three different stages. Stage 1 presents the initial sharply increasing trend of the curves. This sharp increase is because the proportion of outage hours when load demand is satisfied increases if a battery is added to a PV system. The larger the battery capacity, the higher the proportion of outage time when load demand is satisfied. According to Fig. 3.1, the maximum hourly load demand for this example facility is 1,554 kW. Thus, if the

battery capacity exceeds 1,554 kWh, the battery alone can provide enough output for a single hour outage; smaller batteries might fail unless the solar array output is large enough to fill the gap for the first hour (and perhaps successive hours). Then comes stage 2, which shows a reasonably stable linear relationship between battery capacity and the proportion of outage hours when load demand is satisfied. When the battery capacity goes beyond the battery capacity corresponding to the lowest total system cost, the growth trend of the curves becomes clearly flatter.

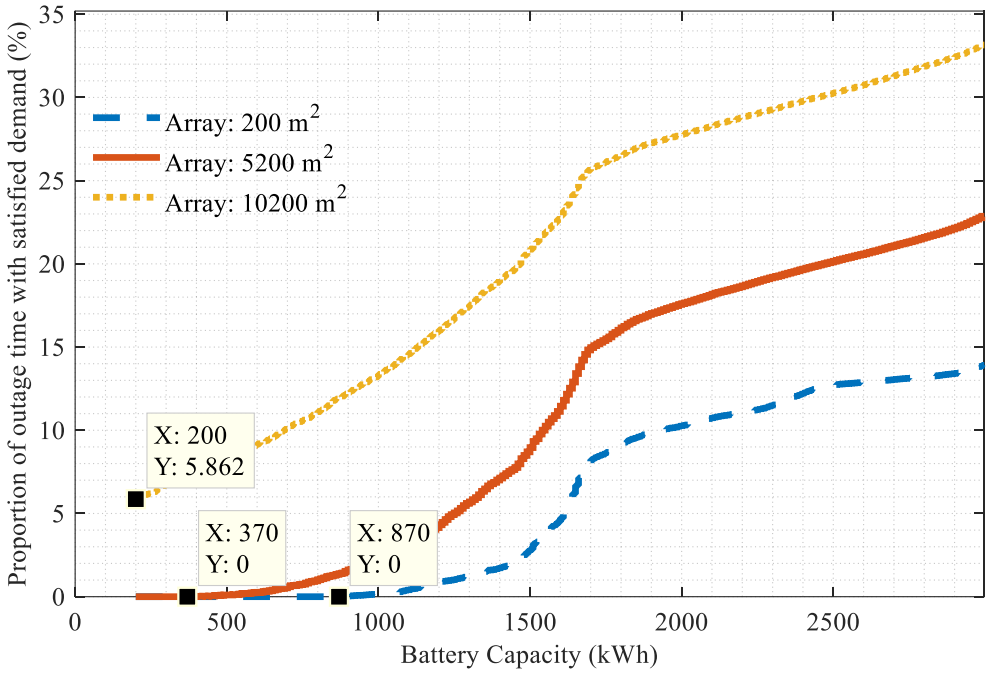

Fig. 3.3 The enlarged view of the initial part of curves in Fig. 3.2

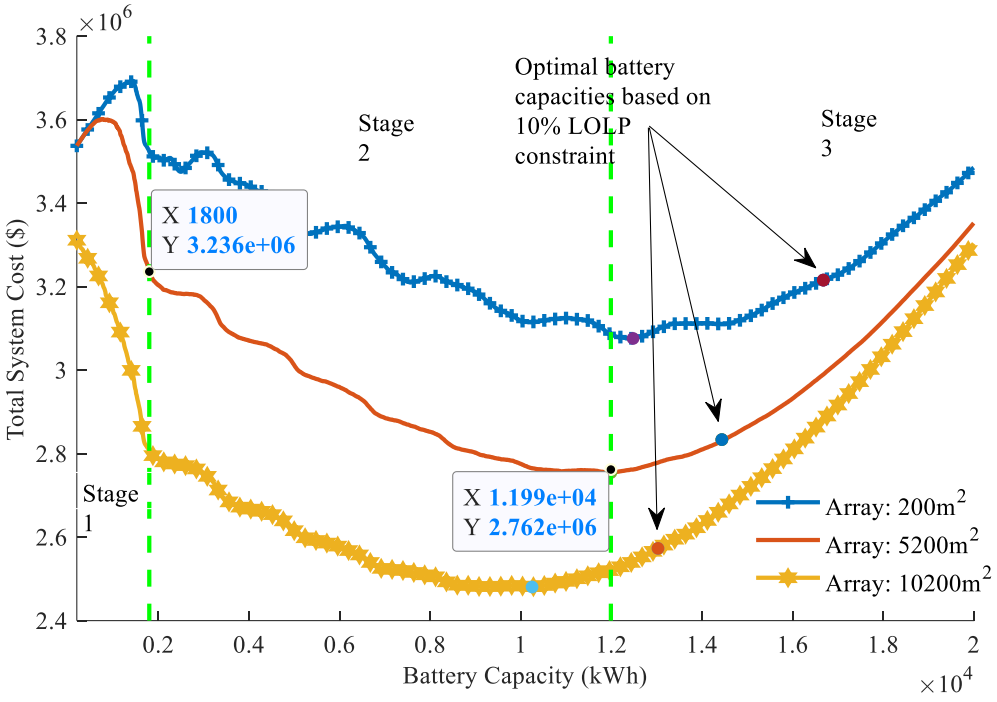

Fig. 3.4 Total system cost for islanding operation of PV + battery systems

In Fig. 3.3, the initial part of Fig. 3.2 can be seen in an enlarged view, highlighting the change in slope after the battery becomes large enough to nearly provide for an average hour's building energy needs; increasing solar array size compensates somewhat for batteries that are not quite big enough. Total system cost is graphed in Fig. 3.4 and also shows the three distinct stages. In stage 1, it can be seen that total system cost for the cases with 200 m^2 and 5200 m^2 array sizes actually increases initially, when a small battery is added to the PV system. It is attributed to the large difference between load demand and PV energy generation as shown in Fig. 3.1. As it can be seen in Fig. 3.3, until battery capacity goes up to around 900 kWh for the array size of 200 m^2 and 400 kWh for the array size of 5200 m^2, the proportion of outage time when load demand is satisfied stays at 0%. Thus, the addition of a small battery does not help to noticeably reduce the unsatisfied load demand, and the increasing investment of battery contributes to a higher total system cost. After battery capacity exceeds these threshold values (as it is shown in Fig. 3.3), the

proportion of outage time with satisfied demand increases as battery capacity increases, and the obtained significant reduction of unsatisfied load demand drives system total cost down sharply. Considering the case with 10200 m^2 array size, the initial total system cost without investment of battery (i.e., only loss of load cost incurred) is much smaller than that of the other two cases, since the PV system with an array size of 10200 m^2 alone can approximately satisfy the load demand of 4% of the outage time. According to Figs. 3.3 and 3.4, the increasing battery capacity continually decreases total system cost in stage 1 for this large-array case.

Considering an array size of 5200 m^2 as an example, after the battery capacity exceeds the threshold value of 1800 kWh, it enters into stage 2, where the decreasing rate of total system cost slows down and there is a linear trend as battery capacity increases. It corresponds to stage 2 indicated in Fig. 3.2, where the proportion of outage time when load demand is satisfied increases linearly with the increase of battery capacity. After the battery capacity exceeds about 12000 kWh, as it is labeled in Fig. 3.4, it comes to stage 3, where the increasing investment cost of adding more battery capacity leads to a higher total system cost instead. This is because increasing battery capacity in stage 3 does not increase the proportion of outage hours when load demand is satisfied as efficiently as it does in stage 2, as it is shown in stage 3 in Fig. 3.2.

Table 3.2 Total system cost of PV + battery system operation in island mode with a combination of battery and array sizing

Total system cost ($\times 10^6$\$)		Array size (m^2)		
		200	5200	10200
Battery Capacity (kWh)	200	3.54	3.54	3.31
	10250	3.11	2.77	**2.48**
	11990	3.09	**2.75**	2.52
	12480	**3.08**	2.76	2.54

Note: Values in bold are the lowest total system cost corresponding to each PV array size.

Table 3.2 presents the total system cost for the three considered array sizes and the battery capacities which correspond to the lowest total system cost for each specific array. The smallest battery tested (200 kWh) is also included as this nearly represents the baseline lost load system cost as a comparison. Based on the results in Table 3.2, it could be seen that the lowest total system cost for a given PV array size decreases as the PV array size increases, and the corresponding battery capacity decreases as well. This is because a larger PV array generates more electricity (subject to the sunlight variability), which contributes to less loss of load cost leading to a lower total system cost and a smaller required battery capacity to address power needs during an outage. All of these minimum cost battery/array configurations yield substantial financial savings compared to the baseline scenario. The optimal battery capacities satisfying 10% *LOLP* constraints are also labeled in Fig. 3.4, which, for the present scenario, are larger than the battery capacities corresponding to the lowest total system cost.

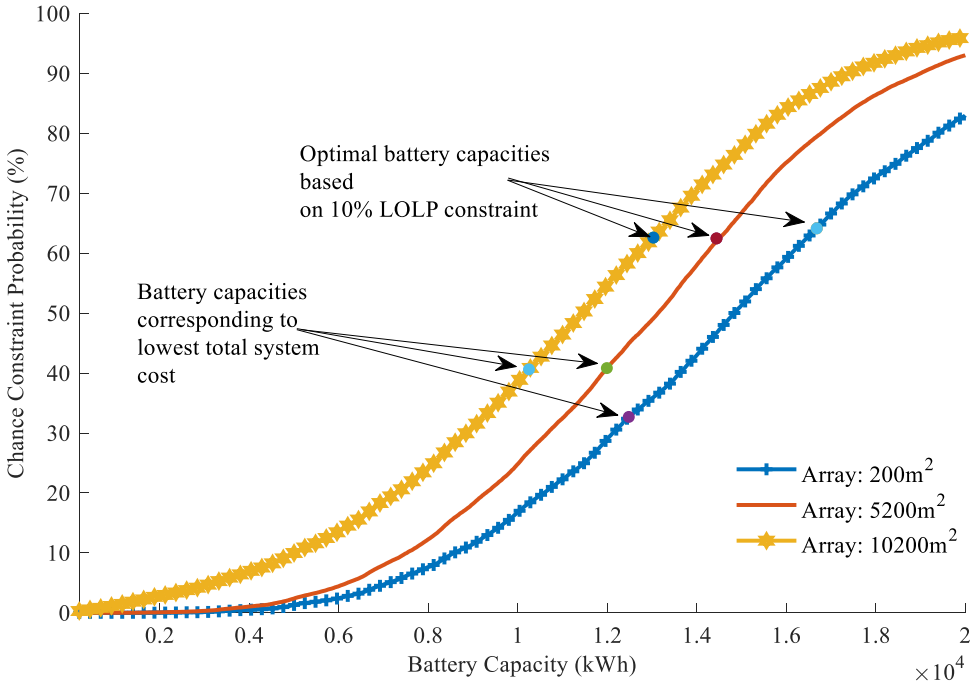

Fig. 3.5 Chance constraint probability for islanding operation of PV + battery system

Table 3.3 Chance constraint probability of PV + battery system islanding operation with different combinations of battery and array sizing

Chance constraint probability (%)		Array size (m^2)		
		200	5200	10200
Battery Capacity (kWh)	10250	18.32	26.90	**40.63**
	11990	28.92	**40.84**	54.65
	12480	**32.69**	44.95	58.57

Note: Values in bold are chance constraint probabilities corresponding to each scenario with the lowest total system cost.

As it was defined in the problem statement, chance constraint probability denotes the proportion of grid outages in which *LOLP* constraint is satisfied by the islanding operation of PV + battery system, i.e., an estimate of the probability that *LOLP* constraint can be met by PV + battery system energy supply during a grid outage.

It can be seen from Fig. 3.5 that the optimal battery capacities which meet the 10% *LOLP* constraint are larger than the battery capacities which achieve the lowest total system cost. Higher chance constraint probabilities, which indicate more reliable islanding energy supply of PV + battery systems, could also be achieved with these optimal battery capacities.

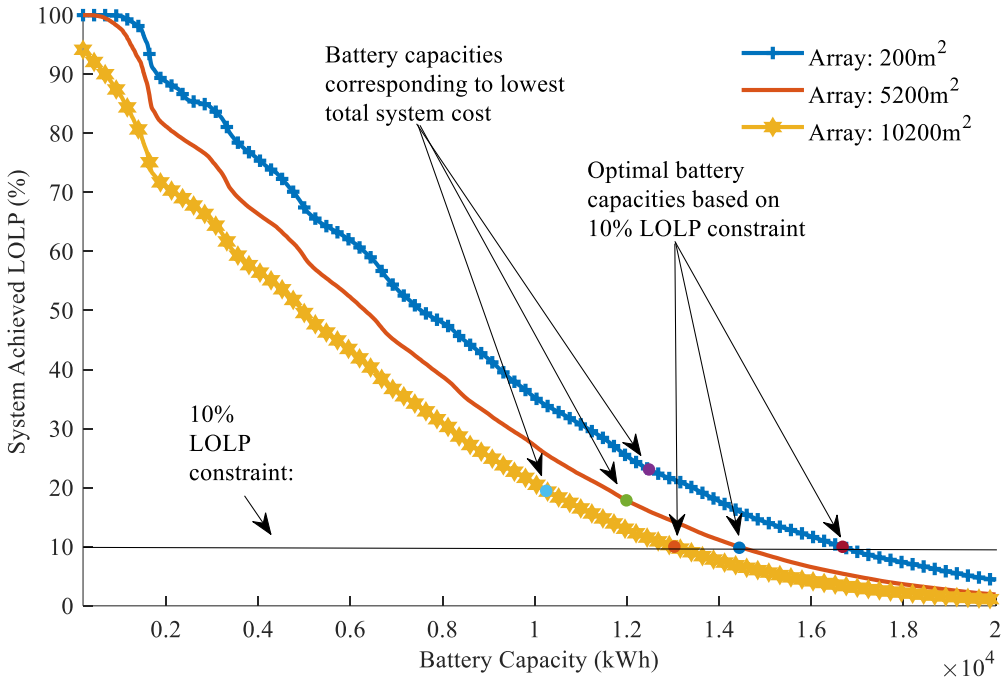

Fig. 3.6 Achieved *LOLP* for islanding operation of PV + battery system

Table 3.4 Achieved *LOLP* of PV + battery system islanding operation with a combination of battery and array sizing

System achieved LOLP (%)		Array size (m^2)		
		200	5200	10200
Battery Capacity (kWh)	10250	33.95	25.67	**19.46**
	11990	25.39	**17.86**	12.95
	12480	**23.09**	16.02	11.44

Note: Values in bold are system achieved *LOLP* corresponding to each scenario with optimal total system cost.

The upper labeled dots in Fig. 3.6 are system achieved *LOLP* corresponding to the lowest total system cost. System achieved *LOLP* denotes the real *LOLP* achieved by islanding mode of PV + battery system during a grid outage. The optimal battery capacities which meet the 10% *LOLP* constraint are also highlighted in the figure, which (in this case) are larger than the battery capacities achieving the lowest total system cost. However, it is worth mentioning here that in another case study conducted by Tsianikas *et al.*, it was proven that it may be feasible to design a PV +

battery system which is less expensive and provides more resilience simultaneously (Tsianikas, Zhou et al. 2019).

A sensitivity analysis has also been conducted in terms of total system cost and the two reliability metrics. The relationship between every 1% decrease of the system achieved *LOLP* (or 1% increase of the chance constraint probability) and the corresponding change of total system cost has been investigated, i.e., how total system cost changes (higher or lower) while the reliability performance of the system is improving. One specific array size is taken as an example. Figs. 3.7 and 3.8 are provided to demonstrate sensitivity analysis results.

Fig. 3.7 shows the results of sensitivity analysis of total system cost vs. system achieved *LOLP*. An example of how Fig. 3.7 should be interpreted is the following; the circle point (13, 12,490) indicates that a positive cost increase of $12,490 is incurred if system achieved *LOLP* decreasing from 13% to 12% is desired because more battery capacity would need to be purchased.

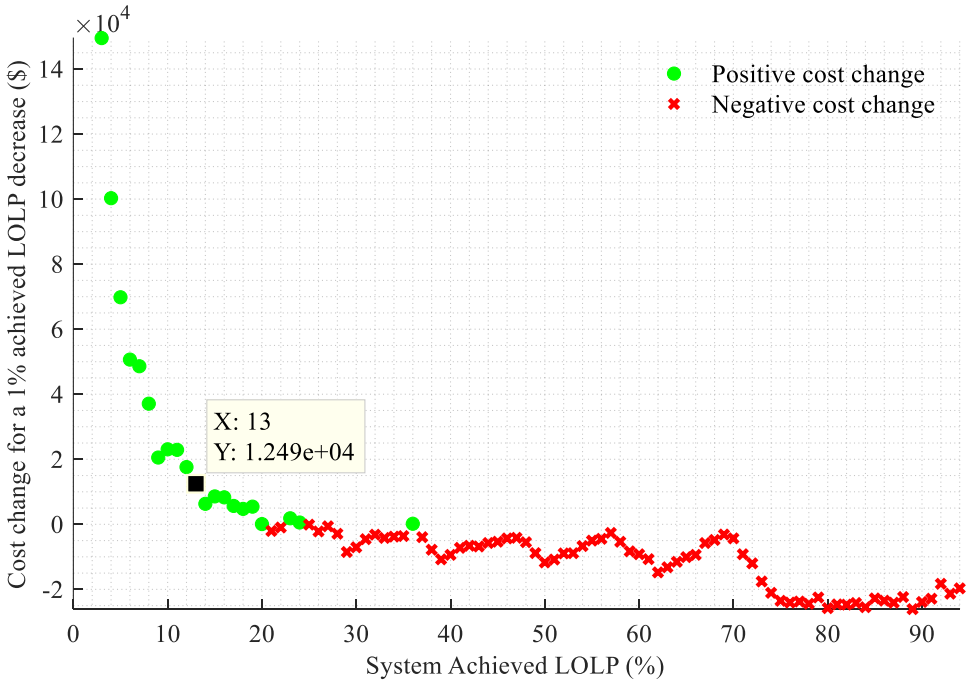

Fig. 3.7 Sensitivity analysis of cost change and achieved *LOLP*

Fig. 3.8a shows the results of sensitivity analysis of total system cost and chance constraint probability. An important characteristic of Figs. 3.7 and 3.8a is the presence of cross points, i.e., points where the system reliability improvement is accompanied by a decrease in total system cost. That is easily justified by simultaneously looking at Figs. 3.8a and 3.8b.

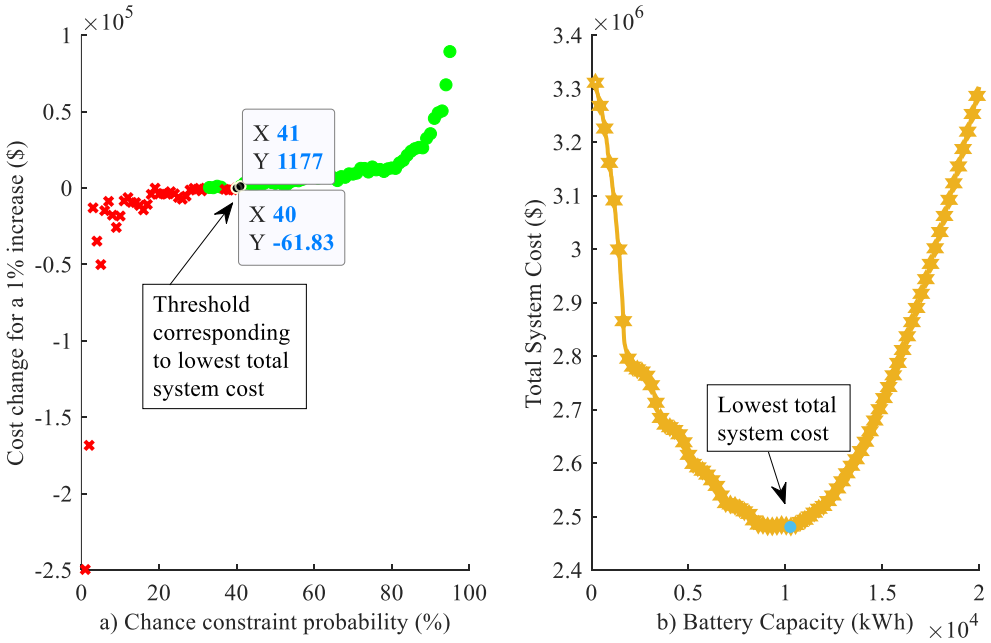

Fig. 3.8 a) Sensitivity analysis of cost change and chance constraint probability, b) Total system cost vs. battery capacity

It can be observed that the lowest total system cost in Fig. 3.8b, which is labeled with an arrow, corresponds to the threshold where cross points end and circle points begin in Fig. 3.8a. It demonstrates the argument that the optimal battery size with respect to system power output performance should be always greater than or equal to this threshold value, given the fact that chance constraint probability is a non-decreasing function of battery capacity. As a result, a battery which is smaller than this threshold value not only causes worse system islanding operation capability but also incurs higher total system cost.

3.1.3 Criticality of several domain-specific parameters

To sum up, in this section of the dissertation a mathematical model is proposed to optimize battery capacity for a PV + battery system starting with a given PV array size. The optimum system is able to meet the load demand of facilities during grid outage for a certain reliability level with minimal cost. The effectiveness of the simulation method is demonstrated by numerical examples using actual data sets of solar irradiation and model facilities' load profiles hourly throughout the year. Overall, the presented methodology is helpful for the future design of grid outage-resilient PV + battery system from both economic and reliability aspects.

At this point, it is considered appropriate to mention one key observation made after completing this work and that is related to the factors that affected the model's results heavily. As was seen, the objective function consisted of two terms: the investment cost and unmet load penalty cost. Therefore, it is safe to assume that the parameters behind these two cost terms, i.e. battery price and *VOLL* are extremely crucial in selecting the optimal battery size for a given PV array. The fact that there is a critical facility (so *VOLL* is high) and also the fact that forecasts predict a sharp decline in battery prices (IRENA 2017) not only strengthen this assumption but also make the need for extensive research in this area more important. This is exactly the research motivation for the work that is presented in Section 3.2.

3.2 Trade-off between investment and unmet load penalty costs

This section investigates the trade-off between two critical factors that could influence or even dictate the rate of adoption of battery systems used to support photovoltaic arrays. The value of lost load (*VOLL*) and battery price greatly influence the economic viability of photovoltaic + battery systems to provide energy resilience during grid outages. The simulation-based optimization method described in Section 3.1 is modified and improved to investigate the effects of *VOLL* and battery price on

the balance between total system cost and system islanding resilience to meet customer demand during a grid outage (Tsianikas, Zhou et al. 2019). One of the great findings of this work is the positive effect that the anticipated decline in battery price could have in enhancing the resilience and effectiveness of renewable energy systems combined with energy storage.

3.2.1 Problem formulation

Before proceeding with the updated problem formulation for this section, it is considered necessary to elaborate more on the approach used to model grid outages. Firstly, N_k is defined as a random set containing all grid outages that occurred in year k, while O_{jk} is defined as a random set of all time intervals of grid outage j in year k. Time, with increments of Δt, is again a discrete parameter. Therefore, t_{ijk} denotes the i^{th} time interval for the j^{th} outage in year k. Finally, the indicator function $\delta(t_{ijk})$ shows whether the demand at the i^{th} time interval for the j^{th} outage in year k is satisfied or not. It takes the value of 1 when the demand is lost for the corresponding time interval, while it is 0 in the opposite case.

Furthermore, it is assumed that the duration of the j^{th} grid outage in year k, T_{jk}, follows a scale adjusted Poisson distribution in which $T_{jk} > 0$. The total grid outage time in year k is $\sum_{j \in N_k} T_{jk}$. K denotes the planning time horizon in years, so the cumulative outage time during the planning horizon is $\sum_{k=1}^{K} \sum_{j \in N_k} T_{jk}$. Based on the definition of $\delta(t_{ijk})$, the outage time with unmet load demand in year k can now be defined as $\sum_{j \in N_k} \sum_{i \in O_{jk}} \delta\left(t_{ijk}\right) \Delta t$, while the cumulative outage time with unmet load demand can be defined as $\sum_{k=1}^{K} \sum_{j \in N_k} \sum_{i \in O_{jk}} \delta\left(t_{ijk}\right) \Delta t$. Similarly, the unmet load demand in

year k is $\sum_{j\in N_k}\sum_{i\in O_{jk}} \delta(t_{ijk})\, D(t_{ijk})\,\Delta t$ and the cumulative unmet load can be calculated as $\sum_{k=1}^{K}\sum_{j\in N_k}\sum_{i\in O_{jk}} \delta(t_{ijk})\, D(t_{ijk})\,\Delta t$. Meanwhile, the starting time of each simulated grid outage is assumed to follow uniform distribution throughout the year.

In the proposed model, there are constraints for energy balance, energy generation and the reliability metrics, i.e., *LOLP* (Loss of Load Probability) and *CCP* (Chance Constraint Probability). The updated formulation is shown as follows (Tsianikas, Zhou et al. 2019):

$$\min \quad B_r b + \sum_{k=1}^{K} VOLL\,(1+\gamma)^{-k} \sum_{j\in N_k}\sum_{i\in O_{jk}} \delta(t_{ijk})\, D(t_{ijk})\,\Delta t$$

$$\text{s.t. } P(t_{ijk}) = \eta\, I(t_{ijk})\, A,\ \forall i \in O_{jk},\ j \in N_k,\ k \in \{1,2,3,...,K\} \tag{3.6}$$

$$B_{\min} = B_r\,(1 - DoD) \tag{3.7}$$

$$Q_B(t_{1jk}) = I_c\, B_r,\ \forall j \in N_k,\ k \in \{1,2,3,...,K\} \tag{3.8}$$

$$Q_B(t_{(l+1)jk}) = \begin{cases} \min\left\{Q_B(t_{ljk}) + P(t_{ljk})\, e,\ B_r\right\}, & \text{if } Q_B(t_{ljk}) + \int_{t_{ljk}}^{t_{ljk}+\Delta t}\left(P(u) - C_p D(u)\right) e\, du < B_{\min} \\ Q_B(t_{ljk}) + \int_{t_{ljk}}^{t_{ljk}+\Delta t}\left(P(u) - C_p D(u)\right) e\, du, & \\ \qquad \text{if } B_{\min} \le Q_B(t_{ljk}) + \int_{t_{ljk}}^{t_{ljk}+\Delta t}\left(P(u) - C_p D(u)\right) e\, du \le B_r & \\ B_r, \qquad \text{if } Q_B(t_{ljk}) + \int_{t_{ljk}}^{t_{ljk}+\Delta t}\left(P(u) - C_p D(u)\right) e\, du > B_r & \end{cases}$$

$$\forall l \in O'_{jk},\ j \in N_k,\ k \in \{1,2,3,...,K\} \text{ where } O'_{jk} = O_{jk} - \{\max O_{jk}\} \tag{3.9}$$

$$\delta(t_{ijk}) = \begin{cases} 1, & \text{if } Q_B(t_{ljk}) + \int_{t_{ljk}}^{t_{ljk}+\Delta t}\left(P(u) - C_p D(u)\right) e\, du < B_{\min} \\ 0, & \text{otherwise} \end{cases} \tag{3.10}$$

$$\forall i \in O_{jk},\ j \in N_k,\ k \in \{1,2,3,...,K\}$$

$$LOLP_{jk} = \frac{\sum_{i \in O_{jk}} \delta(t_{ijk})\, \Delta t}{T_{jk}}$$

$$T_{jk} = \left| O_{jk} \right| \Delta t \tag{3.11}$$

$$\frac{\sum_{k=1}^{K} \sum_{j \in N_k} LOLP_{jk}}{\sum_{k=1}^{K} \left| N_k \right|} \le \alpha$$

$$\forall\, j \in N_k,\ k \in \{1, 2, 3, ..., K\}$$

$$\Pr\left\{ LOLP_{jk} \le \beta \right\} = \frac{\sum_{k=1}^{K} \sum_{j \in N_k} \lambda_{jk}}{\sum_{k=1}^{K} \left| N_k \right|} \ge \zeta, \quad \text{where} \quad \lambda_{jk} = \begin{cases} 1, & \text{if } LOLP_{jk} \le \beta \\ 0, & \text{otherwise} \end{cases} \tag{3.12}$$

$$\forall\, j \in N_k,\ k \in \{1, 2, 3, ..., K\}$$

$$B_{\min} \ge 0, \quad B_r \ge 0, \quad 0 \le \alpha,\ \beta,\ \zeta \le 1$$

The objective function consists again of two different terms, in a similar fashion with the formulation in Section 3.1. The first term, $B_r b$, refers to the battery investment cost. The second term, $\sum_{k=1}^{K} VOLL\,(1+\gamma)^{-k} \sum_{j \in N_k} \sum_{i \in O_{jk}} \delta(t_{ijk})\, D(t_{ijk})\, \Delta t$, refers to the cost of lost load demand of customers during grid outages. These two terms of the objective function compose total system cost (*TSC*), which needs to be minimized. The objective function is stochastic due to the random sets N_k and O_{jk} involved in the second term. As mentioned earlier, N_k is the random set of grid outages for year k and O_{jk} is the random set of time intervals of outage j for year k. According to the modeling procedures described before, the second term in the objective function can be obtained as the expected cost of lost load demand of customers during grid outages.

Eq. (3.6) determines the energy generated by the PV array at time t_{ijk}. Eq. (3.7)

denotes the calculation of $B_{\min}$ based on *DoD*. Eq. (3.8) defines the amount of energy stored in the battery at the beginning of each outage, where I_c is used to decide the initial charging state of the battery. I_c depends on several realistic factors, for example, self-discharge of a battery (%/month), which is an important feature of a battery. As an example, for the case of lead-acid battery, it is equal to 2-5%/month, while for Li-ion battery, self-discharge is 1%/month (Divya and Ostergaard 2009). Eq. (3.9) focuses on the energy balance of PV + battery system during grid outages. When a PV + battery system islands during an outage, non-critical load demand of facilities may be shed, so the PV + battery system only needs to meet the proportion of critical load demand, which is denoted by C_p (Alsaidan, Khodaei et al. 2018). Eq. (3.10) presents the condition under which the indicator function $\delta(t_{ijk})$ equals 1 or 0, which determines whether the load demand at time t_{ijk} is lost or not. Eq. (3.11) indicates the reliability of the islanding generation of the PV + battery system, with α as the upper bound for the expected value of *LOLP*. $LOLP_{jk}$ is defined as the proportion of time when load demand of customers cannot be met by PV + battery system during the j^{th} grid outage for year *k*. The definition of *CCP* is given in Eq. (3.12), which incorporates the *LOLP* constraint.

The simulations here are conducted under different combinations of battery price and *VOLL* in order to investigate the interplay between them. Different battery capacities are evaluated from both economic and reliability aspects according to the simulated grid outages. To obtain the expected values of stochastic functions used in the analyses, it is necessary to introduce some new notation. *S* denotes the total number of simulated grid outages. G_s is the random simulated set of time intervals for the s^{th} simulated outage and t_{is} is the i^{th} time interval of the s^{th} simulated outage. C_s and $LOLP_s$ are the cost of lost load and loss of load probability corresponding to the

s^{th} simulated outage. Note that in the simulation model $\Delta t = 1$, but it could be any other incremental value. The estimated expected values of the cost of lost load, *LOLP* and *CCP* regarding the simulated grid outages are presented as follows:

$$\hat{E}[C] = VOLL \frac{\sum_{s=1}^{S} C_s}{S}, \quad \text{where } C_s = \sum_{i \in G_s} \delta(t_{is})\, D(t_{is}) \tag{3.13}$$

$$\hat{E}[LOLP] = \frac{\sum_{s=1}^{S} LOLP_s}{S}, \quad \text{where } LOLP_s = \frac{\sum_{i \in G_s} \delta(t_{is})}{|G_s|} \tag{3.14}$$

$$\hat{\Pr}\{LOLP_{jk} \le \beta\} = \frac{\sum_{s=1}^{S} \lambda_s}{S}, \quad \text{where } \lambda_s = \begin{cases} 1, & \text{if } LOLP_s \le \beta \\ 0, & \text{otherwise} \end{cases} \tag{3.15}$$

In other important assumptions, the efficiency of PV array conversion in the simulation, η, is again assumed to be 16%. The efficiency of battery charging/discharging process, *e*, is assumed to be 85% and the maximum depth-of-discharge of the battery (DoD), is 70%. Because of the intended use of battery storage as an energy backup system for facility electrification, it should be classified as an energy-oriented and long-duration battery storage system (Padilla 2018). For all the above reasons, charging and/or discharging the battery is not allowed during normal grid operation, which apparently leads to $I_c = 1$, i.e., the battery is considered fully charged at the starting time of each outage. Because of the criticality of a hospital as a test case, C_p is set to be 80%, which determines the proportion of critical load demand of hospital that needs to be met during grid outages. CAIDI and SAIFI are again 8.23 hours/interruption and 1.17 interruptions/year for the Long Island, NY area (Service 2018), where the actual historical solar irradiation data and the hospital load profile are selected (NREL 2013) (NREL 2016). Before proceeding with the test case, it should be mentioned that the main research contribution of this work is to explore,

quantify and visualize the effects and interaction of two important cost-related factors in the optimal configuration of a backup energy system for critical facilities.

3.2.2 Case study and results

In order to consider price fluctuations and variations, it is now assumed that battery price is varying within the range of 20 $/kWh to 220 $/kWh in the simulation. In addition, *VOLL* is varying within the selected 90% confidence level (*CL*) range of 5 $/kWh to 25 $/kWh, which is based on the expectation of *VOLL* level in 2030 (van der Welle and van der Zwaan 2007).

It is needed now to explain why it is the author's belief that a hospital is an excellent and insightful test case to serve as the facility to implement the proposed mathematical models. During the design of energy backup systems, the criticality of buildings is one of the most significant independent variables because it influences the strictness of the system modeler. Therefore, the high-reliability standards required for the operation of a hospital and the high penalty costs for unmet load demand make this specific facility a good test case. Nowadays, the use of electronic equipment and machinery in healthcare is abundant and requires uninterrupted operation with no exceptions at all. Some examples of this type of equipment currently used in hospitals include operating room machinery, life support, nurse calls, blood storage, fire alarm, mass notification, etc. (Padilla 2018). Therefore, robust design of such redundant systems should be made with reliability as the first and top criterion.

For the purposes of this research, a 2-way factorial experimental design is used and analyzed, with battery price and *VOLL* being the two factors. In total, 11 distinct *VOLL* values and 11 distinct battery prices are considered, which means that an 11x11 factorial design with a total of 121 treatments is applied. The measurements are the battery capacities required for the optimal *TSC*, the actual optimal *TSC*, the

expected *LOLP*, and the *CCP*, which are obtained from 1,000 independent simulation runs for each treatment. The photovoltaic array size considered in this analysis is 6,000 m^2. The problem is analyzed under two different scenarios, which are defined as follows:

1) The *unconstrained* case, in which the optimal *TSC* corresponds to the minimum *TSC* of the PV + battery system, without imposing any reliability criteria or constraints.

2) The *constrained* case, in which the optimal *TSC* is the minimum *TSC* of the PV + battery system which also satisfies all the reliability criteria of the problem.

The distinction between the two scenarios is noticeable. The battery sizes contributing to the unconstrained optimal *TSC* are different than that contributing to the constrained optimal *TSC* unless the former ones are also able to satisfy the reliability constraints. The results are presented in both contexts in the following paragraphs.

Figs. 3.9 to 3.16 are 3D figures (presented in subfigures (a)) and contour figures (presented in subfigures (b)) of the experiment results regarding the two factors, i.e., battery price and *VOLL*, being indicated in the horizontal axis and vertical axis, respectively. In the third axis, the presented variables are the battery size, *TSC*, the expected *LOLP*, and the *CCP*, respectively.

1) Unconstrained case

The results shown in Figs. 3.9 to 3.12 correspond to the unconstrained case where the optimal *TSC* is achieved. Battery capacities that contribute to the unconstrained optimal *TSC* under different combinations of battery price and *VOLL* are shown in Fig. 3.9a) (3D figure), and Fig. 3.9b) (2D figure with contours). In Figs.

3.10 to 3.12, the horizontal and the vertical axes are the same with Fig. 3.9, but the third axis is the *TSC*, the expected *LOLP*, and the *CCP*, respectively.

