UNIVERSITY OF ROME "LA SAPIENZA"

SAPIENZA
UNIVERSITÀ DI ROMA

Machine Learning Techniques for Microgrid Energy Management System Modelling and Design

Ph.D. Thesis

Author:

Stefano LEONORI

Supervisor:

Fabio Massimo FRATTALE MASCIOLI

Co-Supervisor:

Antonello RIZZI

POMOS - POlo per la MObilità Sostenibile

DIET-Department of Information Engineering, Electronics, and Telecommunications

February 5, 2019

Abstract

In the last decades, several governments are implementing different energy policies to encompass a sustainable energy future. These policies promote the increasing use of Renewable Energies Sources (RESs), electric vehicle and the application of demand response programs. However, the increasing penetration of distributed RESs and fast charge stations, since are characterized by stochastic and intermittent behaviours, requires a modernization of the whole electric distribution infrastructure for a better control and management of power transients, energy flows oscillations and bidirectionalities. The installation of grid-connected Microgrids (MGs) can be a suitable solution for a bottom-up modernization of the distribution grid into a smart grid. MGs are defined as "groups of interconnected loads and distributed energy resources that act as a single controllable entity with respect to the grid". Their realization requires the equipment of energy storage devices, power converters and ICT infrastructures able to efficiently monitor, control and manage the MG energy flows including the energy exchanged with the grid by means of the design of a suitable Energy Management System (EMS). This thesis is centred on the application of machine learning techniques for the synthesis of a FIS-based MG EMS. The EMS here formulated is in charge to define in real time how to distribute the MG energy flows. For EMS design, FISs have been preferred over other possibilities since rule based inferential systems, other than being featured by low computational cost, allow a better interpretation of the overall decision rule. In this thesis, different algorithms for data driven FIS synthesis have been investigated, relying on Computational Intelligence techniques, and in particular exploiting a hybrid evolutionary-fuzzy approach. Results have been compared with (optimal) benchmark solutions computed by assuming to know a priori the whole time series of the loads and energy generation.

Acknowledgements

Special thanks go to Professors Fabio Massimo Frattale Mascioli, Antonello Rizzi and Giorgio Rizzoni. I am grateful to have collaborated with my wonderful colleagues: Enrico De Santis, Maurizio Paschero and Alessio. Finally, I wish to thank the rest of my Roman and Ohioan labmates: Massimiliano Luzi, Antonello Rosato, Enrico Maiorino, Rosa Altilio, Marco Barbato, Fabriozio Donatoantonio, Matteo Galli, and Nicola Zanardi – each of these exceptionally intelligent people encouraged me and supported me in my research.

Contents

List of Figures

List of Tables

List of Abbreviations

ANFIS	**A**daptive **N**euro **F**uzzy **I**nference **S**ystem
BMS	**B**attery **M**anagement **S**ystem
CR	**C**onsumption **R**ate
CPP	**C**ritical **P**eak **P**ricing
DER	**D**istributed **E**nergy **R**esource
DG	**D**istributed **G**eneration
DR	**D**emand **R**esponse
DP	**D**ynamic **P**rogramming
DSM	**D**emand **S**ide **M**anagement
DSO	**D**istribution **S**ystem **O**perator
EA	**E**volutionary **A**lgorithm
EMS	**E**nergy **M**anagement **S**ystem
ESN	**E**cho **S**tate **N**etwork
ESS	**E**nergy **S**torage **S**ystem
EV	**E**lectric **V**ehicle
FC	**F**ast **C**harge
FIS	**F**uzzy **I**nference **S**ystem
GA	**G**enetic **A**lgorithm
GP	**G**reedy **P**rogramming
HEV	**H**ybrid **E**lectric **V**ehicle
IBR	**I**nclining **B**lock **R**ate
IBP	**I**ncentive-**B**ased **P**rogram
LMS	**L**east **M**ean **S**quare

LMSELR	**L**east **M**ean **S**quare **E**rror **L**inear **R**egression
LP	**L**inear **P**rogramming
MDP	**M**arkov **D**ecision **P**rocess
MF	**M**embership **F**unction
MILP	**M**ixed **I**nteger **L**inear **P**rogramming
MOGA	**M**ulti **O**bjective **G**enetic **A**lgorithm
NN	**N**eural **N**etwork
NSGA-II	**N**on **D**ominated **S**orting **G**enetic**A**lgorithm-**II**
OF	**O**bjective **F**unction
PARC	**P**runing **A**daptive **R**esolution Min-Max **C**lassifier
PBP	**P**rice-**B**ased **P**rogram
PEV	**P**lug-in **E**lectric **V**ehicle
PHEV	**P**lug-in **H**ybrid **E**lectric **V**ehicle
PSO	**P**article **S**warm **O**ptimization
PV	**P**hoto**V**oltaic
RC	**R**ule **C**onsequent
RES	**R**enewable **E**nergy **S**ource
RMSE	**R**oot **M**ean **S**quare **E**rror
RNN	**R**ecurrent **N**eural **N**etwork
RTP	**R**eal-**T**ime **P**ricing
SoC	**S**tate **o**f **C**harge
TOU	**T**ime **O**f **U**se
TR	**T**raining **S**et
TSO	**T**ransmission **S**ystem **O**perator
TS	**T**est **S**et
USM	**U**tility **S**ide **M**anagement
VL	**V**alidation **S**et
WCCS	**W**ithin **C**luster **S**um of **S**quares

WCSD **W**ithin **C**luster **S**um of **D**istance

List of Symbols

Φ	ANFIS Multivariate Gaussian Membership Function and Rule Antecedent	adim
$\mathbf{C}$	Covariance Matrix of the ANFIS Multivariate Gaussian Membership Function	adim
μ	Centroid of the ANFIS Multivariate Gaussian Membership Function	adim
θ	ANFIS Rule Consequent hyperplane	adim
f	ANFIS Rule Consequent Firing Strength	adim
P	Profit	M.U.
$\bar{P}$	Normalized Profit	adim
E^G	Energy Generation	kWh
E^{PV}	Energy Generation from PV	kWh
$E^{PV,opt}$	Optimal Profile Referred to the Energy Generation	kWh
E^L	Energy Demand	kWh
E^{FC}	Energy Demanded by a Fast Charge Station	kWh
$E^{FC,opt}$	Optimal Profile Referred to the Fast Charge Station Energy Demand	kWh
E^{GL}	Prosumer Energy Balance	kWh
$\hat{E}^{GL}$	Prosumer Energy Balance Prediction	kWh
E^S	Energy Exchanged by the Energy Storage System	kWh
E^N	Energy Exchanged between the Microgrid and the Main Grid	kWh
$E^{N,opt}$	Optimal Profile Referred to the Energy Exchanged with the Main Grid	kWh
SoC	ESS State of Charge	p.u.
SoC^{opt}	Optimal Profile Referred to ESS State of Charge	p.u.
C^{sell}	Energy Price in Sale (from MG to Main Grid)	M.U.
C^{buy}	Energy Price in Purchasing (from Main Grid to MG)	M.U.
q	Energy Balance Prediction Trend Line Coefficient	adim

m	Energy Balance Prediction Trend Line Coefficient	adim
m_k^{PV}	Energy generation trend line coefficient	adim
q_k^{PV}	Energy generation trend line coefficient	adim
ε	Input Space-Hyperplane Space Clustering Parameter	adim
λ	MinMax Parameter	adim
$\bar{F}$	Adimensional Generic Objective Function	adim
$\bar{F}^{opt}$	Adimensional Upper Benchmark Objective Function	adim
$\bar{F}^{bmk}$	Adimensional Lower Benchmark Objective Function	adim

Chapter 1

Introduction to Smart Grid

The electric grid is the widest human opera in the world. It can be summarized in three portions: generation, transmission and distribution. Each of these sections has different specific tasks to guarantee. In Figure 1.1 a scheme of a general electric grid is illustrated. The figure distinguishes, besides the energy generators and the loads, the distribution and the transmission lines, their rated voltages and the electric stations where are located the power transformers. These latter can also be viewed as nodes of separation between the portions of the power grid.

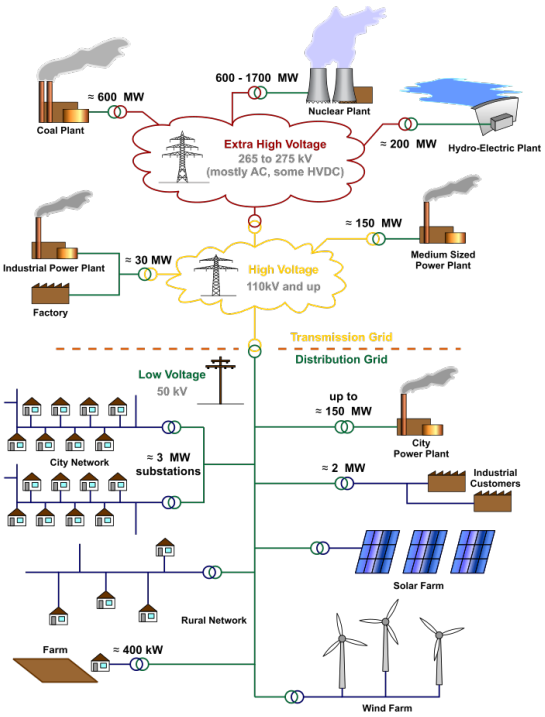


FIGURE 1.1: Scheme of a generic electric grid.

In the upper part of Figure 1.1 are depicted heavy power generators (rate power

higher than 100 MW) usually from fossil fuels (*i.e.* natural gas, carbon, oil) and nuclear fission. Power generators have rated voltages of tens of kVs in order to better preserve the power plants, maximize the power efficiencies and prevent the risk of faults. Together with high density power generation, in figure is depicted the generation from potential energy of dammed water which is mostly used to store or buffer the energy in over-generation or in over-demand. Usually, power generators are installed in isolated or industrial areas. Therefore, a power transmission infrastructure is in charge to deliver the power generated to distribution grids which connect and feed most of the users. Every portion of the grid is characterized by a specific range of voltage or level. In the following table are listed the main voltage levels.

TABLE 1.1: Classification of the voltage levels.

Voltage Bands	Rated Voltage
BT: low voltage	$U \leq 100\ V$
MT: medium voltage	$1 < U \leq 30\ kV$
AT: high voltage	$30 < U \leq 132\ kV$
AAT: very high voltage	$U > 132\ kV$

The transmission grid is featured by AT and AAT voltage levels with the priority to reduce the power losses and drop of voltages. In particular, transmission grids are distinguished in transmission and sub-transmission grids. The transmission grid is featured by AAT voltage and owns a meshed topology. In Italy, see Figure 1.2, it covers the whole country so as to ensure a greater robustness to the electricity system and the supply of energy in case of fault or outage of the power lines.

Sub-transmission grids are restricted to regional extensions and are featured by a lower voltage (AT voltage) since they are closer to urban areas. The sub-transmission topology can be radial, ring or island as shown in Figure 1.3. The radial topology is the most simple, however it is also the least efficient as regards the continuity of service.

The connection between transmission and sub transmission power lines takes places by means of the installation of suitable autotransformer AAT-AT.

FIGURE 1.2: Italian transmission grid topology.

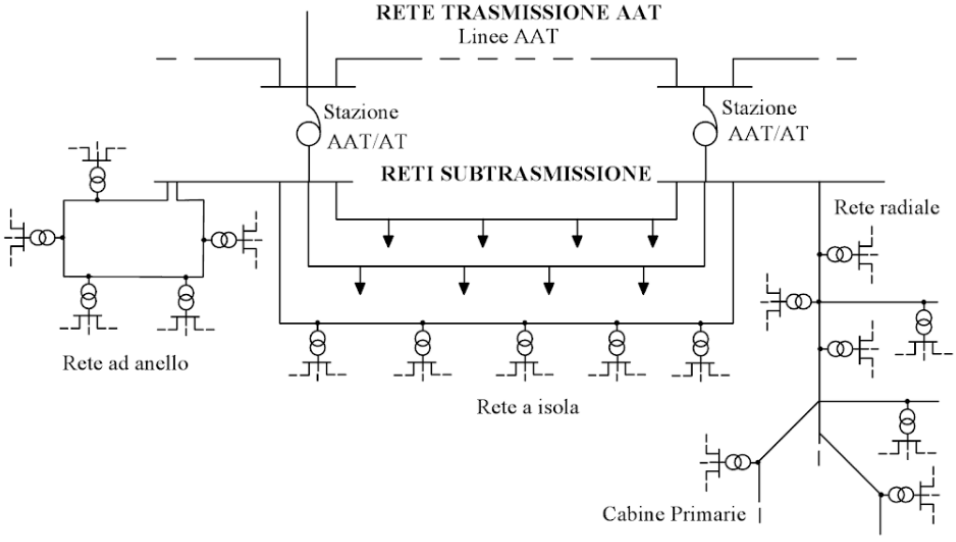


FIGURE 1.3: Italian sub-transmission grid topologies.

Along the transmission, getting closer to urban areas, other power generators can be connected as shown in Figure 1.1. Usually, these are featured by a lower power and have resiliency properties necessary to confer greater robustness and stability to the grid. Such kind of plants are usually gas turbine generators, with rated powers

ranging between tens to few hundreds of MWs. Their more resiliency property is paid with the higher cost of the energy source (natural gas) and less power efficiency (about 30%, high power plants are between $40-60\%$). Also huge loads (*i.e.* big customers) such as industrial factories are connected to the transmission lines. Afterwards, as shown in Figure 1.1, the transmission is connected to the medium voltage distribution by means of primary stations equipped with AT-MT transformers. Following, secondary MT-BT station reduces the voltage rate in proximity of the BT loads as well.

Several grid topologies are adopted in the MT and BT distribution grids as shown in Figure 1.4. Usually, especially in more crowded areas, distribution MT grids are meshed, ring or in doubly fed topology in order to guarantee the continuity of services in case of local faults and outages. It is important to remark that in every case the distribution exercise is always in radial mode for technical, fault detection and safety reasons. Consequently, they must be equipped with dedicated electrical connectors coupled with switchers installed in strategic nodes in order to efficiently accomplish the radial exercise [1]. In BT distribution, the grid topology is often ra-

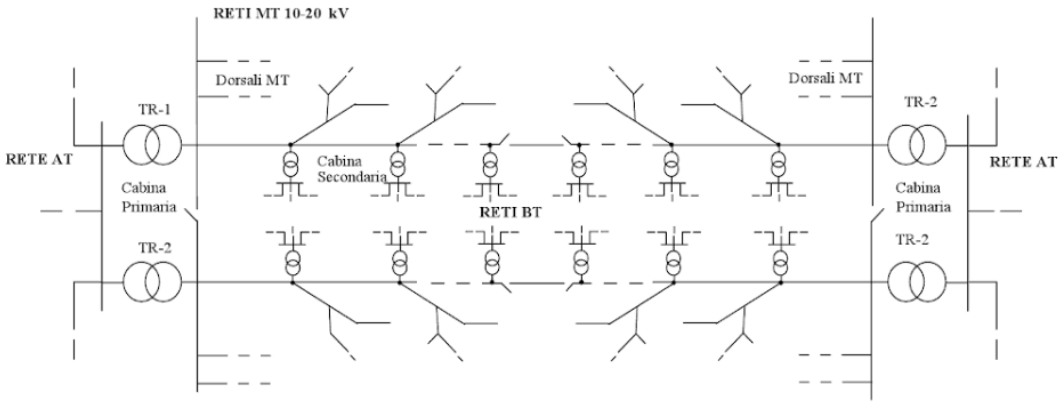


Fig. 1.23 - Rete di distribuzione a MT.

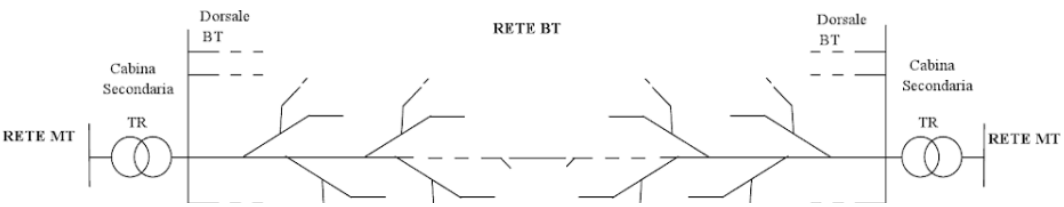


FIGURE 1.4: Ditribution grid topologies.

dial, especially in rural areas. Indeed, unlike the transmission grid, it is more subject to changes and modernizations that are function of the local population growth and development. Therefore, huge investments on the power lines are more discouraged respect with the other grids (*i.e.* generation, transmission and MT distribution) since the distribution grids are characterised by lower depreciation time.

1.1 Distributed Energy Systems

1.1.1 Energy Generation from RESs

RESs are also used to be called distributed due to their low energy and power density with respect to fossil fuels and nuclear generation. In order to better explain it, if it is considered one squared meter of PV panel, the maximum power that can be extracted is about $1000\ [W/m^2] \cdot 0.15 = 150\ [W/m^2]$ where 0.15 is the PV power efficiency and $1000[W/m^2]$ is the solar radiance that can impact on a flat surface in a sunny day. However, by considering a utilization power factor about 0.2, the average daily energy produced is about $150 \cdot 24[h] \cdot 0.2 = 720\ [Wh/m^2]$. This value is much less than the production of the electric energy produced by one normal cubic meter of natural gas which feeds a gas turbine (*i.e.* about few kWh). Not to mention the wind farms, which have power densities much lower than PV systems [1].

The Distributed Generation (DG) from RESs is generally restricted between few kWs to some MWs. However, there are also more extended PV plants and wind farms that range from hundreds to thousands of MW [2] [3] [2].

In the last decades, world countries have been increasingly concerned about the CO_2 global pollution, the efficient use of energy by the customers and the penetration of RES generation into the electric grid. For these reasons, the EU nations, like in most advanced and industrialized countries, advanced the biggest EU research and

[1] https://www.energycentral.com/c/ec/future-energy-why-power-density-matters
[2] en.wikipedia.org/wiki/Photovoltaic power station
[3] en.wikipedia.org/wiki/Wind farm

innovation incentives programme, namely the Horizon 2020 targets [3]. One of the consequences of these types of incentivization programs is that many users have been encouraged to install low power generation plants in proximity of urban areas facilitating the penetration of RES to the distribution grids. A demonstration of this phenomenon is depicted in Figure 1.5 where are shown the PV systems and wind turbines operating in Lazio in 2018.

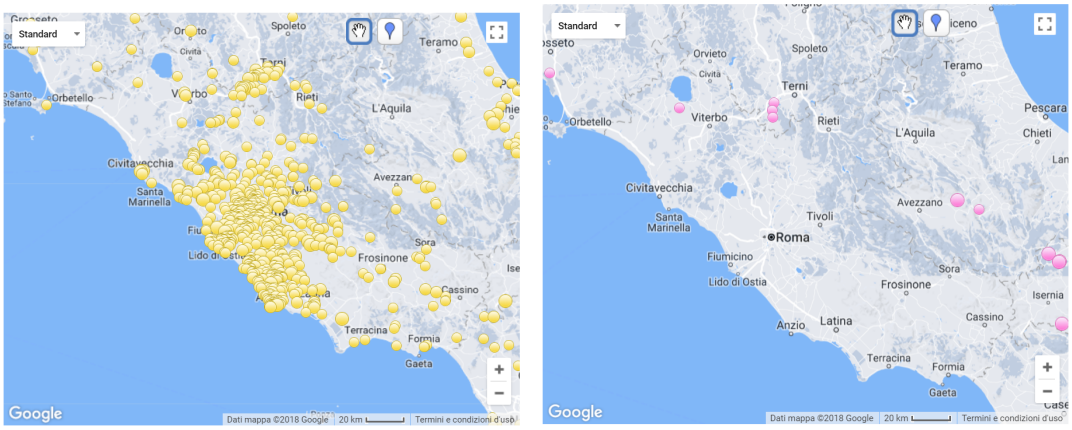


FIGURE 1.5: In the left figure, a map of the PV plants installed in the Lazio region. In the right figure, a map of the wind turbines installed in the Lazio region.
Source: *www.repowermap.org*.

In addition, in Table 1.2 is reported the increase of energy production from RESs in Italy. While the utilization from hydro is stable, the growing of green energies is mainly due to the installation of PV and wind turbine plants. In Figure 1.6 instead is reported the increase of the MW installed from PV plants and wind turbines in Italy with respect to the fossil fuel power plants. It can be noted that the inflection in 2010 coincides with the start of incentive policies in Europe and in Western countries.

TABLE 1.2: Electrical energy production from RESs in Italy [TWh].

Year	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**
Hydro	49,1	51,1	45,8	41,9	52,8	58,5	45,5	42,4
Wind Turbine	6,5	9,1	9,9	13,4	14,9	15,2	14,8	17,7
PV	0,7	1,9	10,8	18,9	21,6	22,3	22,9	22,1
Geothermal	5,3	5,4	5,7	5,6	5,7	5,9	6,2	6,3
Biomass	7,6	9,4	10,8	12,5	17,1	18,7	19,4	9,5

TABLE 1.3: Electrical energy production in Italy [TWh].

Year	2015	2016
Hydro	46,45	43,80
Fossil Fuel	188,7	196,60
Wind Turbine	14,7	17,50
PV	22,6	21,70
Total	272,45	279,6

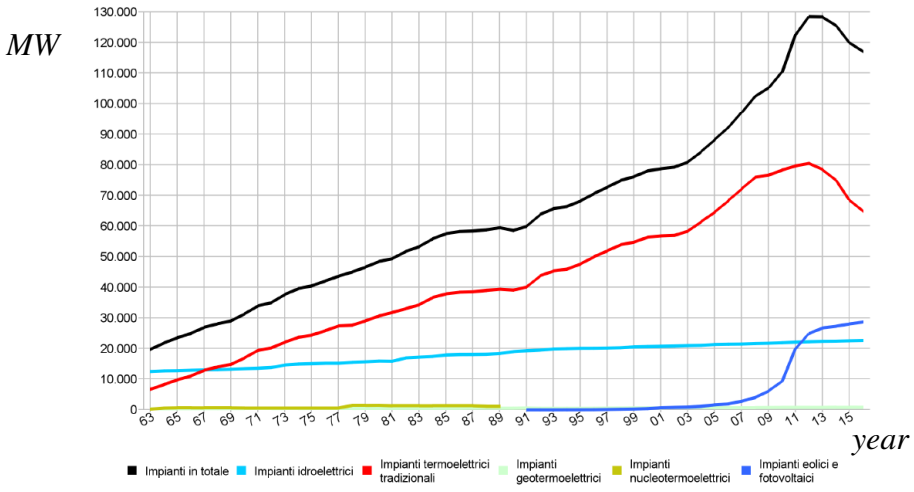


FIGURE 1.6: History of the power installed in Italy per energy source. Cyan: Hydro; Blue: PV; Black: total; Red: Fossi Fuel; Yellow: Nuclear.

Although the generation from RESs does not present any (or negligible) emission, which in the case of biomass is usually considered zero-equivalent since they produce what they absorb, it has a higher cost of the kWh produced and in some cases it is more subject to the opposition by the local population due to their proximity to urban areas. Another important issue to take into consideration, particularly in the case of PV and wind generation, is that the power production from RESs has an intermittent and stochastic behaviour as reported in [4] where the impact of the PV generation on the distribution network and the increase in the cost of the electric kWh is discussed. In particular, the manuscript highlights that, with the coming of the sunset, the PV stops to generate and at the same time the distribution (aggregated) power demand increases until to its second daily power peak. Therefore, gas turbine generators have to be quickly turn on for just few hours in order to reach the energy balance. During the production at noon instead, the grid can generate more energy than that demanded

by local customers, implying the local overproduction and thus reversing the power flow. This means that the generation from RESs, not only involves bidirectional energy flows in the distribution, but can also involve power flows bidirectionalities between the transmission grid and the distribution grid putting the state of the electric grid at risk.

In Germany for instance, it has been registered that in some sunny periods, the overproduction from PV has involved the shutdown of high power stations [4]. This phenomenon is indicated in the literature with the term "Duck Curve" since the particular deformation of the power balance profile in the distribution caused by the excessive PV generation during the day [4].