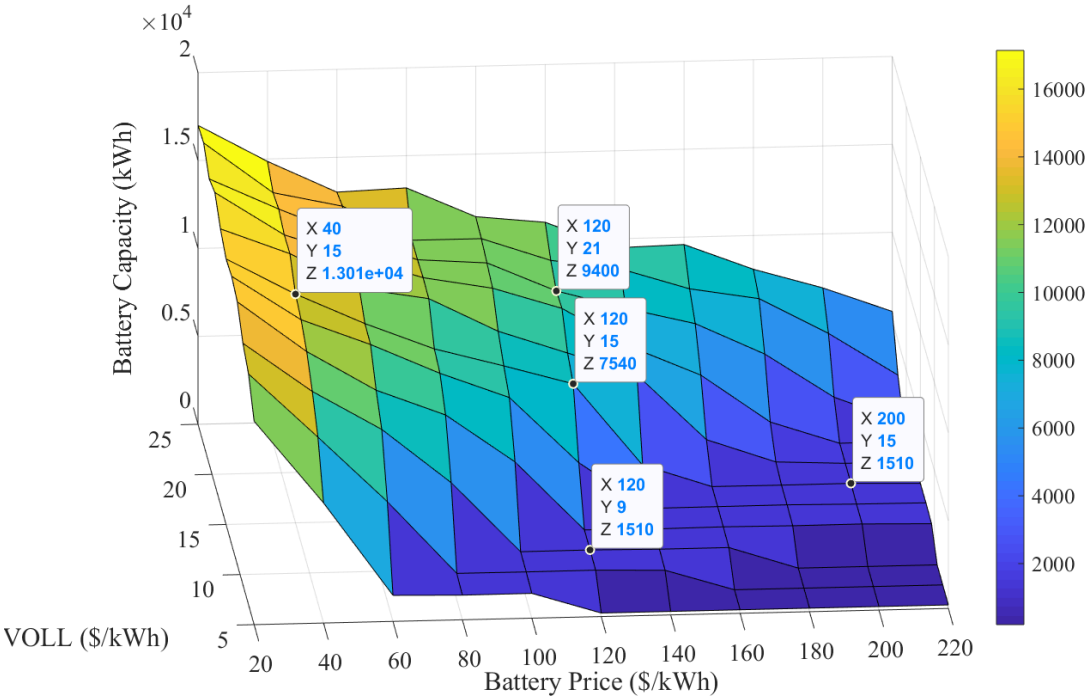

Fig. 3.9a). Battery capacity for the unconstrained scenario as a function of battery price and *VOLL*

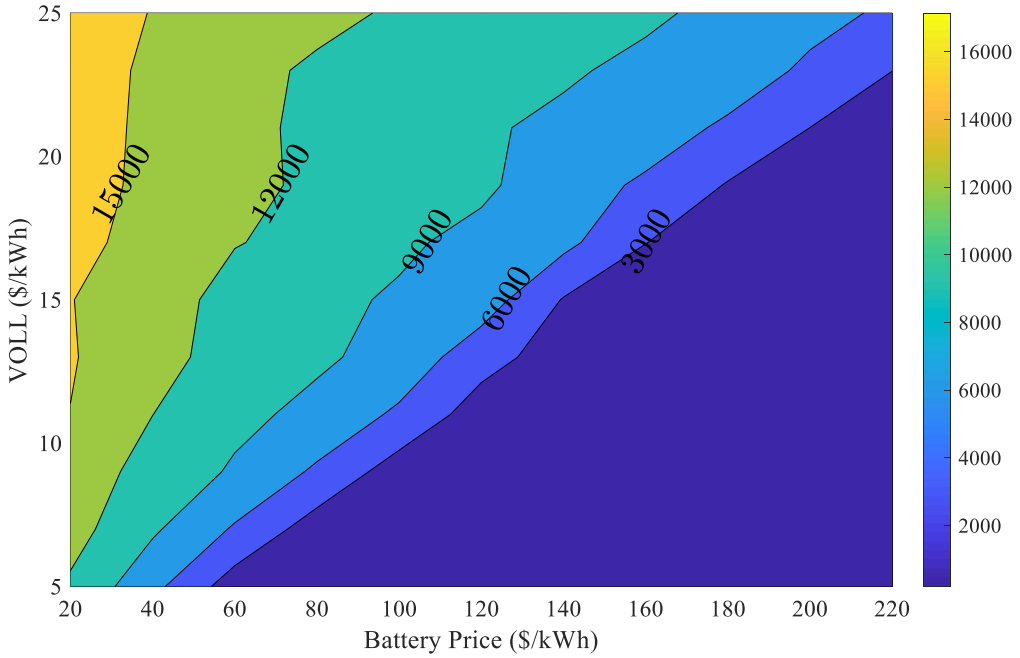

Fig. 3.9b). Contours corresponding to **Fig. 3.9a)**

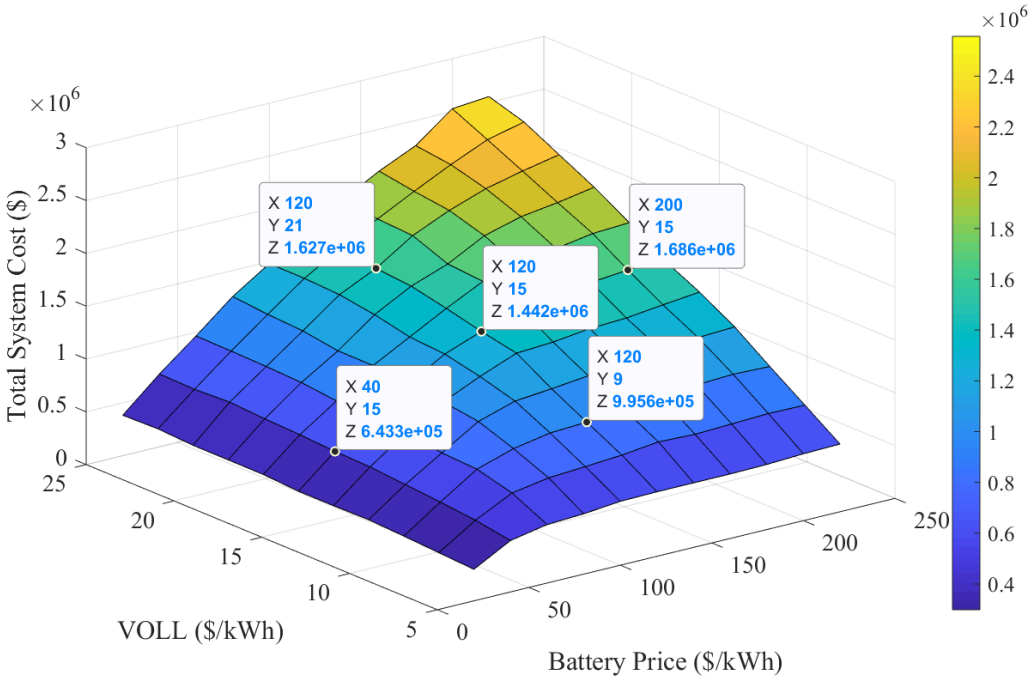

Fig. 3.10a). Optimal *TSC* for the unconstrained scenario as a function of battery price and *VOLL*

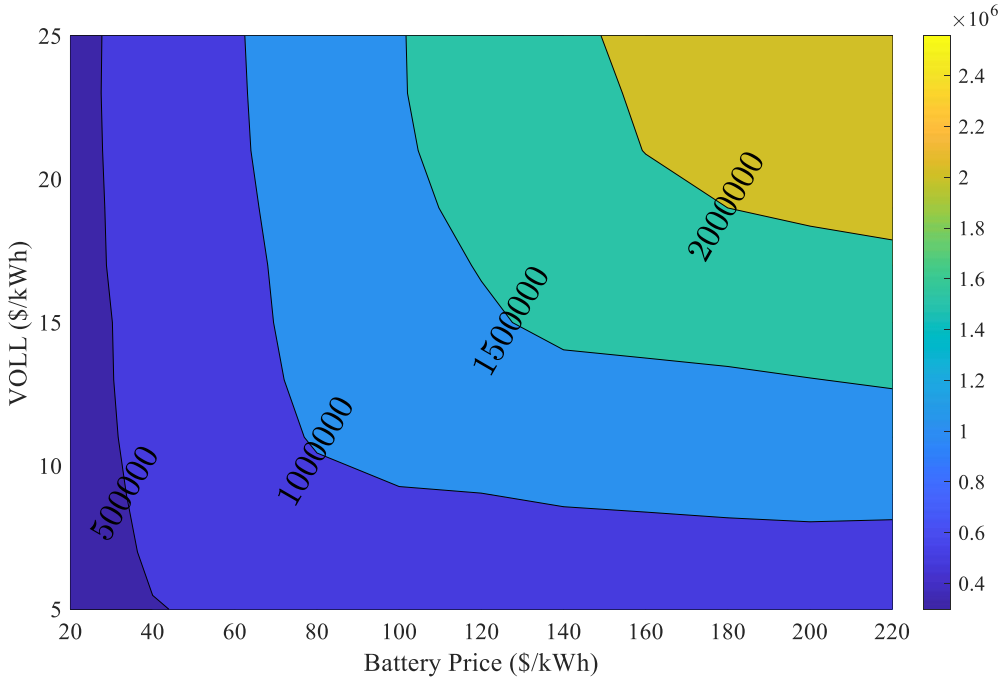

Fig. 3.10b). Contours corresponding to **Fig. 3.10a)**

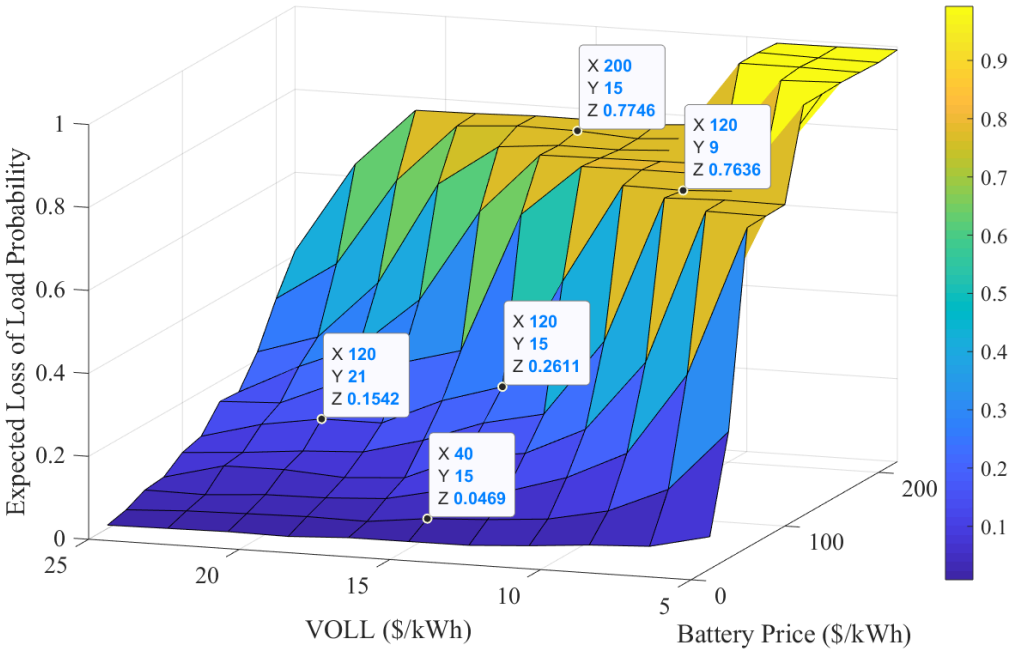

Fig. 3.11a). Expected *LOLP* for the unconstrained scenario as a function of battery price and *VOLL*

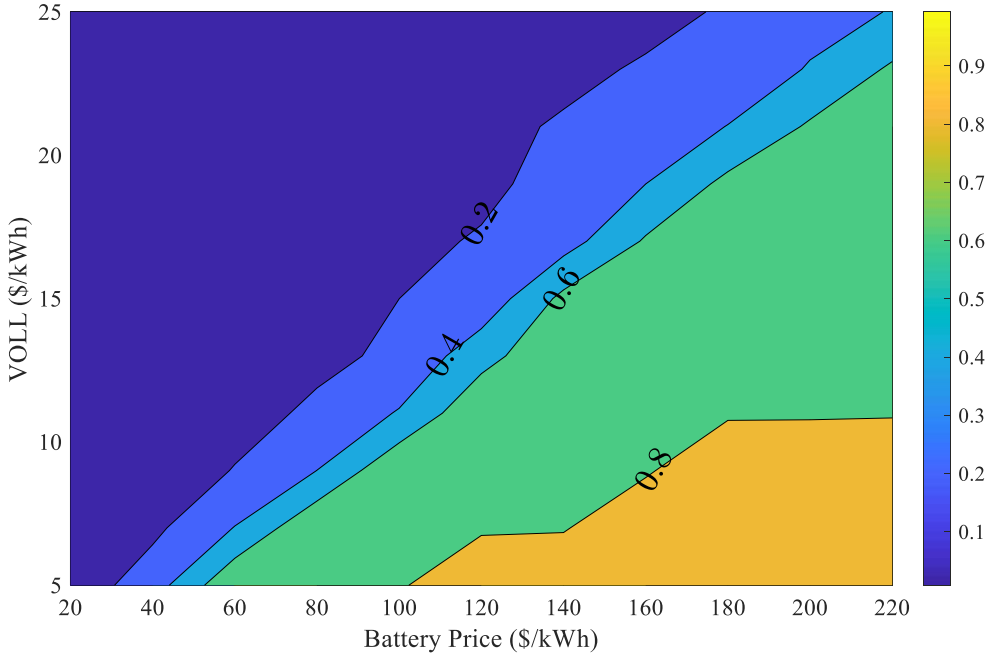

Fig. 3.11b). Contours corresponding to **Fig. 3.11a)**

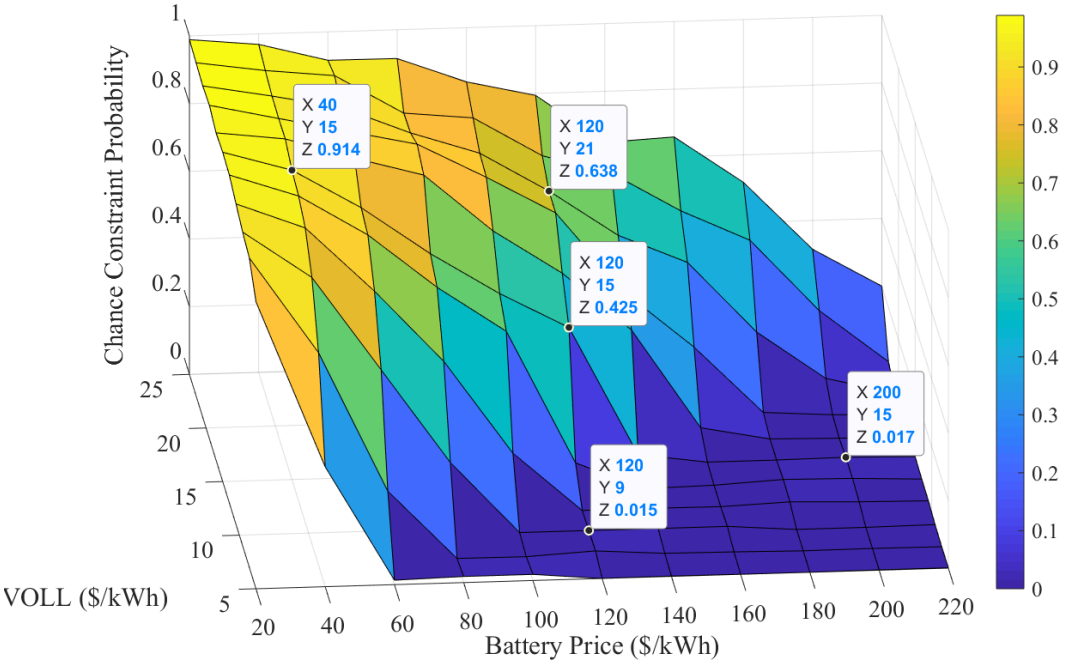

Fig. 3.12a). *CCP* for the unconstrained scenario as a function of battery price and *VOLL*

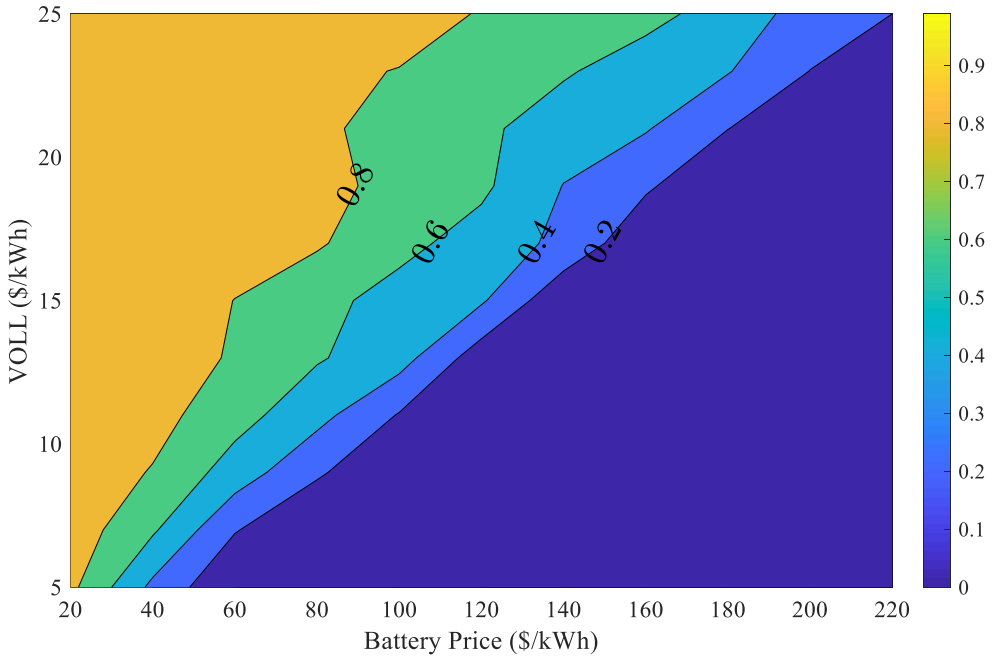

Fig. 3.12b). Contours corresponding to **Fig. 3.12a)**

An important observation from Fig. 3.9a) is that the battery capacities, which are required for the optimal *TSC*, increase when battery price decreases or when *VOLL* increases. As it can be observed, the required optimal *TSC* increases when battery price decreases from 200 $/kWh to 40 $/kWh with the same *VOLL* of 15

$/kWh or when *VOLL* increases from 9 $/kWh to 21 $/kWh with the same battery price of 120 $/kWh. These results come in agreement with the previous findings in Section 3.1 that lower battery price leads to a larger required battery to achieve the optimal *TSC*. When Fig. 3.9a) is analyzed together with Fig. 3.9b), a big plateau is noticeable in the region of low values of *VOLL* and high values of battery price. This plateau means that, if it is too expensive to invest in a battery and simultaneously affordable to allow demand lost during outages, then the actual batteries that correspond to the unconstrained optimal *TSC* are the ones with the lowest sizes. Moreover, as it can be seen from Fig. 3.9b), when *VOLL* decreases, the impact of a change of battery price on the required batteries for the optimal *TSC* becomes larger.

In Fig. 3.10a) it can also be observed that the optimal *TSC* increases when the battery price and/or *VOLL* increases, which is actually what would be expected. However, an important feature of Fig. 3.10b) is that in the upper-left part, the change rate of the optimal *TSC* with respect to battery price is much greater than that of the optimal *TSC* with respect to *VOLL*. Noticing the lower-right part of Fig. 3.10b), the change rate of the optimal *TSC* with respect to battery price becomes smaller than that of the optimal *TSC* with respect to *VOLL*. This phenomenon can be explained by observing Fig. 3.9a) and Fig. 3.9b). When it is economical to invest in batteries, the optimal battery capacities are close to the upper limit required by the facility load profile and outage statistics and are independent of the *VOLL*. It leads to a similar investment cost and outage cost, which is not the case when battery investment costs become higher.

According to Fig. 3.11a), the expected *LOLP* corresponding to the optimal *TSC* decreases only when battery price decreases or *VOLL* increases. On the contrary, it can be observed from Fig. 3.12a) that the *CCP* shows the opposite behavior.

Nevertheless, the most interesting feature that should be highlighted is the two plateaus that exist in Figs. 3.11a) and 3.11b). The first one, which occupies the high *VOLL* and low battery price area, indicates that it is cost-effective to invest in bigger batteries to achieve a very low expected *LOLP*. The other plateau, found in the low *VOLL* and high battery price area, shows the opposite results. If the penalty cost of lost demand is low while the batteries are expensive, it is preferable to endure lost demand during outages instead of investing more in battery storage. These two plateaus can also be observed in Figs. 3.12a) and 3.12b) but in different and opposite regions. The *CCP* is very high when it is economical to invest in battery storage and simultaneously expensive to afford the lost demand during outages, while the *CCP* becomes very low when the opposite conditions are considered.

2) Constrained case

The results shown in Figs. 3.13 to 3.16 correspond to the case where the optimal *TSC* is achieved while adhering to the predetermined reliability constraints. Battery capacities that contribute to the constrained optimal *TSC* with different combinations of battery price and *VOLL* are shown in Fig. 3.13a) (3D figure) and Fig. 3.13b) (2D figure with contours). In Figs. 3.14 to 3.16, the horizontal and the vertical axes are the same as Fig. 3.13, but the third axis of these figures is the constrained optimal *TSC*, the expected *LOLP*, and the *CCP*, respectively.

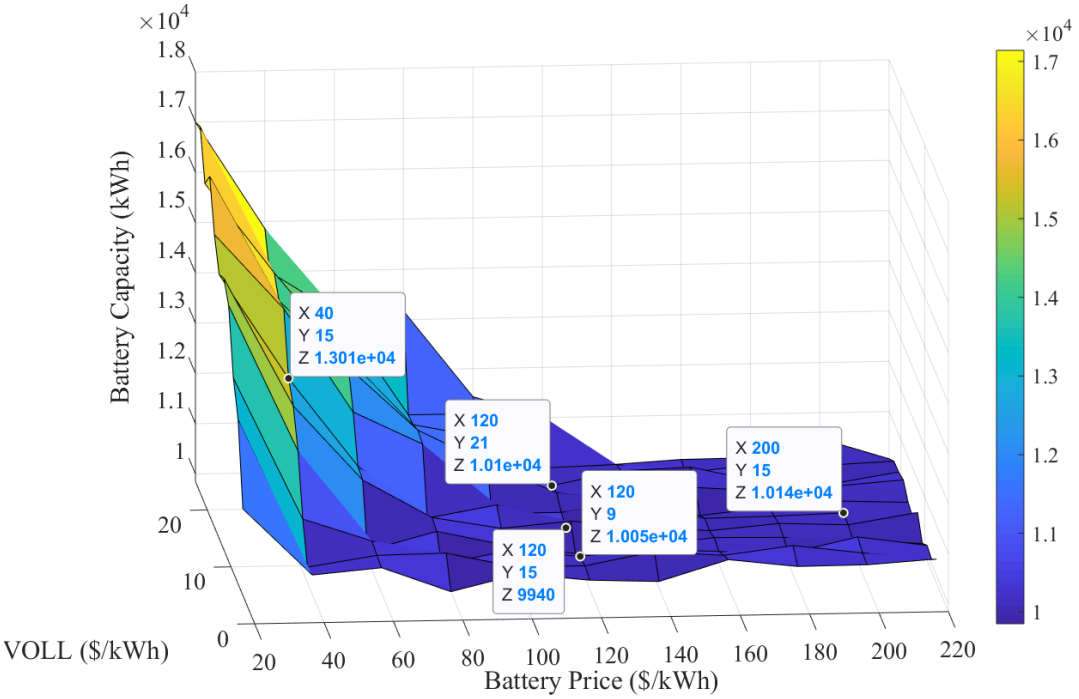

Fig. 3.13a) Battery capacity for the constrained scenario as a function of battery price and *VOLL*

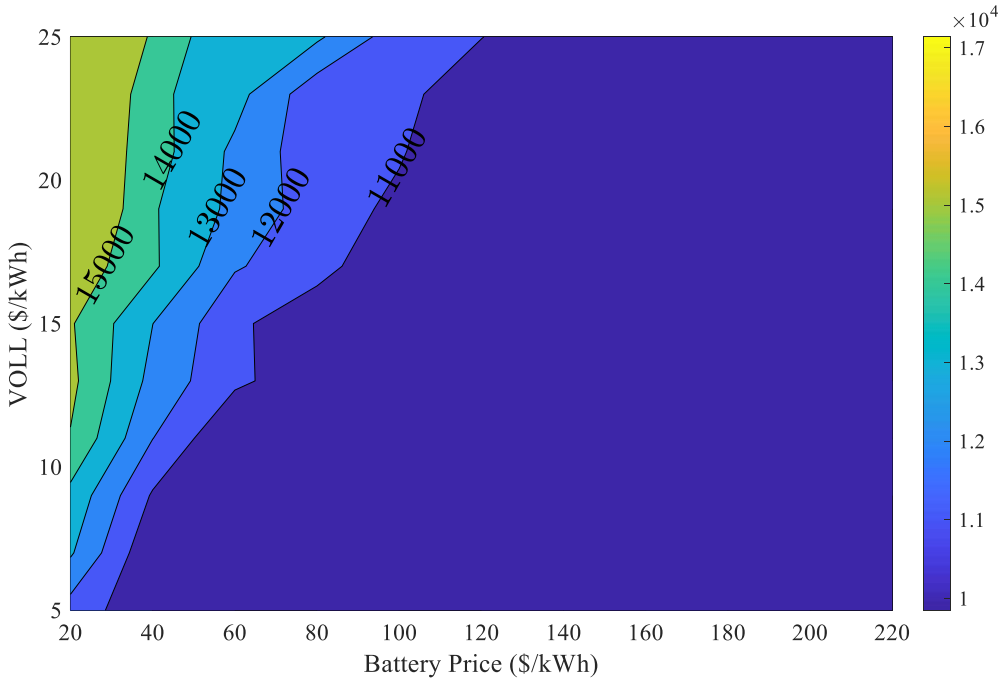

Fig. 3.13b). Contours corresponding to **Fig. 3.13a)**

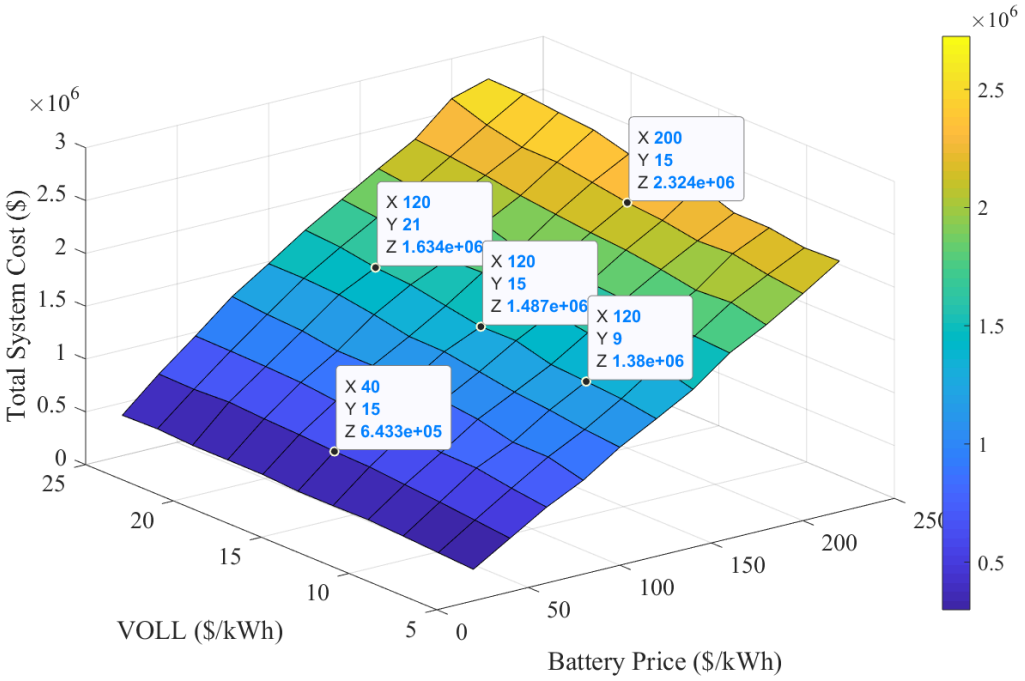

Fig. 3.14a) Optimal *TSC* for the constrained scenario as a function of battery price and *VOLL*

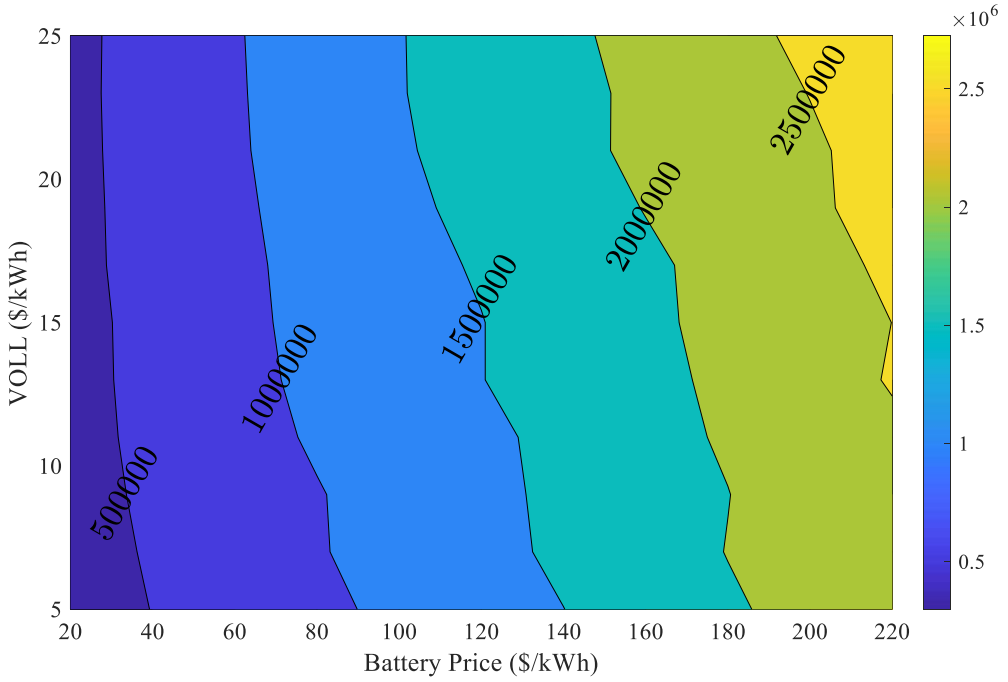

Fig. 3.14b). Contours corresponding to **Fig. 3.14a)**

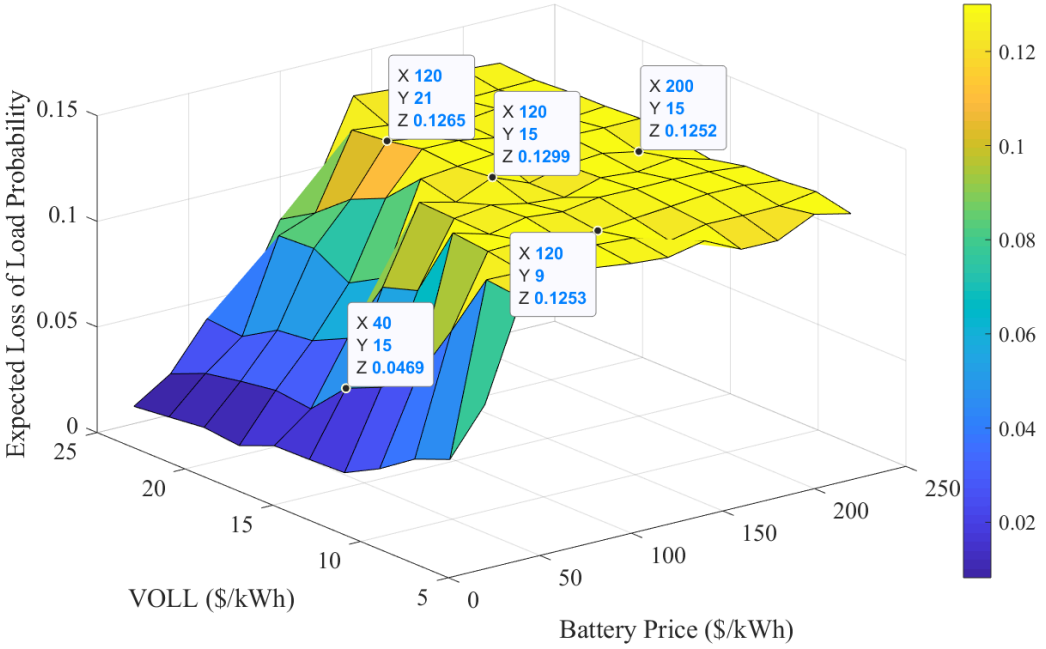

Fig. 3.15a) Expected *LOLP* for the constrained scenario as a function of battery price and *VOLL*

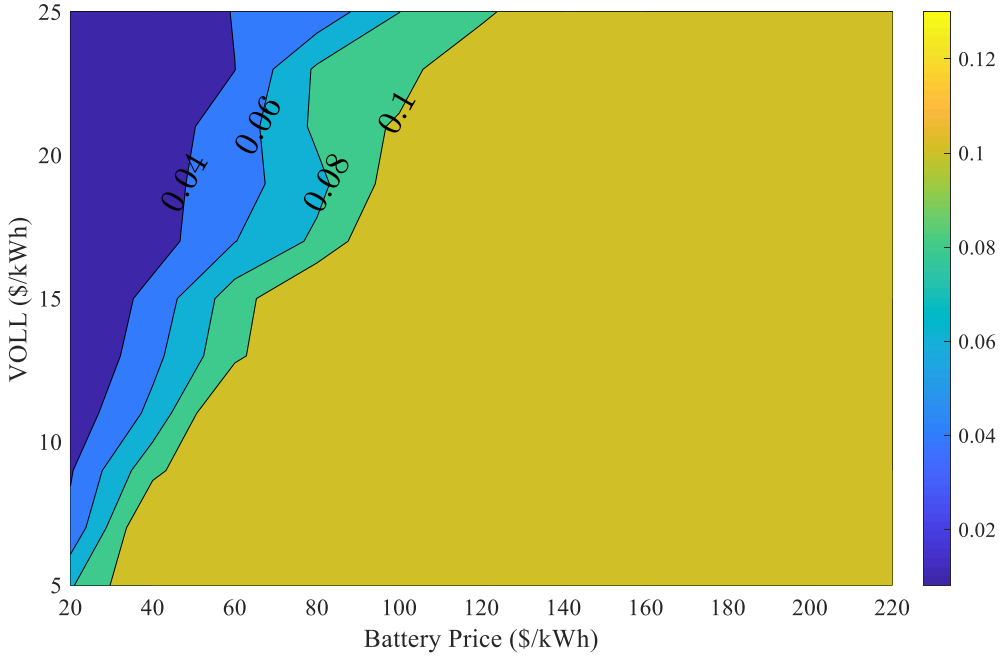

Fig. 3.15b). Contours corresponding to **Fig.3.15a)**

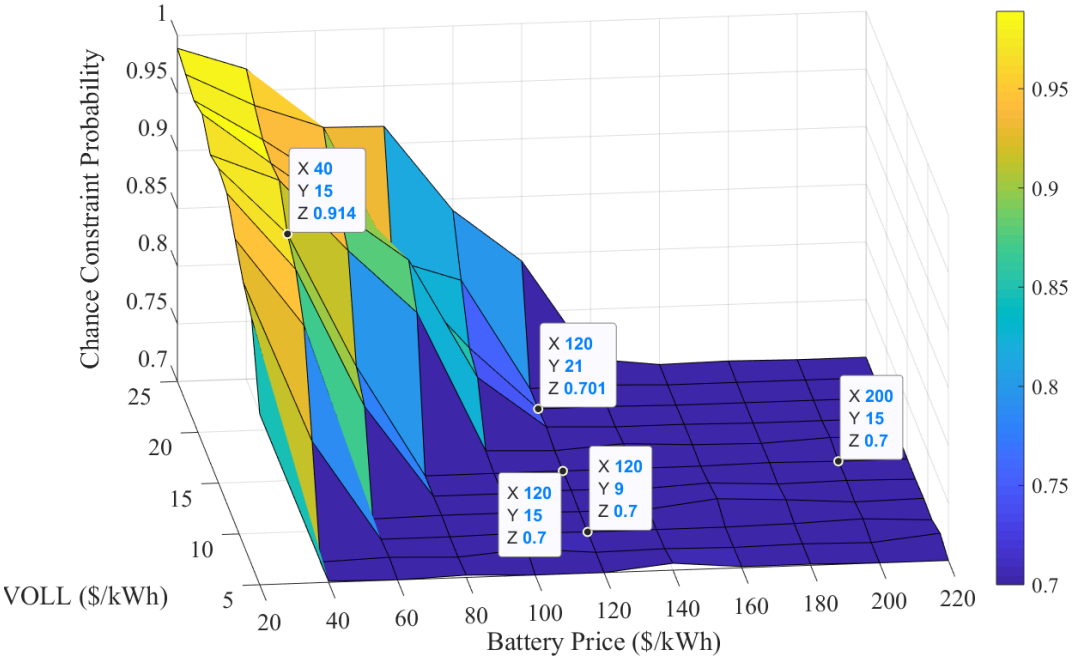

Fig. 3.16a). *CCP* for the constrained scenario as a function of battery price and *VOLL*

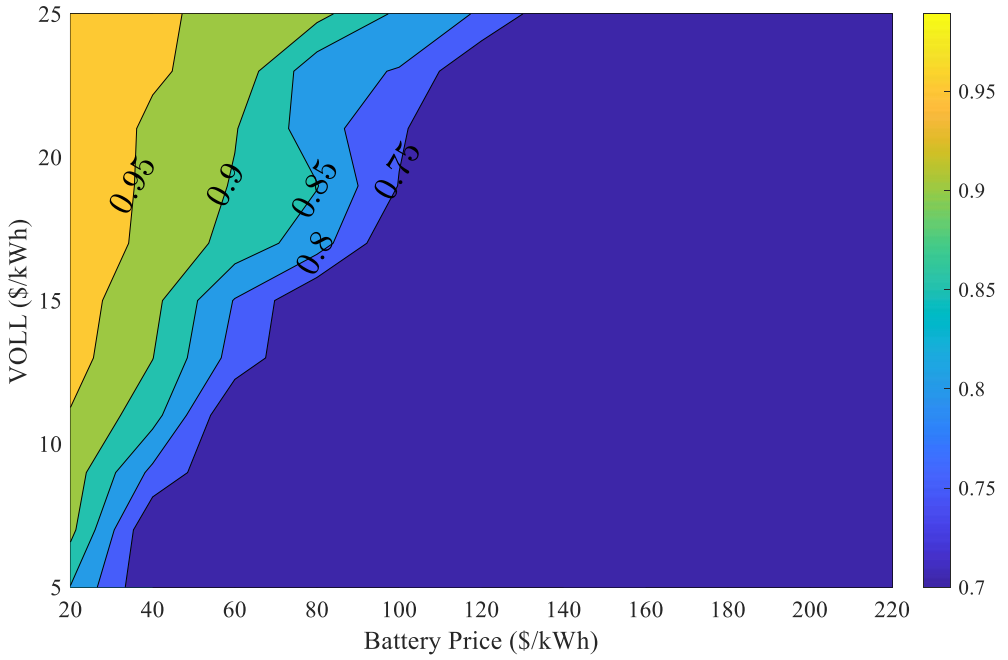

Fig. 3.16b). Contours corresponding to **Fig. 3.16a)**

Fig. 3.13a) indicates that for battery prices higher than 100 $/kWh, the required battery capacities for the optimal *TSC* have very small variations among them. This fact is independent of the value of *VOLL*. Furthermore, the battery

capacity that is required to achieve the optimal *TSC* increases as *VOLL* increases, for example, from 9 $/kWh to 21 $/kWh with a low battery price of 40 $/kWh.

Figs. 3.14a) and 3.14b) present the most consistent and smoothest trend among all figures. In Fig. 3.14b), the change rate of the optimal *TSC* with respect to battery price is greater than the change rate of the optimal *TSC* with respect to *VOLL*. However, the latter rate of change becomes greater in the high battery price area.

Finally, the big plateau appearing in Figs. 3.15a) and 3.15b) should be compared with the plateau in Figs. 3.13a) and 3.13b). When it is expensive to invest in a battery and the economic losses of load demand are relatively low, the required battery capacities barely satisfy the reliability constraints. When compared with Fig. 3.13a), it can be observed that the PV + battery systems with the lowest *TSC* can also satisfy the predetermined *LOLP* constraint when battery price is low. It denotes that the decline of battery price not only decreases the lowest *TSC* but also enhances system output capability in island mode because the system planner can afford to buy a larger battery for the system. Conclusions extracted from the existing plateau in Figs. 3.16a) and 3.16b) are consistent with the previous results. In the region where battery price is high, the *CCP* constraint is almost tight for the optimal battery size.

3.2.3 Key findings

Besides the general conclusions described in Section 3.2.2, it would be valuable if the obtained results are interpreted from the more specific scope of a hospital. It means that the attention should be given to the high *VOLL* and the low battery price based on the forecasted decline in battery prices (IRENA 2017). Furthermore, the discussion focuses only on Figs. 3.13 to 3.16, which correspond to the constrained cases where a relatively strict *LOLP* constraint is satisfied. This is

supported by the high priority that resilience possesses in critical facilities, like a hospital.

There is a pressing need to install relatively large batteries for such critical applications, as can be observed in Fig. 3.13a). This trend is expected to remain the same or even to be intensified in the future when battery prices decrease, which makes energy storage systems more cost-efficient. However, there is also great potential ahead because when battery investment costs decrease, the corresponding *TSC* can be anticipated to be reduced accordingly, as Figs. 3.14a) and 3.14b) suggest. To further strengthen this argument, it can be reminded that the forecasted changing trend of battery price can also bring positive influence on system resilience. As Fig. 3.15a) illustrates, the system designer would be able to decrease the budget and also enhance the energy system resilience.

Finally, there are some interesting relationships between the contour trends shown in Fig. 3.13 and Figs. 3.15 to 3.16, which are collectively presented in Fig. 3.17:

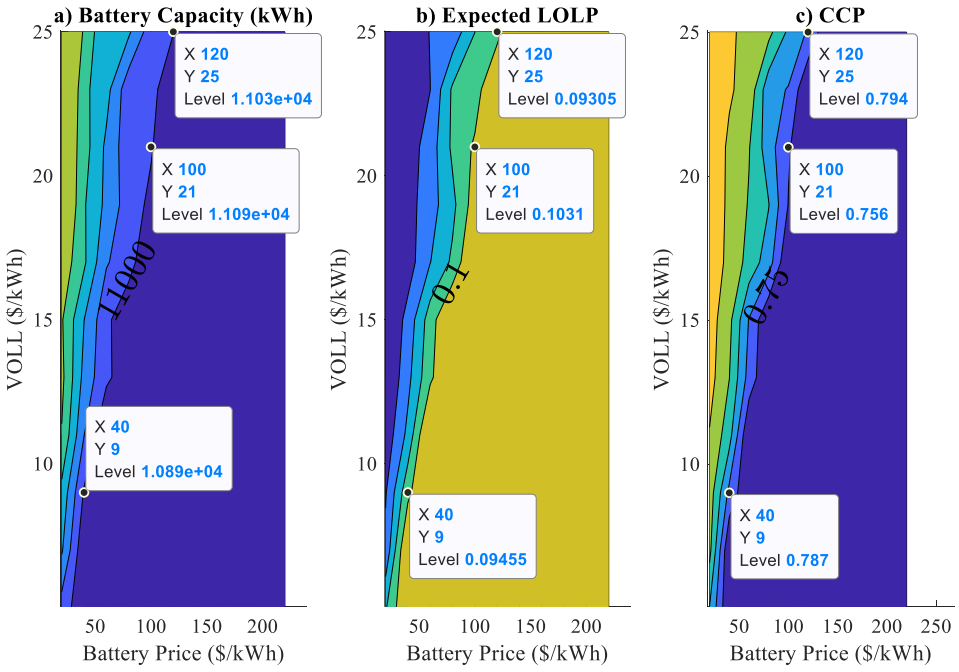

Fig. 3.17 Comparison among battery capacity, expected *LOLP* and *CCP* for the constrained case

There is a plateau that can be clearly seen in the lower-right part of Fig. 3.17a), b) and c). Based on Table 1 and the data labels shown in Fig. 3.17, there are three critical contours in Fig. 3.17a): the required battery is close to 11,000 kWh; b): the expected *LOLP* is close to 0.1; and c): the *CCP* is close to 0.78.

Table 3.5 Selected points in three critical contours with the same battery price and *VOLL* in Fig. 3.17

Evaluation Metrics		Fig. 3.17a)	Fig. 3.17b)	Fig. 3.17c)
[Battery price, *VOLL*]	[40,9]	10890	0.095	0.787
	[100,21]	11090	0.103	0.756
	[120,25]	11030	0.093	0.794

The impacts of *VOLL* and battery price on the required batteries, the expected *LOLP*, and the *CCP* become insignificant after exceeding the critical contours. The explanation for this phenomenon is that there are specific combinations of *VOLL* and battery price, upon which the battery capacities that contribute to the optimal *TSC* make both the expected *LOLP* and the *CCP* constraints not binding. It indicates that even lower *TSC* can be achieved with the desired system islanding output reliability. This is an extremely important research result because it proves that by properly incorporating *VOLL* in the mathematical formulation, it is possible to obtain solutions which are beneficial from the economic and reliability perspectives simultaneously.

Finally, it should be mentioned that the validity of all the above models and results has been verified by querying industry experts of the field.

3.2.4 Importance of exploring more storage types

It is shown in the analyses that the future decline in battery price crucially affects applications where the associated *VOLL* is high. Consequently, it is necessary to combine the latter result with the interaction that exists between *VOLL* and battery price for future optimization of PV + battery system. Although in general it is observed that higher *VOLL* results in higher *TSC*, it is also showed that this

relationship becomes weaker as the battery price decreases. This finding brings great potential for all kinds of applications and especially for those whose energy backup systems need to be highly reliable and efficient.

The natural question that arises at this point is whether battery price is the only characteristic that affects significantly the results of the optimization problem. As it was seen in Section 2.3.1, there are multiple storage types available in the market for microgrid use. Each one of them possesses its own distinct features, such as efficiency, depth-of-discharge (*DoD*), etc. Therefore, it would be worthwhile from a research perspective to study how the selection of a specific storage type tweaks the results that are presented in Sections 3.1 and 3.2.

For this exact reason, Tsianikas *et al.* dived deeper into the problem of investigating the effects that different battery types have on the optimal configuration of photovoltaic (PV) and battery systems (Tsianikas, Zhou et al. 2019). The authors adopted and improved the simulation-based method presented in (Zhou, Tsianikas et al. 2019) and compared different battery types based on their projected characteristics for the year 2030 (IRENA 2017). The four battery types considered are lead-acid, sodium sulphur, vanadium redox, and Li-ion and the tested facility is again a hospital but now located in Orlando, FL. The location-specific datasets used in this work can be found in (NREL 2013) (NREL 2016). The specific values of the characteristics used are provided in Table 3.6 (IRENA 2017):

Table 3.6 Approximate central estimates for year 2030 of cost, efficiency and DoD for four battery types (IRENA 2017)

Parameter / **Battery Type**	Cost ($\$/kWh$)	Efficiency (%)	DoD (%)
Lead-Acid	75	86	55
Sodium Sulphur	165	86	100
Vanadium Redox	120	78	100
Li-ion	224	97	90

In Fig. 3.18, results can be seen for all types of batteries in a dual axis format for *TSC* (left) and *CCP* (right) with a preset *LOLP* constraint of 10%:

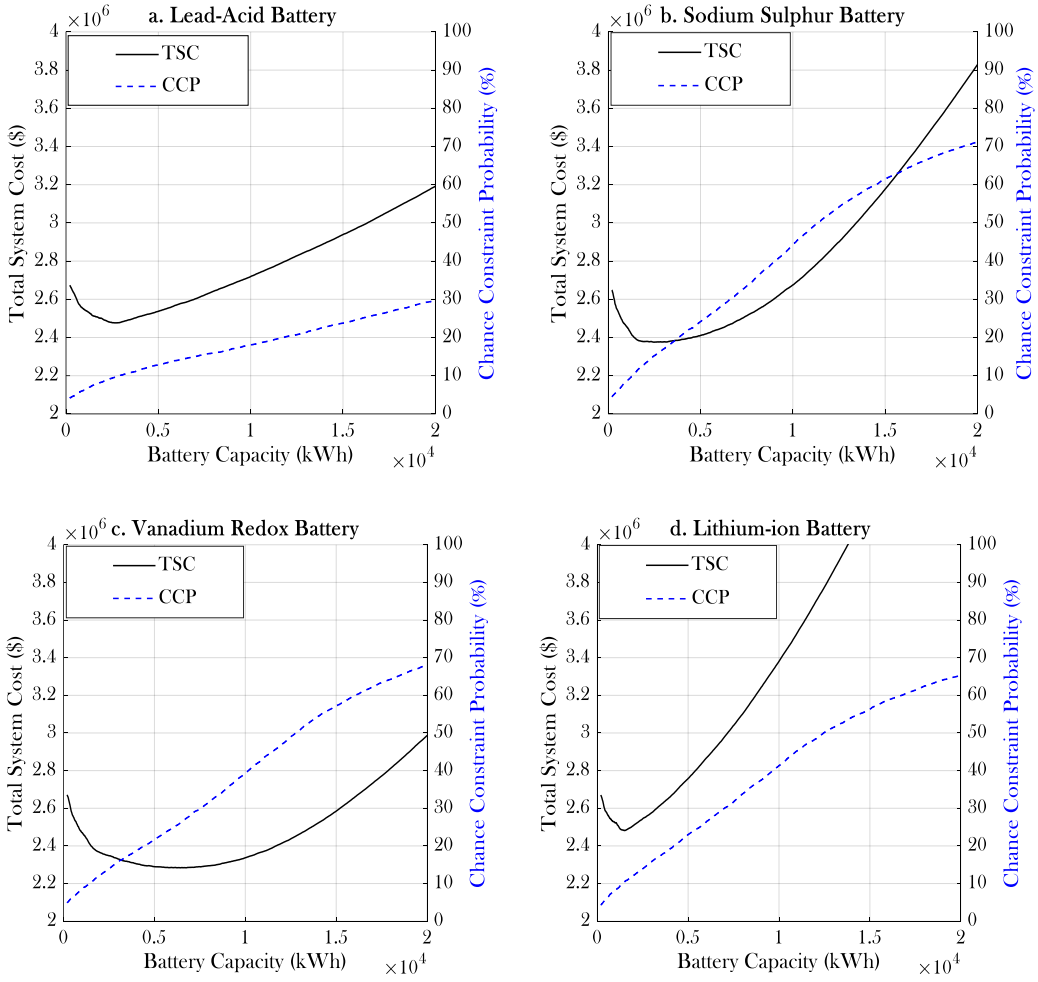

Fig. 3.18 *TSC* and *CCP* vs battery capacity for the four battery types considered

There are some very important and intuitive conclusions that can be drawn from Fig. 3.18, and this can be done through individual or pairwise observations. Firstly, it can be shown that lead-acid battery achieved the worst performance in terms of *CCP* and this is almost exclusively due to its extremely low *DoD*. Li-ion was the most expensive battery and this can be easily justified from the highest comparative *TSC*. On the contrary, it should be noticed that vanadium redox battery is a storage type with great potential because *TSC* stays low as the battery capacity of this type changes in the considered range with a satisfactory level of *CCP*. Lastly, sodium sulphur battery type achieved the highest *CCP*, although the differences are small

when compared with that achieved by vanadium redox and Li-ion battery types. On the contrary, the corresponding *TSC* of sodium sulphur battery lies between that of vanadium redox and Li-ion battery types.

An interesting comparison that should also be examined carefully is between sodium sulphur battery and vanadium redox battery. It should be reminded here that the former type had higher efficiency than that of the latter one, as well as a higher cost. However, their differences in *TSC* are much more significant than their differences in *CCP*. It demonstrates that vanadium redox battery is more cost-effective than the sodium sulphur battery. Another worthwhile comparison is between vanadium redox battery and Li-ion battery. The 10% difference in the *DoD* of these two batteries is able to offset their higher difference of 19% in terms of efficiency, resulting in a very similar *CCP* assessment and therefore making cost the dominant factor for choosing between these two batteries.

Conclusively and based on the simulation results, a general conclusion that can be drawn is the high importance of the three selected battery properties, i.e. cost, efficiency and *DoD* regarding the performance of photovoltaic and battery systems. The results presented here, especially if combined with the insightful research works from the literature on hybrid energy storage systems (HESS) that were presented in Section 2.3.1, are giving birth to a new research need: create temporally dynamic models that integrate different types of storage systems in the same microgrid. This research need is one of the main drivers for the research work that is presented in Section 4.

4. Dynamic Storage Expansion Planning in Microgrids

Herein, a novel mathematical framework that attempts to tackle the dynamic long-term storage expansion planning problem is introduced. In order to do so, the notions of Markov Decision Processes (MDP) and reinforcement learning are heavily utilized, as these were presented in Section 2.5. In the first subsection, the updated framework of the problem is presented. The changes considered in the microgrid formation, as well as changes in the storage modeling, are mentioned there, too. Moreover, the problem formulation as an MDP is provided, along with the necessary mathematical definitions and assumptions. A method which takes advantage of synthetic datasets to mitigate the problem of bias overestimation that the Q-learning algorithm possesses is also proposed. Afterwards, numerical assumptions and case study results are presented, along with the necessary acknowledgment of the limitations of the proposed framework, which however naturally lead to the subsequent research work. In the last subsection, the emphasis is given on how the outage modeling can significantly affect the optimization results.

4.1 Problem framework

Firstly, it is required to present the updated and detailed problem framework that this research work attempts to solve. The objective of the problem is to derive optimal dynamic storage expansion plans for a given microgrid. The main changes from the formulations in Section 3 concern the microgrid formation (power plants, storage options, etc.) and the battery modeling, in terms of investment and operation scheduling. In the end, other necessary assumptions considered in this work are mentioned.

4.1.1 Microgrid formation

In this problem formulation, the microgrid considered belongs to the family of community microgrids, as these are presented in Section 2.2.1. It means that the renewable plants, combined with energy storage units, need to serve a number of facilities in this situation and not a single one. However, it should be mentioned here that the main role of the microgrid is again to provide reliable and uninterrupted backup generation, due to the existence of critical facilities that rely on its electricity supply, such as a hospital. That being said, each facility in the considered microgrid has its own $VOLL^g$ and critical load factor $C_p{}^g$, where $g \in G$ and G is the set of existing facilities. This is a crucially important change, as it affects the way that energy produced by power plants or provided by energy storage systems should be distributed. Concerning this matter, a prioritized scheme is adopted, in which facilities are ranked based on their criticality and need to be served accordingly. Last but not least, the assumption now is that multiple renewable energy plants are incorporated in the microgrid, such as solar panels and wind turbines.

It is clearly established in Section 2.3 of the present dissertation that a mixture of different storage units, resulting in so-called Hybrid Energy Storage Systems (HESS), is leaning towards more beneficial solutions overall for the microgrid. Therefore, it is appropriate to consider not only different storage options in the planning process but also mixing them together in an aggregated fashion, which implies that all of them are together responsible for supplying energy to the microgrid. However, another real novelty of this work is that storage investments do not need to be made at the beginning of each planning period, but they may be made at various decision periods within the planning horizon. This fact could clearly let the system planners leverage the declining projections in storage systems prices in the future.

4.1.2 Storage scheduling and investment scheme

A typical problem existing in all energy systems that contain different types of storage units is the charging and discharging scheduling of the storage systems. While it is an entire optimization problem by itself, it should be mentioned herein that the biggest problem in the presence of multiple ESS units is that a simultaneous discharging of paralleled ESSs would unavoidably result in significant state-of-charge (*SoC*) differences between the various storage units (Semënov, Mirzaeva et al. 2017). These differences, if propagated through several periods, could result in system power drops. The potential solution to this problem depends on whether there is a centralized or decentralized control in the microgrid. In the former case, there is a centralized control unit in the microgrid, which gathers all the necessary information and distribute the amount of energy provided by different ESSs in such a way that guarantees similar *SoC* levels among all the storage units in the system. On the contrary, in the latter case, there are various techniques that can be implemented, such as relating droop coefficients to the levels of *SoC* (Semënov, Mirzaeva et al. 2017). The present implementation is more closely related to the decentralized approach and is based on predetermined contribution ratios that are able to achieve the necessary *SoC* balancing, and actually without utilizing *SoC* real-time information. Towards this direction, it is required to introduce the definitions of these charging and discharging ratios considered in this work. These ratios reflect the proportion of energy that each ESS should contribute while charging or discharging respectively. Therefore, the charging proportions p_c^i and the discharging proportions p_d^i are defined as follows:

$$p_c^i = \frac{B_r^i \, DoD^i}{e^i \sum_{j \in SU} \frac{B_r^j \, DoD^j}{e^j}}, \forall i \in SU \tag{4.1}$$

$$p_d^i = \frac{B_r^i \; DoD^i \; e^i}{\sum_{j \in SU} B_r^j \; DoD^j \; e^j}, \forall i \in SU \tag{4.2}$$

where SU is the set of the various storage units existing in the microgrid, B_r^i is the capacity of the i^{th} storage unit, DoD^i its corresponding depth-of-discharge and e^i the round-trip efficiency. These parameters guarantee similar SoC levels among the different storage units and also guarantee that:

$$\max_{i \in SU} SoC^i = 1 \Rightarrow SoC^i \approx 1, \forall i \in SU \tag{4.3}$$

$$\min_{i \in SU} SoC^i = 1 - DoD^i \Rightarrow SoC^i \approx 1 - DoD^i, \forall i \in SU \tag{4.4}$$

which technically means that all storage units reach simultaneously their maximum and minimum allowed levels of charge.

Concerning the monetary investment in storage units, an amortization model has been adopted where the payments are made annually, and each payment is calculated as follows:

$$P^i{}_{\text{annuity}} = P^i{}_{\text{principal}} \; \frac{IR\left(1+IR\right)^{L^i}}{(1+IR)^{L^i} - 1} \;, \forall i \in SU \tag{4.5}$$

where $P^i{}_{\text{principal}}$ is the principal investment amount of the i^{th} storage unit, IR is the annual interest rate and L^i is the lifetime of the i^{th} storage unit. This amortization model resembles a leasing scheme, in which annual payments and the existence of the storage unit in the system are continued after the lifetime period of the unit expires. This may seem counterintuitive; however, it stems from the fact that storage units cannot be retired under the current approach. Although it would be ideal to include such decisions in the problem formulation, and this would become feasible in the expanded model of Section 5, Eq. (4.5) provides an easy way to incorporate

information about the lifetime of the various storage options in the problem economics, without altering the information about their cost parameters.

4.1.3 Other assumptions

In other assumptions worth mentioning, it should be noted that accurate solar and wind output power calculation has been adopted, based on solar cells per panel, available solar panels, rotor swept area of the wind turbine and the total number of wind turbines (Song, Li et al. 2018). The equations are shown as follows:

$$P_{solar}\left(t_{ijk}\right) = \eta_{solar}\ A_{cell}\ n_{cpp}\ n_{pan}\ I\left(t_{ijk}\right),\ \forall i \in O_{jk},\ \ j \in N_k,\ \ k \in \{1,2,3,...,K\} \tag{4.6}$$

$$P_{wind}\left(t_{ijk}\right) = \begin{cases} 0, \text{if } W\left(t_{ijk}\right) \le W_{in} \\ \frac{1}{2}\eta_{wind}\rho A_{tur}\ n_{tur} W\left(t_{ijk}\right), \text{if } \ W_{in} < W\left(t_{ijk}\right) < W_{out}, \forall i \in O_{jk},\ \ j \in N_k,\ \ k \in \{1,2,3,...,K\} \\ 0, \text{if } W\left(t_{ijk}\right) \ge W_{out} \end{cases} \tag{4.7}$$

where η_{solar} is the solar panel efficiency, A_{cell} is the area of each solar cell, n_{cpp} is the number of solar cells per panel, n_{pan} is the number of solar panels in the system, η_{wind} is the wind turbine efficiency, ρ is the air density, A_{tur} is the rotor swept area of the turbine and n_{tur} is the total number of wind turbines in the system. Moreover, $I\left(t_{ijk}\right)$ and $W\left(t_{ijk}\right)$ denote the solar irradiance and the wind speed at time t_{ijk} accordingly. N_k is defined again as a random set containing all grid outages that occurred in decision period *k*, while O_{jk} is defined as a random set of all time intervals of grid outage *j* in decision period *k* and *K* is the total number of decision periods. Finally, W_{in} and W_{out} are called cut-in and cut-out wind speeds and define the range in which the wind turbine can safely produce energy.

4.2 Problem formulation as a Markov Decision Process

At this point, the definition of the Markov Decision Process (MDP) is given, in which the solver algorithm is based on. Firstly, a detailed definition of the state and action sets and the reward and transition functions are given, following the terminology defined in Section 2.5.1. After that, an approach is proposed to mitigate the overestimation bias problem of the Q-learning algorithm that was described in Section 2.5.3, by using synthetic datasets and metamodeling. Lastly, and before proceeding with case study and results, the final algorithm is presented step-by-step.

4.2.1 MDP formulation

Every MDP, as defined in Section 2.5.1, is actually a tuple of (*S,A,f,R*), meaning that it is fully defined with the state and action sets *S,A*, the transition and reward functions *f, R* and the discount factor γ. Therefore, it is considered necessary to provide the required definitions for these elements, in order to be able to use the appropriate algorithms to derive optimal policies.

Starting with the state space *S* of the problem, it should be mentioned that it consists of three sub-features; time feature, external features, and internal features:

$$
\begin{aligned}
& S = S^{tf} \times S^{ef} \times S^{if} \\
& \text{where:} \quad s^{tf} \in S^{tf} = \{1, 2, ..., K\} \\
& \qquad \mathbf{s}^{ef} = \left(s_{i,j}^{ef}, \forall i \in SU, j \in SC \right) \in S^{ef} \\
& \qquad \mathbf{s}^{if} = \left(s_{i}^{if}, \forall i \in SU \right) \in S^{if}
\end{aligned}
\tag{4.8}
$$

S^{tf} is the time-dependent component of the state space and it simply denotes the current decision period. It should be noted here that it is highly advised for the timing feature to be explicitly included in the state information of the problem. It has been proven that the agent's learning performance is significantly improved when time-awareness of the agent is introduced, by specifically incorporating a time-related space component (Pardo, Takavoli et al. 2018) (Harada 1997). S^{ef} defines the set of

external features of the problem, such as the price, the efficiency and the depth-of-discharge of the storage unit, where *SU* is again the set of storage units and *SC* is the set of storage characteristics included in the formulation. They are called external because the information coming from these characteristics comes from the information that is received from the environment, without the option for the agent to affect them. As an example, an instance of S^{ef} is the vector $\mathbf{s}^{ef}$, consisting of all the elements $s_{i,j}^{ef}$ which denote the value of the j^{th} characteristic of the i^{th} storage unit. Finally, S^{if} is the set of internal features of the problem, such as the storage capacity already installed in the system. They are called internal because the agent is able to affect these components by taking appropriate actions. Therefore, the microgrid's state is defined by a vector such as $\mathbf{s} = \left(s^{tf}, \mathbf{s}^{ef}, \mathbf{s}^{if} \right)$.

Concerning the action set of the problem, it is clearly defined based on the possible actions that the agent can take. In the context of the current problem, the agent should choose between taking no action or deciding to expand the storage capacity of a specific storage unit at one of the available predetermined levels (to align with the discrete time and space assumptions of a DTMC framework). Therefore, it is pretty straightforward that the agent's action can be defined as a vector such as:

$$\boldsymbol{\alpha} = \left(a_{i,l}, \forall i \in SU, l \in SL \right) \in A \tag{4.9}$$

$$\text{s.t.} \quad \sum_{i \in SU} \sum_{l \in SL} a_{i,l} \leq 1$$

$$a_{i,l} \in \{0,1\}, \forall i \in SU, l \in SL$$

In this context, $a_{i,l}$ denotes the binary action of expanding the capacity of the i^{th} storage unit at the l^{th} level, where *SL* is the set of available expansion levels. The

first constraint imposed in (4.9) limits the agent so it cannot take more than one expansion actions per period, while the second constraint guarantees the binarity of the action components.

Proceeding with more definitions, the focus is now given to the state transition function *f*. Given the fact that the state is composed by three parts (time, external and internal component) and by using the notation s for the current state and s' for the next state, the state transition equations are provided below:

$$s^{tf\prime} = f^{tf}\left(s^{tf}\right) = s^{tf} + 1, \quad \forall s^{tf} \in S^{tf} \tag{4.10}$$

$$\mathbf{s}^{ef\prime} = f^{ef}\left(\mathbf{s}^{ef}\right), \text{ where: } \left\{s_{i,j}^{ef}, s^{tf} \in S^{tf}\right\} \text{ is a DTMC with transition matrix } p_{i,j}^{ef}, \forall i \in SU,\ j \in SC \tag{4.11}$$

$$\mathbf{s}^{if\prime} = f^{if}\left(\mathbf{s}^{if}, \boldsymbol{\alpha}\right), \text{ where: } s_i^{if'} = s_i^{if} + \sum_{l \in SL} l\, a_{i,l}, \forall i \in SU \tag{4.12}$$

Eq. (4.10) is the state transition equation for the time feature of the state space and is simply an incremental by-one operation. Eq. (4.11) preserves the Markov property of the external features of the state space; it means that the j^{th} characteristic of the i^{th} storage unit follows a DTMC with the corresponding $p_{i,j}^{ef}$ transition matrix. Finally, Eq. (4.12) is the transition equation for the internal feature of the state space; if it is decided to expand the i^{th} unit's storage capacity at the l^{th} level, the corresponding s_i^{if} is going to be increased appropriately. Overall, the next state could be described by the vector $\mathbf{s}' = \left(f^{tf}\left(s^{tf}\right),\ f^{ef}\left(\mathbf{s}^{ef}\right),\ f^{if}\left(\mathbf{s}^{if}, \boldsymbol{\alpha}\right)\right)$.

The last component of the MDP that needs to be properly defined is the reward function. This is a crucial part since it affects the way that the agent receives signals (i.e. rewards) from the environment. These signals are the main drivers that guide the agent to the derivation of the optimal policies. In a similar fashion with the

previous chapters, but now following the expanded formulation, the reward at the k^{th} decision period can be defined as:

$$r_k(\mathbf{s},\boldsymbol{\alpha}) = -\sum_{i\in SU}\sum_{l\in SL}(K-k+1)y\,a_{i,l}\,P^{i}{}_{annuity} - \sum_{g\in G} VOLL^{g}\sum_{j\in N_k}\sum_{i\in O_{jk}}\delta\left(t_{ijk},g\right)C_p^{g}\,D\left(t_{ijk},g\right)$$

(4.13)

The first term of the reward function is the investment part, which is calculated as $(K-k+1)y$ equal payments of $P^{i}{}_{annuity}$ for each facility i if there is a decision to invest, and where y denotes the number of years in one decision period. The second term of the reward function defines the loss of load cost, for all facilities in the set G of the formulation. Eq. (3.10) presented the definition of δ function for the previous setting, where only one facility was existing. However, herein it is needed to give an updated definition of the δ function, in which it should be assumed that the facilities in the set G are ranked based on the prioritization scheme that is introduced in Section 4.1.1 (i.e. facility 1 is the most critical, facility 2 is the second most critical, etc.):

$$\delta(t_{ijk},g) = \begin{cases} 1, \text{ if } Q^{b}\left(t_{ijk}\right) + \int_{t_{ijk}}^{t_{ijk}+\Delta t}\left(p_c^{b}\left(P_{solar}(u)+P_{wind}(u)\right) - \frac{p_d^{b}}{e^{b}}\sum_{m=1}^{g}C_p^{m}D(u,m)\right)du < B^{b}{}_{\min} \\ 0, \text{ otherwise} \end{cases}$$

for an arbitrary $b\in SU$ and $\forall i\in O_{jk}$, $j\in N_k, k\in K, g\in G$ (4.14)

In simpler terms, $\delta(t_{ijk},g)$ is equal to 1 for the facility g if the energy stored in storage units, defined as Q^{b}, combined with the energy production by renewable plants net the demand of facility g, defined in the second part of the left-hand side of the inequality in the first branch of Eq. (4.14), are lower than the minimum allowed energy level $B_{\min}$. It should be noted here that $B_{\min}$ is determined by the depth-of-

discharge of the storage unit. On the contrary, $\delta(t_{ijk}, g)$ is equal to 0, only when the system (energy in storage units and energy production by renewable plants net the demand of the facility *g*) is able to satisfy the demand for all the facilities up to *g*. Therefore, for the most critical facility (*g*=1), the system needs to be able to meet the demand only for this facility, in order for the load demand to be met. For the second most critical facility (*g*=2), the system needs to meet the demand for facility 1 and facility 2, etc. An important point is that just one arbitrary $b \in SU$ is chosen, in order to determine whether the demand is lost or not for a specific facility. The justification for this comes from Eq. (4.4) which suggests that if one storage unit *b* falls below the minimum allowed level $B_{\min}^{b}$, then the same should apply for the rest of the storage units. Finally, and because of the fact that the system is designed with the main purpose to be backup energy provider, the storage units are utilized only during grid outages. Therefore, it is clear that the larger the storage, the longer the system is able to satisfy the demand of facilities and the fewer are the times that the indicator function δ is equal to 1.

4.2.2 Utilization of synthetic datasets to tackle overestimation bias

In Section 2.5.3 it is shown why the Q-learning algorithm suffers from overestimation bias in highly stochastic environments. Practically, this means that if the agent assumes that there is a chance it would receive an extremely "good" reward if being in a specific state, he may try to transition to that state, even though the average reward is lower than the average rewards from being in other states. In this subsection, it is explained how this phenomenon applies to the examined case and a way that could potentially mitigate this effect is proposed.

Remembering how this overestimation bias problem of the Q-learning algorithm is illustrated in Section 2.5.3, it can be safely assumed that the problem

arises from situations where the agent can get misleading "signals" on how the optimal strategy is structured. In this context, these signals correspond to the rewards that the agent receives in every decision period of the problem. Due to the way that the reward function is defined in Eq. (4.13), it is mainly composed of two negative components: the investment cost and the outage penalty. While the investment is clearly affected solely by the decision to expand storage capabilities, the outage penalty relies heavily on the stochastic events of outages. Considering the scenario of having 0 (or at least very few and/or short-lived) outages in a specific decision period, the agent may consider beneficial for the system to go as is and "suffer" these outages, instead of taking actions to protect against them, i.e. investment actions. This fact results exactly in the misleading "signals" that are mentioned before. In the best possible scenario for us, this phenomenon would just slow down the convergence of the solution algorithm, while in the worst scenario it could result in deriving sub-optimal policies. Consequently, it would be beneficial here to propose a novel approach for mitigating this effect using synthetic datasets and function approximation for the outage cost component of the reward function.

As its name basically reveals, synthetic datasets consist of data observations that are generated programmatically using simulation techniques, and not by real-life experiments and data collection (KDnuggets 2018). In this case, similar simulation techniques to those explained in Sections 3.1 and 3.2 can be utilized in order to generate a synthetic dataset consisting of multiple input features and one output feature, the outage cost. Afterwards, a function approximation technique can be used to map, as closely as possible, the given inputs to the desired output. Therefore, the features needed to "predict" the outage cost form a vector of the following form: $\left(s^{tf},\left(s_i^{if},\forall i\in SU\right)\right)$, meaning that this specific cost component depends on the timing

feature of the state space and the installed capacity of every storage unit in the system. In this context, s_i^{if} denotes the installed capacity of the i^{th} storage unit.

As the first step of this process, a systematic way has to be derived to generate observations to be added in the synthetic dataset. As previously mentioned, each of these observations comes from running *n* individual and independent simulation runs of the system and averaging the obtained results. In order to generate independent observations, a random sample of the input features can be used. That means that each input feature of the dataset (timing feature and installed capacities for all storage units) is arbitrarily picked from specified corresponding ranges. Moreover, for each of these individual simulation runs, outages are generated using again the standard reliability metrics of CAIDI and SAIFI. More specifically, the duration of a specific outage is a shifted Poisson distributed random variable with mean CAIDI and the outage events form a Poisson process with rate SAIFI. After the input features are selected for a specific observation and the outages are generated for each run, the system is simulated *n* times and the output (outage cost) is computed by averaging the results of these *n* simulations. This procedure is iteratively followed until an *S*-sized dataset is created, where *S* is the predetermined desired length of the dataset. Lastly, the random forest algorithm is used as a function approximation for the outage cost, given the synthetic dataset. Therefore, it can now be seen that the reward for the k^{th} decision period can be rewritten as:

$$r_k\left(\mathbf{s},\boldsymbol{\alpha}\right) = -\sum_{i\in SU}\sum_{l\in SL}(K-k+1)\,y\,a_{i,l}\,P^{i}{}_{annuity} - f^{RF}\left(k,\left(s_i^{if},\forall i\in SU\right)\right) \tag{4.15}$$

4.2.3 Final algorithm

At this point, and before proceeding with the numerical case studies and results, it would be useful to provide a schematic and holistic representation of the proposed approach. Although the algorithm used is the classic Q-learning approach, it

should be mentioned that the preprocessing step of synthetic data creation and function approximation is added. The procedure can be seen in Table 4.1 and is the one followed in (Tsianikas, Yousefi et al. 2019):

Table 4.1 Schematical representation of the Q-learning algorithm with preprocessing step

	Algorithm: Q-learning with preprocessing step
1:	**initialization:** random outages for all simulation runs
2:	**for** *every observation* **do:**
3:	select arbitrarily $\left(k,\left(s_i^{if},\forall i \in SU\right)\right)$
4:	**for** *every simulation run* **do:**
5:	simulate system and compute outage cost
6:	**end for**
7:	average over all runs and store observation in the synthetic dataset
8:	**end for**
9:	use random forest to derive f^{RF} from the synthetic dataset
10:	**initialization:** Q table
11:	**for** *every episode* **do:**
12:	**initialization:** starting state *s*
13:	**for** *every decision period* **do:**
14:	select action α based on Q and ε-greedy policy
15:	observe reward $r \sim \left(a, f^{RF}\right)$ and next state $s' \sim$ function f
16:	$Q(s,a) \leftarrow Q(s,a) + \alpha\left[r + \gamma \max_a Q(s',a) - Q(s,a)\right]$
17:	$s \leftarrow s'$
18:	**end for**
19:	**end for**

The first nine lines of the algorithm define the preprocessing step and the last ten lines compose the typical steps of the Q-learning algorithm, adjusted for the current problem. It should be mentioned here, that the main reason for the mitigation of the overestimation bias problem that the Q-learning algorithm imposes comes from the seventh line of the proposed approach. The fact that the average over a large number of simulation runs is used in order to get an estimation of the outage cost, makes the "signal" that the agent perceives much clearer and without unnecessary variance.