1.1.2 Energy Storage Systems

In the last years, the Energy Storage System (ESS) technology increased with the development of new batteries, such as lithium-ion batteries, Fuel Cell etc., and the raise of power converters which reached power efficiencies higher than 95%. It is clear that an efficient distribution of ESSs in the distribution grid could provide more flexibility and control, especially with the spread of RESs. However, their location and sizing is a topic still widely discussed in literature because of the huge extensions of the distribution grids, the chaotic distribution and behaviour of the DG and the high cost of installation of the ESSs. In literature, ESSs are classified per chemistry, costs [money/energy size], power density [kW/kg], energy density [kWh/kg], lifecycle emissions, efficiency and average lifetime (*i.e.* number of cycles) [5], [6], [7]. In [5] a comparison among different technologies is applied to investigate what ESS is more suitable for each kind of application field. In Figure 1.7 and Figure 1.8 are illustrated the ESSs energy densities versus power densities and specific energies versus specific powers, respectively.

In [6] a detailed life cost cycle analysis about ESSs is published. It is mostly based on the analysis of demonstration plants. It shows that, though Lithium-Ion

[4] https://www.ise.fraunhofer.de/en/publications/studies/recent-facts-about-pv-in-germany.html

batteries are the most acclaimed and discussed technology because their high power and specific energies, it is still far from reaching accessible costs in many application fields. In addition, another problem with lithium-ion batteries is the estimate of the current SoC, which is crucial for monitoring the actual battery state of health and energy availability [8].

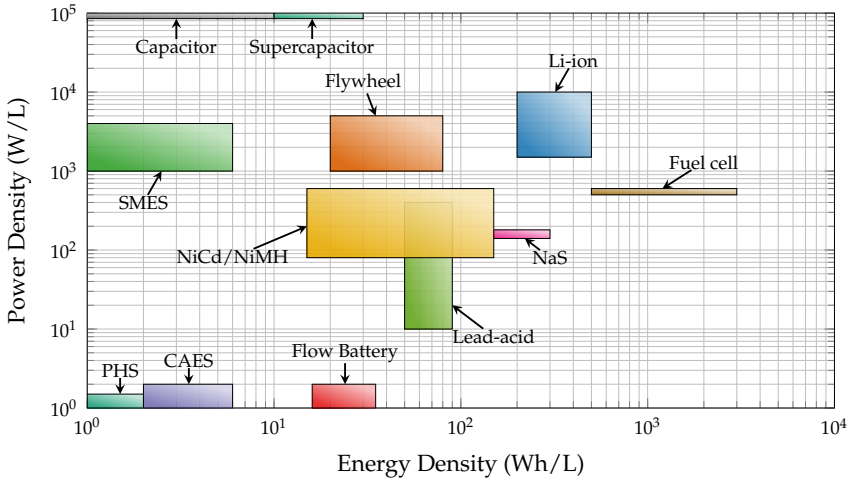


FIGURE 1.7: ESSs energy densities vs power density.

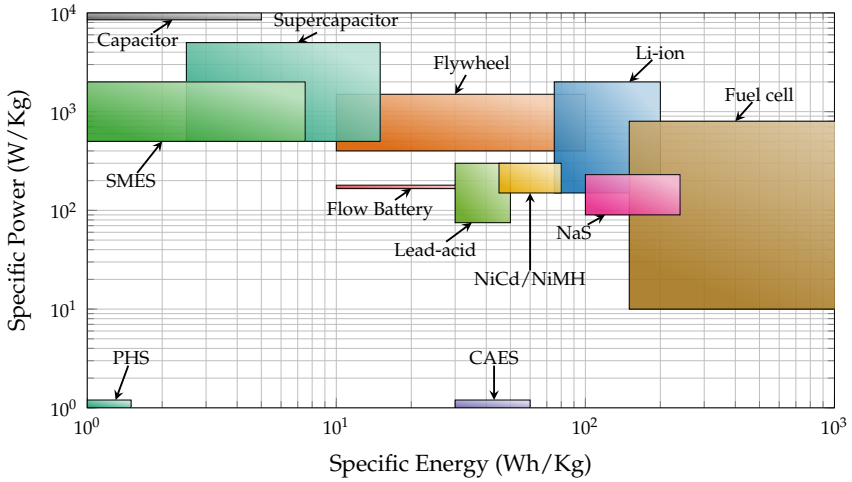


FIGURE 1.8: ESSs specific energies vs specific power.

1.1.3 Fast Charge Stations

The contemporary evolution of energy storage and power converter technologies combined with the need of exploiting green energy production pushed towards the

commercialization of Electric Vehicles (EVs), in particular Hybrid Electric Vehicles (HEVs) and Plug-in Electric Vehicles (PEVs) [9]. Besides the EV technology, it is known that the EV commercialization for replacing conventional Internal Combustion Engine (ICE) vehicles requires the design of suitable infrastructures (*i.e.* charge stations) able to efficiently provide the recharge of PEVs in urban areas and in congestion points.

In a first analysis, it is possible to distinguish two types of PEV charging stations due to their strong difference in terms of customer service, nominal power, demand and infrastructure costs, *i.e.* Fast Charge (FC) stations and slow charging stations [10]. Slow charge stations are featured by few kWs of power transmission (usually less than 10 kW). These are characterised by long waiting times usually higher than one hour for the benefit of reduced installation costs. The power transmission between PEV and the grid can be reduced to a proper cable equipped with a small power converter for the AC-DC conversion. Usually, the slow charge takes place at home by means of the connection to the triphase power line. Otherwise, there can be installed slow charge stations in big centers like malls or parking areas where people are supposed to spend more hours. Because of the long waiting times, the slow charge service can admit stops and the power modulation without compromising the customer service. Moreover, depending on the agreement among the distribution grid, the user and the PEV producers as well, slow charge can also admit the application of Vehicle-To-Grid (V2G) operations (*i.e.* reverse the power flow) according to some needs. For instance, the energy discharged from the PEV can supply the building which the PEV is connected with (public buildings, malls etc.) in order to gain some advantage (*e.g.* peak shaving operations, improve the local autoconsumption etc.) or directly the distribution grid for the application of energy trading operations. The work proposed in [11] for example is focused on the optimal scheduling of PEV home charge with V2G capability in order to maximize the profit generated by the energy exchange with the grid by considering time variable energy prices. In particular, the work is focused on the integration of the battery degradation costs and their

impact on the charging strategies to adopt.

Contrary to the slow charge, the FC service is characterised by short waiting times of the order of minutes. It is clear that the FC service cannot tolerate any kind of delay (*e.g.* load shifting and reshaping operations). FC stations are supposed to be connected to the distribution grid of medium or low voltage, their (DC) power transmission usually ranges between tens to hundreds of kWs. The main advantage of installing public FC stations is the abatement of the so called PEV commuters "range of anxiety". A TEPCO (Tokyo Electric Power Company) study well illustrates such phenomenon [12]. The authors argue that those drivers who are reassured that their vehicle can be recharged in the middle of the trip, are involved in covering greater distances without being stranded. Said in other words, the FC service development is able to eliminate the anxiety discomfort given by the EVs utilization respect with conventional ICE vehicles.

However, it is important to remark that FC service requires a more expensive infrastructure since it must be equipped with suitable power converters and auxiliary power units (*i.e.* ESSs, Supercapacitors etc.) for managing the high power densities and protecting both the connected grid and the EV from power transients. In [13] a broad overview about fast charging and DC charging technical aspects is documented. There are discussed the charging power levels, their different standards and communication protocols. In [14] different FC station topologies for efficiently connect the FC station to the grid are introduced and finally compared.

In addition, an agreement with the DSO for the installation of FC stations is more complicated because of the high power and intermittent behaviours. Indeed, the application of the FC services can have repercussion on more intense drops of voltages, frequency control problems and high power fluctuations to the distribution grid [15]. FC stations have to deal with the energy demand of EV commuters that are featured by a stochastic behaviour difficult to predict as well. Accordingly, it is clear that in case it is planned a capillary expansion of them on a given area, the FC impact on the grid must be accurately investigated in order to adequately strengthen

the distribution grid.

FC stations are suggested to be installed in strategic spots nearby congested areas where they can be reached by a huge number of vehicles in transit like public or commercial areas and along highways. In [16], authors state that 50 kW quick chargers are commercially available to be installed in public parking lots such as universities, hospitals and shopping centers. In these locations, high power densities can be efficiently managed and their impact on the grid is reduced. In [15] it is sustained the commercialization feasibility of electric buses charged by 200 kW FC stations installed in proximity of a university campus. In [17] instead, a stochastic flow location model procedure is applied to investigate the optimal location of FC stations in central Ohio (USA) using real data on commuter habits and congestion nodes.

1.2 Microgrids: a Bottom Up Approach for Smart Grid Realization

In 1999, Italy allowed the liberalization of the electric market with the main objective to encourage the competition on the energy market in order to reduce the energy tariffs, the cost of the energy production and the energy dispatched. Historically, the electric grid was controlled and managed by a single unique operator that was ENEL, Ente Nazionale dell'Energia Elettrica. ENEL was responsible for both the energy production, transmission and distribution. After 1999, with the starting of the actual scenario, three important type of grid operators emerged: the producers, the Transmission System Operator (TSO) and the Distribution System Operator (DSO).

In Italy, the TSO is covered by TERNA, it is responsible for the transmission and dispatching of electric energy throughout the entire Italian territory, and therefore for the safe management of the balance between electricity supply and demand. It has to assure the security of the National Grid, to face events where the local generation may be reduced or disconnected following a system fault in the nearby or for maintenance,

to relieve localized network overloads, to maintain the system stability, to manage the grid voltages and to avoid Black outs.

The DSOs instead has the role to distribute the energy generation to the loads according to a given agreed tariff with the customers. Because of the energy liberalization program of 1999, several DSOs participate in the energy market in order to expand the offer of energy tariffs and services to the customers. Moreover, DSOs have to deal with distributed power generation from RESs when these are connected to the distribution grid.

By considering the spread of generation from RESs and the commercialization of EVs, it is clear that the distribution grid has to manage with the increase of power oscillations and bidirectionalities. These aspects if are not properly managed, are reflected in more risks of outages and faults, the lowering of the quality and continuity of service and a more power plants deterioration since steam turbines and big power plants in general are very sensitive to power transients. For this reason in [18] it is argued the necessity to strengthen the collaboration between DSOs and TSO since greater power generation is moving towards the distribution. In addition, DSOs found in the customer energy demand an element of flexibility to take advantage of. Indeed, also the customers can play an active role in the power grid by controlling, scheduling and efficiently managing their own loads and therefore involving greater flexibility to the grid, especially in presence of the so called *prosumer*, namely customers equipped with local generators and ESSs.

In 2005 the US Department introduced the Demand Response (DR) program. It supports the adoption of energy price tariffs which aim to motivate changes in electric use by end-use customers in response to changes in the price of electricity over time, or even to give incentive payments designed to induce lower electricity use at times of high market prices or when the grid reliability is jeopardized. As remarked in [19], the most important benefit of DR is the improved resource-efficiency of electricity production due to closer alignment between customers electricity prices and the value they place on electricity. Another definition of DR is given in [20] where the authors

define DR programs “as the changes in electricity usage by end-use customers from their normal consumption patterns in response to changes in the price of electricity over time". Namely, DR programs are reflected incentive payments designed to induce lower electricity use at times of high wholesale market prices or when system reliability is jeopardized. DR includes all intentional electricity consumption pattern modifications by end-use customers that are intended to alter the timing, level of instantaneous demand, or total electricity consumption. DR programs, or tariffs, are able to shape the users electricity load profiles in order to improve the reliability and efficiency of the grid [21].

DR programs are mainly divided into the following two branches: Incentive-Based Program (IBP), and Price-Based Program (PBP).

IBP encourages the participation in curtail, reduce the load or using local generation in some critical time periods or, for instance, with the aim to reduce the power peaks. In particular, in literature are distinguished two sub-branches. One is called *classical* while the second *market-based*.

In classical IBP the participating customers receive payments for their participation in the programs. In this case the utilities are allowed to remotely shut down participant equipment on a short notice. Typical remotely controlled equipment includes air conditioners and water heaters. This kind of programs is of interest mainly to residential customers and small commercial customers. In market-based IBP the customers are awarded for their performance, depending on the amount of load reduction during critical conditions.

PBP DR programs instead provide users with different electricity prices at different times. Users are encouraged to use less electricity when the prices are high, that translates into a reduction in demand at peak hours. The program induces the users to dynamically change their energy usage patterns according to the variance of electricity prices.

PBP are classified in four main branches: Time-of-Use (ToU); Real-Time Pricing (RTP); Critical Peak Pricing (CPP); and Inclining Block Rate (IBR) [20].

- ToU: a rate with different unit prices for usage during different blocks of time, usually defined for a 24 hour day. ToU considers also the energy price in selling when the prosumer is allowed to sell energy to the grid.
- RTP: a rate in which the price typically fluctuate hourly or every 15 mins. reflecting changes in the wholesale price of electricity. Customers are typically notified of RTP prices on a day-ahead or hour-ahead basis.
- CPP: a hybrid of ToU and RTP. The basic rate structure is ToU. A much higher CPP price is applied when the demand is very high or the system supply is limited.
- IBR: this tariff is designed with two-level rate structures (lower and higher blocks), such that the more electricity a user consumes, the more he pays per kWh. The price of electricity per energy consumption will rise to a greater value if the hourly (daily or monthly) energy consumption of the user exceeds a certain threshold. IBR incentives the users to distribute their loads among different times of a day to avoid higher rates, helping to reduce the grid's peak-to-average ratio.

In order to accomplish DR programs, the customer or prosumer must be equipped with a suitable device, named Energy Management System (EMS), able to schedule in real time along a given time horizon the customer energy flows.

In literature, two type of EMS are distinguished. The first looks at the customer interests and is called Demand Side Management (DSM). The second type of EMS looks at the power distribution interests, thus is called Utility Side Management (USM). For example, taking into consideration a prosumer MG connected to the grid that is equipped with a suitable ESS, in case the ESS is property of the prosumer, namely it is connected downstream the smart meter, the EMS is demand-side. Otherwise, in case the ESS is connected upstream the smart meter it is property of the grid, therefore the local energy flows are managed for the utility interests (*i.e.* USM).

Three common DR operations performed by the EMS [21] are described and illustrated in Figure 1.9, namely:

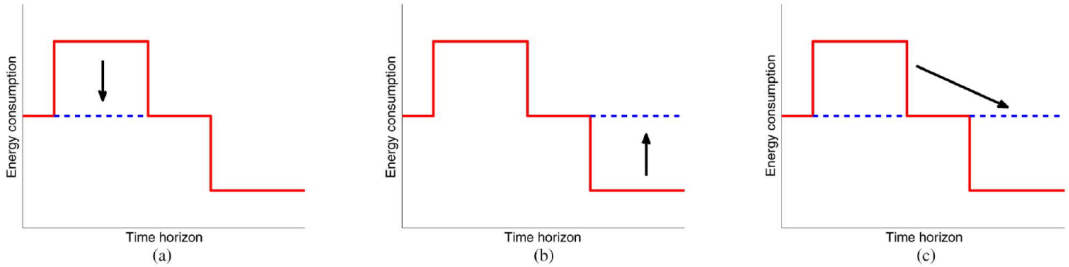


FIGURE 1.9: DR operation over a given time horizon, a) peak clipping; b) valley filling c) load shifting.

- Peak clipping: reduces the peak of energy consumption in order to prevent the load from exceeding the supply capacity of the distribution substations or the thermal limit of the transformers and power supplies. Users would have their satisfaction/comfort reduced since peak clipping cuts down some of their demand unless they can take advantage of internal generation units.

- Valley filling: it promotes the off-peak energy consumption through controlling energy sensitive loads like heat-pumps, boilers or else with the supplying of energy storage devices, such as ESSs and PEVs. The same PEVs recharge may also be subject to valley filling operations.

- Load shifting: it shifts the energy consumption over a given time horizon, *e.g.*, in order to shift the demand from on-peak to off-peak time periods (the combination of peak clipping and valley filling), without reducing the users total energy consumption in a day. Load shifting is suitable for loads characterized by a constant power and where their activation can be postponed without compromising the customer comfort. Dishwashers and washing machine are typical loads that can be subject to loads shifting operation.

In [22] the term smart grid is defined as a "vision" which aims to the "synergistic ability to coordinate multiple assets (devices and systems) to accomplish a specific

function such as peak shaving, feeder load balancing, emergency response, or voltage regulation". The term smart grid was introduced in 2010 by the U.S. Department of Energy which began a series of Communications and Controls Workshops focused on the integration of Distributed Energy Resources (DERs), where the term DER refers to controllable generation, storage, and load connected at the electric distribution level of the system. Smart grid vision was born with the starting of a period of high improvement and commercialization in ICT field and power converters machines, *i.e.* after the development of those technologies able to transmit, control and manage the energy generation from sustainable energy resources. In addition, in the same period most industrialized countries alerted the need of a new modern electrical infrastructure able to increase power efficiencies, minimize the power losses, manage, control and monitor power systems and intermittent bidirectional power flows and protect them from outages, risk of faults and cyber-attacking. Without going in too much details, a smart grid can be briefly summarized as an advanced power grid supported by a suitable ICT infrastructure. From an energy systems management point of view, the smart grid is able to effectively integrate, manage and monitor the power generation from RES, PEVs energy demand, the support of distributed ESS and other forms of DERs. In [23] several technical aspects about smart grids are defined. In particular, there are listed the most important differences between normal grid and smart grid, as well as all the characteristic time frames of each smart grid device.

Although the smart grid vision is clear and deemed necessary for the efficient integration of the DG from RESs and the application of DR programs, actually its realization is not yet commercially feasible. The huge investment costs, the high complexity in controlling and managing the power flows and the need to couple the power network with an ICT infrastructure are the main critical factors. In [24] it is sustained that the smart grid realization can be pursued by means of a bottom up approach. Namely, starting from local restricted areas, buildings, residences and districts in order to realize an interconnected smart and distributed system of systems

on the power distribution level.

More precisely, smart grids realization will pass through the evolution, developments and spreading of the so called MicroGrids (MGs). The U.S. Department of Energy defines MG as a group of interconnected loads and distributed energy resources within clearly defined electrical boundaries, which act as a single controllable entity with respect to the grid. A MG can connect and disconnect from the grid to enable it to operate in both grid-connected and islanded modes [25].

As argued in [26], the MGs development has a key role for the (future) integration of RESs (especially PV and wind turbine), energy efficiency and Green House Gasses emissions.

It is clear that MGs demand the design of robust ICT infrastructures and power converters able to connect each energy system to the MG backbone, namely DG systems, loads and ESSs. MGs must be primarily capable to ensure a certain quality of service to the customers by assuring, beside to the stability of the MG itself, the respect of given constraints and prosumer needs through the efficient monitor, control and management of the local power (energy) flows.

The Microgrid Institute classifies MGs in five categories that are distinguished for their size and functionality. The smallest discrete network is called nanogrid and is restricted to a single building or a single energy domain. When a MG is owned by a local prosumer and is equipped with small or medium DER units, it can be named *residential MG*. More extended MGs are distinguished in district, community and campus as well. These are mainly characterized by different functionalities and typology of user. MGs can also be extended to multi-generation systems, namely on combined heat and electric power generation and storing. In [27] for example, a distributed multi-generation systems optimization management is studied. The work is focused on the optimal management of co-generation units and load shifting for DR programs development.

In residential MGs, the EMS is responsible for the reliable, secure and economical operations in either grid-connected or islanded mode [28]. The EMS is in charge

of scheduling in real time the MG energy flows in function of the prosumers interests and needs. For this reason, it is also called DSM EMS since it makes the interests of the prosumer itself.

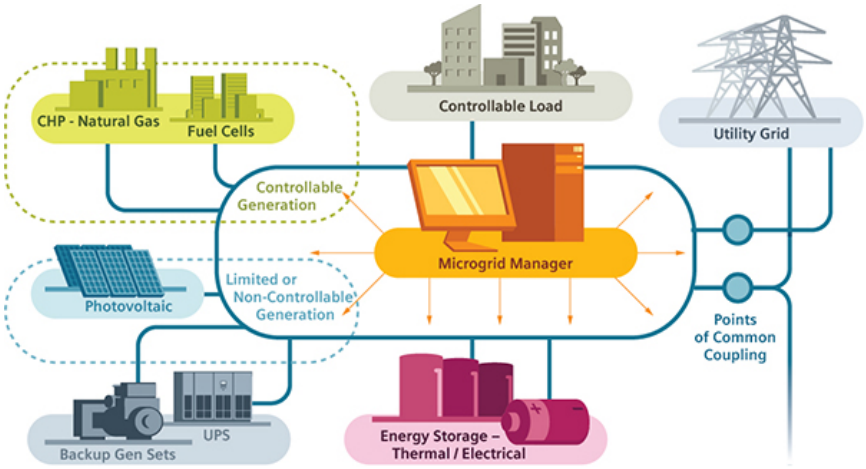


FIGURE 1.10: A MG illustrative scheme.

In control systems, MG EMSs can be placed in the second of a three-layered hierarchical control scheme as discussed in [28] where the authors categorized smart grids and MGs control systems by considering their response speed, the time period in which they operate and the features of the system to control.

In particular, the primary layer is named local control or internal control and is featured by the fastest response. Primary layer control systems are based exclusively on local measurements and require no communication. They keep stable the MG voltage and frequency control and assure the desired exchange of power among the MG energy systems. Power converters that connects the energy systems to the MG backbone operate on this control level.

The secondary level is played by the EMS and is in charge to define the unit commitment, namely how to efficiently assign the MG energy flows in order to satisfy given targets (*i.e.* DR targets). Therefore, the main objective of the EMS consists in finding the optimal (or near the optimal) unit commitment and dispatch of the available DER units so that the assigned tasks are achieved.

Tertiary control is the highest level of control and is responsible for coordinating the operation of multiple MGs interacting with each other in the smart grid system

and communicating needs or requirements from the host grid. This control level typically operates in the order of several minutes, providing signals to secondary level controls at MGs and other subsystems that form the full grid. Tertiary control can be considered part of the host grid, and not the MG itself. Secondary controls instead, coordinate internal primary controls within the MGs and subsystems in the span of a few minutes. Primary controls are designed to operate independently and react instantly in predefined ways to local events.

In [24], [29], [30] are discussed the architectural aspects of DC and AC MGs. In [29] in particular, it is sustained that DC MGs are more efficient and easier to control since the absence of reactive power and frequency control issues and allow a better connection of wind turbine (synchronous) generators as the connection to DC power lines would avoid the use of DC/AC converters [31]. The authors in [29] sustain that, given the convenience in control and the reduction of the conversion steps, the capillary diffusion of the MG would imply the gradual transformation of the distribution grid in DC.

It is clear that according to the scenario here introduced, the smart grid vision can be realized by means of the communication and power exchange between a systems of connected and topologically meshed MG systems.

It is important to note that the smart grid and MGs evolution can also have an important impact also on people's habits. For instance, together with the start of DR programs it would involve the customers to be more and more active in the energy market. The customer habits can be influenced by the same virtual machines designed to predict (and serve) them. The relying on ICT infrastructures for the supplying, controlling and managing of the energy services and the smart grid state would lead to storing a plethora of sensitive data that may put at risk the customers privacy.

1.3 Deterministic Methods in Energy Flow Management

In smart grid and MG energy management deterministic models are used for evaluating the optimal scheduling over a given time horizon that usually is extended to one or more days ahead. The selection of one model with respect to the others is function of the level of abstraction assumed, and the problem (scenario) under analysis. In literature, most of the algorithm used are based on Linear Programming (LP), Mixed Integer Linear Programming (MILP), Dynamic Programming (DP) and Greedy Programming (GP).

About GP, the term greedy is used in computer science to describe any search or decision procedure that selects alternatives based only on local or immediate considerations, without taking into account the possibility that such selection may prevent future access to even better alternatives. Consequently, greedy algorithms describe policies that select actions based only on their short-term consequences [32]. Due to its simplicity, its application is restricted to highly simplified problems.