4.3 Case study

In this section, results are presented from a case study conducted using the methodologies described above. The microgrid considered in this case study consists of several facilities (hospitals, schools, and residential houses) and is located in Westhampton, NY. In Fig. 4.1, the whole area of Westhampton can be seen as a satellite view:

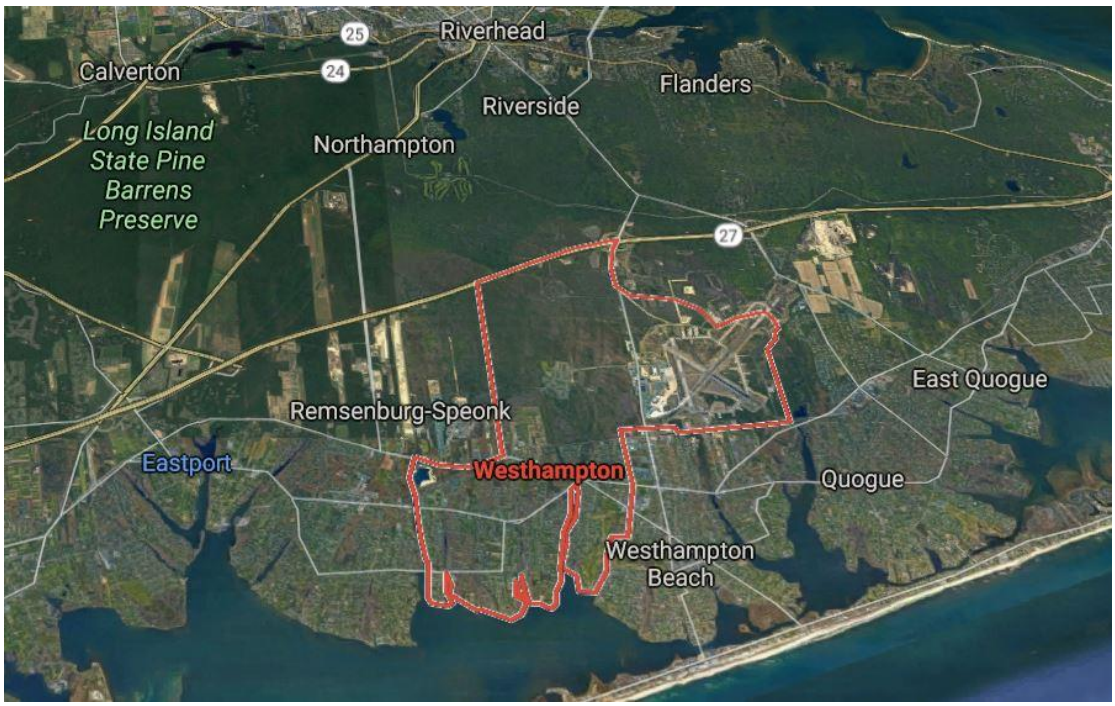

Fig. 4.1 Satellite view of the Westhampton, NY area (Google 2019)

The reason that this particular location is chosen lies in the fact that this is an area with a high wind energy potential, very close to the North Atlantic Ocean. Location-specific demand and meteorological data are again used (NREL 2013) (NREL 2016). Each facility in the microgrid comes with its associated *VOLL* and critical load factor. Concerning the storage options existing in the formulation, four different types of storage technologies are tried: Li-ion battery, lead-acid battery, vanadium redox battery, and flywheel storage system. It is clear from the choices made that it is the author's intent to explore various storage options, including not only electrochemical storage systems but other less common ones. Each storage type has its own characteristics, which of course are expected to affect the results in a

significant fashion. For simplicity and dimensionality reduction purposes, it is assumed that all the storage system characteristics can be described by a different deterministic function of the decision period, except the storage system price which holds its stochastic nature. This stochasticity of the storage price is modeled using Markov Chains, as it is described in the theoretical formulation. For example, the Markov Chain used to model the storage price of the first storage unit (Li-ion) can be seen in Fig. 4.2 below:

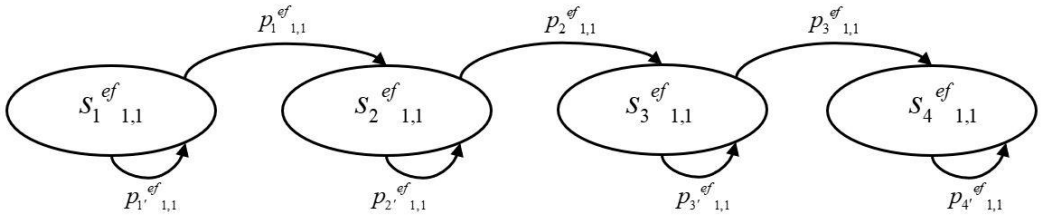

Fig. 4.2 Markov chain for the price of Li-ion storage type

4.3.1 Numerical assumptions

Regarding the numerical assumptions of the case study, a 20-year time horizon is considered, where the decision periods are every 5 years, therefore results in a total of 4 decision periods. The interest rate for storage investment is considered 2% annually. Solar and wind power plants are considered in the microgrid and their production is calculated using Eqs. (4.6) and (4.7) and by considering (Song et al. 2018) (Zhou et al. 2019): $\eta_{solar} = 0.16$, $A_{cell} = 0.0232258\,m^2$, $n_{cpp} = 72$, $n_{pan} = 6000$, $W_{in} = 3\,m^2, W_{out} = 22\,m^2, \eta_{wind} = 0.48, \rho = 1.25\,{kg}/{m^3}, A_{tur} = 1520.53\,m^2, n_{tur} = 10$.

CAIDI and SAIFI are considered 5.122 and 1.155 respectively (Service 2016) (Service 2017). According to the facilities, there are three different types: hospital, school, and residential houses. The assumptions made for these facilities can be seen in Table 4.2 (van der Welle et al. 2007) (Alsaidan et al. 2018a):

Table 4.2 Data related to facilities and their characteristics

Data / Facility	Number	*VOLL*	C_p
Hospital	2	25	0.8
School	5	17	0.6
Residential	300	8	0.4

In the aspect of storage systems, it should be mentioned that it is considered that the agent has the option to choose among three capacity levels for each storage unit and for each decision period. However, it should be noted that the agent is restricted to choose one action at maximum for each decision period, according to Eq. (4.9). The levels used in this case study are 300, 1000 and 3000 kWh. The various storage systems characteristics for each decision period of the problem can be seen in Tables 4.3-4.6 (IRENA 2017):

Table 4.3 Li-ion characteristics for all decision periods

Period / Li-ion	1	2	3	4
State for price MC ($/kWh)	420	310	167	150
Probability for price MC	0.70	0.70	0.70	0
Lifetime (yrs)	12	17	19	20
Efficiency	0.95	0.96	0.97	0.98
DoD	0.90	0.90	0.90	0.90

Table 4.4 Lead acid characteristics for all decision periods

Period / Lead-acid	1	2	3	4
State for price MC ($/kWh)	142	115	77	65
Probability for price MC	0.70	0.70	0.70	0
Lifetime (yrs)	9	11	13	14
Efficiency	0.80	0.81	0.83	0.84
DoD	0.55	0.55	0.55	0.55

Table 4.5 Vanadium redox characteristics for all decision periods

Period / Vanadium redox	1	2	3	4
State for price MC ($/kWh)	385	255	120	95
Probability for price MC	0.70	0.70	0.70	0
Lifetime (yrs)	13	17	20	21
Efficiency	0.70	0.73	0.78	0.79
DoD	1	1	1	1

Table 4.6 Flywheel storage characteristics for all decision periods

Period / Flywheel storage	1	2	3	4

State for price MC ($/kWh)	3100	2600	1950	1700
Probabilityfor price MC	0.70	0.70	0.70	0
Lifetime (yrs)	20	26	30	32
Efficiency	0.84	0.85	0.87	0.88
DoD	0.86	0.86	0.86	0.86

Finally, the algorithmic assumptions considered in this case study should be declared. Firstly, concerning the simulated synthetic data collection, a total of 1000 observations are obtained, where each observation is produced after simulating the system 100 times. For the random forest model that is used to approximate the cost component related to outages, the dataset is split to train/test using a 0.8/0.2 ratio and a total number of 10 forests is used. Finally, concerning the Q-learning algorithm, a total of 10^7 number of episodes is used, $\gamma = 0.9$ and linearly decaying rates α and ε ranging from 1 to 0.02. The decision for a linearly decaying exploration/exploitation tradeoff parameter ε is very important, as it dictates the performance of the algorithm. It basically means, that it would be ideal to explore as much as possible at the initial episodes, while it would be better to just exploit the acquired knowledge and approach convergence at the final episodes.

4.3.2 Results and discussion

As a first step in presenting the results of the case study, it would be interesting to examine the performance of the random forest model for approximating the outage cost. It should be mentioned here that the model obtained an R-squared score of 0.98 on the test set. R-squared is a metric which is used to quantify how well a model fits a set of observations. It is always in the range of [0,1] and a value of 1 technically means that the model is able to explain all the variability of the response variable around the mean. Therefore, it can be fairly sure that the model used explains the data very well. Theoretically, the outage costs are likely to follow a decreasing function of the capacity already installed in the system. However, a decaying rate for this behavior is expected, meaning that the gain from adding more capacity of a

specific storage type in the system is negligible after a point where the already installed capacity is large enough. On the opposite hand, given the distribution of the outage durations, it would also be expected to see an initial phase where adding more capacity does not result in significant outage cost savings. The aforementioned features can be observed schematically in Fig. 4.3 for the storage units of Li-ion and vanadium redox:

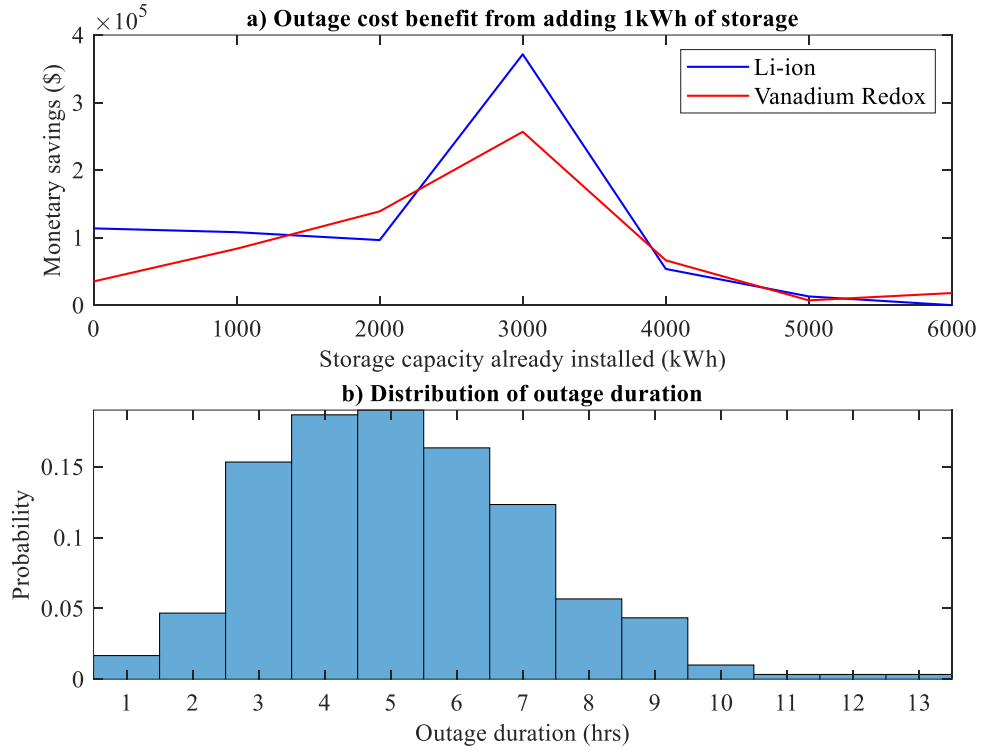

Fig. 4.3a). Outage cost savings b). Distribution of outage duration

Before proceeding with analyzing Fig. 4.3, it should be made clear that Fig. 4.3 corresponds to the initial state of the environment and does not imply any overall superiority of the Li-ion battery over the vanadium redox one. It is clear from observing Fig. 4.3a) that the expected results are obtained. The behavior is similar for both storage types; after the initial phase where the cost savings for adding capacity are small, a peak is observed where the installed capacity is around 3000kWh. After that point, the cost savings are still positive but much more negligible than they were

before. It is also very interesting to examine Fig. 4.3a) in accordance with Fig. 4.3b). It should be reminded here that the duration for each outage simulated follows a shifted Poisson distribution with mean approximately equal to 5.122. Fig. 4.3b) presents the approximate probability that a random outage obtains a value in the range of the horizontal axis. There is a clear threshold around 7hrs where after that, the outage events are really rare. That means, that when a capacity which is fairly enough to satisfy a large number of these outages is reached, it is no more cost-efficient for the planners to expand storage.

After verifying that the random forest model behaves as it should, it is time to observe the optimal policies derived from the proposed methodology. In order to be able to extract optimal policies from the results, it should be remembered that the output of the Q-learning algorithm is the completed Q-table with each field denoting the Q-value of each state-action pair. The amount of knowledge that the Q-learning is able to produce depends heavily of course on the number of episodes that the agent is allowed to experience. In the case study, each state of the environment is actually a tuple of 9 elements: the first element is the timing feature (4 possible values), the next 4 elements are the price states for each storage technology (4 possible values each) and the last 4 elements are the installed capacity again for each storage unit (10 possible values each). Therefore, the total number of states in the system can be calculated to be 2,758,578 states. If they are also multiplied by the number of possible actions in each decision period, the outcome is a total of 35,861,514 number of state-action pairs. Given the large magnitude of the state-action space, the only feasible way for observing the results of the proposed approach would be to derive scenarios for price movement in the Markov Chains and obtain optimal policies for each scenario separately.

Towards this direction, the first 3 scenarios are defined and examined. Scenario 1 refers to the case where the price of each storage unit is declining in every time period. Referring to Fig. 4.2, this scenario corresponds to the case where all the forward transitions are realized. Scenario 2 describes the case where again all storage prices are declining, except the price of vanadium redox unit for the periods 1, 3 (it only declines in period 2). Finally, scenario 3 refers to the case where all storage prices are again declining, except the price of Li-ion battery for periods 1, 3 (it only declines in period 2). Results can be observed schematically in Fig. 4.4:

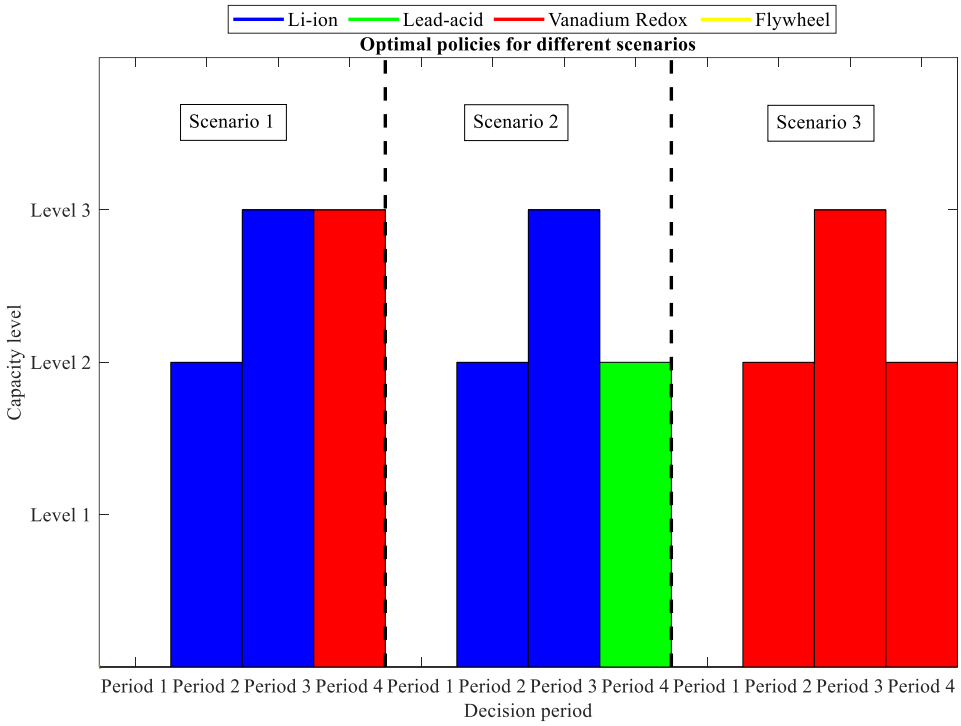

Fig. 4.4 Optimal policies derived for various scenarios

The results in Fig. 4.4 reveal some very interesting trends. The “baseline” scenario 1 presents the optimal policy under which nothing should be done in the first decision period, Li-ion battery should be installed at level 2 (1000 kWh) in the second decision period, again Li-ion battery should be installed at level 3 (3000 kWh) in the third period and finally vanadium redox battery should be installed at level 3 (3000 kWh). These results could be anticipated by looking at Tables 4.3 and 4.5. While the

price difference of these two storage units is statistically negligible in the initial phases, this difference becomes much more significant in the later stages. In these later stages, besides the lower price of vanadium redox battery, its excellent *DoD* plays a crucial role in making this type the preferred choice. These results come in accordance with the findings of Section 3.2.4, where there were already insights that vanadium redox battery holds special potential for future usage. However, in scenario 2 there is a significant difference compared to scenario 1; investment in the last decision period goes to the lead-acid type, removing vanadium redox from the preferred choices. The reason behind this change is straightforward; the vanadium redox price did not decline as sharply as in scenario 1, making it therefore relatively expensive compared to cheaper options. In this situation, lead-acid became the dominant choice, despite its very low *DoD* value. Finally, in scenario 3 the dominance of vanadium redox battery type in all decision periods from 2 and on can be observed. During this scenario, all prices are declining sharply, except for the price of Li-ion battery type. Consequently, vanadium redox took its place and resulted in investments of Level 2, Level 3 and Level 2 again for the decision periods 2, 3, 4 respectively.

At this point, it would be useful to note two more conclusions that can be drawn from Fig. 4.4. Firstly, it is observed that there is a difference between the total installed capacity among the three scenarios; 7000 kWh in scenario 1 and 5000 kWh in scenarios 2, 3. While it would be expected that these values are equal, the results of Fig. 4.3a) should now be revised. There is a certain threshold after which installing more capacity of the same storage type does not result in significant savings. Therefore, that is why in scenario 2 the replacement of the vanadium redox installation at Level 3, is an installation of lead-acid at Level 2 instead of adding more capacity of Li-ion at Level 3. Of course, the same applies to scenario 3 and the case of

vanadium redox battery. However, it is acknowledged that this behavior could be a weakness of the modeling technique and the ways to mitigate these effects are explored in the next section. Lastly, it is also seen that there is no installation of flywheel energy storage system in any scenario. This happened because of the extremely high price of this specific storage type compared to its competitors. In the case study, where critical facilities are located in the microgrid and outages can last several hours, it is clear that someone can find more use in high energy density storage units. As it is seen in Section 2.3.1, flywheel storage systems can be considered as high power density units and low energy density. Of course, these results do not mean in any case that this specific storage type cannot find applications in the microgrid sector. Instead, they would be considered appropriate in situations where fast response is the top criterion for choosing storage options.

To elaborate more on the results obtained concerning optimal policies under various scenarios, it is considered suitable to analyze here a greater number of scenarios. These results are presented in the context of Table 4.7:

Table 4.7 Optimal policies for more potential scenarios

Scenario #	Storage type	Price			Period 2	Period 3	Period 4
Scenario 1	Li-ion	↓	↓	↓	L2	L3	-
	Lead-acid	↓	↓	↓	-	-	-
	VR	↓	↓	↓	-	-	L3
	Flywheel	↓	↓	↓	-	-	-
Scenario 2	Li-ion	↓	↓	↓	L2	L3	-
	Lead-acid	↓	↓	↓	-	-	L2
	VR	→	↓	→	-	-	-
	Flywheel	↓	↓	↓	-	-	-
Scenario 3	Li-ion	→	↓	→	-	-	-
	Lead-acid	↓	↓	↓	-	-	-
	VR	↓	↓	↓	L2	L3	L2
	Flywheel	↓	↓	↓	-	-	-
Scenario 4	Li-ion	→	↓	→	-	L2	L3
	Lead-acid	↓	↓	↓	-	-	-
	VR	→	↓	→	-	-	-
	Flywheel	↓	↓	↓	-	-	-

Scenario 5	Li-ion	→	→	→	-	-	-
	Lead-acid	↓	↓	↓	-	-	-
	VR	↓	→	↓	L2	L3	L2
	Flywheel	↓	↓	↓	-	-	-
Scenario 6	Li-ion	↓	→	↓	L2	L3	-
	Lead-acid	↓	↓	↓	-	-	L2
	VR	→	→	→	-	-	-
	Flywheel	↓	↓	↓	-	-	-
Scenario 7	Li-ion	→	↓	↓	-	L2	L3
	Lead-acid	→	↓	↓	-	-	-
	VR	→	↓	↓	-	-	-
	Flywheel	→	↓	↓	-	-	-
Scenario 8	Li-ion	→	→	↓	-	L1	-
	Lead-acid	→	↓	↓	-	-	L2
	VR	→	→	↓	-	-	-
	Flywheel	→	↓	↓	-	-	-
Scenario 9	Li-ion	↓	↓	→	L2	L3	
	Lead-acid	↓	↓	↓	-	-	L2
	VR	↓	↓	→	-	-	-
	Flywheel	↓	↓	↓	-	-	-
Scenario 10	Li-ion	↓	→	→	L2	L3	-
	Lead-acid	↓	↓	↓	-	-	L2
	VR	↓	→	→	-	-	-
	Flywheel	↓	↓	↓	-	-	-

It should be noted that investment in period 1 is omitted in Table 4.7 because it does not depend on the price scenarios and it was never realized. Table 4.7 contains a total of 10 scenarios: scenarios 1-3 correspond to the ones studied before in Fig. 4.3. The rest of the scenarios in Table 4.7 correspond mainly to various combinations of price movements for the Li-ion and the vanadium redox battery. The reason for that is the other two storage types examined are not able to become the dominant ones unless they gain a competitive advantage against the other two. More specifically, as mentioned before, flywheel energy storage is not chosen under any scenario. One important thing to notice here is that the only scenario in which the total installed capacity at the end of the time horizon is 7000 kWh is the one in which both Li-ion and vanadium redox batteries experience continuous declining trends. In all the other scenarios, the final obtained capacity was 5000 kWh, or even lower; for example, when the two dominant storage types' prices remained steady for the first two periods

(scenario 8), the total installed capacity is way lower than typically. In another aspect, the high penetration of the vanadium redox battery depends heavily on its price movements; in the situations where this type of battery presents a steady behavior for at least two periods, the lead-acid battery is able to surpass it in the decision maker's choices even in cases where its own behavior remained steady for one period, like in scenario 8. Finally, it is obvious that the role of the Li-ion battery in energy systems such as the one examined in this case study is expected to remain crucial for the future. Nevertheless, there is a case where a potential level-off of Li-ion price, combined with a simultaneous decrease in vanadium redox price, could change the things in the hierarchy of these two storage types, as it happened in scenarios 3 and 5.

The final part of the presentation of the research findings contains a check on whether the agent improves its experience with an increasing number of episodes. It should be reminded here that the number of episodes chosen for this experiment is 10^7. However, the question here is how it can be asserted that this number of episodes is enough or not. Given the fact that the exploration/exploitation tradeoff parameter is decaying as a function of the number of episodes, it should always be expected to see improving performance of the agent as time passes by. The answer to this question can originate from running the experiment using a different number of episodes. The results of this procedure can be shown in Fig. 4.5:

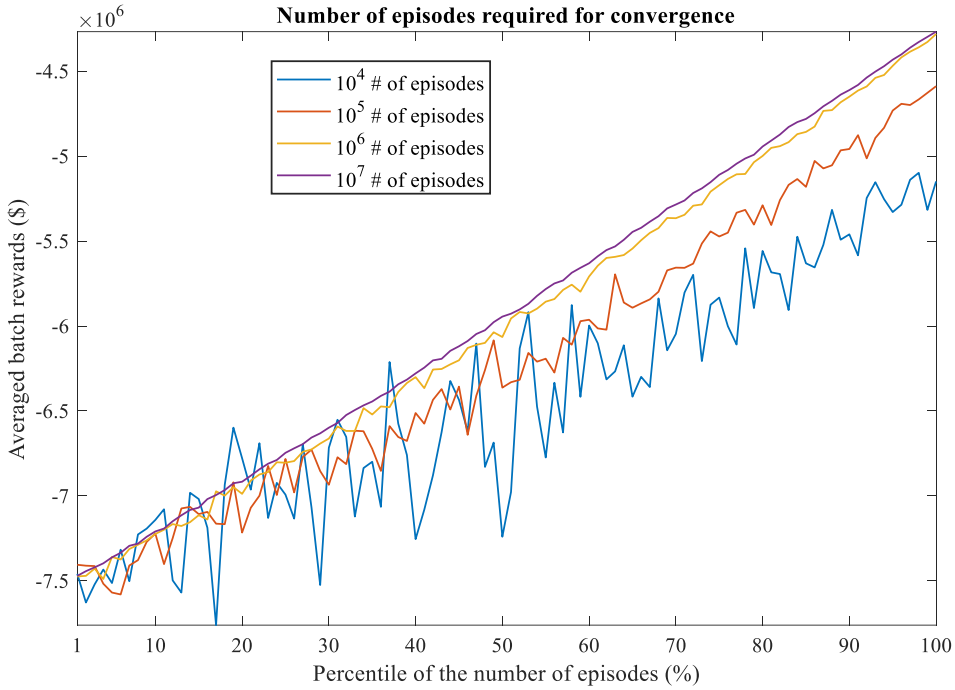

Fig. 4.5 Convergence check for the required number of episodes

In the horizontal axis of Fig. 4.5, the different percentiles of the total number of episodes are shown. In the vertical axis, the averaged total rewards can be seen for the corresponding batch of episodes belonging to that specific percentile. Of course, the exact number of episodes belonging to every percentile depends on the total number of episodes; for example, for the case of 10^6 number of episodes, each batch contains a total of 10^4 number of episodes. This is exactly the reason why different curves present different smoothness levels. However, Fig. 4.5 can prove as a great tool for proving that 10^7 is a well-suited number for this research purposes. In order to see why, specific focus should be given to the last 10 percentiles (90%-100%) in Fig. 4.5; this is exactly the region where the agent starts mostly to exploit its current knowledge and does not explore any more. In other words, the agent's performance becomes there as best as it can get. Therefore, it can be observed that the agent's performance is much worse in the case where 10^4 or 10^5 number of episodes are used. Now comparing the results for the situations of 10^6 and 10^7 number of episodes,

someone could object that the difference is negligible. Nevertheless, given the scale of the problem, even this seemingly small margin constitutes a difference of a few thousand dollars. In hypothetical scenarios where the dimensionality of the problem becomes even higher (which is true in more realistic settings), this difference could become even more significant. On the other hand, by using this logarithmic scale to compare different number of episodes required for convergence, it can definitely be assured that running these experiments for 10^8 or more number of episodes would probably be a waste of computational resources.

4.3.3 The impact of analytical outage modeling

The purpose of this subsection is to highlight the significance of accurate outage modeling when solving optimization problems in the area of expansion planning of energy systems. In order to do so, it is needed to compare two different modeling approaches for the outage events in a microgrid and explore how the optimization results are affected. The problem formulation, the mathematical model and the numerical assumptions are the same as the ones presented in the previous subsections of Section 4.3. In other words, the objective of this problem is again to derive the optimal storage expansion plans for a specific microgrid in a predetermined time horizon. However, the results are now compared under the two outage probabilistic models considered.

Firstly, the current modeling approach used in Sections 3 and 4 up to now should be explained in more detail. Let $\{N(t), t \in [0,+\infty)\}$ be the counting process which defines the outage events in the system. $N(t)$ is considered to be a Poisson Process with rate μ and therefore the number of outages at any given time $\tau > 0$ follows a Poisson distribution with rate $\mu\tau$. Moreover, the duration of each outage T is following a shifted Poisson distribution with a rate κ. It becomes clear from the

definition that all outages are assumed to be independent and identically distributed. While the assumption about the independence of each outage may seem reasonable in some cases, specific attention should be paid to the distribution under which various outages are happening. In order to do that, someone should check the data for the average duration per interruption (CAIDI) that the NY state provided in Table 8, for the case of PSEG-LI (Service 2018):

Table 4.8 CAIDI data provided by NY state for PSEG-LI, years 2012-2017

Year / **Duration**	2012	2013	2014	2015	2016	2017
CAIDI (hrs/int)	22.55	1.65	1.42	1.95	1.46	1.70

By examining the data provided in Table 4.8 it is observed that the outages do not seem to come from the same distribution. There is a huge spike in the average duration per outage in the year 2012 and the most obvious reason for that is the devastating Hurricane Sandy that happened on October 22, 2012, and affected a vast majority of the US Northeast area for prolonged periods. Therefore, it may seem reasonable to propose and test an alternative modeling for outage events, following the approach in (Tsianikas, Yousefi et al. 2020). This alternative scheme is based on the premise that outages can be broadly classified into two categories: regular and severe.

Under this taxonomy, it can still be assumed that each type of events follows a Poisson Process and outages, in general, follow a Poisson Process which is truly a superposition of these two types. Therefore, if $\{N_1(t), t \in [0,+\infty)\}$ with rate μ_1 is a counting (Poisson) process for the regular outage events and $\{N_2(t), t \in [0,+\infty)\}$ with rate μ_2 is a counting (Poisson) process for the severe outage events, $N'(t) = N_1(t) + N_2(t)$ is a superposed Poisson Process with rate $\mu = \mu_1 + \mu_2$. In that case, the probability that a random outage event comes from either of these processes

should also be defined. Therefore, let $\Pr\{Z_n = i\} = \frac{\mu_i}{\mu}$ where Z_n is the type of n^{th} event and of course $i \in \{1,2\}$ in this case. Of course, the duration of the outages can still be assumed to follow a shifted Poisson distribution but in a similar way, two distinct random variables are now defined, such as T_1, T_2 with respective rates κ_1, κ_2. For a better illustration of the abovementioned modifications, Fig. 4.6 shows how the distribution of the outage duration can potentially change:

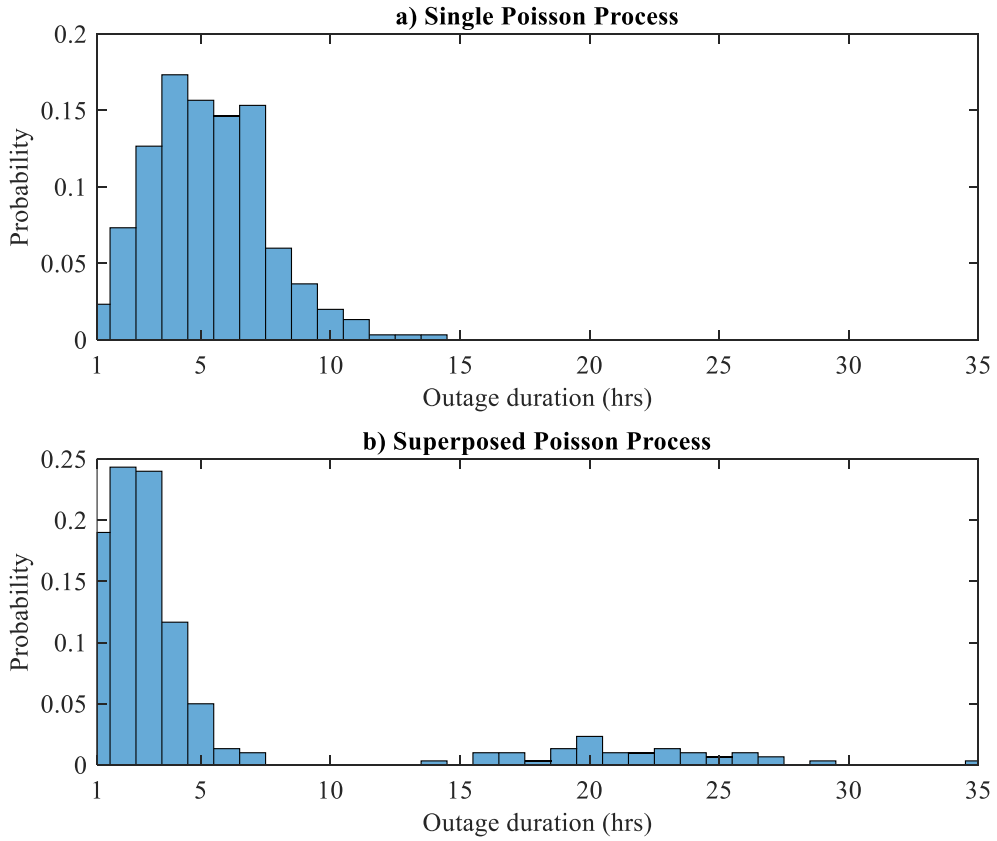

Fig. 4.6 Distribution of outage duration using two different probabilistic modeling approaches

By observing Fig. 4.6, it is more than clear that the distribution of the outage duration is significantly altered, even though the mean duration may have stayed the same.

The optimization results of the two outage modeling approaches can now be presented and compared. As already mentioned, the numerical assumptions are the

same with the ones in Section 4.3.1. The only difference is the number of episodes of the Q-learning algorithm which is now 10^6, for the sake of computational efficiency. The purpose is to showcase that the optimal policies are significantly affected by the outage modeling used. In order to explore and examine the results, specific scenarios need to be defined, in a similar way that they are defined in Section 4.3.2. These scenarios correspond to the price movements in the DTMCs that are used to define the external feature of the state space. By using fixed price trajectories for both approaches, it is possible to compare them in an unbiased way: that means, the difference in the optimal policies can be attributed exclusively to the different outage models. The scenario that is going to be examined here refers to a 420→310→167→150 $/kWh price trajectory for the Li-ion storage unit, 142→115→77→65 $/kWh for the lead-acid battery, 385→255→120→95 $/kWh for the vanadium redox and 3100→2600→1950→1700 $/kWh for the flywheel energy storage system. The results are presented in Table 4.9:

Table 4.9 Optimal policies for both outage models

Outage model / **Decision period**		Single Poisson Process	Superposed Poisson Process
Period 1	State	(0,420,142,385,3100,0,0,0,0)	(0,420,142,385,3100,0,0,0,0)
	Action	Do nothing	Do nothing
Period 2	State	(1,310,115,255,2600,0,0,0,0)	(1,310,115,255,2600,0,0,0,0)
	Action	Add Li-ion at 1000 kWh	Do nothing
Period 3	State	(2,167,77,120,1950,1000,0,0,0)	(2,167,77,120,1950,0,0,0,0)
	Action	Add Li-ion at 3000 kWh	Add Li-ion at 1000 kWh
Period 4	State	(3,150,65,95,1700,4000,0,0,0)	(3,150,65,95,1700,1000,0,0,0)
	Action	Add vanadium redox at 3000 kWh	Add Li-ion at 3000 kWh

It is clear from observing Table 4.9 that the optimal policies obtained from the two outage models differ significantly. The first thing that it should be noticed is the fact that the total storage capacity installed in the microgrid in the single Poisson process scenario is much higher (7000 kWh) than the corresponding total storage capacity installed in the superposed Poisson process scenario (4000 kWh). This

finding could be attributed to the fact that in the latter scenario, the vast majority of outages are not long-lasting and therefore can be handled with a moderate amount of installed storage capacity.

However, this is obviously not the only difference existing in the two optimal policies. In the single Poisson process scenario, it can be observed that the first storage investment happens one decision period earlier than in the superposed Poisson process scenario. It indicates that under this scenario, lost demand during outages is costly enough so it not efficient to just endure the outages and it is better to proceed with storage installation earlier. Finally, it should also be mentioned that in the single Poisson process scenario, vanadium redox battery is chosen in the last decision period, while this is not the case for the superposed Poisson process scenario.

5. Deep Reinforcement Learning for Power and Storage Expansion Planning

At this point, the unified framework for tackling the generalized long-term expansion planning problem should be presented, tested, and analyzed. The purpose of this framework is to provide all the necessary mathematical tools and methodologies for solving real-scale expansion planning problems. It should be mentioned here that although the analysis is again focused in the case of microgrids, the proposed approach could be easily transferred to other applications and domains in which sequential decision making is inherent.

In the first subsection, the focus is given on the main changes in the problem formulation. These changes can be divided into two main categories: firstly, the changes which refer to the updated action space of the problem, and secondly, changes related to the economical rewards associated with the agent. The action space of the problem presented in Section 4 consisted mainly of storage investment actions and this is the exact assumption that is going to be relaxed here. Concerning the reward function, besides the various components that are now added in the current formulation, it is crucial to present the way that reliability is going to be incorporated in the problem. This feature is very important as it is going to further highlight the importance of optimal planning while designing resilient systems serving critical infrastructure. In the second subsection, the details around the algorithmic approach chosen for the given problem are given and explained. This fact is closely connected with Section 5.1.1, due to the fact that the double deep Q-learning algorithm is the exact reason that enables the dimensionality enlargement of the action space. Afterwards, the emphasis is given in the case study, which is going to be presented alongside with results and discussion. The effect of incorporating reliability (in different levels of strictness) into the problem is going to be made clear via Section

5.3. In the fourth subsection, the proposed model is validated against a greedy approach. It is shown illustratively why accounting for delayed rewards is the key to unveiling successful and near-to-optimal investment and retirement policies. Finally, there is going to be a section dedicated exclusively to sensitivity analysis. Various scenarios for the future are explored and presented in accordance with their corresponding findings and discussion.

5.1 Expanded action space and restructured reward function

In this section, the updated action space and reward function are presented. As mentioned above, several components of the MDP formulation presented in Section 4 are shown to be useful in the present section, too. However, these two components (action space and reward function) are significantly altered and therefore new notations and definitions are required. At the end of Section 5.1.1, the updated definitions for the state space and the transition function are also provided.

5.1.1 Action space

The main motivation of this new model is to relax as many simplistic assumptions made in previous sections and be able to provide methodologies and frameworks capable of tackling real-scale expansion problems. Towards this direction, the focus should be given to the main options that the microgrid planner has in terms of MDP actions, as defined in Eq. (4.9). Based on this one, in every decision period, it is possible to invest in any of the available storage units of the set *SU* at any of the available predetermined levels of the set *SL*. Although this provides some flexibility to the microgrid designer, two critical types of options are omitted.

The first type refers to investment in power plants. In Sections 3 and 4, the microgrid entities responsible for local energy production, mainly renewable ones, are considered to be given and therefore there is no option from the design perspective to

affect these. In real cases though, the developed frameworks should incorporate actions related to power plants. As to microgrid entities, in this context, it means the union of power plants and storage units available in the system. Renewable investment costs are also declining and this is the exact same motive that drove the development of a dynamic model for storage expansion. The second type of actions that needs to be added deals with the retirement of existing power plants or storage units. In real cases, there are many occasions (and for many reasons) that it is deemed appropriate to retire an existing unit, no matter if it is replaced by a newer one or not. These reasons may be strictly economical (operations and maintenance costs are higher than decommissioning costs and savings) or environmental.

In order to incorporate actions such as the ones described above, it is necessary to define a new set, *PP*, which refers to the available power plant technologies in the system. Besides, the available expansion levels are now specific for each microgrid entity and defined by the corresponding sets $CL_i, \forall i \in PP \cup SU$. The updated action set is given in Eq. (5.1):

$$\boldsymbol{a} = \left(a_{i,l}, \forall i \in PP \cup SU, l \in CL_i \right) \in A \tag{5.1}$$

$$\text{s.t.} \quad \sum_{i \in PP \cup SU} \sum_{l \in CL_i} a_{i,l} \leq 1$$

$$a_{i,l} \in \{0,1\}, \forall i \in PP \cup SU, l \in CL_i$$

One last thing to be noted here is that in this context, actions refer to bringing the capacity of a specific microgrid entity at a specific level and not expanding it. This is an important difference, as it allows the microgrid planner to take retirement actions for any $i \in PP \cup SU$, if of course $0 \in CL_i$. However, adding capacity on top of an existing one is not permitted under the current model, given the fixed dimension of the state space and the existence of other internal characteristics, too. Therefore,

before proceeding with the reward function, it is required to present the definitions of the state space and transition function, updated with the new set notation introduced here:

$$
\begin{aligned}
& S = S^{tf} \times S^{ef} \times S^{if} \\
& \text{where:} \quad s^{tf} \in S^{tf} = \{1, 2, ..., K\} \\
& \qquad \mathbf{s}^{ef} = \left(s_{i,j}^{ef}, \forall i \in PP \cup SU, j \in EC_i \right) \in S^{ef} \\
& \qquad \mathbf{s}^{if} = \left(s_{i,j}^{if}, \forall i \in PP \cup SU, j \in IC_i \right) \in S^{if}
\end{aligned}
\tag{5.2}
$$

$$
s^{tf\prime} = f^{tf}\left(s^{tf}\right) = s^{tf} + 1, \quad \forall s^{tf} \in S^{tf} \tag{5.3}
$$

$$
\mathbf{s}^{ef\prime} = f^{ef}\left(\mathbf{s}^{ef}\right), \text{ where: } \left\{ s_{i,j}^{ef}, s^{tf} \in S^{tf} \right\} \text{ is a DTMC with } p_{i,j}^{ef}, \forall i \in PP \cup SU,\ j \in EC_i \tag{5.4}
$$

$$
\mathbf{s}^{if\prime} = f^{if}\left(\mathbf{s}^{if}, \boldsymbol{\alpha}\right), \text{ where: } s_{i,j}^{if\prime} = \begin{cases} s_{i,j}^{if}\left(1 - \sum_{l \in CL_i} \alpha_{i,l}\right) + \sum_{l \in CL_i} l\alpha_{i,l}, & \text{if } \ j = 1 \\ s_{i,j}^{if}\left(1 - \sum_{l \in CL_i} \alpha_{i,l}\right) + s_{i,j}^{ef} \sum_{l \in CL_i} \alpha_{i,l}, & \text{if } \ j \neq 1 \end{cases} \tag{5.5}
$$

Eq. (5.2) is the updated definition of the state space, where *EC* is the set of external characteristics of the microgrid entities and *IC* is the set of the internal ones. Some example components of the set *EC* are price, efficiency and lifetime and some example components of *IC* are capacity of the corresponding microgrid entity, current efficiency and current lifetime. The word "current" is used here to denote the characteristic of the entity currently existing in the system and to differentiate with the external characteristics, which correspond to the technologies in general. Each of these sets is again different for each microgrid entity. Eqs. (5.3) and (5.4) are the transition functions for the timing feature and the external feature of the state space, adjusted to account for the inclusion of power plants in the investment decisions. Specific focus is required for Eq. (5.5). The first thing that needs to be mentioned is

that the branching refers to whether it is the state transition of the capacity internal characteristic ($j=1$), or the rest of the internal characteristics ($j\neq 1$). Therefore, if $j=1$, the next state of this feature depends on the level l of the chosen investment action. For all the other internal characteristics (such as efficiency, lifetime etc.), the next state depends on the current state of the corresponding external characteristic, i.e., if an action is taken to invest in a particular technology, the efficiency internal characteristic is dictated by the efficiency state of the external characteristic.

5.1.2 Reward function

The reward function of the MDP formulation in Section 4, as defined in Eq. (4.13), is the sum of two components: the investment cost and the outage cost. There is an inherent trade-off between these two components which was sufficiently analyzed and explained in the previous sections of this dissertation. Herein, the reward function consists of multiple components which are simultaneously affecting and driving the agent towards the desired optimal policies. Therefore, it is deemed appropriate to present these terms in a sequential manner:

1) Investment cost:

Before proceeding with providing the definition of the investment component of the cost, it is required to note that herein there is no reason to take into account the lifetime of the microgrid entities in the investment consideration. The reason why this is the case stems from the fact that lifetime can now be included directly in the state of the problem, as it is explained in Section 5.1.1. Therefore, instead of the leasing scheme considered in Section 4, it is appropriate to consider a lending scheme, where LT refers to the loan term and is measured in years. Using this notation, the annual payment for year n from the start of investment is now:

$$P^{i}_{n,\text{annuity}} = \begin{cases} P^{i}_{\text{principal}} \dfrac{IR\left(1+IR\right)^{LT^{i}}}{\left(1+IR\right)^{LT^{i}}-1}, \text{ if } \quad n \le LT^{i} \\ 0, \text{ if } \quad n > LT^{i} \end{cases}, \ \forall i \in PP \cup SU \tag{5.6}$$

where $P^{i}_{\text{principal}}$ refers again to the principal amount corresponding to a specific microgrid entity. It should be made clear that, for a specific decision period, $P^{i}_{\text{principal}}$ depends of course on $s^{ef}_{i,1}$ where $j=1$ here is the price external characteristic. Finally, the investment cost for the k^{th} decision period is given in Eq. (5.7):

$$C^{inv}_{k} = \sum_{e \in ME} \sum_{n=1}^{y} P^{e}_{n,\text{annuity}}, \text{ where } ME = \left\{ PP \cup SU ; s^{if}_{i,1} > 0, \ i \in PP \cup SU \right\} \tag{5.7}$$

ME is simply a set containing all the microgrid entities existing in the system right now. It should be reminded that *y* is the number of years in a decision period and $s^{if}_{i,1}$ refers to the capacity internal feature.

2) Operational cost:

Investment in storage units is incentivized by the delayed rewards that the agent perceives in terms of loss of load cost savings. In order to incorporate efficiently investment actions for the power plants as well, similar incentives should be provided. One such incentive is the inclusion of operational costs (or savings in this case). Of course, during grid outages, the variable operational costs incurred should be taken into consideration. On the contrary, when there is no outage in the main grid, the total amount of electricity required to be purchased for the operation of the microgrid equals the total (critical) load net the production of these power plants that have variable costs lower than the electricity price at the moment. The rest is the savings incurred by the operation of the distributed energy resources in the microgrid. Finally, the operational cost component for the k^{th} decision period can be given by Eq. (5.8):

$$C_k^{opr} = \sum_{i\in O_{jk}} \sum_{j\in N_k} \sum_{p\in PP} P_p\left(t_{ijk}\right) VC_p - \sum_{i\notin O_{jk}} \sum_{j\notin N_k} \min\left(\sum_{g\in G} C_p^g D\left(t_{ijk}, g\right), \sum_{p\in\left\{PP; VC_p < E\left(t_{ijk}\right)\right\}} P_p\left(t_{ijk}\right) \right) \times\left(E\left(t_{ijk}\right) - eVC\left(t_{ijk}\right)\right)$$

(5.8)

It should be reminded here that N_k is the set of outages in decision period *k*, while O_{jk} is a random set of all time intervals for outage *j* in decision period *k*. VC_p is the variable cost of power plant *p* and *G* is the set of facilities in the microgrid. $E\left(t_{ijk}\right)$ is the electricity price at time t_{ijk} and $eVC\left(t_{ijk}\right)$ is the effective variable cost at time t_{ijk}.

The first term of Eq. (5.8) corresponds to the microgrid operation during outages. It is the variable cost that would be incurred by the production of its distributed energy resources. The second term of Eq. (5.8) is the one that defines the savings that would be incurred by the inclusion of power plants which have lower variable costs compared to the electricity price. It is clear that the higher the production (and the lower the variable cost) of these cost-efficient power plants, the higher the savings that the microgrid is going to receive. To conclude, the reason why the min operator is used in Eq. (5.8) is that in this context it is considered infeasible to sell electricity from the microgrid back to the main grid.

3) Outage cost:

Proceeding with the outage cost, it is now appropriate to examine the slightly modified indicator function for loss load of facility *g*:

$$\delta(t_{ijk}, g) = \begin{cases} 1, \text{ if } Q^b\left(t_{ijk}\right) + \int_{t_{ijk}}^{t_{ijk}+\Delta t} \left(p_c^b \left(\sum_{p\in PP} P_p\left(t_{ijk}\right) \right) - \frac{p_d^b}{e^b} \sum_{m=1}^{g} C_p^m D(u,m) \right) du < B^b{}_{\min} \\ 0, \text{ otherwise} \end{cases}$$

for an arbitrary $b \in SU$ and $\forall i \in O_{jk}$, $j \in N_k$, $g \in G$

(5.9)

By using Eq. (5.9), the outage component of loss load cost can be defined:

$$C_k^{los} = \sum_{g \in G} VOLL^g \sum_{j \in N_k} \sum_{i \in O_{jk}} \delta\left(t_{ijk}, g\right) C_p^g \, D\left(t_{ijk}, g\right) \tag{5.10}$$

This component is simply the second (negated) term of the reward function, as presented in Eq. (4.13). As it is seen later in the present subsection, reliability is considered separately via the method of Lagrange multipliers.

4) Operations and maintenance cost:

The fact that the action space is expanded to include retirement actions, means that in a similar way that power plant investments are incentivized by the inclusion of operational costs (or savings), this type of actions should also properly incentivized. This incentive can be provided by the inclusion of operations and maintenance (O&M) costs for every microgrid entity in the system. Therefore, an entity with high O&M costs and simultaneously low contribution to the system, it can be expected to be retired. This exact cost component can be seen in Eq. (5.11):

$$C_k^{om} = \sum_{i \in PP \cup SU} OM^i CC_k^i \tag{5.11}$$

where OM^i is the O&M rate for entity *i* and CC_k^i is the capacity of entity *i* installed at the system at the k^{th} decision period. Of course, CC_k^i is the $s_{i,1}^{if}$ feature of the state representation for the k^{th} decision period.

5) Retirement cost:

Lastly, the retirement cost needs to be defined. This is done via a single equation, although it is basically the result of decommissioning costs net any salvage value of the retired unit. Herein, it should be reminded that action $a_{i,l}$ refers to the action to invest in entity *i* at level *l*. Therefore if $l = 0$ it constitutes a retirement

action, while if $l \neq 0$, it is actually a replacement action. The retirement cost component can be seen in Eq. (5.12):

$$C_k^{ret} = \sum_{i \in PP \cup SU} \sum_{l \in CL_i} a_{i,l} RT^i CC_k^i \tag{5.12}$$

where RT^i is simply the retirement rate for entity *i* and CC_k^i is again the capacity of entity *i* already installed at the system at the k^{th} decision period.

6) Final reward function including reliability constraint:

Before proceeding with providing the updated reward function, there is one final and crucially important piece that needs to be added: reliability consideration. In the frameworks presented so far in Section 4 and Section 5, reliability is only considered indirectly via the *VOLL*. It should be reminded here that one of the competitive advantages of the microgrids against central grids is the existence of PCC (point of common coupling). This feature enables the microgrid to disconnect from the main grids in disrupting situations and therefore increases its reliability and resilience (Roberts and Chang 2018). It is imperative for this research to take a decisive step to tackle this problem by expanding the current methodologies and properly adding this feature.

Reliability is considered in terms of *LOLP* and *CCP*, as these are presented in Section 3. The *CCP* constraint for facility *g* at the k^{th} decision period is given by:

$$h_k^g(\mathbf{s}, \boldsymbol{\alpha}) = \gamma^g - \Pr\left\{ LOLP_{jk}^g \leq \beta^g \right\} \leq 0 \tag{5.13}$$

where:

$$LOLP_{jk}^g = \frac{\sum_{i \in O_{jk}} \delta(t_{ijk}, g)\, \Delta t}{T_{jk}} \qquad \forall\, j \in N_k \tag{5.14}$$

$$T_{jk} = \left| O_{jk} \right| \Delta t$$

$$\Pr\left\{LOLP_{jk}^{g} \le \beta^{g}\right\} = \frac{\sum_{j\in N_k} \lambda_{jk}^{g}}{|N_k|}, \quad \text{and} \quad \lambda_{jk}^{g} = \begin{cases} 1, & \text{if } LOLP_{jk}^{g} \le \beta^{g} \\ 0, & \text{otherwise} \end{cases} \tag{5.15}$$

where β^{g} is the constraint limiting $LOLP_{jk}^{g}$ and ζ^{g} is the constraint limiting $CCP_{k}^{g} = h_{k}^{g}(\mathbf{s},\boldsymbol{\alpha})$.

It should be noted here that the MDP formulation and reinforcement learning algorithms are well suited to solve unconstrained optimization problems. However, in this case, the maximization problem that needs to be solved, according to Eq. (2.27) and the definitions provided in Eqs. (5.7) - (5.13), is the following:

$$\max\ G_0 = \sum_{k=1}^{K} \gamma^{k-1} r_k(\mathbf{s},\boldsymbol{\alpha}) \tag{5.16}$$

$$\text{s.t.}\quad h_{k}^{g}(\mathbf{s},\boldsymbol{\alpha}) = \zeta^{g} - \Pr\left\{LOLP_{jk}^{g} \le \beta^{g}\right\} \le 0 \tag{5.17}$$

$$\text{where}\quad r_k(\mathbf{s},\boldsymbol{\alpha}) = -C_k^{inv} - C_k^{opr} - C_k^{los} - C_k^{om} - C_k^{ret} \tag{5.18}$$

In order to convert this constrained maximization problem to an unconstrained one, it is appropriate to use an approach called Karush-Kuhn-Tucker (KKT) approach. With this approach, a new function, called the generalized Lagrangian, is introduced to the problem (Goodfellow, Bengio et al. 2016). As the first step in this process, it is needed to define the set *FP*, which contains all the feasible solutions (policies) of the constrained optimization problem:

$$FP = \left\{\pi \,\middle|\, h_{k}^{g}(\mathbf{s},\boldsymbol{\alpha}) \le 0, \forall g \in G, k \in \{1,2,3,...,\mathrm{K}\}\right\} \tag{5.19}$$

Also, by introducing the new variables λ_g, called KKT multipliers, the generalized Lagrangian is now defined as:

$$L(\pi,\boldsymbol{\lambda}) = \sum_{k=1}^{K}\left(\gamma^{k-1} r_k(\mathbf{s},\boldsymbol{\alpha}) - \sum_{g\in G} \lambda_g h_{k}^{g}(\mathbf{s},\boldsymbol{\alpha})\right) \tag{5.20}$$

It is now feasible to solve the constrained maximization problem defined in Eqs. (5.16) - (5.18) using unconstrained optimization in Eq. (5.20). It is proven that, as long as at least one feasible solution exists, then $\max_{\pi} \min_{\lambda,\lambda\geq 0} L(\pi,\boldsymbol{\lambda})$ has the same optimal solution as $\max_{\pi\in FP} \sum_{k=1}^{K} \gamma^{k-1} r_k(\mathbf{s},\boldsymbol{\alpha})$, which was of course the original goal. This follows because:

$$\min_{\lambda,\lambda\geq 0} L(\pi,\boldsymbol{\lambda}) = \begin{cases} \sum_{k=1}^{K} \gamma^{k-1} r_k(\mathbf{s},\boldsymbol{\alpha}), \text{ if } \quad \pi \in FP \\ -\infty, \text{ if } \quad \pi \notin FP \end{cases} \tag{5.21}$$

This of course guarantees that the optimal policy is always within the set of feasible policies.

5.2 Double deep Q-learning for expansion planning

An issue that naturally arises from Section 5.1.1 is that more complex and realistic planning agendas should be accompanied by equally complex and computationally efficient programming techniques. While the action space could theoretically be expanded arbitrarily, it is a requirement to create the necessary tools and frameworks that could successfully accommodate and solve problems of that scale. Towards this direction, techniques and algorithms should be proposed to build models that are not directly affected by the size of the action spaces. In realistic scenarios, the microgrid owners should be able to choose between many investment options and at various levels. However, creating problems of such complexity could make their solutions unreachable; in that case, it is required either to compromise and make simplistic and non-realistic assumptions, or to devise innovative ways to tackle these problems in their entirety. Choosing the latter option makes the research extension of the current section a necessity.

Therefore, it is now time to dig deeper into the exploration of the deep learning-based algorithmic techniques that are presented in Section 2.5.3. More specifically, it is appropriate to solve the expanded long-term planning problem using advanced reinforcement learning approaches, such as double deep Q-learning (van Hasselt, Guez et al. 2016). While the implementation of the double Q-learning algorithm helps to solve definitely the overestimation bias problem of the single Q-learning approach, the inclusion of deep neural networks in this framework accomplishes something even more important: it makes it feasible to disassociate the size of the action space with the derivation of optimally trained Q-tables.

It should be reminded here that, as mentioned in Section 2.5.3, the key elements of deep Q-learning are two: experience replay and target network (Choudhary 2019). Experience replay refers to the way that the deep Q network selects the data to be trained on. The experiences (sequences of state, action, reward, next state) are first stored in a buffer and recovered later in a randomized way. This fact decreases the correlations between training samples. As for the target network, it exists because, if a single network is used both for predictions and targets, the whole training process could be unstable (Choudhary 2019). Consequently, a target network is used separately and has most of the time its' parameters constant. At regular intervals, these are updated by a simple copy of the main deep Q network parameters. Specifically, for the case of the double deep Q-learning algorithm, this two-network approach is very convenient, since the algorithm is already designed this way: the two tables existing in double Q-learning can now serve as the two Q networks in the deep learning version of this approach.

Before proceeding with presenting the proposed approach, it is necessary to explain the role of the simulation-based approach for estimating the outage cost

component, C_k^{los}. In Section 4, synthetic datasets are constructed via the help of the simulation-based method, in order to eventually construct a model that it would be able to produce estimates for the outage cost, given a specific configuration of the system. The whole purpose of the metamodeling technique was to tackle the overestimation bias problem of the single Q-learning algorithm. However, this is no longer needed in the current section, since the double Q-learning approach is specifically designed to deal with this issue (van Hasselt 2010). This fact does not mean that the simulation-based approach is not used in the current context, because it is still needed to provide estimates for the outage cost component, C_k^{los}. Besides, the inclusion of the reliability constraint in Eq. (5.13) requires the existence of a systematic way to assess the reliability performance of different system configurations. The number of simulation runs depends on the desired level of confidence in the estimations, as well as on the level of computational resources. In order to obtain a holistic view of the current algorithmic technique used in this section, the reader can refer to Table 5.1:

Table 5.1 DoubleDQN algorithm with supporting simulation-based approach

Algorithm: Double deep Q-learning with use of simulation-based techniques

initialization: main network Q^A , target network Q^B and experience replay buffer *b*
for *every training episode* **do:**
    initialization: starting state *s*
    for *every decision period* **do:**
        select action α based on Q^A and ε-greedy policy
        simulate the system *n* times and obtain estimates for C_k^{los} and h_k^g
        compute reward $r = \min_{\lambda,\lambda\geq 0}(\gamma^{k-1} r_k - \sum_{g\in G} \lambda_g h_k^g)$ and observe next state s'
        store (s, α, r, s') in experience replay buffer *b*
        if *experience replay* **do:**
            sample experiences from the replay buffer *b*
            for *every experience in sample* **do:**
                $Q^*(s,a) \leftarrow r + \gamma Q^B\left(s', \text{argmax}_a Q^A(s',a)\right)$

13: **end for**
14: train Q^A on sample by minimizing $\left(Q^*(s,\alpha)-Q^A(s,\alpha)\right)^2$
15: **end if**
16: **if** *update target* **do:**
17: copy main network Q^A weights to target network Q^B
18: **end if**
19: $s \leftarrow s'$
20: **end for**
21: **end for**

The simulation approach comes into play in lines 6-7 of the algorithm, where it is required to obtain estimates for the outage component of the cost, as well as for the achieved reliability of the system. Experience replay is performed at predetermined intervals and is shown in lines 9-15. The same applies for the update of the target network parameters which is displayed in lines 16-18. Finally, it should be noted that the actual training of the main Q network happens in line 14.

5.3 Case study

In this section, results are presented from a case study conducted for a microgrid located again in Westhampton, NY. Location-specific demand and meteorological data are used (NREL 2013, NREL 2016). This microgrid is considerably larger in terms of number of facilities and electricity demand (approximately 3.5 MW with a 1% annual growth rate). Furthermore, in this section, there are various external and internal features considered. More specifically, concerning the external variables, the following are included: price, lifetime, efficiency and depth-of-discharge (applicable only for storage units). As for the internal variables, these are: installed capacity, remaining life, current efficiency and current depth-of-discharge (applicable only for storage units). It should be reiterated that external features refer to all these characteristics of the technologies that the agent is allowed only to observe and not to affect. On the contrary, internal features are all the aspects of the problem that the agent can modify with its' actions.

5.3.1 Microgrid formation and numerical assumptions

The facilities that the microgrid is expected to serve are the following: hospitals, outpatient clinics, supermarkets, hotels, offices, schools, restaurants and residential houses. Concerning the power plants that the agent is able to choose amongst, are the following: solar panels, onshore wind turbines, offshore wind turbines, diesel generators and hydro power plants. Lastly, the energy storage options are: Li-ion battery, lead-acid battery, vanadium redox battery, flywheel storage system and pumped-storage hydropower. For the electrochemical storage options, degradation rate is now considered in terms of annual capacity loss (Yang, Xie et al. 2018). The reason why a calendar only, and not a cyclical as well, capacity loss is considered stems from the fact that the energy stored in the system is again intended for backup use during emergency situations, i.e. main grid outages.

By using the deep learning-based methodologies outlined in Section 5.2, it is clear that the state and action space of the problem can be enlarged, incorporating multiple power plant and storage unit options. Due to the increased dimensionality and therefore number of parameters, the data for each facility and microgrid entity considered in this case study, are presented collectively in Appendix A. The yearly data in Tables A.9 - A.18 refer to forecasts based on real past data and are obtained by using Holt-Winters forecasting method. For the Markov Chains of the external features, a 0.8 forward transition probability is assumed for all the decision periods, unless mentioned otherwise. It should be noted that in order to compare results of different test cases in Section 5.3.2 and Section 5.5 in an objective and unbiased way, the optimal policies are derived by using fixed random walks for all the external features of the problem.

Concerning the renewable energy production, solar and wind power are calculated using Eqs. (4.6) and (4.7). Specifically for the case of offshore wind power, a 20% linear reduction of wind speeds from offshore to onshore is assumed (Song, Li et al. 2018). The price of diesel is assumed to be 2.459 $/gallon (Administration 2020) and the electricity price is 0.1386 $/kWh (Administration 2020). It should also be noted that, for algorithmic purposes, the capacity internal feature is represented by the nominal power for the hydro and diesel generators (but without considering the respective efficiencies), while it is represented by the actual number of solar panels/wind turbines for the rest. In order to obtain the nominal power then, the number of solar panels/wind turbines should be multiplied with the nominal power of one solar panel/wind turbine. The horizon of the problem is still 20 years but now the decision periods are 20 (one decision period per year). As to the investment economics, a 2% annual interest rate is assumed with yearly loan payments and a loan term of 10 years.

As for the algorithmic assumptions of the problem, the algorithm of Table 5.1 is run for a total number of episodes equal to 10^6. The buffer b has a total capacity of 20,000 experiences, $\gamma = 0.9$ and the ε-greedy policy refers to a linearly decaying exploration/exploitation rate, ranging from 1 to 0.01. The neural networks (main and target) have a total of 2 hidden layers. The hidden units in each layer are 128. The optimizer is Adam (Kingma and Ba 2014) with a learning rate of 0.01. The experience replay is performed every 32 experiences and the target update every 1,000 episodes. Finally, the number of runs n of the embedded simulation approach of line 6 from Table 5.1 is set to 3.

5.3.2 Results

In order to start exploring the experimental results, a baseline test case needs to be defined. This test case corresponds to using all the technology options mentioned above, but for now assuming $\lambda_g = 0,\ \forall g \in G$. In other words, no reliability constraint has been imposed yet. The optimal policy derived for this test case is depicted in Fig. 5.1:

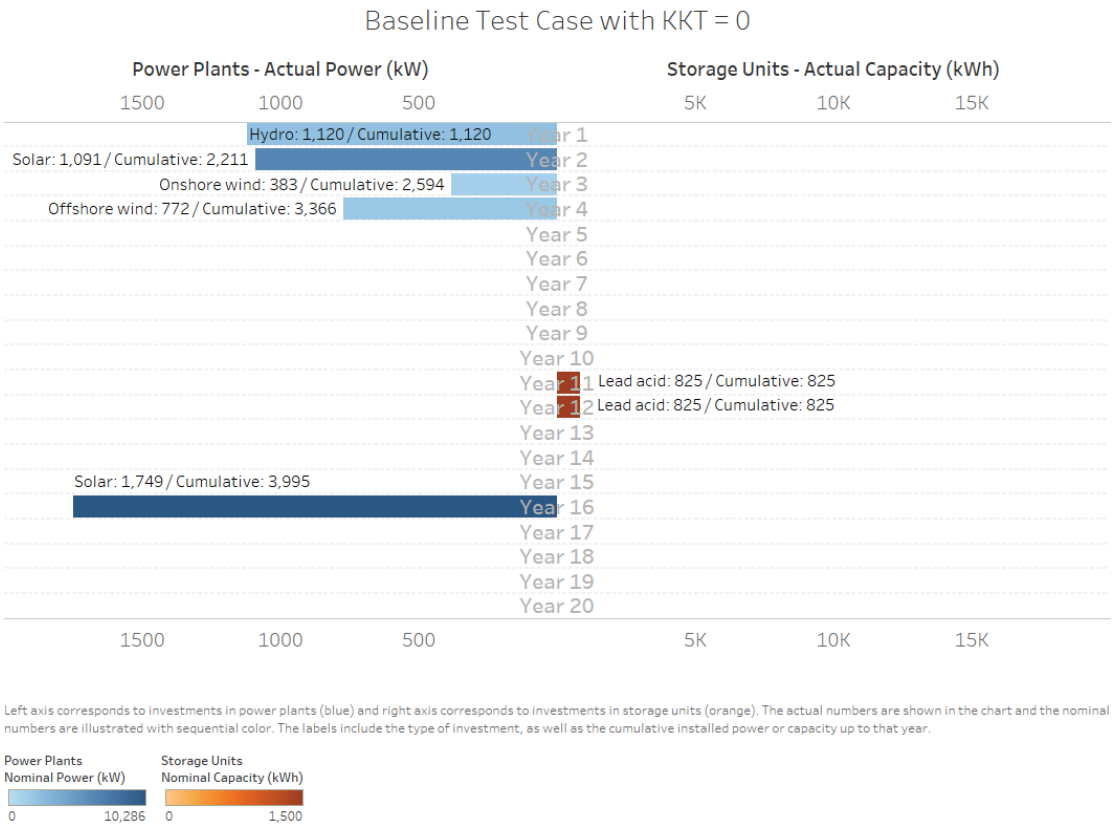

Fig. 5.1 Optimal policy derived for the baseline test case

Fig. 5.1 shows the optimal actions for every year (decision period) of the problem horizon. In the vertical axis, there are the decision periods. The investments in power plants are shown on the left horizontal axis, while the investments in storage units are shown on the right horizontal axis. The actual power corresponds to the power that a specific power plant outputs, by accounting for efficiency losses and its capacity factor. Respectively, the actual capacity corresponds to the capacity of a storage unit, by incorporating its depth-of-discharge. Although the focus should be

given on the actual power and capacity values, another two variables are illustrated in Fig. 5.1 by using color: the nominal power and the nominal capacity. These nominal values correspond to the power and capacity ratings of the power plants and the storage units respectively. In other words, the actual values are the ones someone would actually observe at the system, while the nominal values are the ones which would be obtained if capacity factor and depth-of-discharge are neglected. Finally, the labels show the type and size of investment, as well as the cumulative installed power/capacity in the system up to that year, by following the policy shown.

It can be noted in Fig. 5.1 that investments in power plants are in the first four years, as well as in year 16. This is a result that would be expected from the operational cost component of the reward function and from the fact that the total load demand for year 1 is 3,491 kW, with an annual 1% growth rate since then. The agent is trying to "save" as much as possible by investing in power plants which reduce the need to buy energy from the main grid, according to Eq. (5.8). One interesting feature is the investment in solar in year 16. It should be mentioned that this investment corresponds to a simultaneous replacement of the old solar power plant and it can be attributed to potential technology improvements and price reductions of solar energy, as per Table A.9. However, probably the most important observation coming from Fig. 5.1 stems from the negligible investment in storage; only lead acid is chosen in years 11 and 12 and in a minor size. In fact, in order to further strengthen this observation, the fixed random walk could be slightly modified, by keeping the price of lead acid storage unit constant for all decision periods. This can be achieved by fixing $p_{2,1}^{ef} = 0$, where $i = 2$ corresponds to lead acid unit and $j = 1$ to the price external feature. The optimal policy then can be seen as follows in Fig. 5.2:

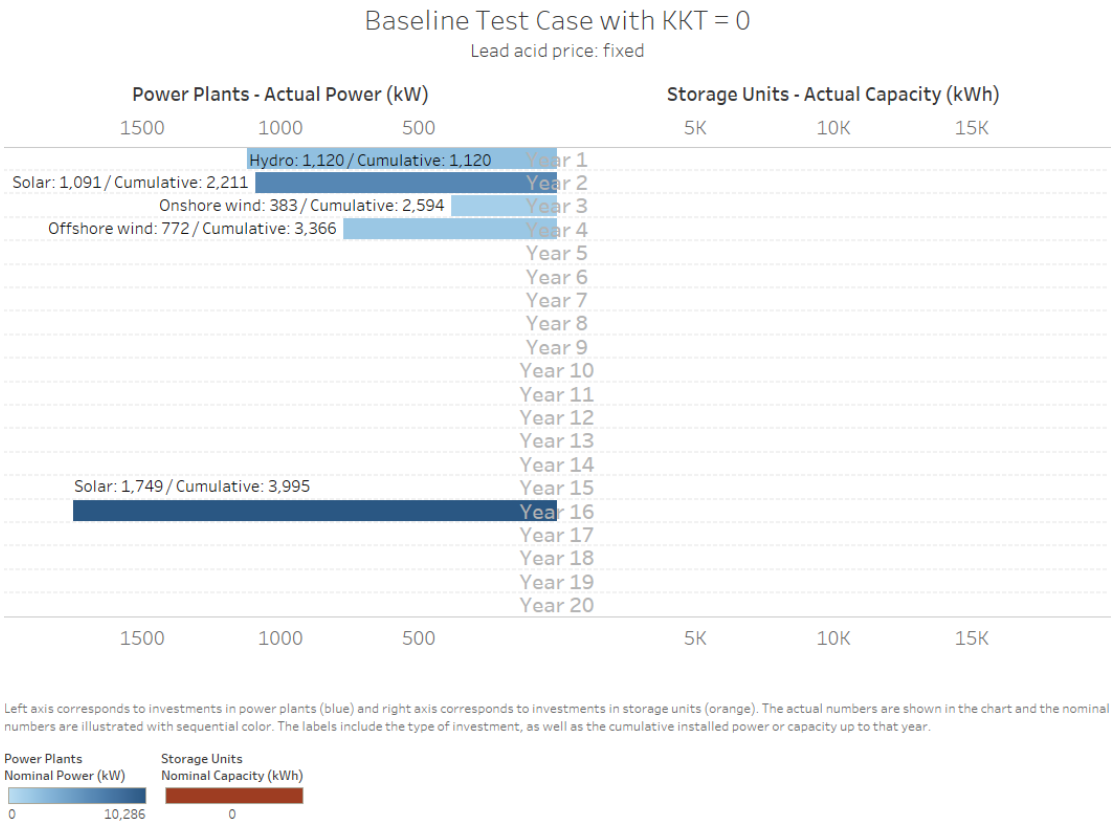

Fig. 5.2 Optimal policy derived for the baseline test case with fixed lead acid price

In this scenario, although the other investments are the same, there is absolutely zero investment in storage. It is therefore clear that the agent seems uninterested in storage investment under this baseline test case. The most obvious explanation for this behavior stems from the fact that there is still no reliability consideration. In other words, although there is a penalty for lost demand incorporated in the reward function, there is no penalty for not meeting the specified reliability criteria, as these are defined in Tables A.1 - A.8. Consequently, the next test case to be examined, is going to test exactly this feature; including reliability consideration by fixing $\lambda_g = 10^5,\ \forall g \in G$. The optimal policy should be observed in contrast with the one obtained in Fig. 5.1. The results are shown in Fig. 5.3:

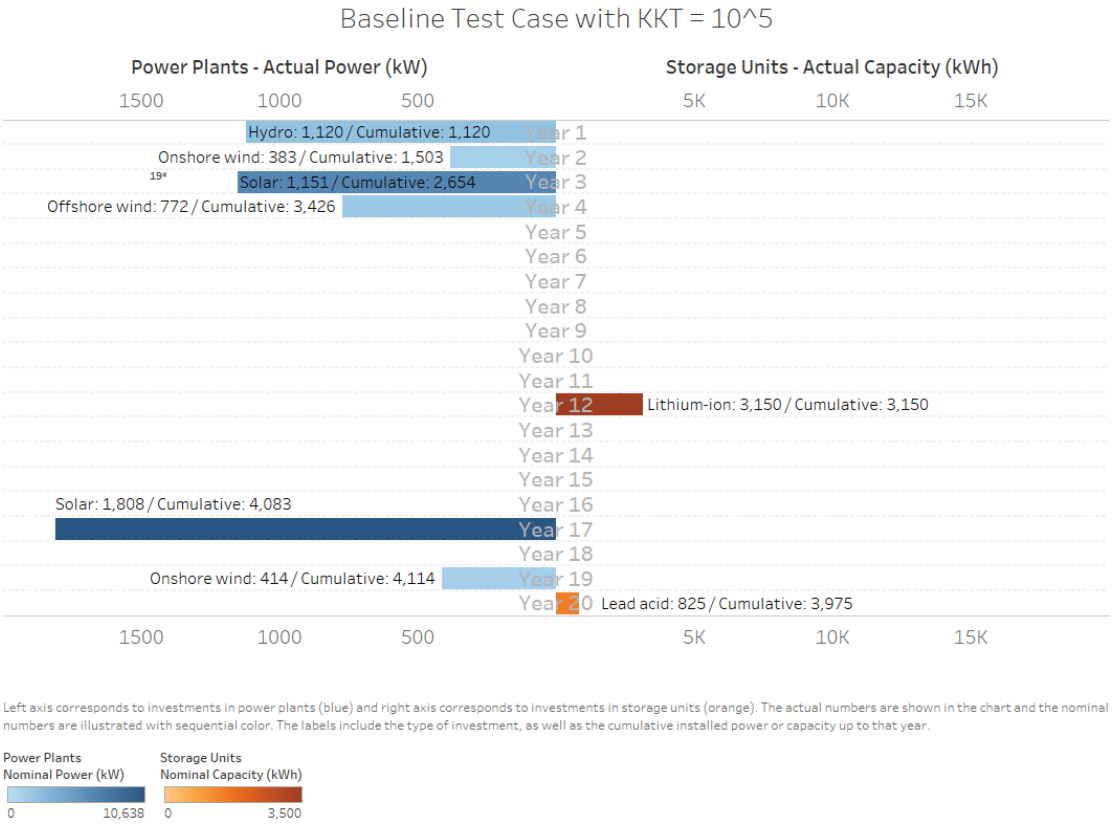

Fig. 5.3 Optimal policy derived for the baseline test case with reliability consideration

By imposing reliability constraints, the investment in storage is now more than three times higher (measured in actual capacity) compared to the original test case in year 12. Moreover, Li-ion technology is now chosen in expense of lead acid, which is of course considered a less efficient option. Last but not least, the cumulative storage capacity installed in the system is close to 4 MW, compared to 0.825 MW for the baseline test case without reliability consideration. According the power plant investments, these are almost the same between the two test cases; however, it should be mentioned that an investment in onshore wind is now taking place in year 19. This investment is due to the fact that the lifetime of the original onshore wind investment expired in year 19, and the agent decided that the optimal action is to replace the retired unit with a newer and more efficient one.