LP optimization is based on the well known *simplex algorithm* [33]. LP based algorithms are able to solve optimization problems featured by linear relationships, *i.e.* defined by linear equations, constraints and OF.

MILP formulation is an extension of LP. MILP based algorithms are able to solve optimization problems featured by linear relationship and discrete variables. They relies on the support of heuristics since the optimization problem is *NP*-complete [34]. The inclusion of discrete variables allow to formulate piecewise-linear function able to approximate non linear functions (*e.g.* efficiency curves of ESSs and diesel generators in function of the power exchanged).

Regarding DP, as explained in [32] where Markov Decision Process (MDP) modelling and reinforcement learning procedures are introduced, the term DP refers to a collection of algorithms that can be used to compute optimal policies given a perfect model of the environment as a MDP, where MDP is meant as a reinforcement

learning task that satisfies the Markov property. As affirmed by the authors, classical DP algorithms are of limited utility in reinforcement learning both due to their assumption of a perfect model and their great computational expense, however they are still important theoretically. In [35] the concept of DP has been introduced from optimization and control point of view. DP is defined as an approach in solving optimization scheduling problems considering a discrete space. It is based on the so called Hamilton-Jacobi-Bellman partial differential equation. DP, though is NP-hard, is much better then apply *brute force* methods since it relies on the optimization of sub-structures by assuming a *divide and conquer* strategy. Though DP based algorithms are more pervasive with respect to the other methods (*i.e.* LP, MILP and GP), they requires an early investigation so as to accurately set the level of discretization (granularity) of the problem under analysis in order to not overload the machine and not end up in local minimums, especially when the problem taken into consideration is continuous.

In the last years, many deterministic optimization tools have been developed which mostly relies on MILP [36]–[38] and DP [39] approaches. In EMS problems, these tools are applied in order to evaluate how to optimally shift, reshape and curtail given known loads and energy generation profiles, store(release) energy to(from) the ESS and buy(sell) energy from(to) the connected grid considering given constraints [40]. The adoption of these tools suggests to think at first that a prosumer grid-connected MG equipped with a DSM EMS is able to perfectly communicate in advance to the DSO the energy needs (or released) along a given time period. Said in few words, the prosumer would be able to agree a day ahead with the DSO at what time and how much energy will be bought or sold so as to achieve a lower tariff. However, due to their deterministic nature, an effective direct application can be tricky, in some cases risky or even impossible. For instance, the need to know a priori the future prosumer energy generation and energy demand, the next energy prices etc. can be infeasible in some cases (*e.g.* unpredictable energy systems, RTP energy policy). Besides, their implementation can be considered as a black box model since

no inferential rule based structure is implemented. In case of a multi-objective optimization problem, the Objective Function (OF) must be formulated as a linear convex combination of cost functions multiplied by suitable meta-parameters which must be properly tuned. Therefore, a first investigation on the tuning of the meta-paramepers is mandatory.

Accordingly, a raw utilization of LP, MILP, GP and DP tools, unless are applied with the support of suitable machine learning algorithms (see next section), they are mostly restricted to solve machine components sizing, observe potential in optimal energy management [11] and test power machines (*i.e.* converters) control system accuracy and performance [41], [42]. This concept is also reiterated in [39] and [43] about HEV energy management[5]. These state that a direct DP-based algorithm application without the supporting of a suitable prediction model on driving cycle information (*e.g.* routes and traffic) cannot be effective since the stochasticity of the problem under analysis. To tackle these problems, the work in [44] is centred on the study of a chance constrained programming formulation that is an extension of the LP formulation in order to strengthen the algorithm tolerance from uncertainties. On the other hand, it is important to note that there are also cases in which the direct application of deterministic optimization tools has been demonstrated effective, as in [27], where the simulation scenario together with a suitable problem simplification allow a direct application of a basic LP. In that work the prosumer MG energy flows scheduling is limited to the current time slot without considering any extended time horizon.

As will be commented in the following section, deterministic optimization algorithms have a key role in MG and smart grid EMSs modelling together with machine learning techniques. Indeed, in this thesis, the LP-MILP and DP tools presented in [38] and [39] are used to evaluate (optimal) benchmark solutions for both supporting the EMS training and validation phases of the models proposed.

[5] In literature HEV are also referred to a particular type of nano-grid

1.4 Energy Flow Management by Machine Learning

Machine learning and data driven approaches have been demonstrated very interesting and successful solutions in solving MG and smart grid real time energy flow management [45]. As discussed in [46], real time EMS modelling are mainly based on the adoption of heuristics, stochastic and soft computing techniques and in many cases are also supported by the application of deterministic optimization algorithms based on GP, LP, MILP or DP.

A type of EMS model very discussed (and very intuitive) in literature consists in the training of prediction algorithms for the implementation of rolling horizon strategies. In real time energy flow management, rolling horizon strategies consist in the development of prediction models for efficiently forecast the future scenario over a given time horizon of about one day in order to run a deterministic optimization algorithm able to evaluate the optimal energy flow distribution over the prediction time horizon itself. This procedure is executed in each EMS time frame and only the first element of the solution found that would be referred to the upcoming time slot is implemented by the EMS to the MG in order to strengthen the EMS performance from the prediction uncertainties. In [47] a rolling time horizon strategy is investigated for the optimal management of a PV system equipped with an ESS considering the profit maximization. The authors show that the solution proposed improves the EMS performance on the profit of about the 13% with respect to a rule based EMS, considering an ideal prediction on the PV system as in [40]. In [48] a two-day-ahead NN-based model is proposed for the prediction of the MG electrical consumption. Instead, a Markov RESs prediction model is applied in [49] for optimally scheduling the MG energy flows.

With the adoption of rolling horizon strategies, there is the risk that the EMS time effort required to calculate a decision can get close to the smart meter sampling time as shown in [49]. Indeed, the computational effort given by the repetition of the prediction algorithms and the optimization algorithm is much greater than a plain

rule-based decision making system. In the first case, it is function of the optimization time horizon, the characteristic time frame length, the number of free variables (*i.e.* the number of energy systems to manage) and, in case of DP optimization, the state variables discretization level set by the operator. In the second case instead, the computational effort is restricted to the rules evaluation that in most of cases are defined by properly tuned hyperplanes or by numbers.

The application of Evolutionary Algorithms (EAs) (*e.g.* Particle Swarm Optimization-PSO, Genetic Algorithms-GAs etc.) for a faster evaluation of the optimal solution and a more robust EMS model are also widely studied in literature. For example, in [50] a differential search algorithm combined with MonteCarlo simulation scenario generation is applied for smart grid optimal scheduling. In [51] a prosumer day-ahead scheduling problem, formulated as a non-linear mixed integer one, has been solved via EAs. In [52] instead, the authors proposed a GA-based EMS able to efficiently define the current exercise configuration topology of a smart grid looking at the power losses minimization.

However, even if the high computational burdens can be solved by means of the adoption of EAs or rather more powerful calculators, these kind of solutions remain highly dependent on their prediction algorithms performance. In addition, these techniques can be considered as black box models, namely featured by a low reliability, since they cannot be referred to a decision making surfaces or even by an inferential rule based systems.

As stated in [46], also the hybridization of heuristics, stochastic and soft computing techniques are pursued in literature. In [53] for instance, Fuzzy Inference Systems (FISs), Artificial Neural Networks (ANNs) and LP formulation are used for optimally scheduling the MG energy flows by considering a multi-objective optimization problem. In [36] a comprehensive residential Energy Management tool built in Python™ is presented. The framework has a modular structure and a MILP-based optimal scheduler is the EMS core. It features a task scheduling logic and a configuration structure to represent different subsystems. The tool is equipped with

fuzzy logic and radial basis function Neural Network (NN) prediction algorithms that are trained by means of a PSO algorithm for the prediction of local loads and the DG.

In literature, fuzzy models like Adaptive Neuro Fuzzy Inference Systems (ANFISs) and FISs, ANNs and MDP are widely explored for load and generation prediction [54] [55] as well as decision making systems modelling [56]. The FIS, ANFIS and NN synthesis can be effectively performed by means of EAs such as GA and PSO, or by clustering algorithms.

Besides being good solution in control, prediction and classification problems, FISs revealed to be very interesting solutions for real time EMS application. The advantage of adopting fuzzy logic with respect to other soft computing techniques and stochastic models (*e.g.* deep NNs, EAs and MDPs) is its higher level of simplicity in implementation and the lower computational burden. Moreover, FISs are characterized by a greater human interpretability given by their "IF-THEN" rule-based inference systems which are able to ensure a higher level of reliability. These aspects are well shown in [56] where the authors examined several energy management strategies based on FISs to smooth the power exchange with the grid of a residential grid-connected MG.

In the following chapters are investigated different FIS and ANFIS modelling techniques for the synthesis of a real time decision making systems applied to a DSM EMS of a grid-connected MG. It is supposed to be equipped with an ESS, a PV plant and a system of aggregated loads. The first studies are focused on the implementation of the EMS by considering Mamdani type FISs modelled by means of a GA which has to tune the FIS parameters considering a specific (mono or multi) OF formulation. Further studies rely on the evaluation of optimal benchmark solutions through DP and LP algorithms in order to formulate a supervised problem (*i.e.* a labelled pattern set) that would permit the synthesis of ANFIS based EMS by means of clustering techniques. Different clustering algorithms based on the plain k-means have been implemented and compared. In particular, the modelling of hierarchical clustering variants in order to reduce the computational burden, the formulation of

different dissimilarity functions and the comparison of different clustering space (*i.e.* input space, the joint input-output space and hyperplane space as well) have been performed. Furthermore, it has been observed that the support of a prediction algorithm upstream the ANFIS decision making system that is in charge to estimate the prosumer energy trends can considerably improve the overall EMS performances.

Since this latter solution proved to be very effective, it has been re-proposed, with the necessary modifications, for the real time energy flow management of a grid-connected FC station MG equipped with a PV system and an ESS. The FC station has been supposed to be installed in the city of Columbus-Ohio for the charging demand fulfillment of local PEV commuters considering the simulation scenario given by the "Smart Columbus" project. Because of the absence of real world data, the energy demand profile of the FC station has been generated by a suitable tool based on a mixed deterministic and stochastic method which was in charge to reproduce the "Smart Columbus" scenario.

Chapter 2

FIS and ANFIS Systems

Fuzzy logic was born with the need of matching the uncertain, vague and imprecise human reasoning with the use of computers for testing and simulating. Its main objective is to try to get rid the high standards of rigour and precision that we have become conditioned to expect from our mathematical analysis of well-structured mechanistic systems, and become more tolerant of approaches which are approximate in nature. If boolean logic is the interpreter of the conventional Aristotelian approach of not-contradiction, fuzzy logic is based on a more tolerant and vague approach that see its counterpart in the Buddha philosophy. The approach proposed by Zadeh has been revealed effective for the analysis of problems which are too complex or too ill-defined for the application of conventional quantitative techniques. Fuzzy logic is based on the introduction of the so called *linguistic variables* whose values are not numbers but words or sentences in a natural or artificial language. In particular, Zadeh started with the introduction of the concept of fuzzy set [57]. It is a labelled class defined by a continuum of grades of membership that ranges in $[0, 1]$ in its universe of discourse. The introduction of fuzzy set allows to provide a natural way of dealing with problems in which the bound of generic classes are uncertain and not well-sharped.

2.1 The Human-Machine Cooperation Issue

In control systems, information and computational science fields, any tool developed especially by means of machine learning techniques, should be designed in order to maximize given properly formulated objective function (or set of objective functions) for efficiently solving the problem under analysis. However, another aspect to take into consideration in problem solving is the human interpretability of the solution found. There are many cases in which it is necessary that the expert operators must be able to provide some insight as to how the solution derives its outputs. Especially in industry and medical applications, tools developed by means of data driven and machine learning approaches, have to be robust and affordable besides being efficient, therefore their interpretability should be strengthen.

In literature are mentioned three type of modelling systems: *black-box*, *white-box* and *grey-box*. White-box modelling assume that everything about the system is known a priori, expressed either mathematically or verbally. In contrast, in black-box modelling, a model is constructed entirely from data using little additional a priori knowledge. For example, in artificial NNs, a structure is chosen for the network and the parameters are tuned to fit the observed data as best as possible. Such parameters are not human interpretable and do not offer any insight about the modelled system. Grey-box modelling is an intermediate approach. It takes into account certain prior knowledge of the modeled system to provide the black-box models with human-interpretable meaning.

Fuzzy modelling techniques can be viewed as grey-box modelling. In fact, with the introduction of the linguistic variables, fuzzy sets, linguistic operators, rule inference systems etc., FIS paradigms allow the modeller to extract and interpret the knowledge contained in it as well as to put some a priori knowledge. Indeed, it is important to remark that fuzzy systems, as well as they were conceived, can be constructed by means of using the knowledge provided by human experts. However, FISs based on expert operators may suffer from a loss of accuracy for ill-known or

data-intensive models that is the main incentive for using fuzzy rules inferred from data. It is then clear the difficulty of improving systems interpretability without losing performance. With the development of machine learning and soft computing techniques, the trade-off of accuracy and interpretability has been the object of active research lines, particularly in fuzzy modelling. In [58] it is affirmed that the utilization of nature inspired methods such as EAs is a valid solution tu pursue in order to preserve some degree of interpretability. The authors in [59] discuss the potential human-computer cooperation and problem solving efficiency in FISs designing. There are investigated the most known and widely adopted techniques based on machine learning for generating the FISs rule base system.

In order to strengthen the human-computer cooperation, the first suggestion for having a better human interpretability of the system that has to be model is the injection of human expert knowledge. However such procedure is not always practicable, it depends mainly on the optimization procedure adopted.

The adoption of the Ockham Razor criterion is another suitable solution to take into consideration. As explained in [60], the Ockham Razor imposes the inclusion of the system optimization complexity in the objective function formulation.

In the following chapters two rule base synthesis procedure are going to be discussed and implemented for the study under analysis. The first is based on the input domain grid partitioning for designing the MFs. The second is based on the adoption of clustering methods for directly generating the Rule Base system.

The first procedure has been considered more interpretable since it is know how the FIS MFs are generated. However, it is also strongly limited since the maximum number of rules increases exponentially with the increasing of the input space dimension (*i.e.* number of linguistic variables). Therefore their synthesis is tricky, especially when it relies on EAs (GAs in our case).

On the contrary, the utilization of clustering algorithms seems to be more practicable and easy to implement in case of the problem complexity increases. However,

as will be better argued in the following chapters, the utilization of clustering algorithms requires to switch to a supervised problem.

2.2 Fuzzy Logic

In [61] a linguistic variable is defined as a variable whose values are words or sentences in a natural or artificial language. The linguistic variable is characterized by the quintuple $(\mathscr{L}, T(\mathscr{L}), U, G, M)$ in which $\mathscr{L}$ is the name of the variable; $T(\mathscr{L})$ is the term set of $\mathscr{L}$, that is, the collection of its linguistic values; U is the universe of discourse; G is a syntactic rule which generates the terms in $T(\mathscr{L})$; and M is a semantic rule which associates with each linguistic value X its meaning. The meaning of a linguistic value X is characterized by a compatibility function, $c : U \rightarrow [0,1]$, which associates with each u in U its compatibility with X.

For example, considering a person with the age of 30 years old, we want to determine if this person is "Old" or "Young" by assuming that the "Age" of an average human being ranges between "0" and "100" (*i.e.* the upper and lower reference limits). In this problem the linguistic variable is the "Age" of the person and it is numerical value is 30. The universe of discourse U referred to the linguistic variable "Age" is then numerically defined in the supposed range $[0;100]$. Moreover, in U is defined a term set composed by two linguistic values, "Old" and "Young". In Figure 2.1 a graphical representation of two possible trapezoidal compatibility functions is illustrated referred to the "Old" and "Young" linguistic values. As shown in figure, the fuzzy reasoning states that a person with an "Age" of 30 years old is "Young" 0.78 and "Old" 0.21. It means that there is a quantified compatibleness with both the term set linguistic values. Therefore, in a first study the person is both "Young" and "Old" with a compatibility value of 0.78 and 0.21, respectively.

From here to the end of the manuscript, linguistic values and their compatibility functions will be referred to as fuzzy sets and Membership Functions (MFs), respectively. It is important to remark that the two computed membership values does not

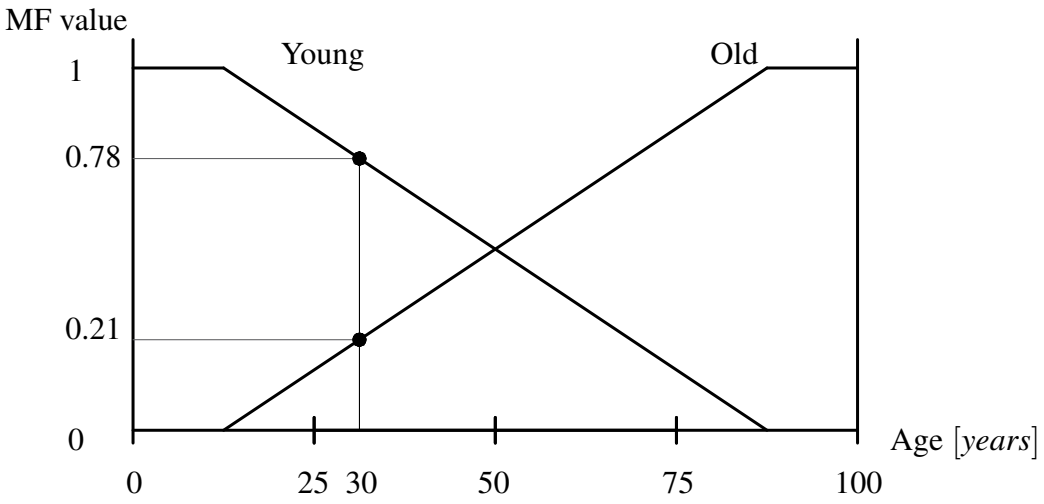


FIGURE 2.1: A Fuzzy term set example composed by two MFs.

express a probability but rather a compatibility (*i.e.* membership) degree. This process of MF attribution is referred with the term *fuzzification*. It consist in reading a crisp input value X (the age of a person in this case) and evaluate the corresponding term set MFs. The domain covered by a MF is known as *support* of the MF. In fuzzy logic, MFs supports of a given term set have to cover the overall universe of the discourse domain in order to preserve it from gaps of knowledge. Moreover, the adoption of fuzzy logic implies that the term set MFs are supposed to be overlapped otherwise the principle of the fuzzy reasoning would not be pursued. Indeed, the not MFs overlapping would be equivalent to approach to boolean logic, therefore the respect of Aristotele's principle of not-contradiction. Normally, the MFs overlap is manually set between the $20-50\%$.

In literature, different MF shapes has been proposed. The most common are triangular, trapezoidal, Gaussian and bell shaped. Gaussian and bell shaped MFs are featured by supports which cover the overall domain. Therefore, unlike triangular and trapezoidal, the covering of the overall fuzzy domain is assured if exists at least one fuzzy set per term set.

The relation between fuzzy sets by means of suitable AND-OR connection operators is able to define a premise of a sentence, that is also known as *rule antecedent*.

It is clear that the formulation of a sentence (or inference) needs both the premise

and the consequent part, that is also known as *rule consequent* (RC). Indeed, if we take into consideration the example in Figure 2.1, it lacks of any utility since the absence of the consequent part for closing the sentence. It lacks of interpretability because of the linguistic values are not coupled with any deduction.

In Fuzzy models, the RC part of a sentence is interpreted by a fuzzy set which belongs to the linguistic variable referred to the fuzzy output space. Depending about the application field, the RC is usually referred to a state, either a class or an action. It is clear that in case the output linguistic variable are more than one, the formulated model is a Multi Input Single Output system.

The matching of the antecedent part with the consequent part generates the so called IF-THEN rule. An example of a generic IF-THEN rule can be the following

$$\text{IF the } \textit{pressure} \text{ is } \textit{high} \text{ THEN the } \textit{volume} \text{ is } \textit{small} \tag{2.1}$$

Where *pressure* and *volume* are linguistic variables of input and output, respectively, *high* and *small* their corresponding linguistic values or fuzzy sets. The rule structure illustrated is called Mandani fuzzy IF-THEN rule. A generic Mamdani rule can be formalized as following

$$\text{IF } X_1 \textit{ is } A_i^1 \text{ AND ... AND } X_n \textit{ is } A_j^n \text{ THEN } Y \textit{ is } B_z \tag{2.2}$$

Where $\mathbf{X} = \{X_1, ..., X_n\}$ is the set of the linguistic variable numerical values. A_j^n is the j-th fuzzy set of the n-th term set and the consequent fuzzy set B_z is referred the z-th MF of the output term set.

Alternatively, as proposed in [62] the RC fuzzy set can be replaced by a mathematical equation as

$$\text{IF the } \textit{velocity} \text{ is } \textit{high} \text{ THEN the } \textit{force} \text{ is } k \cdot (velocity)^2 \tag{2.3}$$

This formulation is called Takagi and Sugeno's fuzzy IF-THEN rule. It is clear

that to switch from Mamadani type to Takagi Sugeno rule B_z (2.2) must be replaced by the RC equation in (2.3).

In the rules formulation, the AND operator defines the conjunction of the fuzzy sets involved in the same rule and therefore the intersection of their respective MFs. The THEN operator is the connector between the rule antecedent part and its consequent part. The construction of a rule can be also formulated by means of the utilization of OR operators in order to consider more fuzzy sets of the same linguistic variable or term set.

The mathematical interpretation of the AND-OR operators between fuzzy set MFs is given by the T-norm operators which operate the intersection between MFs and the S-norm operators which realize the MFs union [63], respectively. There are several T-norm and S-norm operators proposed in literature. Without going in too much details, for T-norm operators it is mostly used the *min* function or the scalar product as for S-norm operators the *max* function. The formulation of more rules in the same model generates the inferential system, *i.e.* the Rule Base system. The AND (intersection) operator is more commonly used with respect to the OR operator for the single rule synthesis. Instead, the OR operator is used to perform operations between rules, namely for the defuzzification process.

2.3 Fuzzy Inference Systems

FISs are also known as fuzzy-rule-based systems, fuzzy models, fuzzy associative memories, or fuzzy controllers when used as controllers. The first type of fuzzy rule base system that deals with real inputs and outputs was proposed by Mamdani (1974) [64], who was able to augment Zadeh's initial formulation in a way that allows it to apply a fuzzy system to a control problem. In the same period, fuzzy modelling and identification were studied by Takagi and Sugeno that have found numerous practical applications in control, prediction and inference modelling [65], [62].

In Figure 2.2 is shown the generic fuzzy rule base system structure introduced by Mamdani [66].

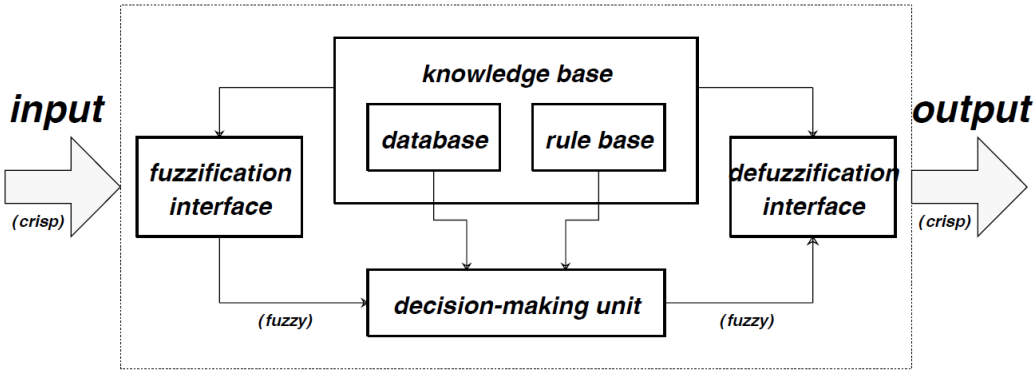


FIGURE 2.2: Generic FIS scheme.

As shown in figure, a FIS is basically composed of five functional blocks

- a rule base containing a number of fuzzy IF-THEN rules;
- a database which defines the MFs of the fuzzy sets used in the fuzzy rules;
- a decision-making unit which performs the rule inference operators;
- a fuzzification interface which transforms the crisp inputs into degrees of match with linguistic values;
- a defuzzification interface which transform the fuzzy results of the inference into a crisp output.

Usually, the rule base and the database are jointly referred to as the knowledge base and the rule base is considered the main core of FISs.

The steps of the fuzzy reasoning performed by FISs are the followings:

- Fuzzification, compare the input variables with the MFs on the premise part in order to obtain the membership values of each fuzzy set.
- Combine through a specific T-norm operator the membership values on the premise part to get firing strength of each rule.

- Generate the qualified consequent (either fuzzy or crisp) of each rule in function of the rule firing strengths.

- Defuzzification, aggregate the qualified consequents to produce a crisp output.

In [65] FISs are summarized in three types. However, for the sake of simplicity, here are illustrated the two type of FISs that are going to be studied in the following chapters, namely Mamdani-based type FISs and Takagi and Sugeno-based type FISs. Their different mechanism of reasoning is described in Figure 2.3 where two similar Multi Input Single Output (MISO) FIS are illustrated. In both cases two fuzzy input x and y are considered, both featured by 2 fuzzy sets and one fuzzy output z which is featured by two fuzzy set. In both cases, the crisp value of z is evaluated through the calculation of the two rules R1 and R2.

Here below are detailed the two type of FISs.

- Mamdani-based type:

  The overall fuzzy output is derived by applying the *max* operator to the qualified fuzzy outputs (each of which is equal to the minimum of firing strength and the output MF of each rule). Various schemes have been proposed to choose the final crisp output based on the overall fuzzy output; for example the centroid of area, bisector of area, mean of maxima and maximum criterion. The defuzzification criteria illustrated in Figure 2.3 is the Centroid of Area, however the adoption of the Center of Gravity is more often preferred because of the easier implementation.

- Takagi and Sugeno-based type:

  In this model Takagi and Sugeno's fuzzy IF-THEN rules are supposed to be used. The same *max* T-norm operator used in Mamadani type is shown in figure. The consequent of each rule is a linear combination of input variables plus a constant term (*i.e.* first order Takagi and Sugeno rule). The final output z is the weighted average of each RC. Most of the differences between this

model and the Mamdani model come from the specification of the consequent part (bell-shaped MFs, or crisp functions). Polynomial or other functions can be adopted as well for the consequent part definition. In case of zero order IF-THEN rules, the FIS becomes again a Mamdani type featured by the Center of Gravity operator.

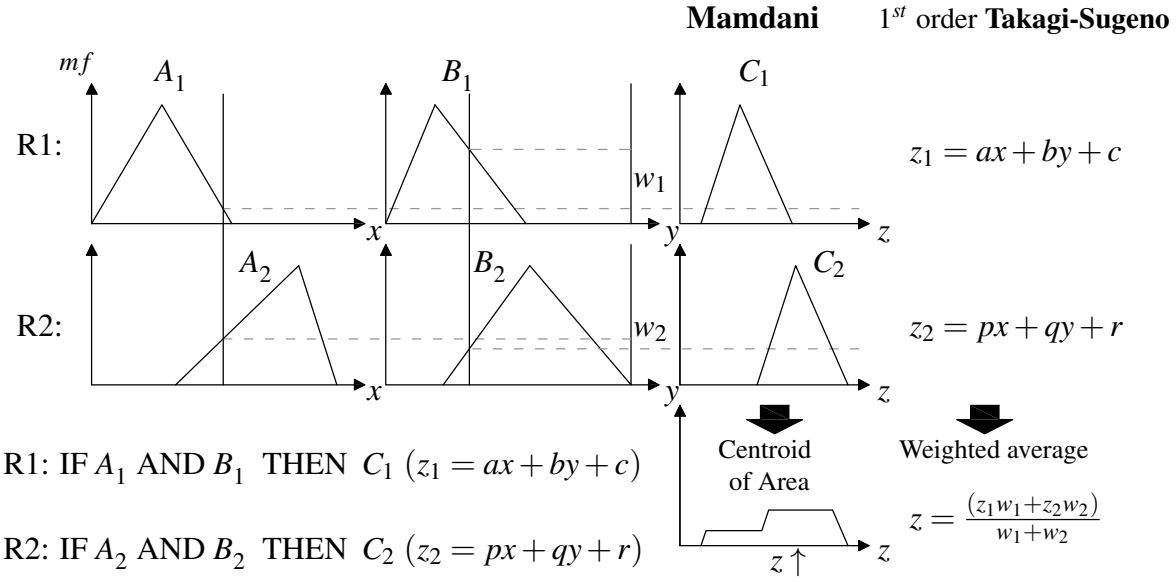


FIGURE 2.3: FIS Process example.

2.3.1 The Mamdani FIS Adopted Model

The MISO FIS here described is modelled in order to be optimized by a suitable evolutionary optimization algorithm. It is a Mamdani type and the fuzzy sets are triangular and trapezoidal at the input domain margins. They are equally spaced and featured by an overlap of the 50%. The rules are constituted by only OR conjunctions defined by the *max* operator. Each rule antecedent part is featured by a number of fuzzy sets equal to the number of linguistic variables. Besides, in the same rule each fuzzy set must belong to a different term set. It is assumed that no knowledge by experts is injected a priori and the number of rules is equal to all the possible combinations between the input fuzzy sets for the rule antecedent definition in order to assure the overall coverage of the input space by the output surface. To each rule is

assigned a rule weight parameter defined in $(0;1]$. It attributes to the corresponding rule its relevance respect with the other rules. The exclusion of the zero value assures to preserve the coverage of the overall fuzzy domains by the support of the MFs. Rule weights parameters can be defined a priori by some expert knowledge or by a suitable optimization algorithm.

2.3.2 ANFIS Model

In in the early 1990*s* in [65] it was introduced a FIS model variant named Adaptive Network-based Fuzzy Inference System that is also called Adaptive Neuro Fuzzy Inference System or ANFIS. ANFIS models are a class of adaptive networks that incorporate both NNs and fuzzy logic principles. The model is mainly based on Takagi–Sugeno FISs. As NNs, ANFIS is a supervised learning algorithm where the inference system has the learning capability to approximate non-linear functions by piece wise linear function approximation. As written in [65] where a comparison between artificial NNs and fuzzy models are listed and discussed, the ANFIS architecture was mainly conceived to model non-linear functions, identify non-linear components on-line in a control system, and predict chaotic time series.

In Figure 2.4 an generic ANFIS model composed by 2 rule is shown. It is equivalent to the 1^{st} order Takagi-Sugeno FIS introduced Figure 2.3 [67].

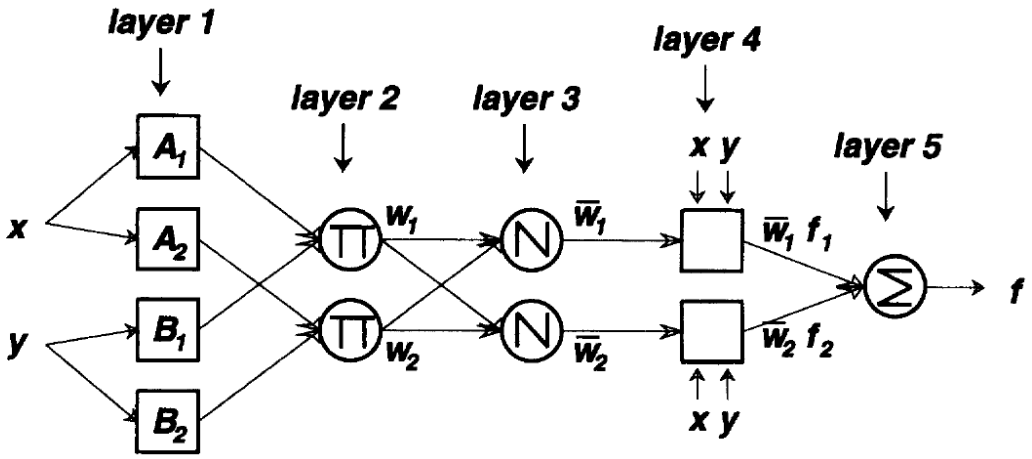


FIGURE 2.4: ANFIS 2 rule based Network example.

As shown in figure, the ANFIS network is composed by 5 layers. The first layer applies the fuzzification process in which all the MFs are calculated. In the second layer are calculated the rule fire strengths and, following, a normalization of them is applied:

$$\bar{\omega}_i = \frac{\omega_i}{\omega_1 + \omega_2} i = 1,2 \tag{2.4}$$

Therefore, in layer 4 each RC hyperplane is evaluated and following their weighted sum is applied for the crisp output evaluation.

In literature, there are proposed several technique of synthesis and optimization. Jang in [65] has formulated the ANFIS model formulation for the application of the gradient descent with the purpose to optimize it similarly to a simple feed forward NN featured by one hidden layer. However, as affirmed by the same author, descendent gradient application is notorious for its slowness and tendency to become trapped in local minima.

Similarly to FIS models, one of the most effective techniques used for the ANFIS synthesis and optimization relies on GAs [66].

As widely explored in some publications, [68] [69], the ANFIS synthesis can pass through clustering applications as well.

The following chapters will consider the modelling of ANFISs using clustering algorithms that consider multidimensional MFs. It is a variant of the model proposed in [68] [69] where the ANFIS MFs are unidimensional. This model can compress the information of the rule antecedent part in the single MF since each MF has the same dimension of the input space. Moreover, the coverage of the input domain by the MFs is always assured thanks to the adoption of multivariate Gaussian MFs regardless of the ANFIS input term set cardinality.

Going into details, a generic multivariate Gaussian MF $\Phi(\bar{\mathbf{u}})$ can be formulated as follows

$$\Phi(\bar{\mathbf{u}}) = e^{-\frac{1}{2}(\bar{\mathbf{u}}-\mu)\cdot\mathbf{C}^{-1}\cdot(\bar{\mathbf{u}}^T-\mu^T)} \tag{2.5}$$

where $\bar{\mathbf{u}}$ is the input vector, μ and $\mathbf{C}$ are the centroid and the covariance matrix of the multivariate Gaussian function.

Instead, the j-th rule is

$$\text{Rule-}j : \text{if } \bar{\mathbf{u}} \text{ is } \Phi^{(j)} \text{ then } y = \mathbf{u} \cdot \theta^{(j)} \tag{2.6}$$

where the the MF value $\Phi(\bar{\mathbf{u}})$ is the rule fire strength. The second term y is the output associated to the j-th rule. It is estimated through the calculation of the associated RC hyperplane, defined by the coefficients $\theta^{(j)}$. It is noteworthy that 2.6 considers the hyperplane coefficients θ that has the same length of the input array and its intercept θ_0. In order to evaluate the hyperplane intercept along with the proper coefficients, it is sufficient to augment $\mathbf{u}$ by appending a first column of 1's, hence $\mathbf{u} := [\mathbf{1}\ \mathbf{u}]$.

In Figure 2.5 it is shown a scheme of the proposed ANFIS architecture. It has an input array $\mathbf{u}$ of dimension n, k rules (MFs) and a unidimensional generic output y. The figure illustrates the fuzzification process with the MFs evaluation, following the hyperplanes evaluation $y^{(i)}$ coupled with their respective the rules firing strengths $f^{(i)}$ and finally the WTA operator. As written before, due to the MF multidimensionality and the absence of AND-OR operators, in this model the rule firing strengths coincide with their respective MF values, *i.e.* $f^{(i)} \equiv \phi^{(i)}$.

Without going into too much details, the proposed ANFIS model is suitable for the synthesis by means of the adoption of clustering techniques. In fact, by considering a labelled pattern set, the application of a clustering algorithm on the pattern set would be able to detect the ANFIS MFs. Therefore, once the MFs are defined (*i.e.* the rule antecedent part is generated), the respective rule hyperplanes (*i.e.* the RC part) can be calculated by relying on the labels coupled with the patterns of each corresponding cluster.

Respect with EAs, clustering algorithms should facilitate the modelling of ANFISs featured by huge input vectors. For example, if it is considered a Mamdani-type

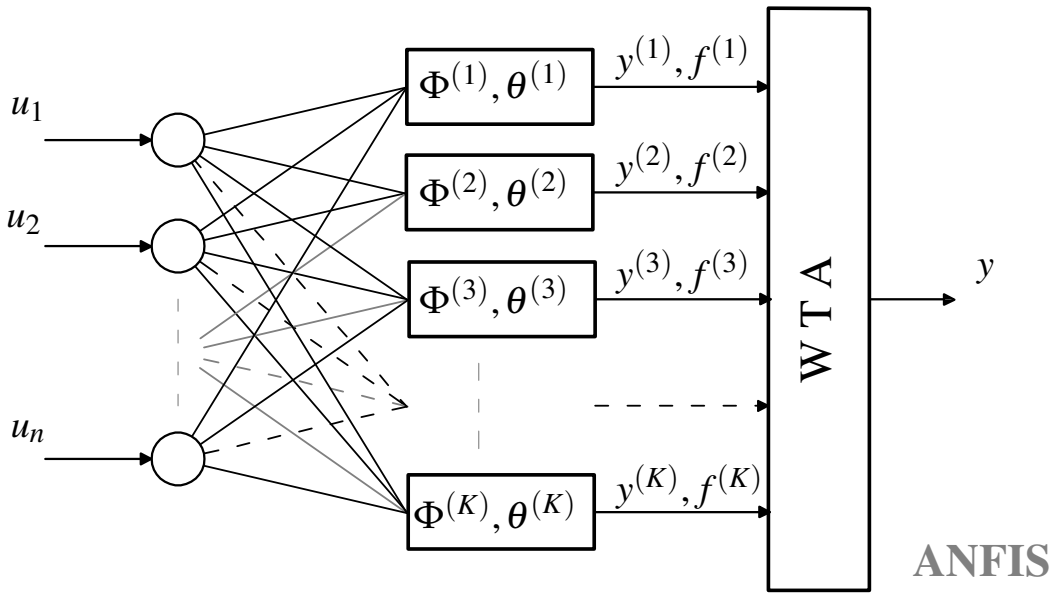


FIGURE 2.5: A general ANFIS scheme.

FIS (see Sec. 2.3.1) which is constituted by 3 inputs each one featured by a term set cardinality equal to 3 (*i.e.* each input variable can be approximated by 3 MFs), since no expert knowledge about the problem under analysis is supposed, the FIS has a total number of rules (*i.e.* rule base cardinality) equal to $3 \cdot 3 \cdot 3 = 3^3 = 27$, namely the term set cardinality elevated by the input array dimension. By considering the same term set cardinality on the output (*i.e.* 3 fuzzy sets), the total number of possible solution that can be referred to the optimization complexity is about $27^3 \cong 19.7 \cdot 10^3$. This value correspond to the total number of rules (*i.e.* 27) elevated by the FIS granularity (*i.e.* 3 MFs for each term set), namely $\left(3^{3 \cdot 3}\right)$. If the number of inputs is increased to 4 and each term set is composed by 3 fuzzy sets, the number of rules increases to $4^3 = 64$. Therefore, the optimization complexity increases to $4^{(3 \cdot 3)} \cong 262 \cdot 10^3$. The example highlights that with the increasing of the input space and the FIS granularity, the optimization complexity increases exponentially. Hence, the utilization of EAs for the FIS optimization can be efficient only when the number of the optimization parameters are properly limited.

Chapter 3

FIS-Hiererchical Genetic Algorithm Paradigm

GA-FIS optimization paradigm has been proven to be an effective tool because of its flexibility and pervasiveness since it is a free derivative solution. In GA-FIS paradigm, the GA is in charge to optimize a FIS by means of the iterative generation of individuals which are featured by chromosomes composed by sequences of genes in which every gene is in charge to tune a given FIS parameter.

In this chapter the FIS model taken into consideration as a reference architecture is described in Chapter 2 Sec. 2.3.1.

About the GA chromosome structure, in nature two class of genes are generally distinguished, namely control and parametric. These are referred in nature to those DNA genes that correspond to regulatory sequences and structural genes. in particular, regulatory sequences are capable to control the activation and deactivation of structural genes. In literature, GAs featured by both type of genes are referred to *hierarchical* GAs. In GA-FIS paradigm, parametric genes correspond to specific FIS parameters while control genes are in charge to activate or deactivate given subsets of parametric genes which code for any protein product [60].

In this chapter, a hierarchical GA optimization procedure is conceived in order to both improve the FIS effectiveness and decrease the number of FIS parameters, namely its complexity and computational burden. In particular, since the FIS-GA paradigm studied has been conceived to be flexible, robust and efficient, it can be

executed optimizing the FIS parameters in sub-steps, or said in other words, by considering more sub-optimization processes developed in series. This procedure has been conceived in order to reduce the overall computational burden of the GA optimization. Moreover, an adaptation of the hierarchical GA for solving Multi Objective (MO) optimization problems is described.

In the following of this chapter the initial FIS reference architecture is described in Sec. 3.1. In Sec. 3.1.1 all the optimization parameters describing the FIS are listed and commented in order to explain their functionality and better understand how the FIS synthesis process works. Following, in Sec. 3.2 it is introduced and detailed the GA adopted, describing each operator performed during the optimization procedure. Finally, in Sec. 3.2 a MO GA adaptation is described.

3.1 FIS Optimization

In order to execute a properly synthesis and optimization of the FIS model under analysis a basic standard architecture must be defined. In particular, for each parameter subject to the optimization process it is necessary to define its range of existence, its initial default value and the type of gene (real, binary etc.).

As stated in Chapter 2 Sec. 2.3.1 the FIS is structured as a Mamdani model featured by trapezoidal and triangular MFs and a *center of gravity* operator.

The FIS optimization parameters here considered are:

- the RCs of each rule
- the MFs position and shape
- the rule weights
- the activation parameters which turn on/off the input MFs

Each set of parameters is defined and distinguished by a specific subset of genes of a specific nature. In particular, those genes dedicated to activate or deactivate the

FIS MFs (*i.e.* MF deletion parameters) are of control nature whereas the rest of them are parametric genes.

The initial structure of the FIS is conceived as following. Each term set is initially composed by N_{mf} symmetric MFs having triangular and trapezoidal (at the domain boundaries) shape characterized by 50% overlap as illustrated in Figure 3.1 for $N_{mf} = 5$.

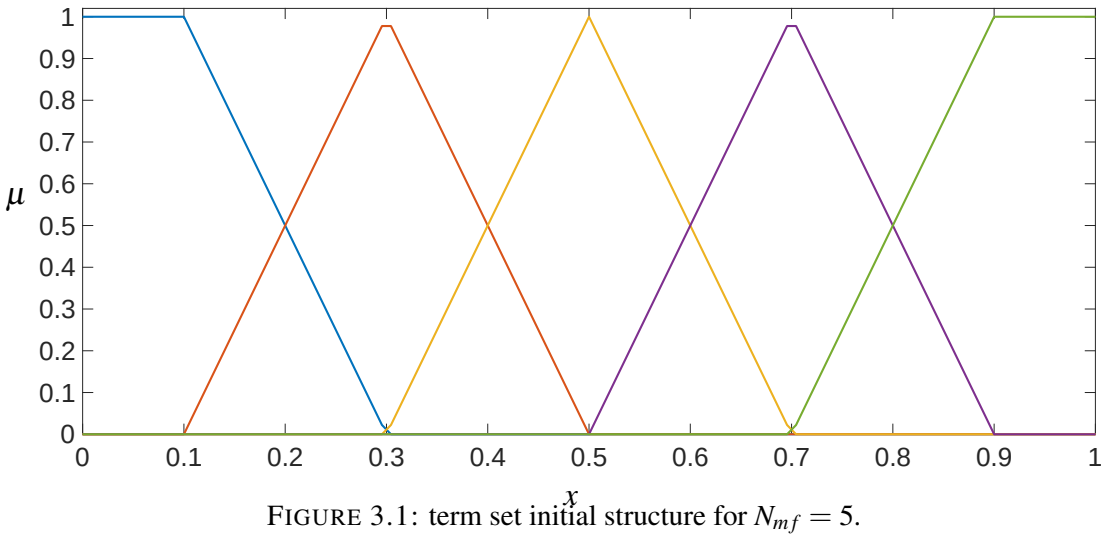


FIGURE 3.1: term set initial structure for $N_{mf} = 5$.

Assuming that the FIS has N_{in} inputs and a single output, the rules are characterized by N_{in} antecedents, each one defined by a MF for each input term set. Indeed, the maximum number of rules N_{rules} is given by all the possible antecedents combination, namely $N_{rules} = N_{mf}^{(N_{in})}$. Therefore, the structure of the i-th fuzzy rule can be written as

$$\begin{gathered}\text{IF } (x_1 \text{ is } A_{i1}) \text{ AND } (x_2 \text{ is } A_{i2}) \text{ AND...AND } (x_{N_{in}} \text{ is } A_{iN_{in}}) \text{ THEN } y \text{ is } B_i, \\ i = 1,2,...,N_{rules}\end{gathered} \tag{3.1}$$

where, x_j, y, A_{ij} and B_i are the j-th input, the output, the MF associated with the j-th antecedent and the MF associated with the consequent of the i-th rule, respectively.

It is assumed that each fuzzy rule must consider only one MF for each term set. Therefore, in case of MF deactivation, the rules including that MF must be deactivated too. For each rule it is attributed a given rule weight $w_i \in (0,1]$ (for

$i = 1, ..., N_{rules}$) in order to state the corresponding rule relevance respect to the others. The defuzzification is applied through the calculation of the center of gravity (see Sec. 2.3.1).

In case of MF deactivation, the deleted MF may leave uncovered space in the input domain since triangular and trapezoidal MFs, unlike bell shaped MFs, are characterized by finite supports. For this reason, it has been imposed that, whenever a MF is deactivated, the nearest base points of the adjacent (active) MFs are crossed and placed under the vertex of the corresponding new adjacent fuzzy set. This ensures the covering of the overall fuzzy domain and a given fixed overlap (50%) of the adjacent MFs. Moreover, in each fuzzy domain the number of MFs in the corresponding term set must be equal or greater than two.

In the following section are described all the sets of the FIS parameters that are subjected to the optimization process.

3.1.1 FIS Optimization Parameters

The set of all the FIS parameters is indicated by Ω. It is composed by four subsets named Θ_{MF}, Θ_{RW}, Θ_{CA} and Θ_{MH}, representing the MFs shape, the rule weights, the RCs association and the MF activation, respectively. The four subsets of parameters are described in details here below.

- *Subset* Θ_{MF}

  The subset of the MFs parameters Θ_{MF} contains the information about the FIS MFs shape and position. According with [70], each triangular MF is described by three parameters each one ranging in a specific interval of its support domain. More precisely, the i-th triangular MF is described by means of l_i, c_i and r_i, representing the abscissas that define the base and the height of the triangle (see Figure 3.2). These points range between their related upper and lower bound named *min* and *max*, respectively, as illustrated in Figure 3.2. The MF

upper and lower bound are formulated in function of l_i, c_i and r_i as stated in (3.2).

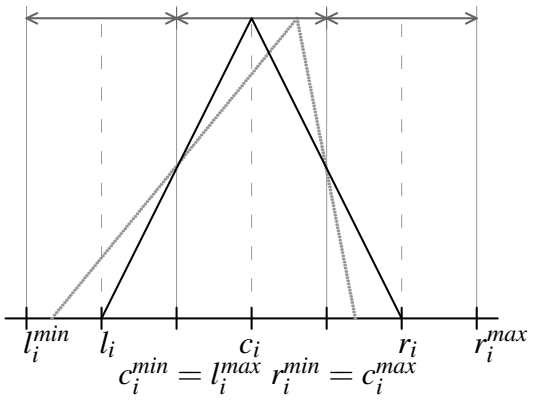


FIGURE 3.2: Parameters and related ranging intervals defining the limits of the triangular MF in black. In dotted an example of its possible mutation.

$$\begin{aligned} l_i^{min} &= l_i - \tfrac{c_i - l_i}{2} \\ c_i^{min} = l_i^{max} &= l_i + \tfrac{c_i - l_i}{2} \\ r_i^{min} = c_i^{max} &= r_i - \tfrac{r_i - c_i}{2} \\ r_i^{max} &= r_i + \tfrac{r_i - c_i}{2} \end{aligned} \tag{3.2}$$

This formulation is extended to all the FIS MFs including the trapezoidal ones. Being each support MF defined in $[0,1]$, those MFs of trapezoidal shape at the edges of the domain can be defined by two points as shown in Figure 3.1.