Continuing the analysis, the focus should be turned to the case of hydro power plants. The investment in this type is happening in the first year, no matter whether

reliability is incorporated or not. The credit for this choice should be given to two factors; firstly, hydro energy is one of the most cost-efficient renewable options of the recent years. Its corresponding LCOE is considered to be the lowest one among all the other options available in this case study and this can be also seen in Fig. 5.4:

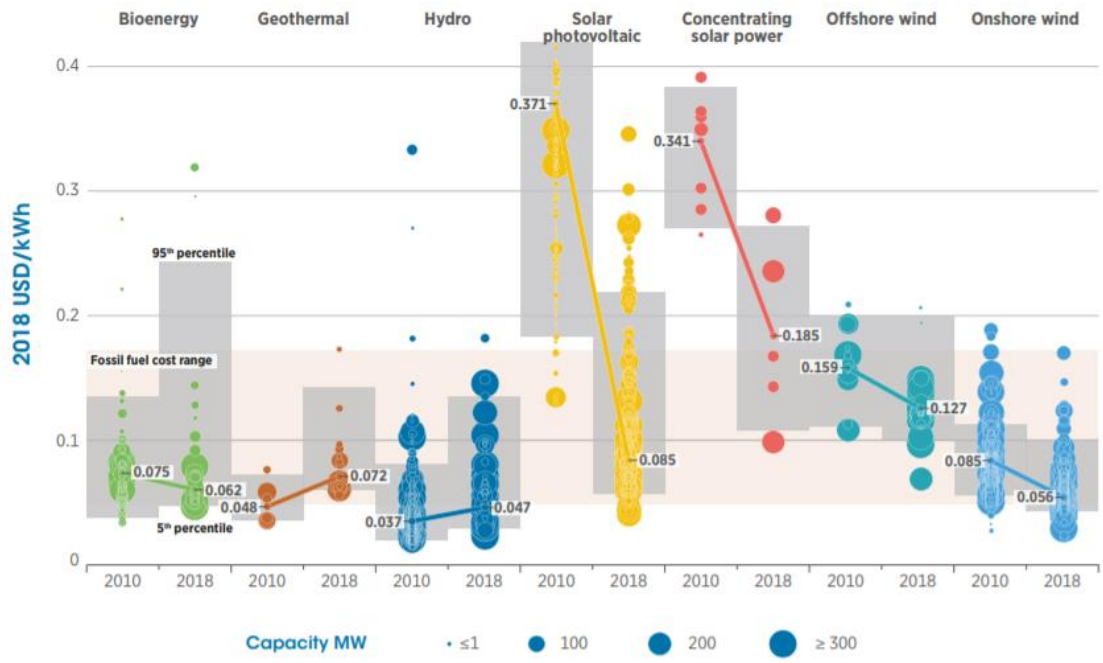

Fig. 5.4 Global LCOE of utility-scale renewable power generation technologies, 2010-2018 (IRENA 2019)

Therefore, the optimal policies obtained in Fig. 5.3 can verify the expected results coming from the industry perspective. However, the attention should be given to one peculiarity of this type of energy compared to its renewable counterparts; hydro energy is much less intermittent than solar or wind energy. The availability factor of hydro energy is close to 1, and therefore, much higher than the corresponding factor of solar or wind energy. This basically means that the agent is deciding to kickstart the investments with hydro power plants, in order not only to reduce the need for buying energy from the main grid, but also to protect the system against potential outages. In simpler terms, it can be said that hydro is taking the

major part in providing reliability for the system, in expense of storage units that could play this role. Although hydropower is a fierce competitor in the field of renewable energy, it should also be mentioned that there are disadvantages when coming to this renewable energy type: firstly, installation of hydro plants is only feasible in certain places where water flow is abundant and secondly, there are associated environmental impacts related to affecting the natural habitat, changing water flow or flooding whole areas. Consequently, the next experiment refers to the baseline test case with reliability consideration in the absence of hydropower option. The results should be compared with the ones of Fig. 5.3 and can be seen illustratively in Fig. 5.5:

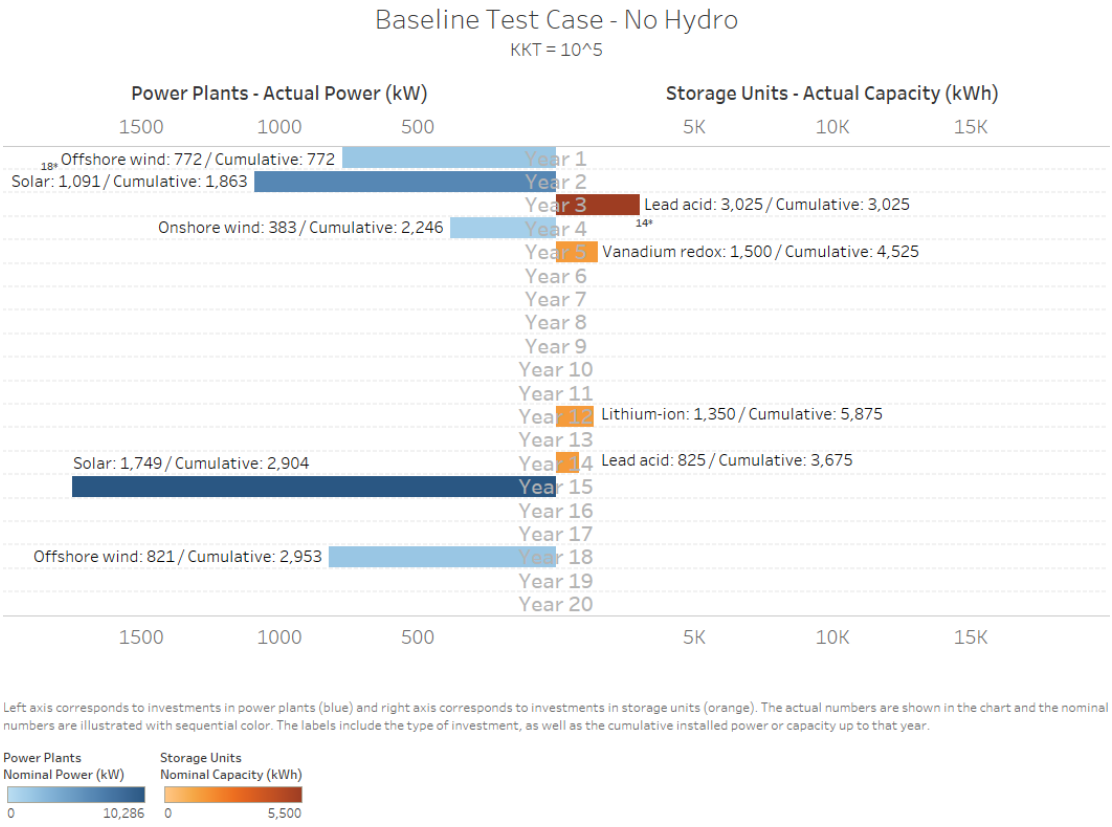

Fig. 5.5 Optimal policy derived for the baseline test case not including hydropower

The effect is drastic. Not only the cumulative storage capacity is higher (with a peak of 5.875 MW compared to 3.975 MW in Fig. 5.3), but more importantly the agent perceives a much more urgent need to invest in storage much earlier; year 3,

compared to year 12 in the original test case. This fact verifies the hypothesis made earlier; when only intermittent renewable power plants are available, the role of energy storage is (and is going to remain) crucial. Moreover, it can be seen in Fig. 5.5 that investment in vanadium redox is chosen before Li-ion in this configuration. This finding may be attributed to the delayed rewards that the agent is looking to receive, by expecting the price of the Li-ion storage unit to further decline, before it decides to invest in that.

As a final test case in this section, it is worthwhile to explore the effect of choosing the appropriate KKT multipliers. These are the way that reliability constraints are imposed in the problem, according to the generalized Lagrangian function as this is defined in Eq. (5.20). However, it is possible that the agent selects policies that occasionally violate the reliability criteria, by risking of course to endure the corresponding penalties. Therefore, in this test case the KKT multipliers are increased for selected facilities ($\lambda_g = 10^6$ for the hospital, $\lambda_g = 3\times10^5$ for the outpatient and $\lambda_g = 10^5$ for the rest). This modification is expected to result in stricter policies, when it comes to the reliability satisfaction of the selected facilities. The results can be seen in Fig. 5.6:

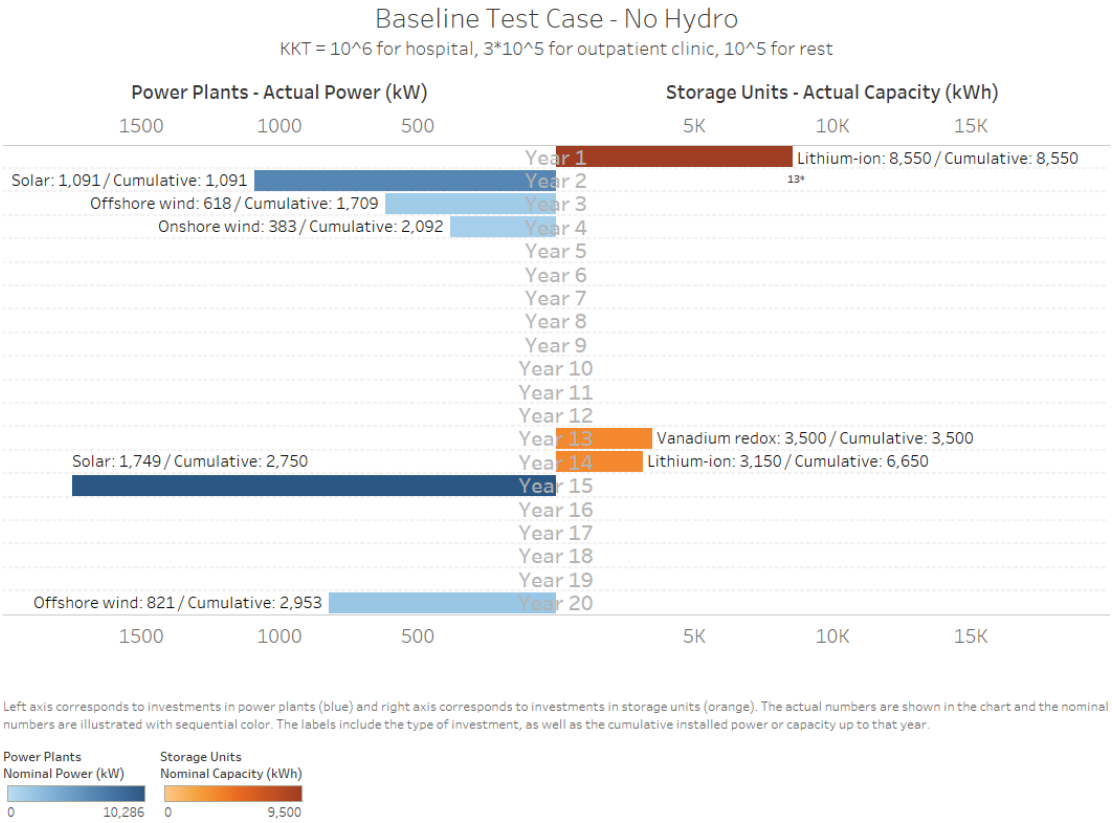

Fig. 5.6 Optimal policy derived when stricter constraints for selected facilities are imposed

The findings can again verify the initial hypothesis. Although this was a seemingly minor change, it altered the optimization results significantly. In this modified test case, the agent decides to start the investment process in storage units for the first time. On top of that, the size of investment is also higher, both in terms of the first-year investment and the cumulative ones. This result can further strengthen the argument that reliability consideration (especially when it comes to critical facilities) should play a major role when designing and testing expansion planning optimization models. For the subsequent Sections 5.4 and 5.5, this last test case of Fig. 5.6 (no hydro and stricter reliability constraints for selected facilities) is considered as the benchmark test case.

5.4 Model validation

In this section, the model is validated against a greedy approach. By greedy approach, it herein means an approach where the agent is simply selecting the action

that seems optimal for the current decision period, discarding any potential delayed rewards. This can be achieved by simply changing line 12 of the algorithm in Table 5.1 from $Q^*(s,a) \leftarrow r + \gamma Q^B\left(s', \text{argmax}_a Q^A(s',a)\right)$ to $Q^*(s,a) \leftarrow r$. It means that the output layer of the neural network does not represent the discounted sum of rewards, but instead only the reward associated with the current decision period. Of course, the use of a second target network is no longer required.

It would be expected that a greedy algorithm in this case, would always produce policies that decide to not invest in any type of power plants or storage units. This would be reasonable to assume if someone considers the high upfront costs associated with investments and that in most of the cases there is a waiting period until the investment reaches the break-even point and starts becoming profitable, if ever. In this case though, this is not exactly true and the reasons are two: firstly, and most importantly, the annual upfront costs are computed using Eq. (5.6), which is basically an amortized lending scheme. This fact, not only helps the agent in its optimal policy discovery, but also makes certain actions economically feasible from year 1. Secondly, there are also other cost components in the reward function which can drive towards that direction. As an example, which falls to this category, there are the operational costs in Eq. (5.8) which, in conjunction with the lending scheme mentioned above, can make investments in power plants profitable immediately. All the above, can be proven by observing the experimental results shown in Fig. 5.7, when a greedy approach is chosen:

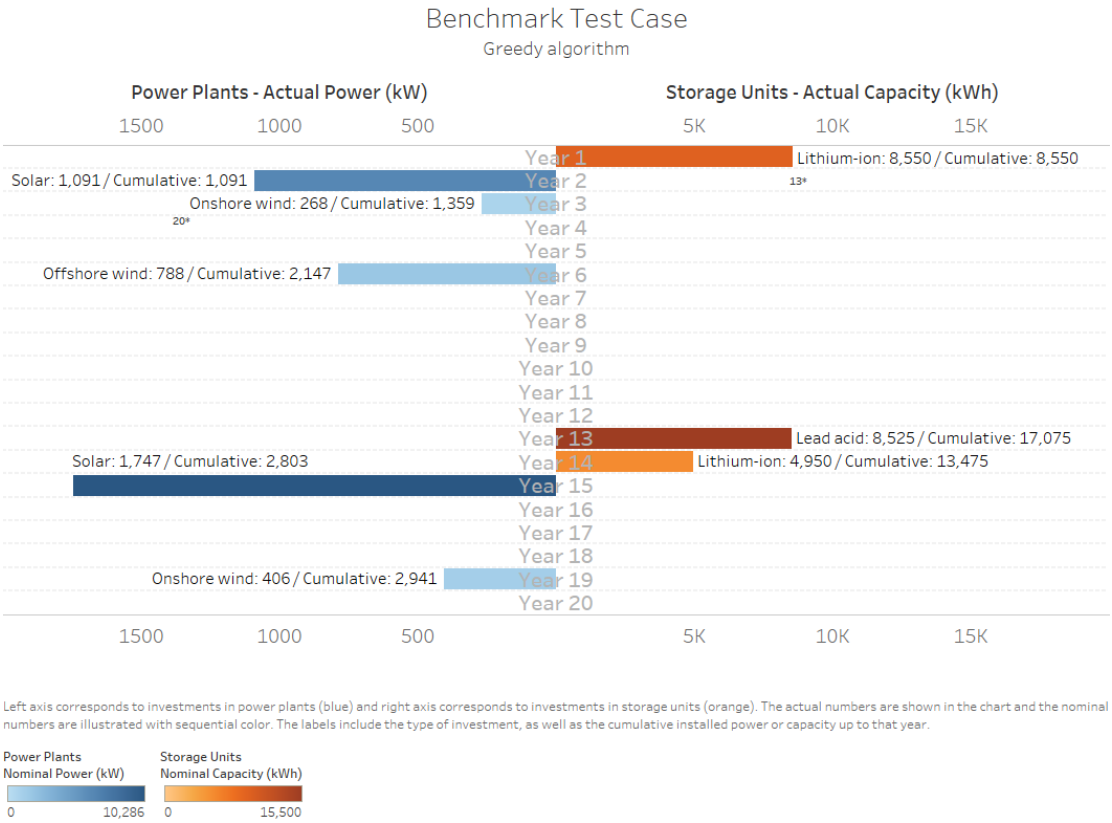

Fig. 5.7 Optimal policy derived when a greedy algorithm is chosen

Although the obtained policy seems similar to the one obtained in Fig. 5.6 using the DoubleDQN algorithm of Table 5.1, there are minor differences that can ultimately play a crucial part: larger storage investments (especially in lead acid technology) or postponing investment in offshore wind are only two of them. But the main purpose here is not to analyze the optimal policy derived from the greedy approach. Instead, the goal is to compare the actual cumulative rewards obtained by this approach and those obtained by the original DoubleDQN algorithm. For this purpose, the two approaches are tested against the same set of 10,000 simulations (random walks of the external features) of the system. In this way, it is possible to compare the two algorithms in an objective and unbiased way. The number of runs in the embedded simulation approach is now 10. The metric chosen for the comparison is the cumulative reward (in $) per random walk and the results can be shown illustratively in Fig. 5.8:

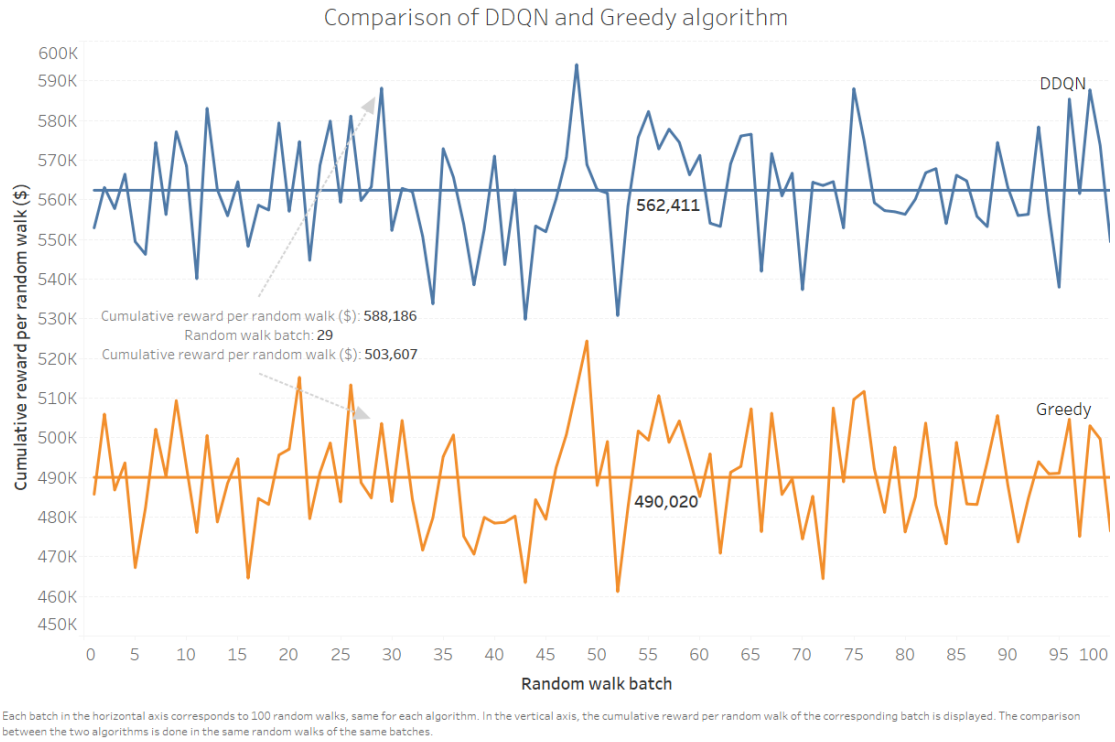

Fig. 5.8 Comparison of DDQN and greedy algorithm

The 10,000 random walks are split into 100 batches, for the sake of smoothing the results. The corresponding random walk batch is on the horizontal axis, while the cumulative reward per random walk (single random walk and not the whole batch) is on the vertical axis. Therefore, this quantity represents the expected cumulative reward for a single simulation of the system. The superiority of the DDQN algorithm can be clearly seen in Fig. 5.8. On average, the difference is slightly more than $72,000, while in some cases it can surpass the amount of $84,000. Ultimately, what is proven here, is the undoubtable dominance of these approaches that account for delayed rewards in their various configurations.

5.5 Sensitivity analysis

In this last subsection of Section 5, several scenarios about the future are explored and sensitivity analysis tests are conducted. The optimal policy derived for the benchmark test case defined in Fig. 5.6 is put in contrast with optimal policies derived for slight modifications of this test case, according to various assumptions

about the environment. Especially in situations where these alternations correspond to scenarios that will bring a significant disruption in the energy industry, it would be insightful to examine them with an even higher level of detail and rigorousness.

5.5.1 Mass adoption of electric vehicles

For example, how the load distribution will be shaped in the future by mass deployment of electric vehicles (EV) is still an open debate and no one will be able to safely predict it until it is seen in motion. However, the best preparation possible is urgently needed and therefore research has already been done in many sub-components of this general problem. Nazemi *et al.* have tried to create several EV charging profiles, depending on the location of the charging dock (Nazemi and Jafari 2019). Their results can be seen in Fig. 5.9:

Fig. 5.9 Several EV charging profiles created by simulation (Nazemi and Jafari 2019)

While the amount of additive electricity demand can be predicted with a satisfactory confidence level, the exact times that this demand will be added to remains still unknown. This fact, combined with the intermittent nature of the renewable energy technologies used in most microgrids, necessitates the analysis

required to be done towards this direction. Therefore, the first component of this sensitivity analysis work corresponds to adding EV loads in certain facilities of the considered microgrid. More specifically, 20 electric vehicles are considered per hotel facility (building 2 type of Fig. 5.9 assumed), 10 electric vehicles per office facility and 2 per residential household. The results are presented in Fig. 5.10:

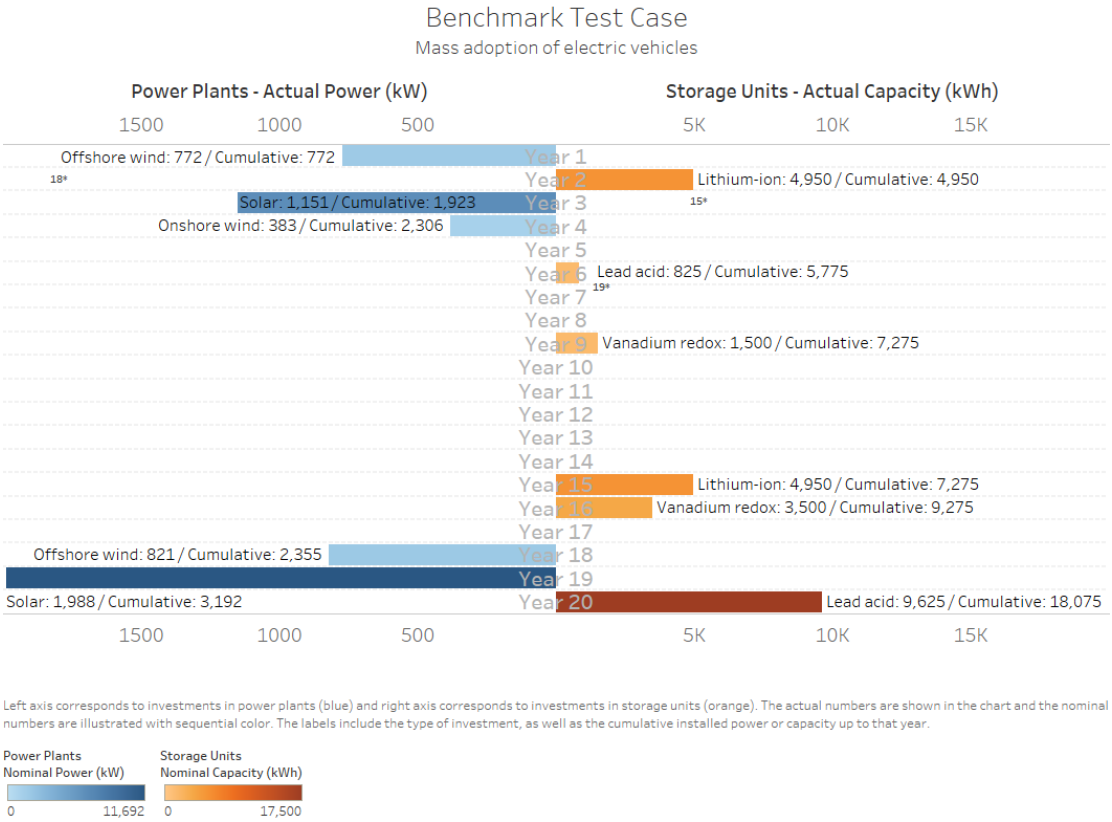

Fig. 5.10 Optimal policy derived when assuming mass adoption of electric vehicles

The first and most obvious finding of Fig. 5.10 has to do with the investment in storage units: not only the cumulative installed capacity is higher for most of the decision periods of the horizon, but also a certain technology (vanadium redox) is replaced in year 16, long before the expiration of the previous unit's lifetime. However, there is another hidden attribute of Fig. 5.10 that may require increased attention. It should be noticed that investment in solar is for the first time coming third in the order of the agent's preferences. Instead, investments in offshore wind and Li-ion storage are preceding. This fact could be attributed to the absence of solar energy

during the day and at the same time the distributional shift to nightly electricity loads according to Fig. 5.9 (vast majority of EVs are found in residential households). Wind energy, in rough terms, present a complimentary profile compared to solar energy: abundant during the night and less significant during the day. Consequently, if such a distributional change is realized, it may be the case that wind energy takes a lead among the other various renewable energy sources.

5.5.2 Value of lost load as a function of outage duration

In most of the analyses so far, the value of lost load $VOLL^g$ for a facility $g \in G$ is considered constant, no matter how many hours g has gone with lost demand. However, in real cases, it may be the case that an outage is increasingly catastrophic for a facility as the time since the start of it goes by. This may be due to critical equipment existent in the facility, damage in food supplies, customers lost etc. The purpose of this section is to relax this exact assumption by relating the actual value of lost load with the outage duration and the consecutive hours with unmet demand. In order to achieve this goal, it is required to define an updated indicator function δ_c, which is a modified version of Eq. (5.9) and is basically a counter of the hours with unmet demand for a given facility:

$$\delta_c(t_{ijk}, g) =$$

$$\begin{cases} \delta_c(t_{(i-1)\,jk}, g) + 1, \text{ if } Q^b\left(t_{ijk}\right) + \int_{t_{ijk}}^{t_{ijk}+\Delta t} \left(p_c^b \left(\sum_{p \in PP} P_p\left(t_{ijk}\right) \right) - \frac{p_d^b}{e^b} \sum_{m=1}^{g} C_p^m D(u,m) \right) du < B^b{}_{\min} \\ 0, \text{ otherwise} \end{cases}$$

$$\text{for an arbitrary } b \in SU \text{ and } \forall i \in O_{jk},\ j \in N_k, k \in K, g \in G \tag{5.22}$$

Having defined an indicator which outputs the number of consecutive hours with unmet demand for a given facility, what is left now is to update the actual outage

cost component of Eq. (5.10) with an equivalent which would represent the actual value of lost load as a function of this newly created indicator. The purpose of this section is not to do a comprehensive analysis on the function families that could play this role, but rather do a sensitivity analysis on what would change in the optimization results with an increasing function of value of lost load. Therefore, an exponential function is chosen and the updated outage cost component for the k^{th} decision period can be found in Eq. (5.23):

$$C_k^{los} = \sum_{j \in N_k} \sum_{i \in O_{jk}} \sum_{g \in G} \delta\left(t_{ijk}, g\right)\left(VOLL^g + e^{z\delta_c\left(t_{ijk}, g\right)}\right) C_p^g \, D\left(t_{ijk}, g\right) \tag{5.23}$$

Taking into account the results of Fig. 4.6, which show the distribution of outage duration under the superposed Poisson process, Eq. (5.23) could give drastically different outage costs compared to its predecessor of Eq. (5.10). It should be noted that z in Eq. (5.23) is simply a multiplier which would control this component to not explode. Table 5.2 shows the value of the quantity $VOLL^g + e^{z\delta_c\left(t_{ijk}, g\right)}$ for various choices of z and different values of $\delta_c(t, g)$, in the case of a facility g with $VOLL^g = 25$ (as the hospital in this present case study):

Table 5.2 The effect of the choice of z in the updated value of lost load function

$\delta_c(t,g)$ / z	15	20	25
0.2	45.09	79.60	173.41
0.3	115.02	428.43	1833.04
0.4	428.43	3005.96	22051.47

Therefore, the optimal policy derived with the updated value of lost load function, and a choice of $z = 0.2$ can be shown in Fig. 5.11:

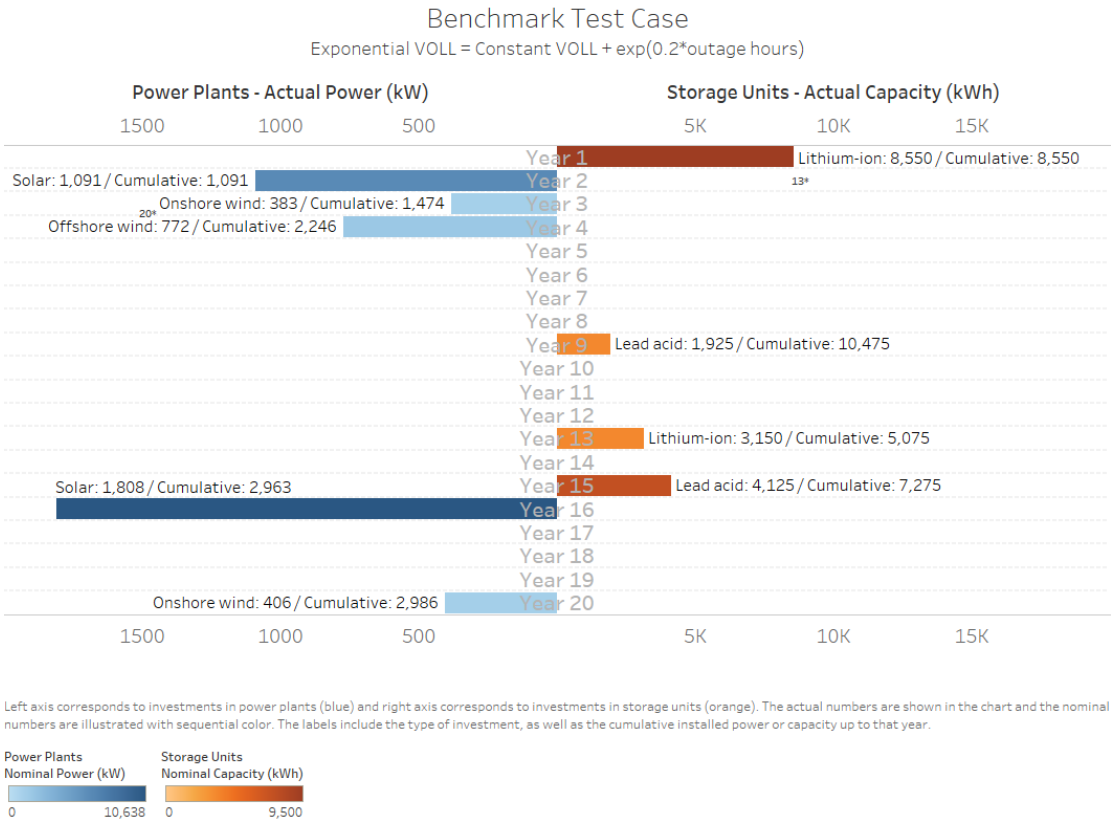

Fig. 5.11 Optimal policy derived assuming exponential VOLL with z = 0.2

The investment in storage in year 1 remains exactly the same, however in year 9 there is an additional investment in lead acid unit, meaning that the need for energy storage is indeed more urgent. However, the differences after year 15 are smoothed out (7.275 MW total capacity installed compared to 6.65 MW in Fig. 5.6). It would be interesting to examine also the case where $z = 0.3$ in Fig. 5.12:

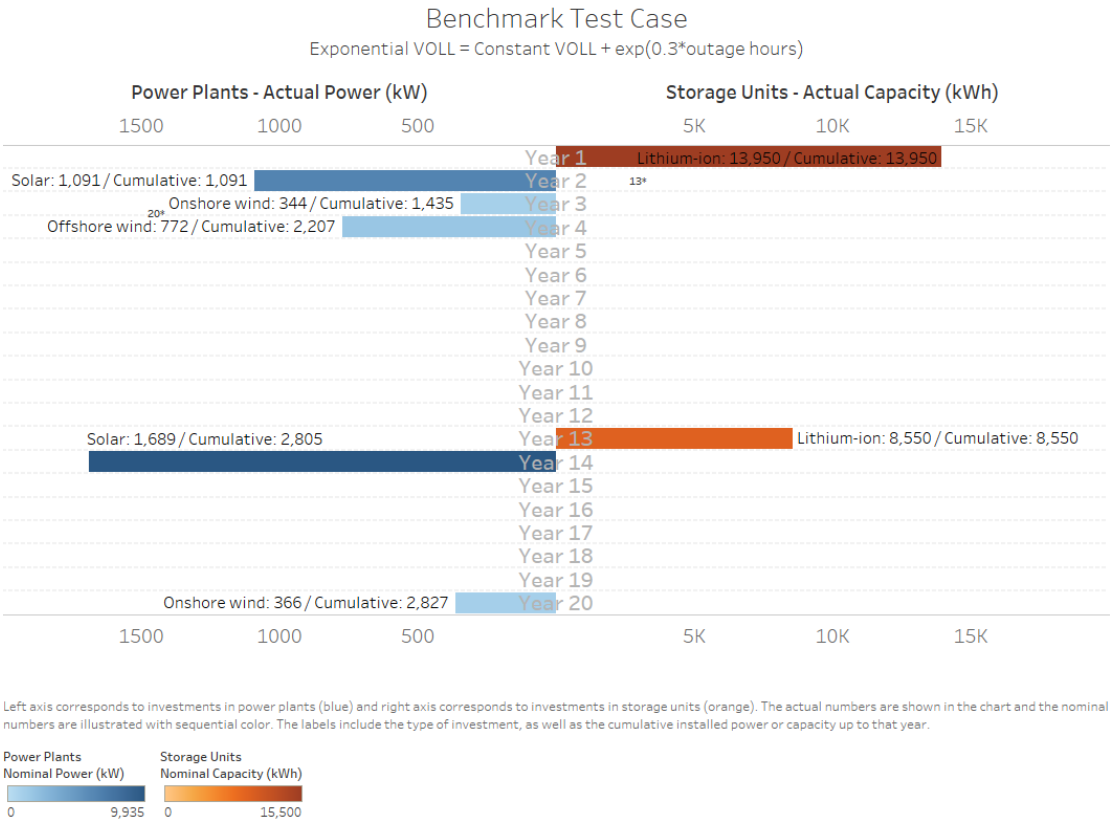

Fig. 5.12 Optimal policy derived assuming exponential VOLL with $z = 0.3$

Fig. 5.12 showcases the effect of the exact choice of exponential function to represent the actual value of lost load. In year 1, the investment in Li-ion storage is the highest ever seen in any experiment before. However, again after year 13, the differences in the cumulative installed capacity are not as significant when compared to the results of Fig. 5.6. This could be another argument on how important energy storage is, especially in the absence of sufficient renewable energy resources. Lastly, it should be noticed that in the case where $z = 0.3$, the renewal of the solar panels takes place 2 periods earlier than when $z = 0.2$ in Fig. 5.11 and 1 period earlier than the original benchmark test case in Fig. 5.6. It means that not only energy storage, but also newer and more efficient renewable energy power plants can help in the battle of fighting potentially catastrophic events.

As a side note, it would be worthwhile to highlight how sensitive is the training process of the algorithm, when exploding rewards take place. In Table 5.2, it

is shown that when $z = 0.4$, the actual value of lost load and consequently the outage cost component C_k^{los} are significantly different than those for the other values of z. In that case, the agent receives exploding rewards in many decision periods that ultimately cause him to not be able to converge in the optimal policy. The training process of course becomes noisy and the output policy is clearly not the optimal one. All the above can be observed in Fig. 5.13:

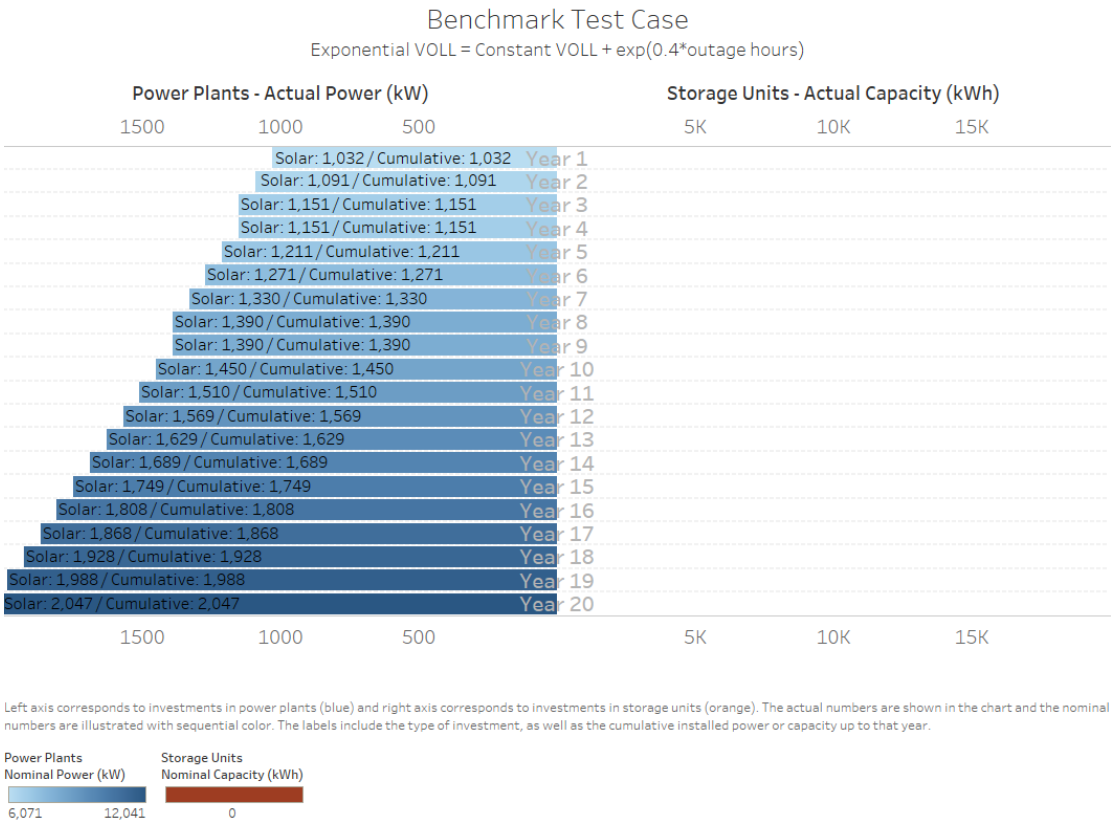

Fig. 5.13 Optimal policy derived assuming exponential VOLL with z = 0.4

The agent is proposing investment in solar energy in every decision period, although this would require to replace the older units every year. Although this would result in increasing the actual power in the microgrid (due to advancements in efficiency of solar energy technology), it should eventually result in significantly lower cumulative rewards when compared to the previous test cases. This is of course due to the retirement cost component as defined in Eq. (5.12) and the total absence of

energy storage units in the proposed policy. Consequently, it can be concluded that this is a case where the algorithm did not converge due to exploding rewards.

5.5.3 Increasing battery prices

Most of the analyses, results, and modeling so far relied on a specific assumption; that, in general, energy storage investment costs are declining and will continue to decline in the future. This assumption is based on industry reports and research works such as (IRENA 2017) which considered that this is the most likely outcome for the future. The adoption of this assumption in the present dissertation can be seen in Tables A.14 - A.18.

Although there is no reason to question the predictions made by field experts, it is more than mandatory to mention that there are also industry reports and news which mention that a price increase is also possible, especially for the Li-ion battery. The price of cobalt, a raw material used in this specific type of batteries, rose 26% in the first quarter of 2018, following 114% annual increase in 2017, according to (Clover 2018). It should be noted here that Sections 5.5.1 and 5.5.3 are somehow correlated, given the fact that Li-ion batteries' price decrease is one of the main drivers of the potential growth of electric vehicles. Another fact that could work as an additive reason for Li-ion batteries' price to increase could be the illnesses that have been reported in cobalt mines in Congo which of course, besides the disastrous humane problems that cause, can affect crucially the production of the metal (Duff 2018). While the issue seems to relate only to this specific type of battery up to now, it seems reasonable that it can cause a more general disruption in the energy storage market. If demand for other types of batteries goes higher, then their associated prices should be expected to go higher, too.