Since each term set is composed by 3 triangular and 2 trapezoidal MFs (*i.e.* $3 \cdot 3 + 2 \cdot 2 = 13$ points) the Θ_{MF} cardinality is equal to 13 times the overall number of FIS term sets. Thus, Θ_{MF} cardinality is $(N_{in} + N_{out}) \cdot 13$ where $N_{out} = 1$ since it is considered a MISO model. More precisely, the subset of MF parameters is given by $\Theta_{MF} = \{l_1, c_1, r_1, l_2, c_2, r_2, ..., l_{N_{MF}}, c_{N_{MF}}, r_{N_{MF}}\}$, where N_{MF} is the maximum number of MFs composing the FIS ($N_{MF} = N_{mf}(N_{in} + 1)$). The Θ_{MF} parameters are initially set so as to generate the FIS term sets of the same form shown in Figure 3.1.

- *Subset* Θ_{RW}

  The Θ_{RW} subset is composed by the rule weights of the FIS. Each element ranges in $(0,1]$. The exclusion of the zero value prevents the possibility that the FIS takes no decision. It is clear that the cardinality of Θ_{RW} is equal to the maximum number of rules N_{rules} of the FIS (*i.e.* the term set cardinality times number of inputs, $N_{MF}^{N_{in}}$), therefore the subset has the form $\Theta_{RW} = \{w_1, w_2...., w_{N_{rules}}\}$ where the i-th element w_i correspond to the rule weight of the i-th rule. All rule weight values are initialized to 1 in order to equally attribute the same importance to each rule.

- *Subset* Θ_{RC}

  The subset Θ_{RC} state the consequent part, namely the output MF of each rule. It has the following form $\Theta_{RC} = \{q_1, q_2, ..., q_{N_{rules}}\}$, where the i-th element is referred to the MF RC of the i-th rule. Each Θ_{RC} element is represented by an integer number that ranges in $[1; N_{mf}]$ where N_{mf} is the cardinality of the output term set (it is equal to the input cardinality).

  When it is not assumed any presence of expert knowledge, the definition of the initial value of Θ_{RC} is tricky. For this reason, in these cases the Θ_{RC} initial state will be randomly generated and then optimized. Moreover, since the rule base consequent part constitutes the FIS knowledge core, its optimization or raw synthesis must be prioritized.

- *Subset* Θ_{MH}

  Θ_{MH} subset includes one binary element referred to each designed fuzzy set in order to establish their activation state. More precisely, the i-th binary element b_i of the subset Θ_{MH} is used to switch on or off the corresponding MF in the related input term set. The Θ_{MH} cardinality is given by the total number of possible antecedentes fuzzy sets $N_{MH} = N_{in}N_{mf}$. Summarizing, the subset of the FIS rules hierarchy parameters is given by $\Theta_{MH} = \{b_1, b_2, ..., b_{N_{MH}}\}$. The

Θ_{MH} initial state foresees the activation of all the fuzzy sets (*i.e.* all the b_i elements $\in \Theta_{MH}$ are initially set equal to 1).

3.2 Genetic Algorithm

The genetic code coincides with the FIS parameters set $\Omega = \{\Theta_{RC}, \Theta_{MF}, \Theta_{RW}, \Theta_{HR}\}$. Accordingly, in the following the two meanings will be attributed to the symbol Ω equivalently. Due to the heterogeneous nature of the elements belonging to the i-th subset Θ_i, the generic gene can be a real, an integer or a binary number. All the genes are constrained by their respective upper and lower bound as described previously in Sec. 3.1.1. It is assumed that the group of genes corresponding to the parameters belonging to each of the four subsets can be optimized simultaneously or sequentially. In the latter approach, the overall genetic variables set is optimized in subsequent optimization steps. During each step the genes belonging to one or more subsets Θ_i are optimized simultaneously, whereas the remaining ones are kept unchanged, until all genes of the genetic variables set are optimized. It is assumed that each subset can be optimized only once during the entire optimization process. Consequently, once the parameters belonging to a given subset Θ_i are optimized, the corresponding genes are frozen for the rest of the GA optimization process. The FIS subsets not yet optimized are set to their initial state defined in the previous section. The number of variables simultaneously optimized during each step depends on the cardinality of the specific subset (or subsets) Θ_i considered in the specific optimization step. Once the initial population G_0 is set, evaluated and sorted according with a generic OF P, each step of the optimization process is subjected to the following sub-steps repeated for each new generation until the GA stops. Once a new generation is created, individuals belonging to the previous generation are no more taken into consideration. It is important to remark that the subset Θ_{RC} must be optimized in the first step in case of a sequential optimization. In fact, the Θ_{RC} form set the knowledge or inferential

core of the model. Indeed, the effective tuning of the other subsets (*i.e.* Θ_{MF}, Θ_{RW} and Θ_{HR}) depends on the Θ_{RC}.

Following, all the details concerning each sub-step of the evolutive optimization process are illustrated. For each sub-step the number of individuals is fixed equal to N_{ind}. Considering the k-th generation G_k, it is subjected to the following operations in order to go to the next generation G_{k+1}:

i The best N_{elite} individuals are selected, the best one is handed down to the next generation G_{k+1} by the $\Gamma_{identity}$ operator, whereas the remaining $N_{elite}-1$ are subjected to a small mutation defined by the Γ_{small} operator.

ii Starting from the current generation, an even number N_{cross} of individuals are chosen by the selection operator Γ_{select} and coupled in $N_{cross}/2$ pairs. Each couple is then subjected to the cross-over operator Γ_{cross}.

iii The best $N_{ind}-N_{elite}-N_{cross}$ individuals are extracted among those unselected of phase-ii. These individuals, which were supposed to be the best among the worst, are subjected to the big mutation operator Γ_{big} before of being integrated into the new offspring.

iv The new offspring is then simulated in order to compute the objective (fitness) function P and sorted.

v Individuals having a fitness greater than a fixed threshold γ_{th} are subject to the deletion operator Γ_{del} which calculates their deletion probability.

vi The N_{del} deleted individuals are substituted by new individuals randomly generated by the operator Γ_{rand}.

Each sub-step of the GA optimization process is arrested when the stop criteria is activated. The stop criteria is denoted with the Γ_{stop} operator.

A schematic representation of the sub-steps performed by the GA in order to evaluate a new generation is given in Figure 3.3.

In the following a detailed description of the genetic operators is given

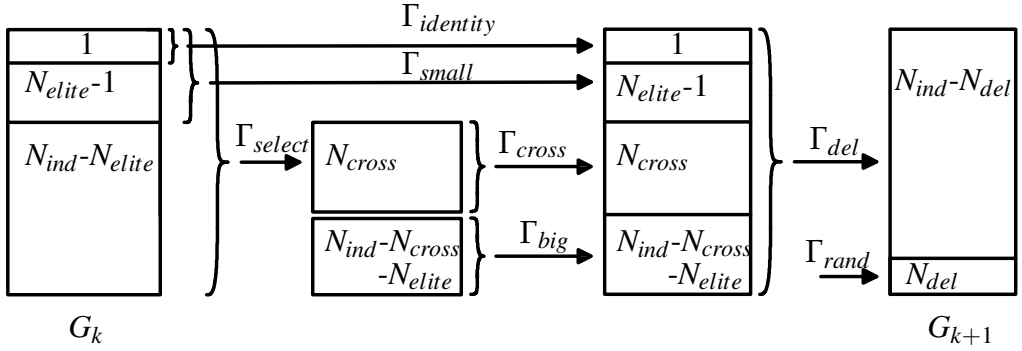


FIGURE 3.3: GA optimization procedure schematic representation.

- *Operator* Γ_{select}

  The Γ_{select} operator is a tournament selection operator. The N_T value defines the number of participants to each tournament.

- *Operator* Γ_{cross}

  The Γ_{cross} operator is a scattered cross-over operator. It is applied to a couple of individuals randomly chosen by those that are previously selected. This procedure is repeated until all the selected individuals have been crossed.

- *Operator* $\Gamma_{identity}$

  The $\Gamma_{identity}$ operator copies individuals in the next generation without altering their genes

- *Operators* Γ_{small} *and* Γ_{big}

  When an individual is subject to Γ_{small} or Γ_{big} operator their genes have a distributed probability of being mutated equal to γ_{freq}. The mutation consists in altering the selected gene by adding a random real number extracted from a uniform distribution between the interval $[\gamma_{dist}, -\gamma_{dist}]$ in case of Γ_{small}. For Γ_{big} the interval is extended by an integer scale factor γ_{scale} in order to enhance the exploration. The γ_{dist} value is expressed in percentage with respect to the difference between the respective upper and lower bound of each gene. Integer variables are applied to the same mutation operators and are successively rounded. Mutation rates, both big and small, decreases generation by

generation thanks to a specific exponential factor γ_{damp} applied to Γ_{small}. The mutation distances (*i.e.* γ_{dist}) are then exponentially reduced and at the N_{stop}-th generation they vanish. This procedure aims at encouraging the exploitation at the expense of the exploration capability with the increasing of the generations.

- *Operator* Γ_{del}

  The operator Γ_{del} assigns the probability p_{del} to delete an individual characterized by a fitness γ_{fit}, as follows:

$$p_{del} = \begin{cases} 1 - e^{-\frac{|\gamma_{th} - \gamma_{fit})|}{\gamma_{th}}} & \text{if } \gamma_{fit} > \gamma_{th} \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

  Therefore, the N_{del} deleted individuals are replaced by other N_{del} new individuals randomly generated. The new individuals are spawn by choosing the value of each gene considering a uniform distributed probability between their respective upper and lower bounds.

- *Operator* Γ_{stop}

  It defines the stop conditions of each GA sub-step optimization. If after N_{conv} consecutive generations it is not observed any improvement on the best fitness greater than a certain percentage γ_{conv} of the maximum number of allowed generations N_{stop}, the GA stops in advance the sub-optimization process. The process stops however after N_{stop} generations.

3.3 NSGA-II for MO Problems

Many real-world engineering problems are characterized by multiple objectives to be minimized (or maximized), such as the overall costs, energy losses, comfort, reliability etc. A very simple solution is to sum the OFs by means of a convex linear

combination in order to return to a single OF formulation

$$F = \sum \lambda_i \cdot f_{i i=1} | \sum \lambda_{i i=1} = 1, \lambda_i \in [0,1] \forall i \tag{3.4}$$

However, this trick needs to tune the OF meta-parameters λ_i in order to state which OF is more important or relevant compared to the others. It is clear that in this case the normalization in $[0,1]$ of every OF is mandatory for a correct definition of the meta-parameters. Although it is always suggested to apply the OF normalization, there are cases in which an accurate estimation of its upper and lower limits is not trivial as the a priori evaluation of the meta-parameters. To avoid this, several solutions have been proposed in the literature that do not need to establish a priori how to combine and order the OFs.

A successful MO variant of the GA is the well known NSGA-II (Non-dominated Sorting Genetic Algorithm) [71]. The MO-GA proposed wants to privileges those not-dominated individuals in a generation by introducing a suitable criterion based on the Pareto front evaluation.

3.3.1 The Non-Dominated Sorting Procedure

The non-dominated sorting procedure is based on the computation of how many individuals dominate another given individual by considering a MO optimization problem. In [71] it is defined as the *population count*. The population count consists in calculating the number of solutions n_p in a given population G that dominate the solution p. Given a population, the best set of individuals are those which owns the lowest population count value.

The proposed dominance criterion can be explained as follows: given two individual vectors X and Y defined by their respective OFs value $X = (x_1, x_2...)$ and $Y = (y_1, y_2...)$, the vector X dominates Y in case of:

$$X \prec Y \; when: \; x_i \leq y_i \; \forall \; i = 1,2... \; AND \; if \; \exists \; j \; \in \; \{1,2...\} \; | x_j < y_j \tag{3.5}$$

where $\prec$ is the dominance operator. Therefore, $X \prec Y$ only in case of all the OFs are lower or equal than the compared solution OFs and, in addition, at least one solution must be lower than its counterpart. It is clear that for each solution is necessary to calculate its respective population count. Said in other words, each solution must be compared with the rest of the solutions in order to find if it is dominated or not.

This procedure requires a number of comparisons that is approximately the population size times the number of objectives. At this stage, all individuals in the first non-dominated front are found. In order to find the individuals in the next non-dominated front, the solutions of the first front are discounted temporarily and the above procedure is repeated and so on until all the solutions are ranked in their corresponding front. The overall sorting procedure is summarized in Alg. 1

It is important to note that each front evaluated can be populated by more individuals. Hence, it is necessary to define an inner hierarchy in order to univocally rank each front and thus the entire population. In [71] it is discussed the utilization of the *crowding distance* function that has been demonstrated to be very effective in the literature. Given a non-dominated set of solutions (*i.e.*, the individuals assigned to the same front), the crowing distance evaluates the level of distance among the individuals. In this manner, the selection of similar individuals is avoided. The crowding distance computation requires to sort the population according to each OF value in ascending order of magnitude. Thereafter, the solutions with the smallest or largest OF values are assigned an infinite distance value. All the other (intermediate) solutions are assigned a distance value equal to the absolute normalized difference in the function values of two adjacent solutions. This calculation is applied to each OF. The overall crowding distance value is then calculated as the sum of individual distance values corresponding to each OF.

The pseudocode in Alg. 2 outlines the crowding-distance computation procedure of all the solutions in an non-dominated set $\mathscr{I}$. Here, $\mathscr{I}[i].m$ refers to the m-th OF value of the i-th individual in the set $\mathscr{I}$ and the parameters f_m^{max} and f_m^{min} are the maximum and minimum values of the m-th OF. After all members of the population

Algorithm 1 GA optimization procedure schematic front assignment representation.

```
1: procedure FAST NON DOMINATED SORT (G)
2:     for p ∈ G do                                   ▷ for each individual
3:         S_p = ∅
4:         n_p = 0
5:         for q ∈ G do                               ▷ for each individual
6:             if (p ≺ q) then                        ▷ if p dominates q
7:                 S_p = S_p ∪ {q}                    ▷ add q to the set of solution dominated by p
8:             else if (q ≺ p) then
9:                 n_p = n_p + 1                      ▷ increment the domination counter of p
10:            end if
11:        end for
12:        if (n_p = 0) then                          ▷ p belongs to the first front
13:            p_rank = 1
14:            F_1 = F_1 ∪ {p}
15:        end if
16:        i = 1                                      ▷ initialize the front counter
17:        while F_i ≠ ∅ do
18:            Q = 0                                  ▷ used to store the members of the next front
19:        end while
20:        for p ∈ F_i do                             ▷ for each p ∈ F_i
21:            for p ∈ S_p do                         ▷ for each q ∈ S_p
22:                n_q = n_q − 1
23:                if (n_q = 0) then                  ▷ q belongs to the next front
24:                    q_rank = i + 1
25:                    Q = Q ∪ {q}
26:                end if
27:            end for
28:        end for
29:        i = i + 1
30:        F_i = Q
31:    end for
32: end procedure
```

$\in \mathscr{I}$ are assigned to a metric distance, we can compare each solution in function of the proximity extension respect with the adjacent solutions. A solution with a smaller value of the distance measure is assumed to be more crowded respect with the other solutions. This is exactly what we compare in the proposed crowded-comparison operator that is described below.

Algorithm 2 Crowding distance assignment considering in an non-dominated set $\mathscr{I}$.

```
1: procedure FAST NON DOMINATED SORT (G)
2:     l = |I|                                   ▷ number of solutions in I
3:     for each i, set I[i]_distance = 0 do
4:         set I[i]_distance = 0                 ▷ initialize distance
5:     end for
6:     for each objective m do
7:         I = sort(I, m)        ▷ sort using each objective value so that boundary
       points are always selected for all other outputs
8:         I[1] = I[l]_distance = ∞
9:         for i = 2 to l − 1 do
10:            I[i]_distance = I[l]_distance + (I[i+1].m − I[i−1].m) / (f_m^max − f_m^min)
11:        end for
12:    end for
13: end procedure
```

By indicating with $\prec_n$ the crowded-comparison operator, it leads to the selection process at the various stages of the algorithm towards a uniformly spread-out Pareto-optimal front. Assuming that every individual i in the population has two attributes:

1. non-domination rank (i_{rank})
2. crowding distance ($i_{distance}$)

we now define a partial order $\prec_n$ as

$$i \prec_n j \text{ if } (i_{rank} < j_{rank}) \text{ or } (i_{rank} < j_{rank} \text{ and } (i_{distance} > j_{distance})) \tag{3.6}$$

that is, between two solutions with differing non-domination ranks, we prefer the solution with the lower (better) rank. Otherwise, if both solutions belong to the same front, then we prefer the solution that is located in a lesser crowded region. Once this

procedure has been completed for each front, the individuals of the given population are finally ranked and can be sorted.

Afterwards, the MO-GA can proceed with the definition of the next generation by applying the GA operators as introduced in 3.2. This operation is repeated for each generation until the the stop criteria is activated.

Because of the multiplicity of the OFs, also the stop criteria needs to be revisited. In particular, the early stop given by the operator Γ_{stop} is activated if it is not registered any improvement of the best individual greater than the threshold γ_{conv} for N_{conv} generations considering all the OFs. Otherwise, the GA stops in case of is reached the maximum number of generation N_{stop}.

Chapter 4

ANFIS Synthesis by Clustering

In literature, several methodologies for training FISs and ANFISs when expert knowledge leaks are proposed. In Mamdani-type FISs, the most straightforward approach consists in grid partitioning both input and output spaces in order to formulate the rule antecedents [72] and employ heuristics for selecting the RCs [73]–[75]. However, methods based on grid partitioning suffer from two major drawbacks: first, there is no guarantee that all fuzzy sets (and rules) will be fired; second, the synthesis complexity combinatorially increases with the grid resolution and the input dimensionality [72] as demonstrated in Sec.2.3.2 of Chapter 2. Further, finding the suitable grid granulation is another challenging task which is strictly problem-related.

In order to overcome these problems, a second approach based on data clustering is often preferred by adopting Takagi-Sugeno-type FISs or more in general ANFISs. Training such systems consists in performing a piece-wise linear approximation of the MISO function to be modelled. To this aim, clustering procedures can be used to determine the rule base system. The utilization of clustering algorithms implies a consistent reduction of the number of fuzzy sets (rules), hence a reduction of the model complexity [72]. Each cluster generated would be used to determine both position and shape of the multivariate-Gaussian MFs (*i.e.* the rule antecedent part), and the output hyperplanes (*i.e.* the RC part) are evaluated by considering all the input-output pairs included in each cluster using an ordinary least squares procedure.

At first glance, one might cluster the input or the output space in order to build

the MFs. However, these techniques rely on the strong assumption that similar pattern in the input space are also similar in the output domain, which is rarely satisfied. A possible solution is to consider the joint input-output space by means of a linear convex combination between dissimilarities in the input and output spaces. However, this technique, by privileging compactness in both spaces, does not assure building hyperplanes which well-approximate the resulting clusters. Hence, one can perform a proper hyperplane clustering, which should conceive well-approximating hyperplanes, at the risk of neglecting clusters' compactness in the input space. Therefore, two main strategies can finally be pursued: clustering the joint input-hypeplane space or, as suggested in [69], using a Neurofuzzy Min-Max classifier after hyperplane clustering.

In this chapter, several clustering algorithms are introduced and successively compared for the EMS synthesis problem. By starting from the widely-known k-means algorithm [76], [77] applied in the input space, the models studied are mainly re-adaptations and/or extensions of it (still, well-known in literature) in order to explore different (dis)similarity measures, investigate about the best clustering space to consider (input-hyperplane space) [69] and implement hierarchical clustering variants for accelerating the clustering process as well.

The following chapter is organized as follows. In Sec. 4.1 it is introduced the basic k-means algorithm followed by their variants on the (dis)similarity measures. Sec. 4.2 is focused on how to deal with the joint input-hyperplane and pure hyperplane space of clustering and therefore how to reformulate the problem in Sec. 4.1. In Sec. 4.3 a practical application of a MinMax classifier in supporting the the hyperplane clustering procedure is introduced. The chapter end with Sec. 4.4 where two model of hierarchical clustering, agglomerative and divisive, are respectively described.

4.1 Partitional k-means Clustering

k-means is an hard partitional clustering algorithm which, given a dataset $S = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{N_P}\}$, returns k non-overlapping groups (clusters), *i.e.* $S = \{S_1, ..., S_k\}$, such that $S_i \cap S_j = \emptyset$ if $i \neq j$ and $\cup_{i=1}^{k} S_i = S$, such that objects in the same cluster are more similar to each other than to those in other clusters. In order to find the k clusters, k-means aims at minimizing the following OF, namely the Within Cluster Sum of Squares (WCSS):

$$WCSS = \sum_{i=1}^{k} \sum_{\mathbf{x} \in S_i} \|\mathbf{x} - \mathbf{r}(i)\|_2^2 \tag{4.1}$$

where $\|\mathbf{x} - \mathbf{r}(i)\|_2^2$ is the squared Euclidean distance between pattern $\mathbf{x}$ and the i-th cluster representative $\mathbf{r}(i)$, usually known as *centroid*, defined as the component-wise mean amongst patterns in cluster S_i. Minimizing (4.1) is, however, an NP-hard problem [78] and what is commonly known as k-means is actually a heuristic which, as such, does not guarantee to find an optimal solution.

k-means is based on the Voronoi iteration or, equivalently, an Expectation-Maximization algorithm which works as follows:

i Select k initial centroids according to some heuristics (*e.g.* randomly);

ii Assignment (Expectation) Step: assign each pattern to nearest cluster (closest centroid);

iii Update (Maximization) Step: update the clusters centroids;

iv Loop ii–iii until a given stopping criterion is met (e.g. maximum number of iterations is reached or centroids' update is below a given threshold).

4.1.1 k-medians

A commonly used variant of the k-means algorithm consists in changing the (dis)similarity measure from squared Euclidean distance to 1-norm (also known as Manhattan, Taxi-Cab or CityBlock distance), leading to the so-called k-medians problem [79].

The 1-norm (dis)similarity measure implies to consider the Assignment Step in Sec. 4.1 and the OF (4.2) with no squares involved, but considering the absolute value only. Therefore, the k-medians OF shall be referred to as, more generally, the Within Clusters Sum of Distances (WCSD):

$$WCSD = \sum_{i=1}^{k} \sum_{\mathbf{x} \in S_i} \|\mathbf{x} - \mathbf{r}(i)\|_1 \tag{4.2}$$

k-medians still works by the Expectation-Maximization steps introduced in Sec. 4.1. However, in this case the cluster representative is the *median*, evaluated by taking the component-wise median rather than the mean amongst patterns in clusters. k-medians, thus, ensures that single medians attributes come from the dataset at hand, condition that is not guaranteed in k-means. Due to the minimization of the 1-norm rather than squared 2-norm, k-medians is more robust to noise and outliers with respect to k-means; indeed, the median is not (so-much) skewed in presence of (few) very low or very high values.

4.1.2 k-medoids

In k-medoids [80] the cluster representative (known as *medoid* or *MinSOD*[1]) is the cluster datapoint which minimizes the sum of distances within the cluster itself. Conversely to k-means and k-median, in k-medoids clusters representatives are actual members of the dataset at hand by definition. k-medoids, due to the representatives definition, can ideally deal with any (dis)similarity measures. The k-medoids problem has been solved by means of the implementation proposed in [81], which is based to the same Voronoi iterations at the basis of k-means and k-medians. Therefore, its OF can generally be defined as

$$WCSD = \sum_{i=1}^{k} \sum_{\mathbf{x} \in S_i} D(\mathbf{x} - \mathbf{r}(i)) \tag{4.3}$$

[1] Minimum Sum Of Distances

where $D(\cdot,\cdot)$ is the (dis)similarity measure. In this model, the adopted (dis)similarity measure for k-medoids is the Mahalanobis distance [82], defined as following:

$$d(\mathbf{x},\mathbf{r}(i)) = \sqrt{(\mathbf{x}-\mathbf{r}(i))^T \cdot \mathbf{C}_i^{-1} \cdot (\mathbf{x}-\mathbf{r}(i))} \tag{4.4}$$

where $\mathbf{C}_i$ is the covariance matrix for the i-th cluster and $\mathbf{r}(i)$ is its representative (*i.e.* the medoid).