Conclusively, all the above facts constitute some very insightful reasons to conduct further research and explore what would happen under the realization of these scenarios. It should be mentioned here that it is considered mandatory that these scenarios should not be explored without the inclusion of reliability constraints in the dynamic model. Given higher energy storage prices, acquiring new storage for a given microgrid could become economically infeasible, unless there are strong drivers from the reliability perspective. Therefore, the results should again be put in contrast with the ones obtained from the benchmark test case in Fig. 5.6. Finally, the last experiment explores the effect on the optimal policy when the price states of Li-ion battery are actually an increasing function of time, instead of a decreasing one. For the specific data, the reader should refer to Table A.14. The optimal policy derived in this test case can be seen in Fig. 5.14:

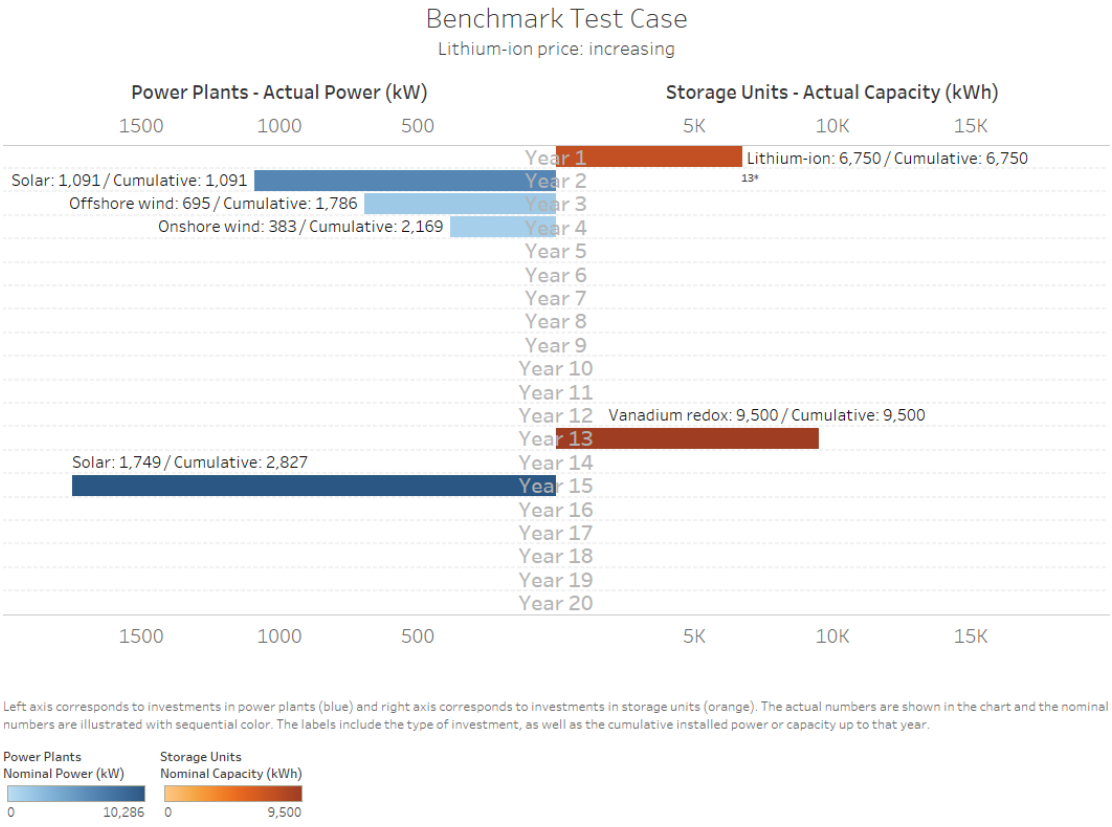

Fig. 5.14 Optimal policy derived when Li-ion price states are increasing

Fig. 5.14, compared with 5.6, presents a similar storage investment (although in a smaller size) in Li-ion technology in year 1. This is an expected outcome, since the price of Li-ion is considered equal in both test cases initially, no matter whether it is an increasing or decreasing function of time. However, the crucial difference is found in year 13. That is the time when in Fig. 5.14 a significantly larger investment in vanadium redox is chosen (9.5 MW compared with 3.5 MW in Fig. 5.6) in order to replace the old Li-ion unit. In year 14, the agent chose not to install additional storage capacity. On the contrary, in the original benchmark test case, the smaller vanadium redox investment is followed by another investment in Li-ion storage. The above result comes to verify the findings of Section 4, in which it is explained in detail how it is highly probable that these two technologies are likely to be fierce competitors in the energy storage industry in the future. Of course, as it is illustrated in Fig. 5.14, which technology will gain the upper hand depends heavily on their corresponding future costs.

To conclude with this subsection, it is necessary to highlight the importance of accurate forecasting when studying expansion problems such as the one in the current research work. It is shown in Fig. 5.14 that minor changes in the price forecasting of just one microgrid entity can significantly alter the optimization results. Of course, the same applies for any other problem parameter (such as energy load for the various facilities) whose forecast is used in the optimization frameworks and algorithms. In the next section of the dissertation, the focus is turned on the short-term energy load and how accurate forecasts for this quantity can be produced.

6. Short-term Energy Load Forecasting

Energy load forecasting is important for the efficient and reliable operation of grids, especially smart grids, which leads to uninterrupted power supply to consumers. From the perspective of the practical application, energy load forecasting can be divided into four distinct bins: (a) very short-term load forecast when the prediction horizon is in the order of the upcoming few minutes (vSTLF), (b) short-term load forecast (STLF), usually from several minutes to one week ahead, (c) medium-term forecast, usually from a week to a year ahead, and (d) long-term forecast, usually longer than a year ahead. In this section of the dissertation, the analysis is restricted to short-term load forecasting mainly for residential household demands, and ranging from several minutes up to few hours into the future. Such short-term demand prediction is often required for various applications, such as scheduling power system operation, energy balancing or energy market trading. In a nutshell, STLF plays a significant role in improving the power systems planning and operations within a power grid setup.

The purpose of the current section can be summarized in two main points. Firstly, it is explained how analytical and accurate forecasting techniques can be used in the context of the models presented in previous sections of the present dissertation and why this is another crucial step towards relaxing simplistic assumptions. In this aspect, it is also elaborated why neural network-based techniques are seemingly superior than their competitors at the moment, when assigned such tasks. Secondly, there is a detailed comparison and analysis among the most popular neural network-based techniques based on three metrics: their degree of success to predict short-term energy load, their ability to capture inherent uncertainties of typical residential load datasets and also their computational efficiency. Finally, the effect of incorporating

weather predictors in the forecasting tools is studied, as well as some peculiarities of energy load datasets at the residential level.

6.1 Motivation behind the usage of forecasted energy load

The main driver behind the need for development accurate and analytical tools for energy load prediction is the fact that in many cases (especially as the size of the grid gets larger) it has become inefficient to use real data from the control perspective (Hernandez, Baladron et al. 2014). This fact could cause discrepancies on how the proposed models are trained (if trained with real data) and tested. Additionally, the new players that are introduced in the electrical systems (Electric Vehicles, Smart Customers, Renewable Energy) are expected to provide an extra motive for demand forecasting.

Computational techniques that utilize past data and forecasted weather parameters or incorporate daily activities and appliance usages, have demonstrated reasonable forecasting accuracy for some residential and commercial building level data. These techniques range from simple linear and nonlinear regression type models, decision trees, support vector machines (SVM), autoregressive moving average models, fuzzy-logic based approaches, gradient boosting methods and artificial neural networks to name a few. Essentially, conventional models try to estimate the parameters from an assumed sequential model. However, these models are generally not expressive enough for describing the dynamics behind complex real-life data such as household level electricity consumption. With the advent of modern computing architectures such as graphical processing units (GPUs) and cheap cloud computational power, it has now become common practice to deploy machine learning-based models.

Indeed, over the past few years, deep learning techniques such as DNN, DBN and RNN have become an active research field within STLF for consumption at different levels. These deep learning techniques are mainly adopted with the task to learn layers of meaningful latent representations. Therefore, in the next subsection, a thorough comparison of neural network-based approaches is attempted to facilitate a better understanding about the relationship between STLF accuracy and architecture, weather predictors and size of facilities.

6.2 Comparison of neural network approaches

Leveraging the theoretical methodologies outlined in Section 2.4, it is considered appropriate to dive in a practical case study, following the methodology in (Tsianikas, Xie et al. 2020). To evaluate and compare the different neural network-based approaches, benchmark electricity energy load dataset and corresponding weather dataset are used. These energy load datasets are comprised of half-hourly collected observations. In the case of weather dataset with hourly data interval, it is appropriate to use interpolation techniques to fill the missing observations.

In the first subsection, more information is given on the datasets initially, the computational techniques afterwards and the evaluation metrics at last. Afterwards, the results comparison for the classical neural network approaches is conducted, where some space is also devoted to highlight the importance of utilizing weather data as well as to showcase the peculiarities of residential energy datasets.

6.2.1 Experimental setup and evaluation metrics

The chosen benchmark dataset arises from energy consumption readings for a sample of 5,567 London Households that took part in the UK Power Networks (UKPN) led Low Carbon London project (Networks 2014). This archive contains 36,460 measurements gathered between November 2011 and February 2014.

Readings are taken at half hourly intervals, i.e. values are reported in kWh/hh (per half hour). The customers in the trial are recruited as a balanced sample representative of the Greater London population. The consumption of a single household, identified with ID: MAC000002 (Network 2015) belonging to the CACI Acorn group (CACI 2014), utilizing a Standard tariff is used in this study. Additionally, meteorological service provided by Meteoblue (Meteoblue 2012) is used for easy access of consistent weather data in hourly resolution during the studied period. However, it is considered appropriate that the comparison between the neural network-based methodologies to be conducted in an additional dataset, coming from a different location. For this reason, three more energy consumption datasets are used, provided by the Office of Energy Efficiency & Renewable Energy for all TMY3 locations in United States (NREL 2013). The readings in these datasets are provided in kW (hourly). Also, meteorological data are obtained by NREL using their interactive data viewer tool (NREL 2016). The location chosen for these datasets is Central Park, NY, USA. The characteristics of all facilities used in this research work can be seen in Table 6.1:

Table 6.1 Characteristics of facilities used in this research work

Facility ID	Facility Type	Facility Location	Min Demand (kW)	Max Demand (kW)	Avg Demand (kW)
1	Residential	London	0.130	5.988	0.485
2	Residential	NYC	0.456	3.564	1.418
3	Hotel	NYC	102.997	475.391	283.637
4	Hospital	NYC	523.416	1546.208	1,136.961

In terms of the classical neural network approaches that are used and tested in this section, these are mainly four: firstly, a vanilla feedforward neural network (FFNN) is used, with densely connected layers and dropout layers between them. Afterwards, RNNs are tested and more specifically the LSTM and the GRU architectures. It should be noted here that both single and stacked versions of these networks were tested. Finally, a CNN is used where the convolutional and the pooling

layers are followed by two densely connected layers, including the output. This architecture is similar with the one used in (Van Zaen, El Achkar et al. 2018). The input window is kept fixed at 24 hours (48 half-hourly observations) and the output window varies between 0.5 hours (1 observation) and 6 hours (12 observations). The reason why different output windows are tested is that it is appropriate to examine whether the predicting performance of the neural networks deteriorates as the horizon for prediction gets longer. Furthermore, the weather predictors appended in the dataset include three variables: temperature (measured in °C), precipitation (measured in mm) and sunshine duration (measured in min). Due to the different units of the various electricity and weather variables, it is considered appropriate that the datasets are preprocessed by scaling each feature to a default range between 0 and 1. Of course, predicted results should be postprocessed by inverse transforming them to the original units. Finally, the datasets comprise observations for one whole year, which is split to train/validation using the following rule: the first 345 days are used for training purposes, where the last 20 days are used for testing. The results presented refer mostly to the performance of the algorithms in the test set.

Herein, it is useful to present the evaluation metrics used to compare the various methodologies. The main metrics are three: normalized root mean squared error (NRMSE), maximum error (maxERR) and runtime. The NRMSE corresponds to the whole testing dataset and therefore provides an objective baseline for comparison among the selected methods. In order to obtain the NRMSE, the RMSE is computed initially and then normalized by dividing with the range of the energy consumption values. On the other hand, in this context the maxERR is defined as the mean absolute error (MAE) for the top 10% of absolute errors in the testing dataset. The reason why this specific metric is chosen is because of a well-known problem when predicting

electricity demand in household level: due to the high variability of electricity usage in a household, there are certain “peaks” and “valleys” that many popular forecasting techniques are unable to capture. Therefore, it is considered appropriate to use maxERR so the neural network-based methodologies can be evaluated on this aspect, too. The maxERR is also normalized using the range of values in the whole dataset. Finally, in order to measure the computational efficiency of the different approaches, the runtime is defined as average runtime (in sec) per training epoch of the neural networks until convergence.

6.2.2 Results and discussion

In this section, the results of the classical neural network approaches are presented. Firstly, the overall comparison of the four approaches for facility 1 from Table 6.1 is done, using all three evaluation metrics mentioned in Section 6.2.1. Afterwards, it is considered necessary to highlight the importance of incorporating weather variables in the predictive models, by showing in contrast the results with and without the weather predictors. Finally, the effect of facility size (or level of aggregation) in the prediction results is shown. For this task, facilities 2-4 from Table 6.1 are used and NRMSE and maxERR are plotted with respect to the size of the facilities.

Before diving deeper into the numerical results, it should be mentioned that proper hyperparameter tuning is needed for all four neural network approaches. The final hyperparameter values that showed the best performance for each neural network architecture can be seen in Table 6.2:

Table 6.2 Selected hyperparameter values for the four approaches

	Number of layers	Units	Dropout	Filters	Kernel size	Pool size
FFNN	1	256	0.2	-	-	-
LSTM	1	128	0.3	-	-	-

GRU	1	128	0.3	-	-	-
CNN	1	256	0.2	128	2	2

It should be noted that number of layers refers to: the number of neural network layers without considering the output layer for the FFNN, the number of stacked RNN layers for LSTM and GRU and the number of 1D convolutional layers that preceded two dense layers for the case of CNN. Different values (ranging from 1-5) were tried to determine the optimal number of layers and although some deeper networks outperformed their shallower counterparts, the differences were negligible and that is the reason why the shallow networks are finally chosen. The optimizer used is Adam algorithm with a learning rate of 0.001.

As a first step in the attempted comparison, it is required to present the overall results of the aforementioned neural network techniques. Results are presented for both a short-term prediction window (0.5 hours) and a long-term prediction window (6 hours). The input window is kept fixed at 24 hours. The facility under consideration is facility 1 from Table 6.1. Results can first be seen schematically in Fig. 6.1:

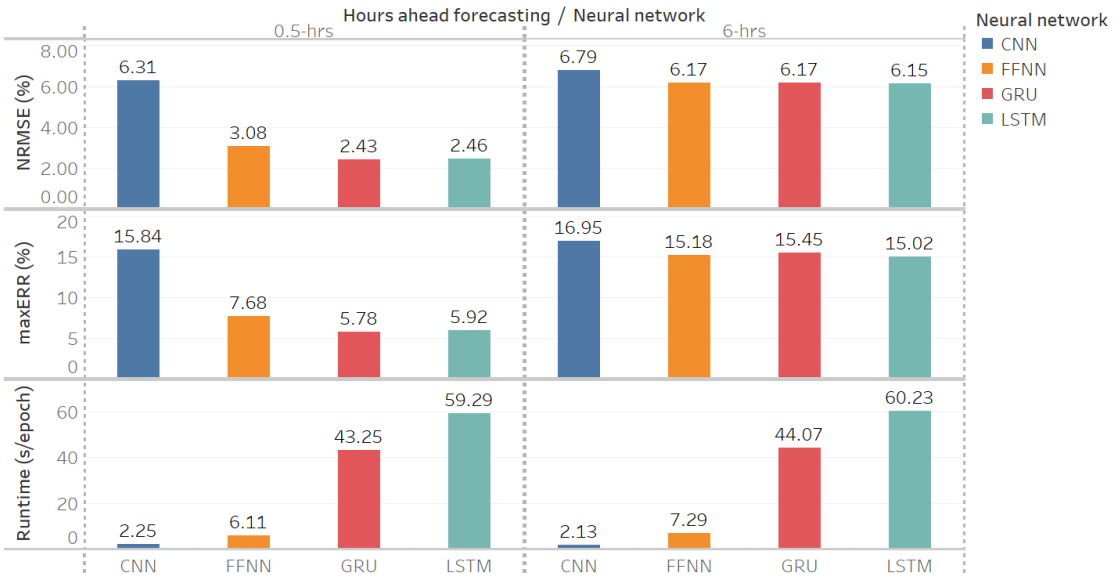

Fig. 6.1 Comparison of CNN, FFNN, GRU and LSTM using three evaluation metrics

There are several interesting observations that can be made from Fig. 6.1. First of all, the differences between the various approaches are much smaller in the long-term prediction window than in the short-term prediction window. That basically means that when the output window interval is narrowed down, there are specific methodologies that can significantly outperform others, while this is not the case for a larger output window.

Therefore, in order to compare the four methodologies, it is necessary to focus the attention to the short-term prediction window (0.5 hours). In this aspect, it seems that CNN performed worse in terms of NRMSE and maxERR, although it was the fastest approach used. The regular FFNN achieved the second lowest runtime among all approaches tested. This can be reasoned due to the number of parameters need to be optimized compared with the other types of neural networks. However, the emphasis should probably be given on the RNN methodologies (LSTM, GRU). They outperformed their competitors in both NRMSE and maxERR, although they are also slower. Based on the results presented in Fig. 6.1, it would be valid to claim that GRU achieved the best performance among all approaches tested, since it scored the lowest for NRMSE and maxERR and additionally it was notably faster than LSTM.

Proceeding with results, it is worth mentioning and proving the effect that using weather predictors has in the prediction accuracy of these models. It is important to show that this performance gain from utilizing weather information is invariant of the specific model architecture used and applies to various methodologies. In order to do so, NRMSE is presented for facility 1 from Table 6.1 for all neural network approaches by using bars for both scenarios: the NRMSE obtained without using weather data and the NRMSE obtained with using weather data. The results can be seen in Fig. 6.2:

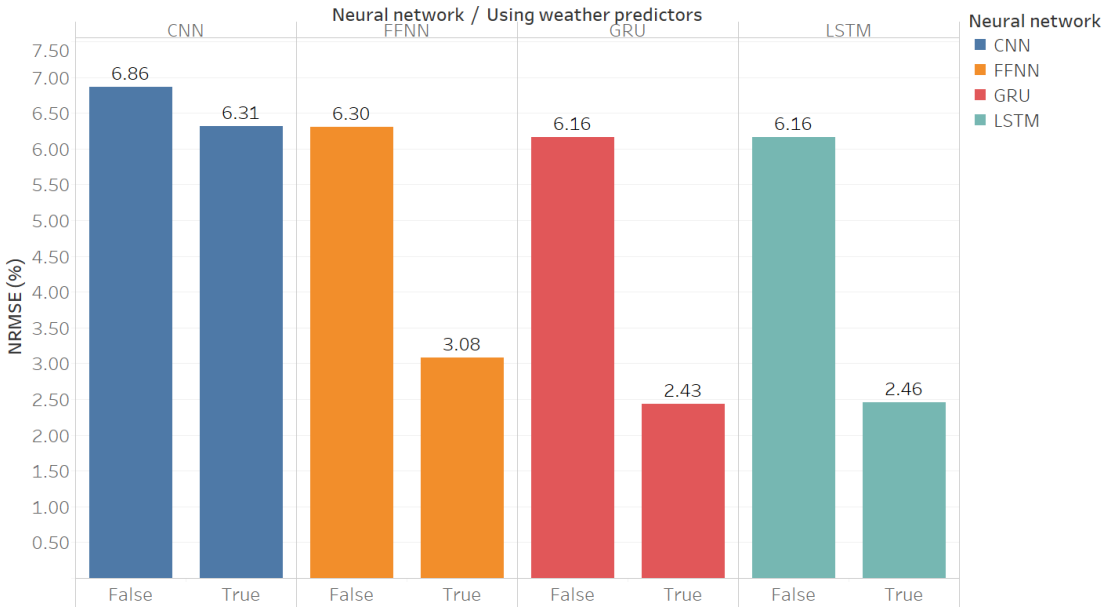

Fig. 6.2 The effect of using weather predictors in NRMSE

From Fig. 6.2, it can be seen that the hypothesis explained earlier is now verified. There is a performance gain occurring for all types of neural networks used. However, it may be useful noting that this performance gain seems to correlate with the prediction power of the neural network approach. Methodologies that achieved the best results according to Fig. 6.1 (LSTM, GRU) seemed to also present the highest performance increase when weather data are also fed to their pipelines.

To conclude with this section, it is worth digging deeper into why forecasting residential energy load is notably more difficult than forecasting for other types of facilities. For this purpose, the attention is given to the NYC-based facilities 2-4 of Table 6.1. These are three types of buildings with increasing size, in terms of average energy consumption (residential, hotel and then hospital). NRMSE and maxERR results are presented in the vertical and horizontal axis respectively. It should be mentioned here that in order for the results to be comparable amongst the various facilities, both NRMSE and maxERR are now normalized by dividing not with the range but with the average demand of facilities. The results can be observed schematically for all neural network architectures in Fig. 6.3:

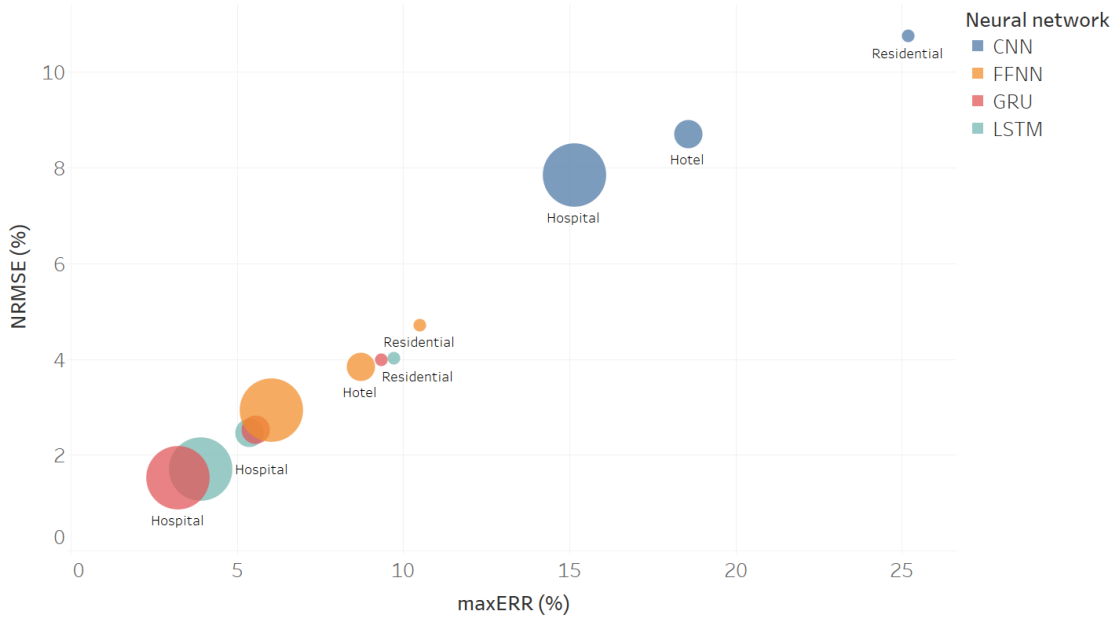

Fig. 6.3 NRMSE and maxERR for three different types of facilities

In Fig. 6.3, the size of the circle is used to denote the type of facility. Smaller circles are for residential facility, moderate-sized circles are for the hotel and the larger circles are for the hospital. Moreover, color is used in a similar fashion with Figs. 6.1 and 6.2, to denote the neural network architecture. Therefore, the purpose of Fig. 6.3 is to showcase that it is common (although not always true) to observe lower (normalized) errors as the size of the facility under consideration becomes larger. This finding applies to all four methodologies considered in this work. It justifies that usually residential facilities are more difficult to predict by nature and require developing accurate and detailed approaches for dealing with them. As a final note in Fig. 6.3, there is one observation which could be considered an outlier: the performance of the residential facility for the CNN architecture. Compared to the other datapoints, it is lying in the far upper-right corner of Fig. 6.3, shrinking the rest of them to the bottom-left.

Conclusions and Research Extensions

To sum up, the current research suggests a unified framework for the long-term microgrid expansion planning problem. By placing renewable energy sources and energy storage systems at the center of the formulation and modeling, the examined problems are solved by utilizing simulation-based techniques and reinforcement learning algorithms. The work that has been done is significant from both theoretical and applied research contributions' perspective and can successfully pinpoint to potential research extensions that could follow. The current work includes but is not limited to presenting an approach to determine optimal battery sizing with given photovoltaic arrays, investigating the relationships between critical problem parameters and developing dynamic expansion models considering various renewable energy plants and energy storage technologies. Lastly, a study on forecasting short-term residential energy load is conducted and presented. All of the above serve together the purpose of tackling successfully the multi-period, stochastic and dynamic problem of power and storage expansion planning in microgrid setting.

At this point, it would be beneficial to reiterate on the highly important research contributions. The most notable of these is definitely the formulation of dynamic expansion models, tailored specifically to microgrid use cases, which have not been extensively studied in the literature, to the best of the author's knowledge. Moreover, the introduction of modern and advanced reinforcement learning algorithms to the field of long-term energy planning is another significant novelty of the present work. Q-learning and its' neural network-based counterparts have recently been used in the energy sector, but mostly in short-term planning problems, such as unit commitment and battery scheduling. The incorporation of these techniques in the current research work have definitely a strong impact from a research perspective, as

they facilitate and enhance the process of rigorous formulation of such complex types of problems. Moreover, developing simulation-based techniques that are taking a closer look at several problem parameters, while utilizing real and location-specific data, can be considered as another novelty of this research work. Last but not least, studying thoroughly the problem of forecasting short-term residential energy load with neural network-based techniques is found to be missing from the literature up to this point. The analytical and detailed consideration and incorporation of stochastic modeling for several aspects of the problem is able not only to further strengthen the research contributions from the theoretical perspective, but also to create eventually a more realistic problem formulation compared to the ones found in the literature. However, this last task is one task for which the work should never stop and of course the goal of this research is to guide through that direction.

Towards this idea of never-ending research, it would be interesting to present some potential research extensions of the current work. These extensions can be broadly categorized in the following:

- Quantification and analysis of *VOLL*
- State space: Expand-and-extract
- Investigation of other promising RL techniques
- Incorporation of forecasting in dynamic models

According the first research extension, it should be reiterated that the current study attempts to study and model as many stochastic parameters of the problem as possible. One of them, *VOLL*, is specifically modeled as an exponential function of the outage duration and a sensitivity analysis test is conducted in Section 5. However, there is clearly a lot more room for research in that space. Proper quantification with detailed mathematical models is not only encouraged but required, if such studies in

the future aim to produce results as realistic as possible. The correlation between the monetary penalties associated with lost load and the duration of outages should be further studied and expanded. *VOLL* is one of the most critical parameters of models such as the ones presented in this research, and even slight modifications in the way it is perceived are expected to bring significant changes.

In Section 5, there is a specific focus on expanding the action space of the problem, by incorporating new types of actions into the model. However, there is an equally critical from a research perspective need to elaborate more on what is considered the ideal state representation in such problems. From one side, it is required to expand the state space of the environment, in order to account for more internal and external characteristics and to add more functionality into the model. On the other side, the attention should be given to extracting the most important of these characteristics, that would help the agent to converge faster. Convergence time is crucial from the practical perspective, since optimizing on this one would allow to solve efficiently problems of realistically large scale.

In terms of the reinforcement learning algorithms studied, the focus is clearly given to Q-learning and its' neural network-based versions. However, it is imperative that other RL techniques are also examined and compared with the ones already implemented. There is recently an increasing attention to actor-critic methods in RL, such as Deep Deterministic Policy Gradient (DDPG). Actor-critic methods are able to simultaneously estimate value functions (critic) and update accordingly the policy distribution (actor). Especially DDPG has the additive advantage that it handles continuous action spaces, which would be of course a great feature in the models. Finally, it is worthwhile mentioning Monte Carlo Tree Search (MCTS), another algorithm that is being studied and implemented recently mostly in game-based RL.

The basic idea of MCTS is to build a tree on all the possible scenarios of the simulation but explore only those that are the most promising ones. Overall, it should be mentioned that exploring new ideas, testing new algorithms and comparing them to ones that have been already implemented can only bring positive value to the subject.

The last research extension refers to how forecasting (of energy load or other problem parameters) could be used as an addition to the present models. In Section 6, there is a thorough comparison of neural network-based techniques on their forecasting short-term residential energy load capabilities. In a similar manner, many researchers are constantly developing modern and efficient techniques for forecasting other parameters related to the problem under examination. However, the specific issue of how to incorporate these forecasts in the dynamic expansion models is still to be investigated. The successful completion of such a task would further help towards the direction of relaxing as many simplistic assumptions as possible and therefore increase the robustness of the optimization models and frameworks.

The ultimate goal of this research is to highlight the awareness and attention that needs to be paid in the optimal design of microgrids and the future grids in general, incorporating renewable energy sources and storage systems. It is safe to assume that by developing analytical and detailed methodologies, such as the ones presented in this work, a decisive step towards this direction can be taken. Further research should always be encouraged and motivated and it can be stated that the current methodologies and frameworks can serve as a great driver for that purpose.

Appendix A: Supplemental data for facilities, power plants and storage units

Table A.1 Hospital facility parameters (van der Welle and van der Zwaan 2007)

Count	*VOLL* ($/kWh)	Critical load C_p	*LOLP* constraint β	*CCP* constraint ζ
2	25	0.8	0.2	0.8

Table A.2 Outpatient clinic facility parameters (van der Welle and van der Zwaan 2007)

Count	*VOLL* ($/kWh)	Critical load C_p	*LOLP* constraint β	*CCP* constraint ζ
2	19	0.8	0.2	0.8

Table A.3 Supermarket facility parameters (van der Welle and van der Zwaan 2007)

Count	*VOLL* ($/kWh)	Critical load C_p	*LOLP* constraint β	*CCP* constraint ζ
3	10	0.6	0.5	0.5

Table A.4 Hotel facility parameters (van der Welle and van der Zwaan 2007)

Count	*VOLL* ($/kWh)	Critical load C_p	*LOLP* constraint β	*CCP* constraint ζ
3	9	0.5	0.5	0.5

Table A.5 Office facility parameters (van der Welle and van der Zwaan 2007)

Count	*VOLL* ($/kWh)	Critical load C_p	*LOLP* constraint β	*CCP* constraint ζ
5	8	0.5	0.5	0.5

Table A.6 School facility parameters (van der Welle and van der Zwaan 2007)

Count	*VOLL* ($/kWh)	Critical load C_p	*LOLP* constraint β	*CCP* constraint ζ
3	7	0.4	0.5	0.5

Table A.7 Restaurant facility parameters (van der Welle and van der Zwaan 2007)

Count	*VOLL* ($/kWh)	Critical load C_p	*LOLP* constraint β	*CCP* constraint ζ
7	6	0.9	0.5	0.5

Table A.8 Residential house facility parameters (van der Welle and van der Zwaan 2007)

Count	*VOLL* ($/kWh)	Critical load C_p	*LOLP* constraint β	*CCP* constraint ζ
300	5	0.3	0.5	0.5

Table A.9 Solar energy parameters (NREL 2016, Raimi 2017, IRENA 2019)

Year # / Variable	Price state ($/kW)	Lifetime state (yrs)	Efficiency state
1	1183.51	33	0.17
2	1157.02	33	0.18
3	1130.53	34	0.19
4	1104.05	34	0.20
5	1077.56	35	0.21
6	1051.07	35	0.22

7	1024.58	36	0.23
8	998.09	36	0.24
9	971.60	37	0.25
10	945.11	37	0.26
11	918.63	38	0.27
12	892.14	38	0.28
13	865.65	39	0.29
14	839.16	39	0.30
15	812.67	40	0.31
16	786.18	40	0.32
17	759.69	41	0.33
18	733.21	41	0.34
19	706.72	42	0.35
20	680.23	42	0.36
Action levels (solar panels)	2000, 4000, 6000, 8000, 10000, 12000, 14000, 16000, 18000, 20000		
Retirement cost ($/solar panel)	17.23		
O&M annual cost rate (%)	0.63		
Capacity factor (%)	17		

Table A.10 Onshore wind energy parameters (NREL 2016, Raimi 2017, Song, Li et al. 2018, IRENA 2019)

Year # / Variable	**Price state ($/kW)**	**Lifetime state (yrs)**	**Efficiency state**
1	1484.54	17	0.48
2	1453.04	17	0.48
3	1421.54	18	0.48
4	1390.04	18	0.48
5	1358.54	19	0.49
6	1327.05	19	0.49
7	1295.55	20	0.49
8	1264.05	20	0.49
9	1232.55	21	0.49
10	1201.05	21	0.50
11	1169.56	22	0.50
12	1138.06	22	0.50
13	1106.56	23	0.50
14	1075.06	23	0.51
15	1043.56	24	0.51
16	1012.07	24	0.51
17	980.57	25	0.51
18	949.07	25	0.52
19	917.57	26	0.52
20	886.07	26	0.52
Action levels (onshore wind turbines)	1, 2, 3, 4, 5, 6, 7, 8, 9, 10		
Retirement cost ($/onshore wind turbine)	5737.94		
O&M annual cost rate (%)	0.70		
Capacity factor (%)	34		

Table A.11 Offshore wind energy parameters (NREL 2016, Raimi 2017, Song, Li et al. 2018, IRENA 2019)

Year # / Variable	Price state ($/kW)	Lifetime state (yrs)	Efficiency state
1	4290.92	17	0.48
2	4251.77	17	0.48
3	4212.61	18	0.48
4	4173.46	18	0.48
5	4134.30	19	0.49
6	4095.14	19	0.49
7	4055.99	20	0.49
8	4016.83	20	0.49
9	3977.68	21	0.49
10	3938.52	21	0.50
11	3899.37	22	0.50
12	3860.21	22	0.50
13	3821.05	23	0.50
14	3781.90	23	0.51
15	3742.74	24	0.51
16	3703.59	24	0.51
17	3664.43	25	0.51
18	3625.28	25	0.52
19	3586.12	26	0.52
20	3546.97	26	0.52
Action levels (offshore wind turbines)	1, 2, 3, 4, 5, 6, 7, 8, 9, 10		
Retirement cost ($/offshore wind turbine)	36385.23		
O&M annual cost rate (%)	1.40		
Capacity factor (%)	45		

Table A.12 Diesel generator parameters (Kozlowski 2002, Raimi 2017, Ericson and Olis 2019)

*without considering efficiency

Year # / Variable	Price state ($/kW*)	Lifetime state (yrs)	Efficiency state
1	800	20	0.39
2	800	20	0.40
3	800	20	0.41
4	800	20	0.42
5	800	20	0.43
6	800	20	0.44
7	800	20	0.44
8	800	20	0.44
9	800	20	0.44
10	800	20	0.44
11	800	20	0.45
12	800	20	0.45
13	800	20	0.45
14	800	20	0.45
15	800	20	0.45
16	800	20	0.46
17	800	20	0.46
18	800	20	0.46
19	800	20	0.46

20	800	20	0.46
Action levels (kW)	100, 400, 700, 1000, 1300, 1600, 1900, 2200, 2500, 2800		
Retirement cost ($/kW)	31		
O&M annual cost rate (%)	4.38		

Table A.13 Hydro power plant parameters (Oldham 2009, IRENA 2019, Association 2020, Company 2020) **without considering efficiency

Year # / Variable	Price state ($/kW**)	Lifetime state (yrs)	Efficiency state
1	1518.20	40	0.80
2	1544.40	40	0.80
3	1570.60	40	0.80
4	1596.80	40	0.80
5	1623.00	40	0.80
6	1649.19	40	0.80
7	1675.39	40	0.80
8	1701.59	40	0.80
9	1727.79	40	0.80
10	1753.99	40	0.80
11	1780.19	40	0.80
12	1806.39	40	0.80
13	1832.59	40	0.80
14	1858.79	40	0.80
15	1884.99	40	0.80
16	1911.19	40	0.80
17	1937.38	40	0.80
18	1963.58	40	0.80
19	1989.78	40	0.80
20	2015.98	40	0.80
Action levels (kW)	100, 400, 700, 1000, 1300, 1600, 1900, 2200, 2500, 2800		
Retirement cost ($/kW)	303.64		
O&M annual cost rate (%)	2.50		
Capacity factor (%)	50		

Table A.14 Li-ion parameters (Unterreiner, Julch et al. 2016, IRENA 2017, Marchi, Pasetti et al. 2017, Yang, Xie et al. 2018, Cole and Frazier 2019) ***right sub-column corresponds to increasing prices (Section 5.5.3), left sub-column to the baseline scenario (rest of Section 5)

Year # / Variable	Price state ($/kW***)		Lifetime state (yrs)	Efficiency state	DoD state
1	470.00	470.00	12	0.94	0.90
2	449.47	478.95	13	0.94	0.90
3	428.95	487.89	14	0.94	0.90
4	408.42	496.84	15	0.94	0.90
5	387.89	505.79	16	0.95	0.90
6	367.37	514.74	16	0.95	0.90
7	346.84	523.68	17	0.95	0.90
8	326.32	532.63	17	0.95	0.90
9	305.79	541.58	18	0.96	0.90
10	285.26	550.53	18	0.96	0.90
11	264.74	559.47	19	0.96	0.90
12	244.21	568.42	19	0.96	0.90
13	223.68	577.37	20	0.97	0.90

14	203.16	586.32	20	0.97	0.90
15	182.63	595.26	21	0.97	0.90
16	162.11	604.21	21	0.97	0.90
17	141.58	613.16	22	0.98	0.90
18	121.05	622.11	22	0.98	0.90
19	100.53	631.05	23	0.98	0.90
20	80.00	640.00	23	0.98	0.90
Action levels (kWh)	1500, 3500, 5500, 7500, 9500, 11500, 13500, 15500, 17500, 19500				
Retirement cost ($/kWh)	220				
O&M annual cost rate (%)	2.50				
Degradation annual rate (%)	1.71				

Table A.15 Lead acid parameters (Unterreiner, Julch et al. 2016, IRENA 2017, Marchi, Pasetti et al. 2017, Yang, Xie et al. 2018, Cole and Frazier 2019)

Year # / Variable	Price state ($/kW)	Lifetime state (yrs)	Efficiency state	DoD state
1	260.00	9	0.80	0.55
2	248.95	10	0.80	0.55
3	237.89	11	0.80	0.55
4	226.84	12	0.80	0.55
5	215.79	13	0.81	0.55
6	204.74	13	0.81	0.55
7	193.68	14	0.81	0.55
8	182.63	14	0.81	0.55
9	171.58	15	0.82	0.55
10	160.53	15	0.82	0.55
11	149.47	16	0.82	0.55
12	138.42	16	0.82	0.55
13	127.37	17	0.83	0.55
14	116.32	17	0.83	0.55
15	105.26	18	0.83	0.55
16	94.21	18	0.83	0.55
17	83.16	19	0.84	0.55
18	72.11	19	0.84	0.55
19	61.05	20	0.84	0.55
20	50.00	20	0.84	0.55
Action levels (kWh)	1500, 3500, 5500, 7500, 9500, 11500, 13500, 15500, 17500, 19500			
Retirement cost ($/kWh)	88			
O&M annual cost rate (%)	2.50			
Degradation annual rate (%)	1.71			

Table A.16 Vanadium redox parameters (Unterreiner, Julch et al. 2016, IRENA 2017, Marchi, Pasetti et al. 2017, Yang, Xie et al. 2018, Cole and Frazier 2019)

Year # / Variable	Price state ($/kW)	Lifetime state (yrs)	Efficiency state	DoD state
1	400.00	13	0.70	1.00
2	383.16	14	0.71	1.00
3	366.32	15	0.72	1.00
4	349.47	16	0.73	1.00
5	332.63	17	0.74	1.00
6	315.79	17	0.74	1.00
7	298.95	18	0.75	1.00

8	282.11	18	0.75	1.00
9	265.26	19	0.76	1.00
10	248.42	19	0.76	1.00
11	231.58	20	0.77	1.00
12	214.74	20	0.77	1.00
13	197.89	21	0.78	1.00
14	181.05	21	0.78	1.00
15	164.21	22	0.79	1.00
16	147.37	22	0.79	1.00
17	130.53	23	0.80	1.00
18	113.68	23	0.80	1.00
19	96.84	24	0.81	1.00
20	80.00	24	0.81	1.00
Action levels (kWh)	1500, 3500, 5500, 7500, 9500, 11500, 13500, 15500, 17500, 19500			
Retirement cost ($/kWh)	300			
O&M annual cost rate (%)	2.50			
Degradation annual rate (%)	1.71			

Table A.17 Flywheel storage parameters (Amiryar and Pullen 2017, IRENA 2017, Marchi, Pasetti et al. 2017, Yang, Xie et al. 2018, Cole and Frazier 2019)

Year # / Variable	Price state ($/kW)	Lifetime state (yrs)	Efficiency state	DoD state
1	3100.00	20	0.83	0.86
2	2989.47	21	0.83	0.86
3	2878.95	22	0.84	0.86
4	2768.42	22	0.84	0.86
5	2657.89	23	0.85	0.86
6	2547.37	23	0.85	0.86
7	2436.84	24	0.86	0.86
8	2326.32	24	0.86	0.86
9	2215.79	25	0.87	0.86
10	2105.26	25	0.87	0.86
11	1994.74	26	0.88	0.86
12	1884.21	26	0.88	0.86
13	1773.68	27	0.89	0.86
14	1663.16	27	0.89	0.86
15	1552.63	28	0.90	0.86
16	1442.11	28	0.90	0.86
17	1331.58	29	0.91	0.86
18	1221.05	29	0.91	0.86
19	1110.53	30	0.92	0.86
20	1000.00	30	0.92	0.86
Action levels (kWh)	1500, 3500, 5500, 7500, 9500, 11500, 13500, 15500, 17500, 19500			
Retirement cost ($/kWh)	50			
O&M annual cost rate (%)	2.50			
Degradation annual rate (%)	0			

Table A.18 Pumped-storage hydropower parameters (Oldham 2009, IRENA 2017, Marchi, Pasetti et al. 2017, Yang, Xie et al. 2018, Cole and Frazier 2019, Association 2020)

Year # / Variable	Price state ($/kW)	Lifetime state (yrs)	Efficiency state	DoD state
1	1000.00	60	0.80	0.90

2	989.47	60	0.80	0.90
3	978.95	60	0.80	0.90
4	968.42	60	0.80	0.90
5	957.89	60	0.80	0.90
6	947.37	61	0.81	0.90
7	936.84	61	0.81	0.90
8	926.32	61	0.81	0.90
9	915.79	61	0.81	0.90
10	905.26	61	0.81	0.90
11	894.74	62	0.82	0.90
12	884.21	62	0.82	0.90
13	873.68	62	0.82	0.90
14	863.16	62	0.82	0.90
15	852.63	62	0.82	0.90
16	842.11	63	0.83	0.90
17	831.58	63	0.83	0.90
18	821.05	63	0.83	0.90
19	810.53	63	0.83	0.90
20	800.00	63	0.83	0.90
Action levels (kWh)	1500, 3500, 5500, 7500, 9500, 11500, 13500, 15500, 17500, 19500			
Retirement cost ($/kWh)	200			
O&M annual cost rate (%)	2.50			
Degradation annual rate (%)	0			

References

Abas, N., A. Kalair and N. Khan (2015). "Review of fossil fuels and future energy technologies." Futures **69**: 31-49.