Since the k-medoids algorithm minimizes the sum of pairwise distances rather than the sum of squares, it is more robust to noise and outliers with respect to k-means.

4.2 Input-Hyperplane & Pure Hyperplane Clustering

Albeit clustering is an unsupervised problem by definition, it can also be considered a dataset at hand that consists in labelled patterns; thus, having the form

$$S = \{(\mathbf{x}_1,y_1),(\mathbf{x}_2,y_2),...,(\mathbf{x}_{N_P},y_{N_P})\} \tag{4.5}$$

where $\mathbf{x}_i \in \mathbb{R}^{N_F}$ and $y_i \in \mathbb{R}$, for $i = 1,...,N_P$, which we refer to as *input* and *output* space(s), respectively.

In this section it is considered a clustering problem in the joint input-hyperplane space and its restriction to the only hyperplane space. In order to consider it, the clusters representatives must be re-defined. Indeed, if one has to work in the input space (*i.e.* with unlabelled patterns), the clusters representatives as defined in Sec. 4.1 and Sec. 4.1.2 suffice. Conversely, as concerns joint input-output spaces, each cluster will be described by:

- its original representative $\mathbf{r}$ (either mean, median or medoid – depending on the algorithm at hand)
- its covariance matrix $\mathbf{C}$
- a set of $N_F + 1$ coefficients θ

The former two quantities will be used in order to build the ANFIS MFs according to (2.5), whereas the latter will be used in order to define the hyperplane which locally approximates the input-output mapping according to (2.6). Specifically, it can be evaluated using the Least Mean Squares (LMSs) estimator:

$$\theta_i = \left(\mathbf{X}_i^T \mathbf{X}_i\right)^{-1} \mathbf{X}_i^T Y_i \tag{4.6}$$

where $\mathbf{X}_i$ is the set of input patterns lying in the i-th cluster and Y_i is the set of corresponding output values (ground-truth). It is worth noticing that patterns in $\mathbf{X}_i$ will be augmented by appending a heading 1 such that their dimension is $N_F + 1$: in this manner $\theta_i \in \mathbb{R}^{N_F+1}$, as it also considers the hyperplane's intercept (see (2.6)). In order to fully consider both the input and output spaces, the (dis)similarity measure has been readapted, regardless of the specific adopted clustering algorithm. The pattern-to-cluster (dis)similarity measure is defined as a convex linear combination between the point-to-representative distance (input space) and the approximation error given by the interpolating hyperplane (output space):

$$\hat{d}(\mathbf{x}, \langle \mathbf{r}, \mathbf{C}, \theta \rangle) = \varepsilon \cdot d(\mathbf{x}, \mathbf{r}) + (1 - \varepsilon) \left(y - \theta^T \cdot \mathbf{x}\right)^2 \tag{4.7}$$

where $d(\cdot, \cdot)$ is one of the given (dis)similarity measures (either squared Euclidean, Manhattan or Mahalanobis), the triad $\langle \mathbf{r}, \mathbf{C}, \theta \rangle$, as introduced, defines the fuzzy rule and, finally, $\varepsilon \in [0, 1]$ is a trade-off parameter which tunes the linear convex combination. It is worth noting that the leftmost term, by considering patterns and centroids and by not considering output values or hyperplanes, can be seen as the distance in the *input space*. Conversely, the rightmost term considers output values and hyperplanes in lieu of centroids, therefore can be seen as the distance in the *hyperplanes space*. Joining the two terms, makes (4.7) a suitable *joint input-hyperplanes space* distance measure [69], [83] with ε in charge of weighting the input and hyperplanes spaces contributions.

It is worth noticing that if $\varepsilon = 1$, the rightmost term in (4.7) will not be considered, thus the formulation collapses into a standard clustering problem (see Sec. 4.1).

Conversely, by setting $\varepsilon = 0$ only the hyperplanes space is considered, thereby leading to a proper *hyperplane clustering*, and (4.7) becomes

$$\hat{d}(\langle \mathbf{x}, y \rangle, \theta) = (y - (\theta^T \mathbf{x} - \theta_0))^2 \tag{4.8}$$

By matching (4.7) and (4.8) it is clear that in hyperplane clustering centroids have no impact in the distance measure. Indeed, only the hyperplane coefficients vector θ can be considered as the proper cluster representative. In turn, hyperplanes are *affine subspaces* as they also include the fixed term, namely the intercept, as stressed by the term θ_0 in (4.7) and (4.8).

The rationale behind hyperplane clustering is that a first-order Sugeno FIS performs a piecewise linear approximation of the unknown function to be modelled. Thus, instead of searching for clusters in the input, output or joint input-hyperplane space [83], a much more effective way to determine the fuzzy rules is to gather input-output pairs that are "well approximated" by the same hyperplane. By following this approach, the cluster representative is just the coefficients vector θ, and the distance measure between a generic input-output pair $\langle \mathbf{x}, y \rangle$ with respect to a cluster is defined as the approximation error of the corresponding hyperplane on $\mathbf{x}$ (as in Eq. (4.8)).

However, it is important to remark the importance of the $\mathbf{r}$ and $\mathbf{C}$ evaluation also in hyperplane clustering as well as evaluating θ in case of input clustering. The triads $\langle \mathbf{r}, \mathbf{C}, \theta \rangle$ in output from the clustering process are necessary for the ANFIS modelling.

4.3 MinMax Classification Algorithm for Improving the ANFIS Synthesis

Regardless of the specific clustering algorithm adopted amongst the ones presented in the previous section, whether clusters can be described by means of their representatives μ, their covariance matrices $\mathbf{C}$ and their hyperplane coefficients θ, they can be used in order to synthesize ANFISs. Indeed, by recalling (2.5)–(2.6), it is clear that representatives and covariance matrices are useful for building the MFs (2.5), whereas the hyperplanes are useful for building the Takagi-Sugeno RCs (2.6).

However, it is possible that two (or more) sets of patterns well-separated in the input space are approximated by the same hyperplane. Such ambiguous scenarios make the ANFIS MFs definition harder as heavy overlaps between MFs may happen.

In order to overcome this problem, a classifier can be placed right after the clustering phase before generating the corresponding FIS as proposed in [69], in order to refine the hyperplane clustering solution. In particular, a (Pruning) Adaptive Resolution Min-Max Classifier (PARC) [84] is proposed as classification system in order to avoid large MF superpositions.

The PARC classifier aims at covering the training set patterns with a set of hyperboxes where each hyperbox is marked with a given label. If all patterns belonging to a given hyperbox share the same label (say, l), then the hyperbox is 'pure' since it will be univocally associated with label l, as remarked in [69]. Otherwise, if a given hyperbox contains patterns belonging to different labels, then such hyperbox is said 'hybrid'. The PARC training phase considers two lists, L_H and L_P, containing hybrid and pure hyperboxes, respectively. At the beginning of the training procedure, all patterns belong to a unique hybrid hyperbox. Recursively, the hybrid hyperbox covering the highest number of patterns is selected from L_H and cut with a suitable hyperplane in order to try to generate pure hyperboxes. Whether new pure hyperboxes are generated, they are added to L_P and formerly hybrid hyperboxes are removed from L_H. Pure hyperboxes belonging to the same class are (if possible)

merged together, leading to the minimum size pure hyperbox containing the pure hyperboxes selected for merging. This helps in reducing the number of hyperboxes, thus the model complexity. This process is repeated until L_H is empty. It is clear that such training phase might lead to a huge number of hyperboxes, especially for not-easily separable classes. According to basic learning theory[2] [85], during its training phase, the PARC classifier aims at minimizing the following cost function:

$$F = (1-\lambda)\mathscr{E} + \lambda\mathscr{C} \tag{4.9}$$

where $\mathscr{E}$ takes into account the model performances on the training set (percentage of misclassified patterns) and $\mathscr{C}$ takes into account the model complexity (percentage of hyperboxes with respect to the training set size). The parameter $\lambda \in [0,1]$ tunes the linear convex combination between the two factors: as $\lambda \to 1$, the PARC model will (ideally) feature a low structural complexity with low performances on the training set. Conversely, as $\lambda \to 0$, the PARC model will perform well on the training set, but will also feature a high number of hyperboxes (high structural complexity, with risk of overfitting).

Since the PARC classifier is in charge of avoiding large MF superpositions possibly returned by the clustering algorithm, the dataset definition must be revisited. Nonetheless, since classification is a supervised problem by definition, the dataset will again consist of labelled patterns. Specifically, each input pattern $\mathbf{x}$ will be labelled with an integer in range $[1,k]$, namely the cluster (hyperplane) to which $\mathbf{x}$ has been assigned to by the clustering algorithm. As described above, PARC will design proper non-overlapping hyperboxes which, at the end of the training phase, will be 'pure'. Hyperboxes, thus, lead to a more accurate partitioning, the union of which can be seen as an 'improved' clustering solution. Since hyperboxes can be treated as clusters, it is therefore possible to evaluate their respective representatives, covariance matrices and the hyperplane coefficients vectors. These 'improved' clusters or,

[2] The widely-known Ockham's Razor Criterion

specifically, their representatives, covariance matrices and hyperplane coefficients, will finally be used for synthesizing the ANFIS network (see (2.5)–(2.6)).

The benefits of applying the PARC classifier after the hyperplane clustering phase are depicted in Figure 4.1 and Figure 4.2 through the description of two toy-examples. In both the examples the (hyperplane) clustering algorithm is set with a number of cluster to detect equal to 3, but only in Figure 4.2 is applied the PARC classifier.

Figure 4.1 depicts the undesired scenario described above where the upper part of the figure sketches the hyperplane clustering solution, whereas the lower part sketches the resulting MFs. In this example, a single blue hyperplane approximates two clusters well separated in the input space, with a third cluster between them approximated by the orange hyperplane. A fourth compact cluster in the input space is approximated by the green hyperplane. This means that the regions pertaining the two MFs (the orange and blue Gaussian curves) are mostly overlapped, leading to ambiguity in the selection of the winning rule for a wide interval of the input axis.

Conversely, Figure 4.2 shows the results after the PARC execution. PARC has been able to design hyperboxes around these four clusters: albeit the first and third cluster shared the same label (cf. the blue hyperplane from Figure 4.1), PARC detected the second cluster as an in-between and in order to avoid overlappings, the first and third clusters (from left) have been assigned to different hyperboxes (though sharing the same label – *i.e.* the original, blue, hyperplane). The four resulting hyperboxes have been treated as new clusters, thus their respective centroids, covariance matrices and hyperplanes have been re-evaluated from scratch. As shown in the upper part of Figure 4.2, each cluster now features its own hyperplane with no more ambiguities. Absence of hyperplanes ambiguities, in turn, reflects in a non-ambiguous ANFIS MFs, as shown in the lower part of Figure 4.2.

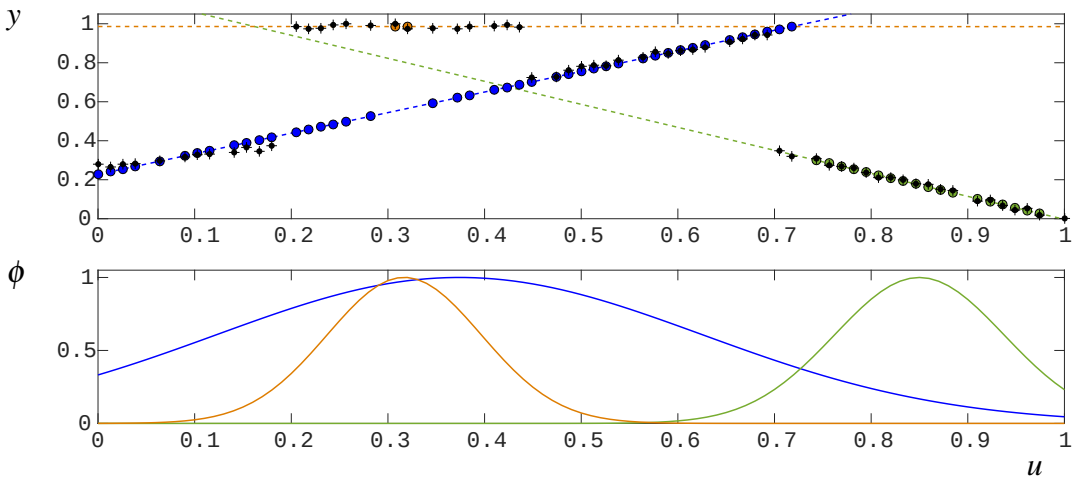


FIGURE 4.1: SISO ANFIS synthesis function approximation by k-means clustering without Min-Max PARC classification set to the detection of 3 clusters. Above, the training input-output samples as crossed black dots and in dashed lines the generated RC hyperplanes. The ANFIS output is represented by coloured dots, the same colour of the RC on which they lay on. Depicted below, the MFs with the same colour of their corresponding RCs.

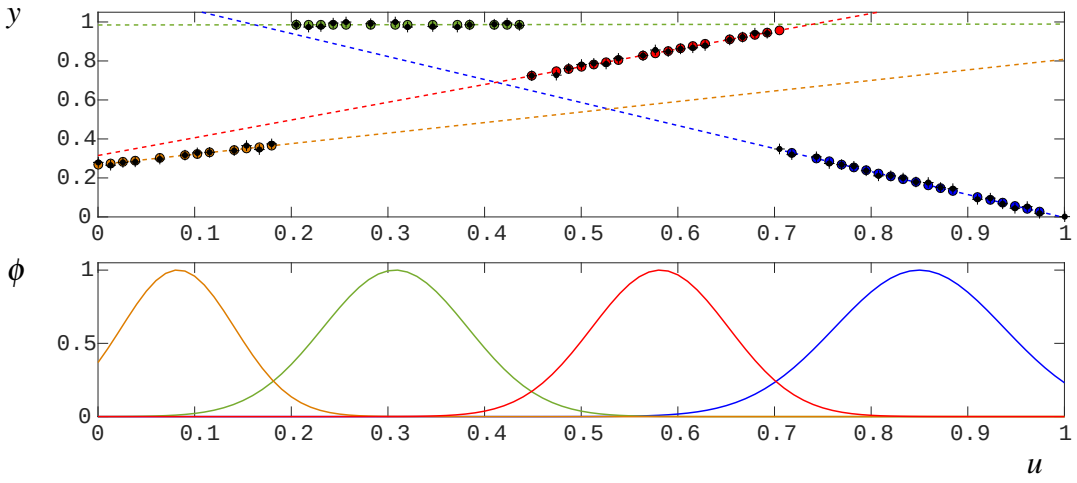


FIGURE 4.2: SISO ANFIS synthesis function approximation by k-means clustering supported by Min-Max PARC classification.

4.4 Hierarchical Clustering

Partitional clustering defined in Sec. 4.1 allows to calculate one solution considering a given exact number of k-clusters. It is clear that in order to detect the right number of clusters the first suggestion is to vary the k-clustering in a given range and therefore launch the partitional clustering a number of times equal to the range extension.

However, looking at the computational burden, this procedure can be critical especially in case of huge datasets and when are set huge numbers of clusters to detect. For this reason, in this section are proposed two hierarchical clustering procedures based on partitional k-clustering. These are able to highly reduce the computational burden and preserve the clustering effectiveness. In particular, starting from a given solution found with partitional clustering, these increment (divisive clustering) or decrease by one (agglomerative clustering) the number of clusters by respectively cutting or merging the cluster or those two clusters appropriately selected. The operation of cutting or merging is repeated until the whole clustering range is explored.

As in partitional clustering, the hierarchical algorithms can be re-adapted to the k-means variants (*i.e.* k-medoids and k-medians) and by considering the joint input-hyperplane space and pure hyperplane space of clustering as well.

4.4.1 Agglomerative

The agglomerative k-means-based hierarchical clustering procedure builds a binary tree in a bottom-up approach by starting from an initial set of $k_{\max}$ leaves (clusters) computed in an initial step by the partitional variant as described in Sec. 4.1. The merging procedure, used to generate higher level clusters in the hierarchy, operates as follows:

1. considering all leaf clusters at the k-th step, evaluate:

   - the centroid, defined as usual as

   $$\mu_i = \frac{1}{|\mathscr{S}_i|} \sum_{j=1}^{|\mathscr{S}_i|} \mathbf{x}_j, \text{ for } i = 1,\ldots,k \tag{4.10}$$

   - the within-cluster sum of (4.8)

   $$e_i = \sum_{j=1}^{|\mathscr{S}_i|} \hat{d}(\langle \mathbf{x}_j, y_j \rangle, \theta_i), \text{ for } i = 1,\ldots,k \tag{4.11}$$

namely the cluster dispersion, computed as the within-cluster error due to the i-th hyperplane approximation

- the pairwise squared Euclidean distance matrix between centroids

$$\mathbf{D}_{i,j} = \|\mu_i - \mu_j\|_2^2, \text{ for } i,j = 1,\ldots,k \tag{4.12}$$

- the pairwise error matrix by considering the approximation error obtained if any two clusters had to be merged

$$\mathbf{E}_{i,j} = d(\langle \widetilde{\mathbf{X}}, \widetilde{\mathbf{y}} \rangle, \widetilde{\theta}), \text{ for } i,j = 1,\ldots,k \tag{4.13}$$

where

$$\begin{cases} \widetilde{\mathbf{X}} = \begin{bmatrix} \mathbf{X}_i \\ \\ \mathbf{X}_j \end{bmatrix} \quad \widetilde{\mathbf{y}} = \begin{bmatrix} \mathbf{y}_i \\ \\ \mathbf{y}_j \end{bmatrix} \\ \widetilde{\theta} = (\widetilde{\mathbf{X}}^T \widetilde{\mathbf{X}})^{-1} \widetilde{\mathbf{X}}^T \widetilde{\mathbf{y}} \end{cases} \tag{4.14}$$

- the overall error matrix, function of both the inter-clusters and the intra-clusters approximation errors

$$\bar{\mathbf{E}}_{i,j} = \mathbf{E}_{i,j} - (e_i + e_j)/2 \tag{4.15}$$

- the overall score matrix, which considers both the inter-clusters separation (*i.e.* distance between centroids in the input space) and the overall approximation errors

$$\mathbf{S} = (\bar{\mathbf{E}} + \mathbf{D})/2 \tag{4.16}$$

2. select the i, j pair corresponding to the maximum value in $\mathbf{S}$ as the two clusters to fuse

3. merge the two clusters, as in (4.14), then evaluate the new centroid, covariance matrix and hyperplane

4. Loop 1–3 until a single root node (*i.e.* the entire dataset) is obtained

A typical merging procedure is sketched in Figure 4.3 for $k_{\max} = 3$ clusters. By supposing that clusters 1 and 2 (in red) are selected for merging (that is, $\mathbf{S}_{1,2} > \mathbf{S}_{1,3}$ and $\mathbf{S}_{1,2} > \mathbf{S}_{2,3}$), an additional parent node (in green) is built by merging these clusters. The merging procedure creates at each iteration a new cluster in the hierarchy, and thus a layer on its dendrogram representation, herein defined as the set of leaf nodes (clusters). As instance, the layer considered at the second iteration ($k = 2$) consists in the green and blue clusters. Further merging these two clusters returns a single root node and the algorithm halts.

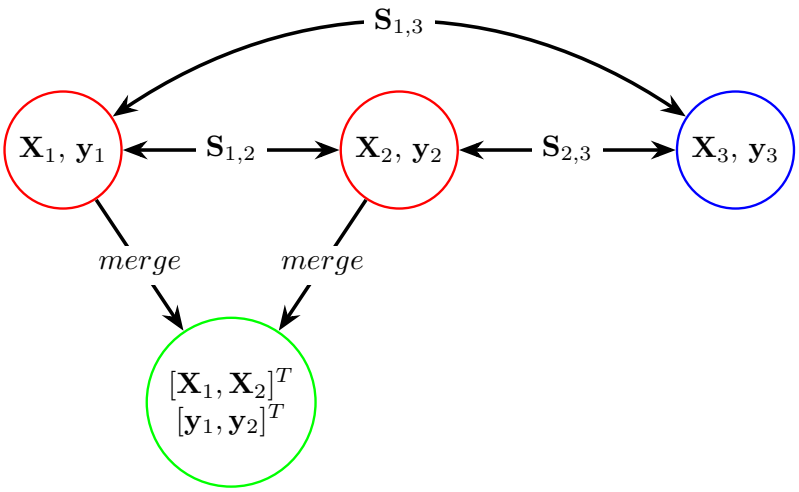


FIGURE 4.3: Hierarchical k-means agglomerative merge procedure.

4.4.2 Divisive

The divisive k-means variant works in a dual fashion with respect to the agglomerative counterpart. Indeed, rather than by starting with $k_{\max}$ leaf nodes, it starts by considering a single root node corresponding to the entire dataset. Successively, it creates two new leaf nodes (clusters) by running a 2-means on the former. Subsequent leaf nodes are created as follows:

1. considering all clusters within the k-th layer (thus, the set of leaf nodes after k splitting), evaluate their respective within-cluster error e (cf. Step #1 of the agglomerative variant, specifically (4.11))

2. select the cluster with maximum local approximation error as the one to split

3. run a 2-means in order to create two new clusters. Then evaluate their respective hyperplanes and centroids

4. Loop 1–3 until a maximum depth $k_{\max}$ is reached

A typical splitting procedure is sketched in Figure 4.4. A 2-means is run on the whole dataset (in black), leading to two leaf nodes (blue and green). By supposing that $e_1 > e_2$, the blue cluster acts as the node to split and a 2-means is performed, leading to further two leaf nodes (in red). In Figure 4.4, a layer for $k = 2$ can be extracted by considering the blue and green clusters, whereas a layer for $k = 3$ can be extracted by considering the green and the two red clusters. Further, by supposing that $k_{\max} = 3$, the algorithm halts.

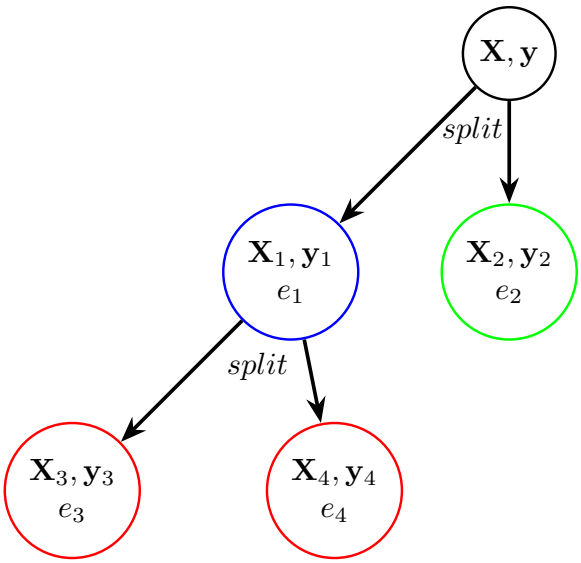


FIGURE 4.4: Hierarchical k-means divisive split procedure.

Chapter 5

Prosumer Energy Balance Prediction

This chapter is focused on the prediction of a prosumer energy balance profile by considering a model based on a particular type of recurrent NNs named Echo State Networks (ESNs).

Recurrent Neural Networks (RNNs) are demonstrated in literature very effective models for both the prediction of the energy generation from RESs and the energy demand as reported in [86] and [87], respectively.

In particular, ESNs, which are single layer RNNs, are widely adopted in forecasting problems because of their high simplicity of implementation and low computational burden compared to multi layered NNs. Indeed, ESNs do not need to apply error back-propagation procedure for the training. However, while being practical and conceptually simple, ESNs require some experience and insight to achieve the hailed good performance [88].

This chapter studies a paricular ESNs-based prediction model with the aim to improve the efficacy of a fuzzy-based decision making system that is supposed to be implemented on an EMS of a prosumer grid-connected MG.

5.1 ESN Based Prediction Model

Considering a generic RNN with a supervised training, it is optimized on a training input-output sampling time series data of the form

$$D = \{u(1), y^*(1); u(2), y^*(2); ...; u(M), y^*(M)\} \tag{5.1}$$

The time series data is defined by the input vector $u \in \mathbb{R}^K$ and the target (ground truth) output $y^* \in \mathbb{R}^L$. Therefore, the prediction task is to train a RNN starting from an input-output time series in order to reproduce y^* minimizing the following error function

$$E(y, y^*) = \frac{1}{L} \sum_{i=1}^{L} \sqrt{\frac{\sum_{j=1}^{N} (y_i(j) - y_i^*(j))^2}{M}} \tag{5.2}$$

where y is the RNN signal set of the time series.

For the sake of clarity, from here to the end of the chapter the input-output series will be referred to as the Training Set (TR).

The ESN architecture is composed by three layers: the input layer which state is denoted by u; the reservoir denoted by x; and the output layer y. To each layer corresponds a matrix where each entry corresponds to the connections weight between a pair of neurons. The weights matrix $\mathbf{W^{in}} \in \mathbb{R}^{\mathbf{K} \times \mathbf{N}}$ defines the connections between the input layer of dimension K and the reservoir that is characterized by N neurons. The square matrix $\mathbf{W} \in \mathbb{R}^{\mathbf{N} \times \mathbf{N}}$ defines the connections among the reservoir neurons themselves. $\mathbf{W^{out}} \in \mathbb{R}^{\mathbf{N+K} \times \mathbf{L}}$ defines the connections between the input layer concatenated with the reservoir (*i.e.* $[u(n); x(n)]$) and the output layer $y(n)$, where the term n is referred to a generic time step.

Therefore, being $u(n)$ the signal input, the state update equations are formulated as following:

$$\tilde{x}(n+1) = \tanh(\mathbf{W}x(n) + \mathbf{W^{in}}u(n+1)) \tag{5.3}$$

$$x(j+1) = \alpha\tilde{x}(n+1) + (1-\alpha)x(n) \tag{5.4}$$

where $x(n)$, of size K, is the reservoir state, tanh is the adopted reservoir activation

function, and α is the leak rate factor. It is clear that the updated reservoir state does not depend by its previous state if $\alpha = 1$, therefore $x(n+1) = \tilde{x}(n+1)$.

The ESN output is calculated by the readout equation of the form

$$y(n) = g(\mathbf{W^{out}}[x(n); u(n)]) \tag{5.5}$$

where g is the identity activation function.

As exhaustively expounded in [89], both $\mathbf{W}$ and $\mathbf{W^{in}}$ are randomly generated by considering few parameters that must be set empirically or possibly by adopting some heuristics. Once $\mathbf{W}$ and $\mathbf{W^{in}}$ are set, $\mathbf{W^{out}}$ is determined by minimizing (5.2) through a Least Mean Square Error Linear Regression (LMSELR).

As observed in [90] and [88], the parameters to define must be constrained in a given range with the aim to comply with the Echo State property. The parameters here considered for the training of the ESN are the following:

- The reservoir dimension defined by N.
- The reservoir spectral radius λ, as a limit to the reservoir connection weights.
- The reservoir connectivity ρ which defines the matrix sparsity.
- The leak rate factor α of (5.4).
- The input scaling and input shifting vectors represented by $\bar{\psi}$ which is a $2 \times K$-dimensional vector that defines how to scale and shift each input.

In order to efficiently treat the input scaling and input shifting vectors and evaluate the ESN performance, the ESN input vector is always considered subjected to a normalization $[0,1]$ as the ESN output has to be de-normalized. However for the sake of convenience, operations of normalization and de-normalization, though are performed, these are not highlighted in the ESN model.

5.2 Prediction Algorithm Synthesis

As shown in Figure 5.1 the ESN architecture is designed as a MISO RNN (*i.e.* $L = 1$) which is designed to predict the energy produced (or demanded) by a given energy system in a specific time frame. Therefore, considering a prediction algorithm featured by a time horizon T of N_T time slots, the number of ESNs that needs to be tuned is equal to the number of time slots N_T multiplied by the number of energy systems to predict. It is clear that, by considering this kind of architecture, each ESN works independently without interacting with each other.

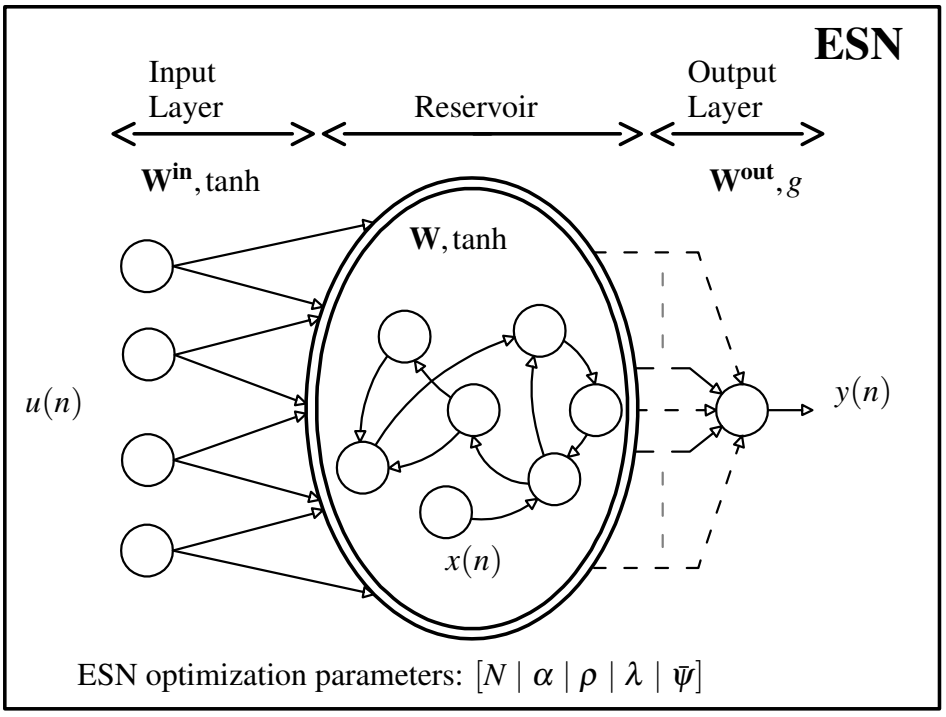


FIGURE 5.1: ESN scheme. The solid arrows indicate fixed and randomly generated connections whereas dotted arrows trainable connections. The circles represent the neurons which hold a specific activation function.

Although the model is ineffective from a computational point of view [1], it permits to individually tune each ESN without changing (freeze) those already tuned, thus maintaining their performances. For the sake of clarity, in case it is increased the

[1] as introduced in Sec. 5.1 ESNs and NNs in general can be modelled as MIMO systems. Hence, a first variant can foresee an overall number of ESNs reduced to R (*i.e.* the number of object to predict). However, this is not the only strategy that can be pursed. Indeed, the overall prediction system can be reduced to a single MIMO NN (or ESN) as well. Nevertheless, the more the system is going to be simplified, the more useful information are going to be lost.

prediction time horizon (N_T) by 1 time frame, it is necessary to model a number of new ESNs equal to the number of energy systems R taken into account. This strategy has been mainly adopted with the aim to identify the most suitable prediction time horizon extension for supporting the fuzzy-based decision making system of the MG EMS.

In Figure 5.2 a generalized ESNs based prediction model is shown. There are considered R energy systems to predict. To each energy system is associated an array of N_T ESNs, *i.e.* equal to the number of ESNs time frames which cover the prediction time horizon. Given the i-th energy system, their corresponding ESNs read the given $u^i(n)$ input array and output the value $y^i(n)$ at the time-step n. Therefore, the sum of the overall outputs per each energy system ($\sum y^i_j(n)_i^R \; \forall \; j = 1,..,N_T$) returns the future (prosumer) predicted energy balance profile $\{y(n)_1, y(n)_2, ..., y(n)_{N_T}\}$ of length N_T. In this study the number of energy systems subject to the prediction model R are 2, namely a system of aggregated loads and a PV system.

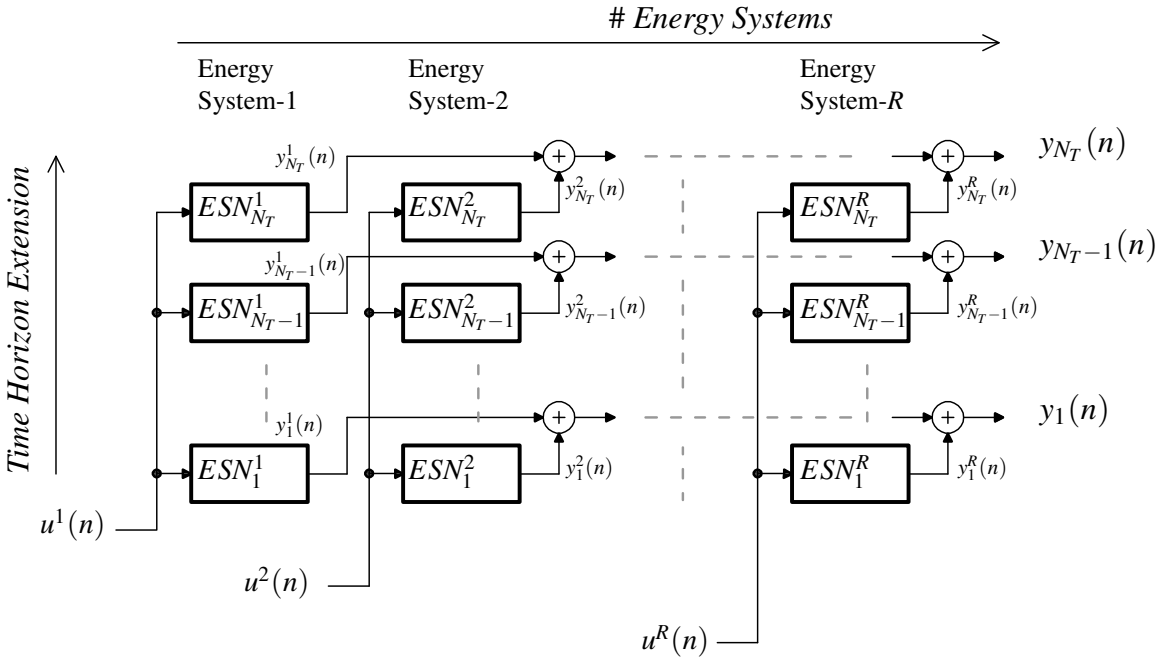


FIGURE 5.2: A generalized ESN-based prediction model scheme where is considered a prediction time horizon of N_T time frames and R energy systems.

Going into details of the single ESN architecture, as illustrated in Figure 5.2 it is featured by an input array of 4 elements.

Considering the prediction on the PV system at the time slot n, the ESNs input array $u^G(n)$ is defined by the last lecture of the energy generation $E^G(n)$ given by its energy meter, the current day $d(n)$, the current hour $h(n)$ and a rough daily estimation of the weather forecast defined by $F^G(d(n))$ (*e.g.* sunny day, cloudy day etc.), namely

$$u^G(n) = \left\{E^G(n), d(n), h(n), F^G(d(n))\right\} \tag{5.6}$$

Considering the prediction on the energy demand, the input array is

$$u^L(n) = \left\{E^L(n), d(n), h(n), F^L(d(n))\right\} \tag{5.7}$$

where $E^L(n)$ is supposed to be the last lecture of the energy demand given by the load meter and $F^L(d(n))$ the rough daily forecast about the prosumer activity.

The choice of the input $F^G(n)$ is motivated by the fact that, as stated in [91], the application of a prediction algorithm based on NNs for PV generation needs information about the weather forecast so as to work efficiently since it is strictly related to the PV production. This information has been summarized in 3 states: high, medium and low production with respect to the maximum daily PV production. About the load, it is assumed that the prosumer can notify the EMS about its presence or absence during the day. Therefore, the daily energy demand estimation $F^L(n)$ is summarized in 2 states: presence or not of the prosumer activity.

Every ESN is modelled separately by implementing a GA on a given TR-Validation Set (VL) pair by means of the tuning of the parameters introduced in Sec. 5.1, namely λ, N, ρ, α and $\bar{\psi}$. Therefore, each individual chromosome contains the information about the ESN parameters which define the ESN architecture. Once the ESN individual is generated, its optimization proceed on the TR whereas the fitness defined in (5.2) is evaluated on the VL. Further details about the GA and ESN procedure of synthesis can be found in [75].

Chapter 6

MG Energy Systems Flexibility

In this chapter a novel approach based on graph theory for MG energy flow analysis and management is proposed. The study is focused on the efficient representation of how much energy demand and energy production of the energy systems connected to the MG should be exchanged in real time compared to what can be reshaped, shifted or more in general controlled. In particular, a novel formulation focused on the representation of the MG elements degree of flexibility is proposed in order to support the design of the MG EMS decision making system formulation.

Since this chapter is dedicated to the MG energy flows representation and management, a properly level of abstraction is considered. In particular, the energy production and energy demand are discretized in time frames equal to the sampling time of the MG smart meter. In each time frame, each power profile is considered constant. Power losses and reactive powers are neglected as system of controls and power converters are ideally formulated.

6.1 Controllable and not Controllable Prosumer Energy Flows

In the formulation proposed the prosumer energy systems together with the distribution grid are treated as *MG elements* since they are supposed to be connected to the MG backbone by means of their respective power converters.

Therefore, considering a generic grid-connected MG, it can be summarized by the following elements:

- Local loads: these are unidirectional energy systems; in particular, some loads can be scheduled, shifted or reshaped, while others must be fulfilled without tolerating any late, *i.e.* in real time.

- Local DG units: DG units can be wind turbines and PV systems or small-sized controlled generators like diesel units or microturbines.

- PEV charge stations: there are mainly two type of charge stations, fast and slow. In slow charge the power requested from PEVs can be reshaped within a certain time period. In some cases, slow charge can involve the participation to V2G as well. Conversely, the FC foresees the real time fulfillment of the PEV energy demand.

- ESSs: these are in charge to improve the MG flexibility and autonomy.

The MG system can be summarized as the set of four aggregated elements

$$MG =: \{PEVs, ESS, DGs, Loads\} \tag{6.1}$$

They are mostly characterized by unidirectional power flow orientation with the exception of the ESSs and, in case of V2G service, the connected PEVs. In Figure 6.1-a) a graphical representation is described. There, the connected grid has been represented as the 5^{th} additional element. It has the role of balancing the whole MG power requirements.

In the proposed formulation it is assumed to know in advance for each element how much energy must be fulfilled in the next time slot with respect to how much can be controlled by considering the ESS and the management of special loads.

As show in Figure 6.1-b), at the k-th time slot for each i-th aggregated MG element, it has been distinguished how much energy must be satisfied in the same time

slot ($E^{NC}_{j,k}$) with respect to how much is exchanged ($E^{C}_{j,k}$) by means of the application of DR operations (*i.e.* load shifting, load reshaping, etc.). Moreover, their summation ($E^{NC}_{k} + E^{C}_{k}$) defines the energy exchanged with the main grid (E^{N}_{k} in red).

$$E^{NC}_{k} = \sum E^{NC}_{k,i\ i} \tag{6.2}$$

$$E^{C}_{k} = \sum E^{C}_{k,i_i} \tag{6.3}$$

$$E^{N}_{k} = E^{C}_{k} + E^{NC}_{k} \tag{6.4}$$

To remark this concept in Figure 6.1-b) the two energy values $E^{NC}_{k,i}$, $E^{C}_{k,i}$ of each aggregated element are graphically represented by orange circles and green circles, respectively, where their amplitude is proportional to their corresponding value. The grid element is distinguished in red so as to highlight that it operates as an element of equilibrium.

The summation of the non-controllable terms and controllable terms are depicted in Figure 6.1-c) where the orange circle can be meant as a energy constraint whereas the green circle is the energy exchange by the ESS, the controlled generation, the flexible loads and the connected PEV on slow charge station by means of an intelligent management of the aggregated elements. Therefore, it represents the overall energy systems flexibility exploited at the given k-th time slot. The proposed formulation brings to the definition of a graph of three nodes which represent the status of the MG at the k-th time slot (see Figure 6.1-d)) where the vertices define how much energy is exchanged by the tree nodes whereas the edges point out their orientation and distribution.

This formulation points out that once E^{C}_{k} is determined the problem is solved in case of E^{NC}_{k} is supposed to be known. Therefore, this formulation also emphasizes that the EMS must be supported with an algorithm capable of intelligently distinguishes and evaluate the non controllable terms respect with how much energy can be effectively managed by the lecture of the MG elements and the application of

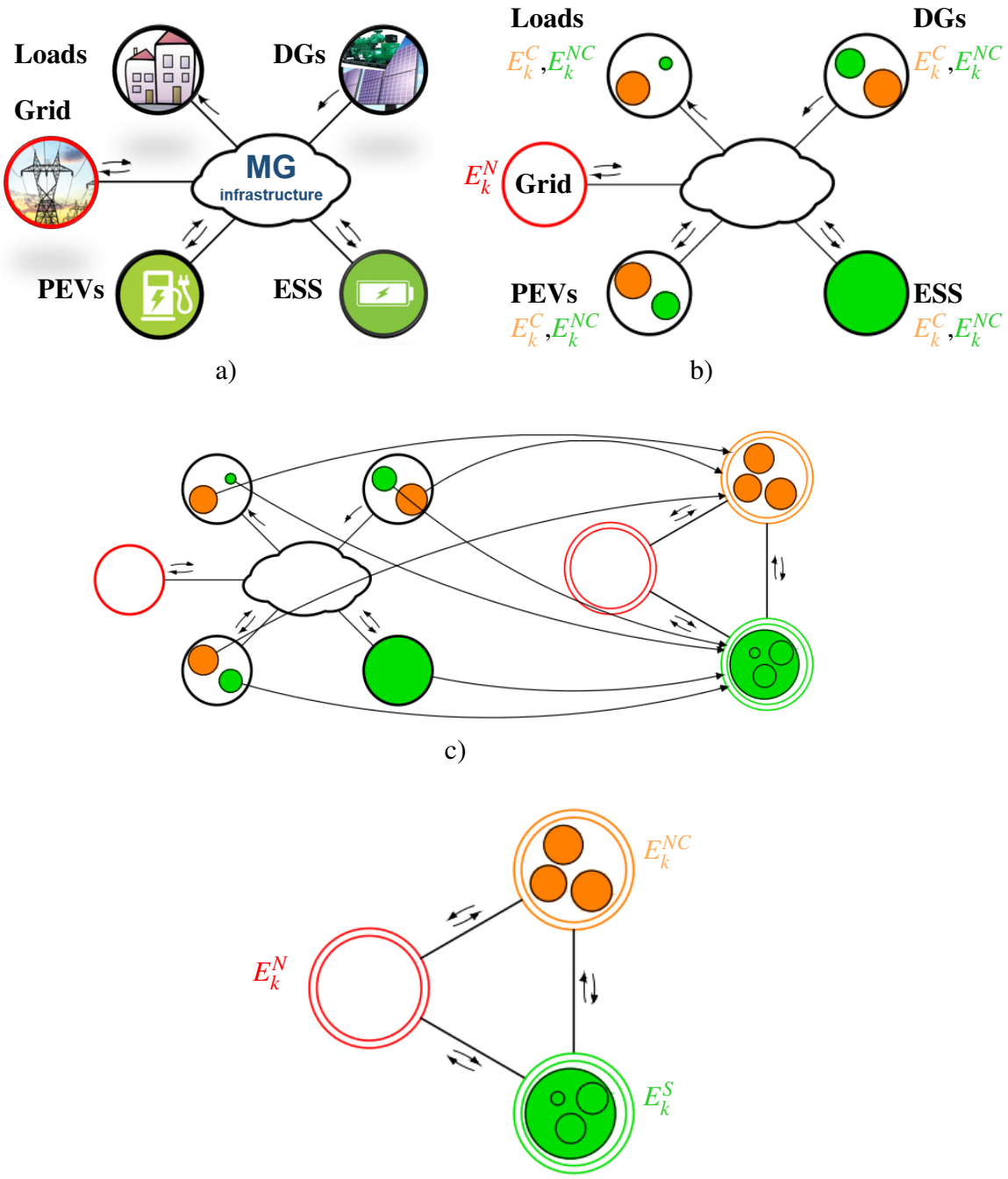


FIGURE 6.1: Representation of the MG energy flows and management, formulation of controllable and non controllable terms.
a) A MG energy System grouped per aggregated elements.
b) Graphical representation of MG elements, the non controllable terms in orange and the controllable terms in green.
c) Transition to a more generalized formulation.
d) The MG EMS must be able to define in real time the controllable term and split it among the MG energy systems.

given constraints defined by the prosumer. Therefore, the EMS should be able to detect the MG flexibility in order to properly evaluate E_k^C that, subsequently, has to be split and redistribute between each MG elements and sub-elements.

The proposed formulation suggests the modelling of a Dynamic-ESS featured by a dynamic capacity which indicates what is the MG energetic flexibility, namely the current range domain of E_k^C in that given time slot. However, it is clear that in case of the absence of flexible loads and controllable DG units the prosumer energy balance coincides with E^{NC} whereas E^C is equal to the energy exchanged by the only ESS and the range domain of E^C is function of the current ESS SoC and capacity.

Chapter 7

Microgrid Reference System for the EMS Modelling

7.1 Problem Simplifications

The studies introduced in next Chapters deal with the synthesis of a decision making system applied to a DSM MG EMS. In particular, it has been considered a prosumer equipped with an ESS unit and a PV generator. The EMS is in charge to define how to split in real time the MG energy flows among the customer, the PV generation, the ESS and the grid by considering one or more OFs. Here below are introduced the main problem assumptions and simplifications in order to define a useful level of abstraction to correctly place the problem under analysis:

- The EMS time frame is equal to its MG smart meter. According with the literature and smart metering in general, it has been chosen a time frame of 15 minutes.
- Together with the EMS time frame, also the MG energy systems have been sampled and discretized in a 15 minutes time frame.
- Along each time frame, the prosumer power production and power demand have been considered constant.

- The demanded energy must be actually consumed and the main grid has an unlimited energy capacity, while low level operations such as voltage and reactive power control are not considered.
- The power transmission losses within the MG are considered negligible. The online control module ensures that the power balance is achieved during the real-time operation.
- The EMS has a sample time considerably greater than the characteristic time of the ESS power control, therefore the ESS inner loop time characteristic has been neglected.
- The ESS is equipped with a BMS capable of perfectly control the battery SOC.
- Low level operations such as voltage and reactive power control are not considered.

In case of some of the listed problem simplifications are not respected in specific studies these are properly denoted and motivated. It is clear that, taking into consideration the formulation proposed in Sec. 6, the sum of the current energy production and the current energy demand corresponds to E^{NC} since they cannot be shifted or reshaped and the energy exchange by the ESS is E^{C}.

7.2 MG Architecture

In Figure 7.1 the MG architecture scheme is represented. It remarks the presence of power converters control systems and the ICT infrastructure that enable to monitor the MG state and the external signals which can be referred to the energy prices, the weather forecasts etc. The ICT infrastructure communicate to the power converters how to respect the energy balance between the MG power systems and the grid by means of the EMS. The EMS input array will be specified case by case since it changes in function of the problem under analysis.