ADL (2018) "A brief introduction to reinforcement learning."

Administration, U. S. E. I. (2020). Electric Power Monthly.

Administration, U. S. E. I. (2020). Gasoline and Diesel Fuel Update.

Aggarwal, V. (2019) "What are the most efficient solar panels on the market? Solar panel efficiency explained."

Akorede, M. F., H. Hizam and E. Pouresmaeil (2010). "Distributed energy resources and benefits to the environment." Renewable & Sustainable Energy Reviews **14**(2): 724-734.

Alsaidan, I., A. Khodaei and W. Gao (2018). "A comprehensive battery energy storage optimal sizing model for microgrid applications." IEEE Transactions on Power Systems **33**(4): 3968-3980.

Alsaidan, I., A. Khodaei and W. Z. Gao (2016). "Distributed Energy Storage Sizing for Microgrid Applications." 2016 Ieee/Pes Transmission and Distribution Conference and Exposition (T&D).

Alsaidan, I., A. Khodaei and W. Z. Gao (2018). "A Comprehensive Battery Energy Storage Optimal Sizing Model for Microgrid Applications." Ieee Transactions on Power Systems **33**(4): 3968-3980.

Ambrosone, G., S. Catalanotti, U. Coscia and G. Troise (1985). "Comparison between Power and Energy Methods of Analyses of Photovoltaic Plants." Solar Energy **34**(1): 1-8.

Amin, M. (2008). "Challenges in Reliability, Security, Efficiency, and Resilience of Energy Infrastructure: Toward Smart Self-healing Electric Power Grid." 2008 Ieee Power & Energy Society General Meeting, Vols 1-11: 69-73.

Amiryar, M. E. and K. R. Pullen (2017). "A Review of Flywheel Energy Storage System Technologies and Their Applications." Applied Sciences-Basel **7**(3).

Arun, P., R. Banerjee and S. Bandyopadhyay (2009). "Optimum sizing of photovoltaic battery systems incorporating uncertainty through design space approach." Solar Energy **83**(7): 1013-1025.

Association, I. S. (2012). IEEE Guide for Electric Power Distribution Reliability Indices. IEEE.

Association, I. S. (2014). "APPA distribution system reliability & operations survey report."

Association, N. H. (2020) "Affordable."

Atilgan, B. and A. Azapagic (2015). "Life cycle environmental impacts of electricity from fossil fuels in Turkey." Journal of Cleaner Production **106**: 555-564.

Bakirtzis, G. A., P. N. Biskas and V. Chatziathanasiou (2012). "Generation Expansion Planning by MILP considering mid-term scheduling decisions." Electric Power Systems Research **86**: 98-112.

Bhattacharyya, S. C. (2012). "Energy access programmes and sustainable development: A critical review and analysis." Energy for Sustainable Development **16**(3): 260-271.

Birnie, D. P. (2014). "Optimal battery sizing for storm-resilient photovoltaic power island systems." Solar Energy **109**: 165-173.

Birnie, D. P. (2016). "Analysis of energy capture by vehicle solar roofs in conjunction with workplace plug-in charging." Solar Energy **125**: 219-226.

Bocklisch, T. (2015). "Hybrid energy storage systems for renewable energy." 9th International Renewable Energy Storage Conference, Ires 2015 **73**: 103-111.

Borovykh, A., S. Bohte and C. W. Oosterlee (2017) "Conditional time series forecasting with convolutional neural networks."

Bucciarelli, L. L. (1986). "The Effect of Day-to-Day Correlation in Solar-Radiation on the Probability of Loss-of-Power in a Stand-Alone Photovoltaic Energy System." Solar Energy **36**(1): 11-14.

Cabral, C. V. T., D. Oliveira, A. S. A. C. Diniz, J. H. Martins, O. M. Toledo and L. D. B. M. Neto (2010). "A stochastic method for stand-alone photovoltaic system sizing." Solar Energy **84**(9): 1628-1636.

CACI (2014). Acorn User guide.

Careri, F., C. Genesi, P. Marannino, M. Montagna, S. Rossi and I. Siviero (2011). "Generation Expansion Planning in the Age of Green Economy." Ieee Transactions on Power Systems **26**(4): 2214-2223.

Caruana, C., A. Sattar, A. Al-Durra and S. M. Muyeen (2015). "Real-time testing of energy storage systems in renewable energy applications." Sustainable Energy Technologies and Assessments **12**: 1-9.

Casares, F. J., R. Lopez-Luque, R. Posadillo and M. Varo-Martinez (2014). "Mathematical approach to the characterization of daily energy balance in autonomous photovoltaic solar systems." Energy **72**: 393-404.

Cesena, E. A. M., T. Capuder and P. Mancarella (2016). "Flexible Distributed Multienergy Generation System Expansion Planning Under Uncertainty." Ieee Transactions on Smart Grid **7**(1): 348-357.

Chauhan, A. and R. P. Saini (2014). "A review on Integrated Renewable Energy System based power generation for stand-alone applications: Configurations,

storage options, sizing methodologies and control." Renewable & Sustainable Energy Reviews **38**: 99-120.

Chen, C., S. Duan, T. Cai, B. Liu and G. Hu (2011). "Smart energy management system for optimal microgrid economic operation." Iet Renewable Power Generation **5**(3): 258-267.

Chen, H. S., T. N. Cong, W. Yang, C. Q. Tan, Y. L. Li and Y. L. Ding (2009). "Progress in electrical energy storage system: A critical review." Progress in Natural Science-Materials International **19**(3): 291-312.

Chen, Z. and Z. XI (2018). Reliability-based optimal design of a micro-grid system under natural disasters. ASME 2018 International Mechanical Engineering Congress and Exposition.

Cho, K., B. Van Merriënboer, D. Bahdanau and Y. Bengio (2014) "On the properties of neural machine translation: Encoder-decoder approaches."

Choudhary, A. (2019) "A Hands-On Introduction to Deep Q-Learning using OpenAI Gym in Python."

Chung, J., C. Gulcehre, K. Cho and Y. Bengio (2014) "Empirical evaluation of gated recurrent neural networks on sequence modeling."

Clover, I. (2018) "Lithium-ion battery prices could rise as cost of cobalt shoots up, warn analysts."

Coit, D. W., S. Selcuklu, N. Chatwattanasiri and N. Wattanapongsakorn (2015). "Stochastic Multiple Objective Electric Generation Expansion Planning." 2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (Ecti-Con).

Cole, W. and A. W. Frazier (2019). Cost projections for utility-scale battery storage. NREL.

Colonnese, A. (2017). "Microgrid Business Strategy: An Evolutionary Industry Perspective."

Company, W. V. I. (2020) "How Hydropower Works."

Covert, T., M. Greenstone and C. R. Knittel (2016). "Will We Ever Stop Using Fossil Fuels?" Journal of Economic Perspectives **30**(1): 117-137.

Curry, C. (2017). Lithium-ion Battery Costs and Market. Bloomberg New Energy Finance.

Dahmardeh, M. and Z. Xi (2019). "State-of-Charge Uncertainty of Lithium-ion Battery Packs Considering the Cell-to-Cell Variability." ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems.

Diaf, S., M. Belhamel, M. Haddadi and A. Louche (2008). "Technical and economic assessment of hybrid photovoltaic/wind system with battery storage in Corsica island." Energy Policy **36**(2): 743-754.

Dimeas, A. L. and N. D. Hatziargyriou (2010). "Multi-Agent Reinforcement Learning for Microgrids." Ieee Power and Energy Society General Meeting 2010.

Divya, K. C. and J. Ostergaard (2009). "Battery energy storage technology for power systems-An overview." Electric Power Systems Research **79**(4): 511-520.

DoD, U. S. (2015). Naval Facilities Engineering Command. Technology transition final public report: smart power infrastructure demonstration for energy reliability and security (SPIDERS).

DoE. (2014). "How microgrids work." from https://www.energy.gov/articles/how-microgrids-work.

Duff, L. (2018). TrendForce Expects Prices of Lithium-ion Batteries to Increase by 5~15% in 3Q18 Due to Rising Costs of Materials. TrendForce.

Dufo-Lopez, R. and J. L. Bernal-Agustin (2008). "Multi-objective design of PV-wind-diesel-hydrogen-battery systems." Renewable Energy **33**(12): 2559-2572.

eia. (2018). "How electricity is generated." from https://www.eia.gov/energyexplained/index.php?page=electricity_generating.

El-Khattam, W. and M. M. A. Salama (2004). "Distributed generation technologies, definitions and benefits." Electric Power Systems Research **71**(2): 119-128.

Ellabban, O., H. Abu-Rub and F. Blaabjerg (2014). "Renewable energy resources: Current status, future prospects and their enabling technology." Renewable & Sustainable Energy Reviews **39**: 748-764.

EnergySage. (2019). "How much do solar panels cost in the U.S. in 2019?".

Ericson, S. and D. Olis (2019). A comparison of fuel choice for backup generators. NREL, Joint Institute for Strategic Analysis.

Faisal, M., M. A. Hannan, P. J. Ker, A. Hussain, M. Bin Mansor and F. Blaabjerg (2018). "Review of Energy Storage System Technologies in Microgrid Applications: Issues and Challenges." Ieee Access **6**: 35143-35164.

Farzan, F., K. Mahani, K. Gharieh and M. A. Jafari (2015). "Microgrid investment under uncertainty: a real option approach using closed form contingent analysis." Annals of Operations Research **235**(1): 259-276.

Firmo, H. T. and L. F. L. Legey (2002). "Generation expansion planning: An iterative genetic algorithm approach." Ieee Transactions on Power Systems **17**(3): 901-906.

François-Lavet, V., D. Taralla, D. Ernst and R. Fonteneau (2016). Deep reinforcement learning solutions for energy microgrids management. European Workshop on Reinforcement Learning.

Göğüş, Y. (2009). "Mechanical energy storage." Energy Storage Syst.

Goodfellow, I., Y. Bengio and A. Courville (2016). "Deep Learning." Deep Learning: 1-775.

Google. (2019). "Google Maps." from https://www.google.com/maps.

Hajipour, E., M. Bozorg and M. Fotuhi-Firuzabad (2015). "Stochastic Capacity Expansion Planning of Remote Microgrids With Wind Farms and Energy Storage." Ieee Transactions on Sustainable Energy **6**(2): 491-498.

Hakimi, S. M. and S. M. Moghaddas-Tafreshi (2014). "Optimal Planning of a Smart Microgrid Including Demand Response and Intermittent Renewable Energy Resources." Ieee Transactions on Smart Grid **5**(6): 2889-2900.

Harada, D. (1997). Reinforcement learning with time. AAAI, AAAI-97 Proceedings.

Hatziargyriou, N., H. Asano, R. Iravani and C. Marnay (2007). "Microgrids." Ieee Power & Energy Magazine **5**(4): 78-94.

Helm, D. (2016). "The future of fossil fuels-is it the end?" Oxford Review of Economic Policy **32**(2): 191-205.

Hemmati, R., R. A. Hooshmand and A. Khodabakhshian (2013). "Comprehensive review of generation and transmission expansion planning." Iet Generation Transmission & Distribution **7**(9): 955-964.

Hemmati, R., R. A. Hooshmand and A. Khodabakhshian (2016). "Coordinated generation and transmission expansion planning in deregulated electricity market considering wind farms." Renewable Energy **85**: 620-630.

Hemmati, R., H. Saboori and P. Siano (2017). "Coordinated short-term scheduling and long-term expansion planning in microgrids incorporating renewable energy resources and energy storage systems." Energy **134**: 699-708.

Hernandez, L., C. Baladron, J. M. Aguiar, B. Carro, A. Sanchez-Esguevillas and J. Lloret (2014). "Artificial neural networks for short-term load forecasting in microgrids environment." Energy **75**: 252-264.

Hernandez, L., C. Baladron, J. M. Aguiar, B. Carro, A. J. Sanchez-Esguevillas, J. Lloret and J. Massana (2014). "A Survey on Electric Power Demand Forecasting: Future Trends in Smart Grids, Microgrids and Smart Buildings." Ieee Communications Surveys and Tutorials **16**(3): 1460-1495.

Hines, P., J. Apt and S. Talukdar (2008). "Trends in the History of Large Blackouts in the United States." 2008 Ieee Power & Energy Society General Meeting, Vols 1-11: 4545-+.

Hirsch, A., Y. Parag and J. Guerrero (2018). "Microgrids: A review of technologies, key drivers, and outstanding issues." Renewable & Sustainable Energy Reviews **90**: 402-411.

Hochreiter, S. and J. Schmidhuber (1997). "Long short-term memory." Neural Computation **9**(8): 1735-1780.

Hontoria, L., J. Aguilera and P. Zufiria (2005). "A new approach for sizing stand alone photovoltaic systems based in neural networks." Solar Energy **78**(2): 313-319.

Huang, S. (2018) "Introduction to Various Reinforcement Learning Algorithms. Part I (Q-Learning, SARSA, DQN, DDPG)."

Huang, Y., P. M. Pardalos and Q. P. Zheng (2017). Electrical power unit commitment: deterministic and two-stage stochastic programming models and algorithms.

iea. (2018). "World Energy Outlook." from https://www.iea.org/weo/.

Inglesi-Lotz, R. (2016). "The impact of renewable energy consumption to economic growth: A panel data application." Energy Economics **53**: 58-63.

IRENA. (2017). "Electricity storage and renewables: Costs and markets to 2030."

IRENA (2019). Renewable power generation costs in 2018.

Jakhrani, A. Q., A. K. Othman, A. R. H. Rigit, S. R. Samo and S. A. Kamboh (2012). "A novel analytical model for optimal sizing of standalone photovoltaic systems." Energy **46**(1): 675-682.

Jing, W. L., C. H. Lai, W. S. H. Wong and M. L. D. Wong (2017). "Dynamic power allocation of battery-supercapacitor hybrid energy storage for standalone PV microgrid applications." Sustainable Energy Technologies and Assessments **22**: 55-64.

Kabir, E., P. Kumar, S. Kumar, A. A. Adelodun and K. H. Kim (2018). "Solar energy: Potential and future prospects." Renewable & Sustainable Energy Reviews **82**: 894-900.

Kaldellis, J. K., D. Zafirakis and E. Kondili (2010). "Optimum sizing of photovoltaic-energy storage systems for autonomous small islands." International Journal of Electrical Power & Energy Systems **32**(1): 24-36.

Kannan, S., S. M. R. Slochanal and N. P. Padhy (2005). "Application and comparison of metaheuristic techniques to generation expansion planning problem." Ieee Transactions on Power Systems **20**(1): 466-475.

Kazem, H. A. and T. Khatib (2013). "A novel numerical algorithm for optimal sizing of a photovoltaic/wind/diesel generator/battery microgrid using loss of load probability index." International Journal of Photoenergy **2013**.

Khan, A. Z., Y. Y. Sun and A. Ashfaq (2014). "Generation Expansion Planning Considering Externalities for Large Scale Integration of Renewable Energy." 2014 Ieee International Conference on Intelligent Energy and Power Systems (Ieps): 135-140.

Khayatian, A., M. Barati and G. J. Lim (2018). "Integrated Microgrid Expansion Planning in Electricity Market With Uncertainty." Ieee Transactions on Power Systems **33**(4): 3634-3643.

Khodaei, A., S. Bahramirad and M. Shahidehpour (2015). "Microgrid Planning Under Uncertainty." Ieee Transactions on Power Systems **30**(5): 2417-2425.

Khodaei, A. and M. Shahidehpour (2013). "Microgrid-Based Co-Optimization of Generation and Transmission Planning in Power Systems." Ieee Transactions on Power Systems **28**(2): 1582-1590.

Kingma, D. P. and J. Ba (2014) "Adam: A method for stochastic optimization."

Klein, S. A. and W. A. Beckman (1987). "Loss-of-Load Probabilities for Stand-Alone Photovoltaic Systems." Solar Energy **39**(6): 499-512.

Klinger, C., O. Landeg and V. Murray (2014) "Power Outages, Extreme Events and Health: a Systematic Review of the Literature from 2011-2012."

Knowledge, M. (2016) "The Growing Role of Energy Storage in Microgrids."

Koutroulis, E., D. Kolokotsa, A. Potirakis and K. Kalaitzakis (2006). "Methodology for optimal sizing of stand-alone photovoltaic/wind-generator systems using genetic algorithms." Solar Energy **80**(9): 1072-1088.

Kozlowski, D. (2002) "Onsite Options."

Kuznetsova, E., Y. F. Li, C. Ruiz, E. Zio, G. Ault and K. Bell (2013). "Reinforcement learning for microgrid energy management." Energy **59**: 133-146.

Lantero, A. (2014). "How Microgrids Work." from https://www.energy.gov/articles/how-microgrids-work.

LAZARD (2018). Lazard's Levelized Cost of Energy Analysis.

Leahy, E. and R. S. J. Tol (2011). "An estimate of the value of lost load for Ireland." Energy Policy **39**(3): 1514-1520.

Lewis, N. S. (2016). "Research opportunities to advance solar energy utilization." Science **351**(6271).

Li, F. D., M. Wu, Y. He and X. Chen (2012). "Optimal control in microgrid using multi-agent reinforcement learning." Isa Transactions **51**(6): 743-751.

Li, S. Y., D. W. Coit and F. Felder (2016). "Stochastic optimization for electric power generation expansion planning with discrete climate change scenarios." Electric Power Systems Research **140**: 401-412.

Lipton, Z. C., J. Berkowitz and C. Elkan (2015) "A critical review of recurrent neural networks for sequence learning."

Lucio, J. H., R. Valdes and L. R. Rodriguez (2012). "Loss-of-load probability model for stand-alone photovoltaic systems in Europe." Solar Energy **86**(9): 2515-2535.

Luz, T., P. Moura and A. de Almeida (2018). "Multi-objective power generation expansion planning with high penetration of renewables." Renewable & Sustainable Energy Reviews **81**: 2637-2643.

Ma, T., H. X. Yang and L. Lu (2014). "A feasibility study of a stand-alone hybrid solar-wind-battery system for a remote island." Applied Energy **121**: 149-158.

Mahani, K., F. Farzan and M. A. Jafari (2017). "Network-aware approach for energy storage planning and control in the network with high penetration of renewables." Applied Energy **195**: 974-990.

Mahani, K., M. A. Jamali, D. Nazemi and M. Jafari (2020). Economic and Operational Evaluation of PV and CHP combined with Energy Storage Systems considering Energy and Regulation Markets. 2020 IEEE Texas Power and Energy Conference (TPEC).

Mahani, K., Z. L. Liang, A. K. Parlikad and M. A. Jafari (2019). "Joint Optimization of Operation and Maintenance Policies for Solar-Powered Microgrids." Ieee Transactions on Sustainable Energy **10**(2): 833-842.

Mahani, K., D. Nazemi, M. A. Jamali and M. Jafari (2020). "Evaluation of the behind-the-meter benefits of energy storage systems with consideration of ancillary market opportunities." The Electricity Journal.

Maleki, A. and F. Pourfayaz (2015). "Optimal sizing of autonomous hybrid photovoltaic/wind/battery power system with LPSP technology by using evolutionary algorithms." Solar Energy **115**: 471-483.

Mandelli, S., C. Brivio, E. Colombo and M. Merlo (2016). "Effect of load profile uncertainty on the optimum sizing of off-grid PV systems for rural electrification." Sustainable Energy Technologies and Assessments **18**: 34-47.

Mandelli, S., C. Brivio, E. Colombo and M. Merlo (2016). "A sizing methodology based on Levelized Cost of Supplied and Lost Energy for off-grid rural electrification systems." Renewable Energy **89**: 475-488.

Marchi, B., M. Pasetti and S. Zanoni (2017). "Life cycle cost analysis for BESS optimal sizing." International Scientific Conference - Environmental and Climate Technologies, Conect 2016 **113**: 127-134.

Mariam, L., M. Basu and M. F. Conlon (2016). "Microgrid: Architecture, policy and future trends." Renewable & Sustainable Energy Reviews **64**: 477-489.

Mbuwir, B. V., F. Ruelens, F. Spiessens and G. Deconinck (2017). "Battery Energy Management in a Microgrid Using Batch Reinforcement Learning." Energies **10**(11).

McKinsey. (2018). "The potential impact of electric vehicles on global energy systems."

Mellit, A., S. A. Kalogirou, L. Hontoria and S. Shaari (2009). "Artificial intelligence techniques for sizing photovoltaic systems: A review." Renewable & Sustainable Energy Reviews **13**(2): 406-419.

Melo, F. S. (2001). "Convergence of Q-learning: A simple proof." Institute for Systems and Robotics.

Meteoblue (2012). Hourly historical weather simulation data since 1985.

Mnih, V., K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra and M. Riedmiller (2013). Playing atari with deep reinforcement learning, arXiv.

Muralitharan, K., R. Sakthivel and R. Vishnuvarthan (2018). "Neural network based optimization approach for energy demand prediction in smart grid." Neurocomputing **273**: 199-208.

Nandi, S. K. and H. R. Ghosh (2010). "Prospect of wind-PV-battery hybrid power system as an alternative to grid extension in Bangladesh." Energy **35**(7): 3040-3047.

Nazemi, D. and M. A. Jafari (2019). EV Charging Profiles.

Nemati, H., M. A. Latify and G. R. Yousefi (2018). "Coordinated generation and transmission expansion planning for a power system under physical deliberate attacks." International Journal of Electrical Power & Energy Systems **96**: 208-221.

Network, U. P. (2015). SmartMeter Energy Consumption Data in London Households.

Networks, U. P. (2014). Low Carbon London.

NREL (2008). Power System Planning: Emerging Practices Suitable for Evaluating the Impact of High-Penetration Photovoltaics.

NREL (2013). Commercial and Residential Hourly Load Profiles for all TMY3 Locations in the United States. OpenEI (Ed.).

NREL (2016) "Distributed generation renewable energy estimate of costs."

NREL (2016). PSM Global Horizontal Irradiance. Viewer, N.D. (Ed.).

O'Brien, G. and A. Hope (2010). "Localism and energy: Negotiating approaches to embedding resilience in energy systems." Energy Policy **38**(12): 7550-7558.

Office, W. E. T. (2014) "How Do Wind Turbines Work?".

Oldham, K. (2009) "Decommissioning dams - costs and trends."

Padilla, F. (2018) "Critical power: hospitals and data centers."

Pardo, F., A. Takavoli, V. Levdik and P. Kormushev (2018) "Time limits in reinforcement learning."

Parhizi, S., H. Lotfi, A. Khodaei and S. Bahramirad (2015). "State of the Art in Research on Microgrids: A Review." Ieee Access **3**: 890-925.

Park, J. B., J. H. Kim and K. Y. Lee (2002). "Generation expansion planning in a competitive environment using a genetic algorithm." 2002 Ieee Power Engineering Society Summer Meeting, Vols 1-3, Conference Proceedings: 1169-1172.

Pereira, A. J. C. and J. T. Saraiva (2011). "Generation expansion planning (GEP) - A long-term approach using system dynamics and genetic algorithms (GAs)." Energy **36**(8): 5180-5199.

Pereira, S., P. Ferreira and A. I. F. Vaz (2017). "Generation expansion planning with high share of renewables of variable output." Applied Energy **190**: 1275-1288.

Prehoda, E. W., C. Schelly and J. M. Pearce (2017). "US strategic solar photovoltaic-powered microgrid deployment for enhanced national security." Renewable & Sustainable Energy Reviews **78**: 167-175.

Raimi, D. (2017). Decommissioning US Power Plants: Decisions, Costs and Key Issues.

Rajesh, K., K. Karthikeyan, S. Kannan and C. Thangaraj (2016). "Generation expansion planning based on solar plants with storage." Renewable & Sustainable Energy Reviews **57**: 953-964.

Raju, L., S. Sankar and R. S. Milton (2015). "Distributed Optimization of Solar Micro-grid using Multi Agent Reinforcement Learning." Proceedings of the International Conference on Information and Communication Technologies, Icict 2014 **46**: 231-239.

Roberts, D. and A. Chang (2018) "Meet the microgrid, the technology poised to transform electricity."

Rodgers, M., D. Coit, F. Felder and A. Carlton (2019). "Assessing the effects of power grid expansion on human health externalities." Socio-Economic Planning Sciences **66**: 92-104.

Rodgers, M., D. Coit, F. Felder and A. Carlton (2019). "A Metamodeling Framework for Quantifying Health Damages of Power Grid Expansion Plans." International Journal of Environmental Research and Public Health.

Rodgers, M. D. (2016). Simulation-based Optimization Models for Electricity Generation Expansion Planning Problems Considering Human Health Externalities. Doctoral, Rutgers University.

Sadeghi, H., M. Mohammadian, A. Abdollahi, M. Rashidinejad and S. M. Mahdavi (2014). "Renewable-Based Generation Expansion Planning Considering Environmental Issues Using GSA." 2014 Iranian Conference on Intelligent Systems (Icis).

Sadeghi, H., M. Rashidinejad and A. Abdollahi (2017). "A comprehensive sequential review study through the generation expansion planning." Renewable & Sustainable Energy Reviews **67**: 1369-1394.

Salloum, Z. (2018) "Double Q-Learning, the Easy Way."

SEIA (2019) "Solar Industry Research Data."

Semënov, D., G. Mirzaeva, C. D. Townsend and G. C. Goodwin (2017). A battery storage control scheme for AC microgrids. 20th International Conference on Electrical Machines and Systems (ICEMS).

Service, D. o. P. (2018). 2017 ELECTRIC RELIABILITY PERFORMANCE REPORT.

Sirikum, J. and A. Techanitisawad (2006). "Power generation expansion planning with emission control: a nonlinear model and a GA-based heuristic approach." International Journal of Energy Research **30**(2): 81-99.

Solutions, V. (2017) "Introduction to Reinforcement Learning."

Song, S. L., Q. Li, F. A. Felder, H. G. Wang and D. W. Coit (2018). "Integrated optimization of offshore wind farm layout design and turbine opportunistic condition-based maintenance." Computers & Industrial Engineering **120**: 288-297.

statista (2019). Wind energy - cumulative installed capacity in the United States 2017.

Su, W. C., J. H. Wang and J. Roh (2014). "Stochastic Energy Scheduling in Microgrids With Intermittent Renewable Energy Resources." Ieee Transactions on Smart Grid **5**(4): 1876-1883.

Sutton, R. and A. Barto (2015). Reinforcement Learning: An Introduction, The MIT Press.

Swartz, J., A. Ghofrani and M. Jafari (2017). "Sizing Methodology for Combined Renewable Energy Systems." 2017 Ieee Power & Energy Society Innovative Smart Grid Technologies Conference (Isgt).

Tekiner-Mogulkoc, H., D. W. Coit and F. A. Felder (2012). "Electric power system generation expansion plans considering the impact of Smart Grid technologies." International Journal of Electrical Power & Energy Systems **42**(1): 229-239.

Tekiner-Mogulkoc, H., D. W. Coit and F. A. Felder (2015). "Mean-risk stochastic electricity generation expansion planning problems with demand uncertainties considering conditional-value-at-risk and maximum regret as risk measures." International Journal of Electrical Power & Energy Systems **73**: 309-317.

Tekiner, H., D. W. Coit and F. A. Felder (2010). "Multi-period multi-objective electricity generation expansion planning problem with Monte-Carlo simulation." Electric Power Systems Research **80**(12): 1394-1405.

Telegraph, T. (2019). "How do we harness renewable energy sources?", from https://www.telegraph.co.uk/education/stem-awards/electrical/microdgrid-technology/.

Ton, D. and J. Reilly (2017). "Microgrid Controller Initiatives: An Overview of R&D by the US Department of Energy." Ieee Power & Energy Magazine **15**(4): 24-31.

Tsianikas, S., X. Xie, S. R. Puri, A. K. Parlikad and D. Coit (2020). Comparison of Neural Network Based Approaches for Short-term Residential Energy Load Forecasting. Submitted manuscript to: "Energy and Buildings".

Tsianikas, S., N. Yousefi, J. Zhou and D. Coit (2020). The impact of analytical outage modeling on expansion planning problems in the area of power systems. IISE Annual Conference.

Tsianikas, S., N. Yousefi, J. Zhou, D. Coit and M. Rodgers (2019). A Sequential Resource Investment Planning Framework using Reinforcement Learning and Simulation-Based Optimization: A Case Study on Microgrid Storage Expansion. Submitted manuscript to: "Production and Operations Management".

Tsianikas, S., J. Zhou, D. P. Birnie and D. Coit (2019). Economic trends and comparisons for optimizing grid-outage resilient photovoltaic and battery systems. Applied Energy.

Tsianikas, S., J. Zhou, D. P. Birnie and D. Coit (2019). Techno-economic optimization of a PV + battery system: A case study for a hospital in Orlando, FL. IISE Annual Conference.

Tsianikas, S., J. Zhou, N. Yousefi and D. Coit (2019). Battery selection for optimal grid-outage resilient photovoltaic and battery systems. IISE Annual Conference.

Unterreiner, L., V. Julch and S. Reith (2016). "Recycling of Battery Technologies - Ecological Impact Analysis Using Life Cycle Assessment (LCA)." 10th International Renewable Energy Storage Conference, Ires 2016 **99**: 229-234.

Valinejad, J., M. Marzband, M. F. Akorede, T. Barforoshi and M. Jovanovic (2017). "Generation expansion planning in electricity market considering uncertainty in load demand and presence of strategic GENCOs." Electric Power Systems Research **152**: 92-104.

van der Welle, A. and B. van der Zwaan (2007). An Overview of Selected Studies on the

Value of Lost Load (VOLL). Energy research Centre of the Netherlands (ECN).

van Hasselt, H. (2010). Double Q-Learning. Advances in Neural Information Processing Systems.

van Hasselt, H., A. Guez and D. Silver (2016). Deep Reinforcement Learning with Double Q-Learning. Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16).

Van Zaen, J., C. M. El Achkar, R. E. Carillo and A. Hutter (2018). Detection and Classification of Refrigeration Units in a Commercial Environment: Comparing Neural Networks to Unsupervised Clustering.

Wagar, A., S. R. Wang, S. M. Dawoud, T. Tao and Y. D. Wang (2015). "Optimal Capacity Expansion-Planning of Distributed Generation in Microgrids considering Uncertainties." 2015 5th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (Drpt 2015): 437-442.

Wang, Z. J., Y. Chen, S. W. Mei, S. W. Huang and Y. Xu (2017). "Optimal expansion planning of isolated microgrid with renewable energy resources and controllable loads." Iet Renewable Power Generation **11**(7): 931-940.

Willis, K. G. and G. D. Garrod (1997). "Electricity supply reliability - Estimating the value of lost load." Energy Policy **25**(1): 97-103.

World, R. E. (2018) "Making Modern Microgrids Work."

Xi, Z., J. R., L. C. and H. M. (2016). Recent Research on Battery Diagnostics, Prognostics, and Uncertainty Management. Advances in Battery Manufacturing, Service, and Management Systems.

Xie, X., A. K. Parlikad and R. S. Puri (2019). "A Neural Ordinary Differential Equations Based Approach for Demand Forecasting within Power Grid Digital Twins." 2019 Ieee International Conference on Communications, Control, and Computing Technologies for Smart Grids (Smartgridcomm).

Yang, F., Y. Y. Xie, Y. L. Deng and C. Yuan (2018). "Predictive modeling of battery degradation and greenhouse gas emissions from US state-level electric vehicle operation." Nature Communications **9**.

Yang, H. X., L. Lu and W. Zhou (2007). "A novel optimization sizing model for hybrid solar-wind power generation system." Solar Energy **81**(1): 76-84.

Yousefi, N., S. Tsianikas and D. W. Coit (2020). "Reinforcement learning for dynamic conditionbased maintenance of a system with individually repairable components." Quality Engineering.

Zhang, F., Z. C. Hu and Y. H. Song (2013). "Mixed-integer linear model for transmission expansion planning with line losses and energy storage systems." Iet Generation Transmission & Distribution **7**(8): 919-928.

Zheng, J., C. C. Xu, Z. Zhang and X. H. Li (2017). "Electric Load Forecasting in Smart Grids Using Long-Short-Term-Memory based Recurrent Neural Network." 2017 51st Annual Conference on Information Sciences and Systems (Ciss).

Zhou, J., N. Huang, D. W. Coit and F. A. Felder (2018). "Combined effects of load dynamics and dependence clusters on cascading failures in network systems." Reliability Engineering & System Safety **170**: 116-126.

Zhou, J., S. Tsianikas, D. P. Birnie and D. W. Coit (2019). "Economic and resilience benefit analysis of incorporating battery storage to photovoltaic array generation." Renewable Energy **135**: 652-662.

Zhou, J., S. Tsianikas, N. Yousefi and D. W. Coit (2019). "Sizing optimization of solar power projects using machine learning techniques for time series forecasting (in progress)."

Zhou, W., C. Z. Lou, Z. S. Li, L. Lu and H. X. Yang (2010). "Current status of research on optimum sizing of stand-alone hybrid solar-wind power generation systems." Applied Energy **87**(2): 380-389.