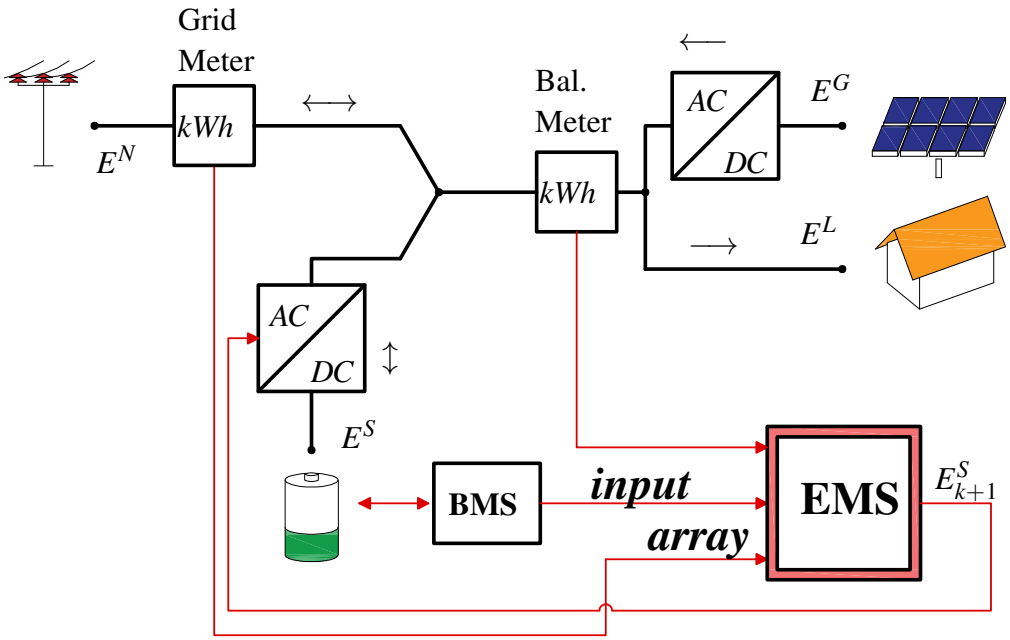


FIGURE 7.1: MG architecture. Signal wires in red, power lines in black.

Here followings a proper problem formulation is illustrated.

Given the prosumer energy generation and energy demand at the k-th time slot, E_k^G (≥ 0) and E_k^L (≤ 0), respectively, it is possible to compute the prosumer energy balance as

$$E_k^{GL} = E_k^G + E_k^L \qquad n = 1,2,... \tag{7.1}$$

where $E_k^{GL} > 0$ in case of over-production and $E_k^{GL} < 0$ in case of over-demand. The respect of the energy conservation law implies that the sum of the energies exchanged among the MG energy systems and the grid is equal to 0, hence

$$E_k^{GL} = -E_k^N - E_k^S \qquad n = 1,2,... \tag{7.2}$$

where E_k^S is the energy exchanged with the ESS ($E_k^S > 0$ in case of discharge) and E_k^N is the energy exchanged with the grid ($E_k^N > 0$ when the grid sells energy to the prosumer).

The synthesis of the MG EMS can be focused on the minimization of different OFs such as the operational costs, the reduction of the stress generated to the grid

(*i.e.*, peak shaving and power fluctuation minimization), the achievement of some energy incentives and the maximization of the incomes generated by the energy trade with the grid. For example, in case it is considered a ToU price policy, it is possible to express the profit P generated by the energy trade with the main grid in a given time period composed by N_S time slots as

$$P = \sum_{n=1}^{N_S} P_k \quad \text{where} \quad P_k = \begin{cases} E_k^N \cdot C_k^{buy} & \text{if } E_k^N > 0 \\ E_k^N \cdot C_k^{sell} & \text{if } E_k^N \leq 0 \end{cases} \tag{7.3}$$

where C_k^{sell} and C_k^{buy} are the corresponding energy costs (at time slot n) in sale and purchase, respectively, expressed in Monetary Unit per kWh $[M.U./kWh]$.

In most of the studies proposed [1] it has been excluded the direct exchange of energy between the ESS and the grid in order to prevent that the EMS exploits the ESS for buying and selling energy so as to produce high energy fluctuations. This assumption is respected by considering the following constraint

$$|E_k^{GL}| = |E_k^N| + |E_k^S| \qquad n = 1,2,... \tag{7.4}$$

The EMS decision making systems here investigated are MISO models featured by a low number of inputs (see Figure 7.1) which allows to design an efficient system of easy implementation and low complexity. The main difference between each case of study is about the simulation scenario under analysiss, the particular EMS decision making system architecture, the EMS training procedure and the OF(s) formulation.

[1] Studies that do not consider these constraints have been appropriately commented

7.3 Prosumer Dataset and ToU Energy Prices

The prosumer dataset referred to the PV system and the aggregated loads power demand shown in Figure 7.2 have been provided by *AReti S.p.A*, the electric distribution utility that manages the power grid in the area of Rome. The aggregated loads and PV time series are related to measurements made in the course of one year with a sampling rate of 15 minutes. Depending on the problem formulation and the EMS

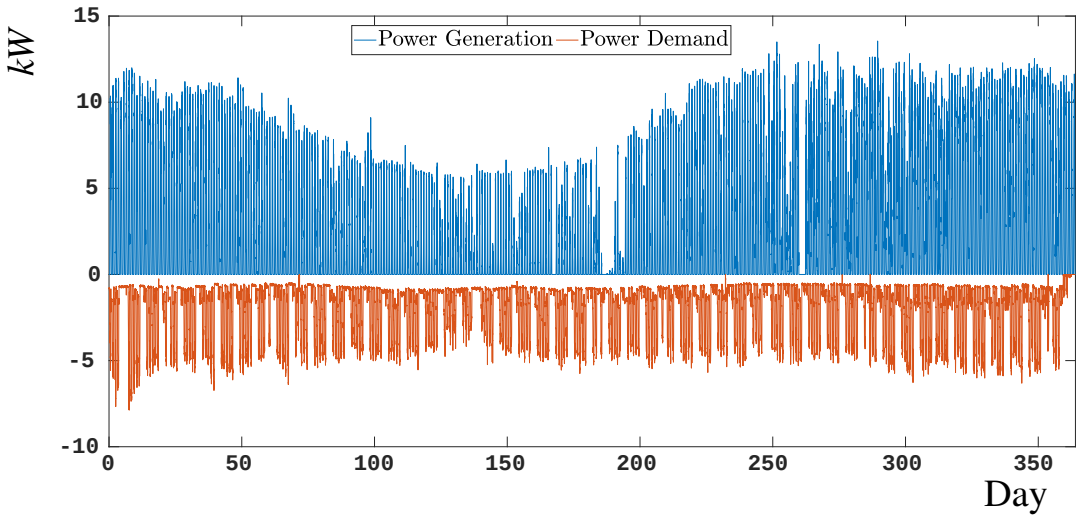


FIGURE 7.2: Prosumer reference dataset.

synthesis procedure adopted in the following cases of study, the overall dataset, or only one interval of it, has been split in subsets in order to build some TR-VL pairs and the test set TS, or rather, a unique TR and TS. The reason for selecting a limited data range over the whole year is due to the PV generation trend that changes with the seasons.

In Figure 7.3 instead are shown the daily C^{buy} C^{sell} energy price series of the adopted ToU energy price policy [92].

7.4 EMS Modelling Benchmark Solution Formulation

As introduced in Sec. 1.3 of Chapter 1, with the increasing of interest in DG and (local) energy flow optimization, deterministic optimization tools based on LP, MILP

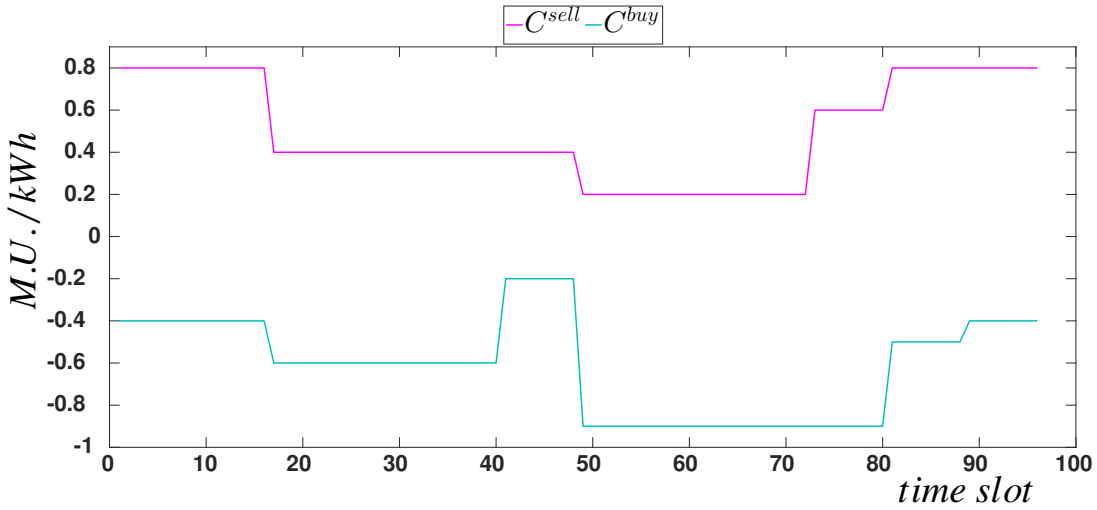


FIGURE 7.3: Daily ToU energy price series.

and DP started to be widely adopted in research for MG and local energy scheduling. Indeed, considering a fair (piece-wise) linearization and simplification of the problem under analysis, these are able to find the optimal MG energy flow scheduling over a given time horizon in order to maximize the incomes, preserve the stability of the grid and guarantee the quality of service. In particular, LP based algorithms are applied when the problem is subjected to linear equations, constraints and OF. MILP algorithms instead, are able to include integers variables as well. DP algorithms can solve more complex problems through a fair discretization of the problem under analysis at the expenses of higher computational costs. Furthermore, its application requires an early investigation so as to accurately set the degree of discretization (granularity) of the problem under analysis in order to not overload the machine and not end up in local minimums.

In our case, the application of deterministic optimization tools is used to properly normalize the EMS OF in $[0;1]$ and to generate a labelled pattern set with the purpose to switch to a supervised problem. In particular, the labelled pattern set is realized through the evaluation of the (optimal) profiles referred to the energy exchanged with the grid $E^{N,opt}$ and the ESS SoC SoC^{opt}. Furthermore, for the sake of completeness, a lower benchmark solution is formulated in order to define the corresponding lower OF limit rather than normalize it by considering the only upper benchmark solution.

It is formulated by considering the MG paradigm defined in Sec. 7.2 without the support of the ESS. Such assumption implies that the current energy balance corresponds to the energy exchanged with the grid (*i.e.* considering (7.2), $E_k^{GL} = -E_k^N$ and $E_k^S = 0 \; \forall k$). Therefore, taking into consideration the minimization of a generic OF F and the corresponding upper and lower benchmark solution performance values F^{opt} and $F^{w/o}$, respectively, the normalized OF value $\bar{F}$ can be formulated as follows

$$\bar{F} = -\frac{F - F^{opt}}{F^{w/o} - F^{opt}} \tag{7.5}$$

In case it is considered the maximization of the profit given by (7.6), it can be rewritten as

$$\bar{P} = -\frac{P - P^{opt}}{P^{w/o} - P^{opt}} \tag{7.6}$$

where P, P^{opt} and $P^{w/o}$ are the profits obtained by the synthesised fuzzy-based decision making system, the profit generated by the upper benchmark solution (*i.e.* by the optimization tool) and the lower benchmark solution, respectively. It is noteworthy that while the profit P in (7.3) should be maximized, the normalized performance $\bar{P}$ in (7.6) should be minimized. Indeed, the more $\bar{P}$ is close to 0 the more the EMS manages the MG energy flows efficiently, reaching a profit close to the optimal one. Conversely, when $\bar{P}$ is close to 1, the ESS management can be considered inefficient.

It is important to remark that, considering the MG energy management problem described in Sec. 7.2, the OF formulation should looks at more aspects of the MG. For instance, although the main objective can be the profit function (7.3), a fair utilization of the ESS should be taken into consideration by avoiding that its SoC stops for long periods to its upper and lower limits. For this reason, the OF formulation should be a suitable combination of more (normalized) cost functions $f^i \in [0;1]$ of the form

$$\bar{F} = \sum \lambda^i \bar{f}^i, \sum \lambda^i = 1 \tag{7.7}$$

where λ^i are those meta-parameters which establish the level of importance of each corresponding cost function f^i.

7.4.1 A LP Application Example

This section describes a LP formulation applied to the evaluation of the optimal MG energy flow profiles in order to maximize the profit function defined in (7.3). The problem under analysis assumes that the energy balance must be split between the ESS and the grid. Hence, it must be imposed that the ESS cannot directly exchange energy with the grid in order to avoid too high E^N values. To properly accomplish it, the following problem formulation has been considered.

For each time slot the energy balance E_k^{GL} has been distinguished in the energy balance over production (E^{GL+}) and over demand (E^{GL-})

$$E_k^{GL-} = \begin{cases} E_k^{GL} & \text{if } E_k^{GL} < 0 \\ 0 & \text{otherwise} \end{cases}, \quad E_k^{GL+} = \begin{cases} E_k^{GL} & \text{if } E_k^{GL} \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{7.8}$$

According to the problem simplification introduced in Sec. 7.1 and the equations introduced in Sec. 7.2, the following constraints are considered

$$E_k^{GL} = E_k^{GL+} + E_k^{GL-}, \ \forall\, k = 1,...,N \tag{7.9}$$

$$E_k^{S} = -(E_k^{S+} + E_k^{S-}), \ \forall\, k = 1,...,N \tag{7.10}$$

$$E_k^{N-} = E_k^{GL-} - E_k^{S-}, \ \forall\, k = 1,...,N \tag{7.11}$$

$$E_k^{N+} = E_k^{GL+} - E_k^{S+}, \ \forall\, k = 1,...,N \tag{7.12}$$

$$0 \geq E_k^{S-} \geq E_k^{GL-}, \ \forall\, k = 1,...,N \tag{7.13}$$

$$0 \leq E_k^{S+} \leq E_k^{GL+}, \ \forall\, k = 1,...,N \tag{7.14}$$

$$E^{dch-max} \leq -E_k^S \leq E^{ch-max}, \ \forall\, k = 1,...,N \tag{7.15}$$

$$E^{min} - E^{S0} \leq \sum_{i=0}^{i=k} -E_i^S \leq E^{max} - E^{S0}, \ \forall\, k = 1,...,N \tag{7.16}$$

where E^{S0} is the initial energy of the ESS and E_k^S is the energy exchanged by the ESS at the k-th time slot. It must be within $E^{dch-max}$ and E^{ch-max}. The E^S series must be equal to the sum of E^{S+} and E^{S-} in order to distinguish the the energy discharge and charge, respectively. E^{min} and E^{max} are the ESS state of energy upper and lower limits, respectively.

Finally, taking into consideration (7.3), the OF is

$$P^{opt} = \sum_{k=0}^{k=N} \left(E_k^{N-} \cdot p_k^{buy} + E_k^{N+} \cdot p_k^{sell} \right). \tag{7.17}$$

Chapter 8

Graph Representation of a MG

8.1 Microgrid Configuration Representation Study

8.1.1 Notation and Definitions

Assuming a grid connected MG equipped with a generation system, an aggregated load and an ESS representing a general prosumer, it can be modelled and represented through a graph composed by 4 vertices and 6 edges as shown in Figure 8.1. The

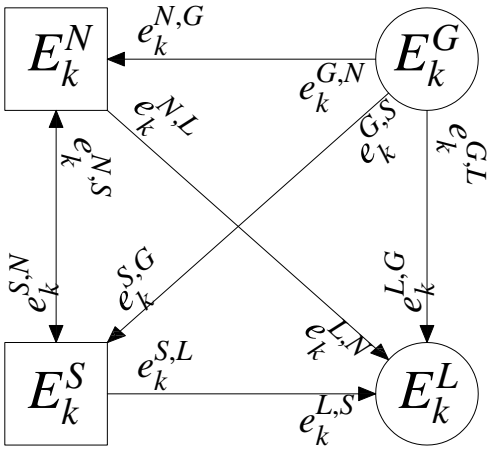


FIGURE 8.1: Graph representation of the MG.

4 vertices represent the aggregated load, the aggregated generation, the aggregated ESS and the *Network* (*i.e.* the Main Grid). These vertices will be referred to with the symbols L, G, S and N, respectively and will be labelled with the symbols E_k^G, E_k^L, E_k^S and E_k^N, which represent the amount of energy exchanged in the k-th time slot respectively by the aggregated generation, the load, the storage and the network. The

energy exchanged by each vertex is assumed to be positive when the corresponding aggregated system (L, G, N and S) behaves as a generator. Therefore, the load energy E_k^L and the generated energy E_k^G will be always negative and positive, respectively. For this reason the G and L vertices will be said to be unidirectional, on the contrary S and N will be said to be bidirectional. As shown in Figure 8.1, the unidirectional and bidirectional vertices are represented with circles and squares, respectively. Edges connected to a unidirectional vertex present an a priori orientation, consequently the only edge that can transfer energy in both directions is the one joining the two bidirectional vertices N and S. This edge is used to model the energy trade between the storage system and the network. Edge labels will be used to describe the fraction of the energy exchanged between the two interconnected vertices. Two different labels can be defined for each of the 6 edges shown in Figure 8.1. These two labels named $e_k^{i,j}$ and $e_k^{j,i}$ are related to each other by the following constraint equation

$$e_k^{i,j} E_k^i = -e_k^{j,i} E_k^j \quad k = 1,2,... \tag{8.1}$$

More precisely, the symbol $e_k^{i,j}$ represent the fraction of the total energy E_k^i exchanged by the vertex i with the vertex j in the k-th time slot. In other words, (8.1) states that the fraction of energy delivered from the vertex i toward the vertex j is equal to the fraction of energy received by the vertex j from the vertex i. Consequently, it should be noted that for each vertex the sum of the edge labels referred to the edges connected to the same vertex must be unitary, as expressed by the following constraint equations

$$\begin{array}{ll} \text{(Vertex } \mathbf{N}\text{)} & e_k^{N,S} + e_k^{N,L} + e_k^{N,G} = 1 \\ \text{(Vertex } \mathbf{S}\text{)} & e_k^{S,N} + e_k^{S,L} + e_k^{S,G} = 1 \\ \text{(Vertex } \mathbf{G}\text{)} & e_k^{G,N} + e_k^{G,L} + e_k^{G,S} = 1 \\ \text{(Vertex } \mathbf{L}\text{)} & e_k^{L,N} + e_k^{L,G} + e_k^{L,S} = 1 \end{array}, \qquad k = 1,2,... \tag{8.2}$$

Finally, it should be noted that, the combination of the 10 constraint equations (4 vertex and 6 edge constraint equations (8.2) and (8.1) respectively) implies the conservation of the total energy exchanged for each time slot. In fact, by multiplying each terms of (8.2) by the corresponding vertex energy, adding the results and applying (8.1), it is obtained:

$$E_k^N + E_k^S + E_k^G + E_k^L = 0 \qquad k = 1,2,... \tag{8.3}$$

8.1.2 Configurations Enumeration and Classification

In this study it is assumed that each of the six edges of the graph can be enabled or disabled within the k-th time slot. Moreover, for each enabled edge it is important to consider the energy flow orientation. As previously said the orientation of an edge which connects at least one unidirectional vertex is well defined a priori, whereas only the edge interconnecting the two bidirectional vertices S and N can assume both the possible orientations. Consequently there are 5 edges that can assume two states only, whereas the remaining edge can assume three states. Therefore the MG can assume potentially $3^1 \cdot 2^5 = 96$ topological configurations.

Assuming that the MG is connected with the main grid through a single meter, all the potential configurations allowing simultaneous positive and negative energy exchange with the grid must be discarded. In other words the MG cannot sell and buy energy from the grid at the same time. Similarly, the ESS cannot be charged and discharged simultaneously and consequently all the potential configurations that allow this operation must be discarded. Considering both the previously described requirements 62 among the 96 potential configurations must be discarded due to their unfeasibility. The remaining 34 configurations, shown in Figure 8.2, will be said feasible. In Figure 8.2 the vertices disconnected by the MG are drawn in grey in order to highlight that they are inactive (*i.e.* they do not exchange energy). It is useful to remark that some of these configurations describe limit situations which may be

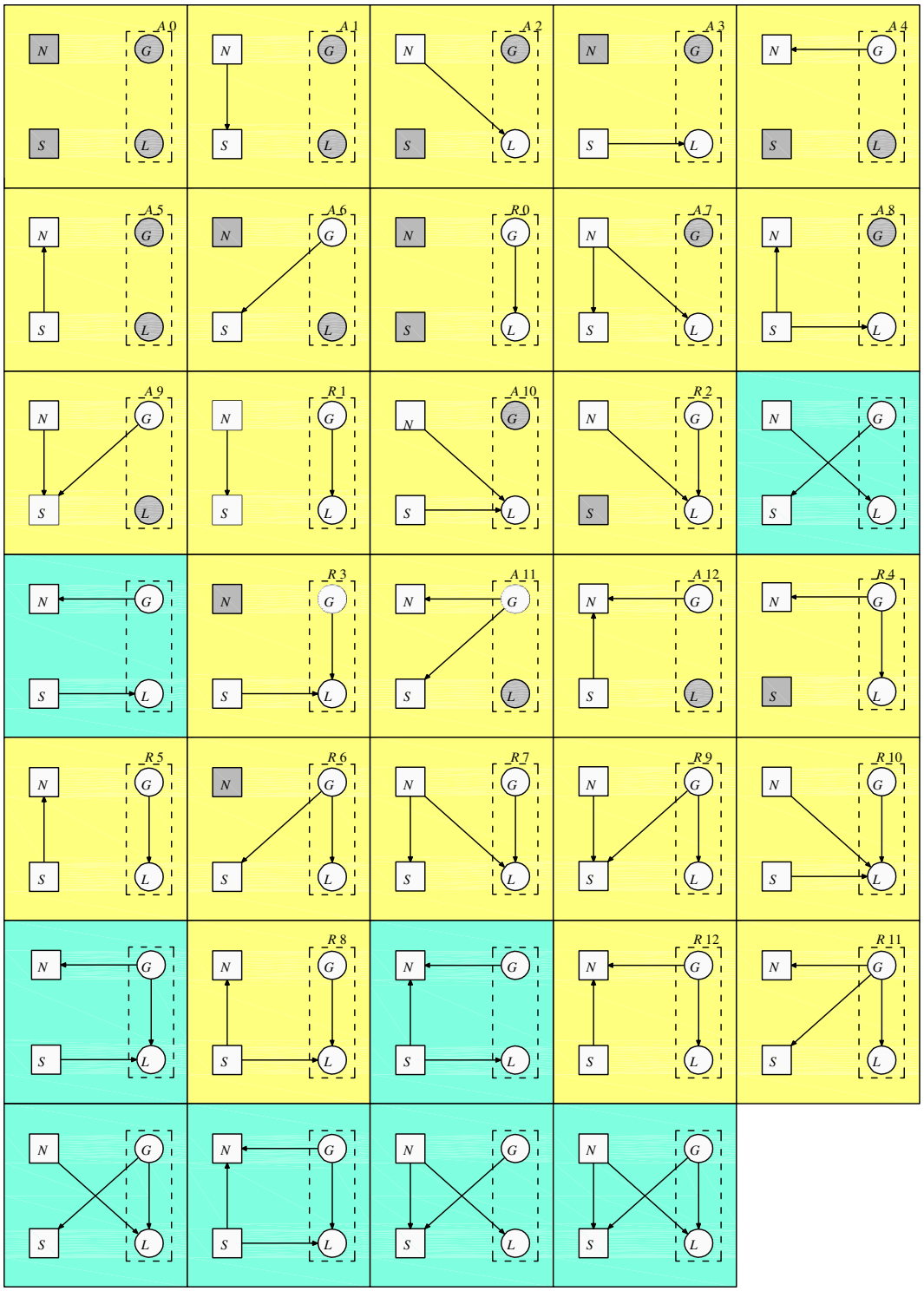


FIGURE 8.2: Summary of the 34 feasible configurations of the MG.

very rare in practical situation but, for the completeness of the study, they must be listed. For example, the configuration at the upper left corner of Figure 8.2 describes the situation in which there is no generation, no demand and no energy trade between the ESS and the grid. In the following this configuration will be referred to as trivial.

At this point, it is useful to introduce the local consumption assumption. Based on this hypothesis it is assumed that the energy E_k^G produced by the aggregated generation will be primarily used to meet the load requirements. This assumption allows to define the prosumer energy balance as

$$E_k^{GL} = E_k^G + E_k^L, \qquad k = 1,2,... \tag{8.4}$$

At topological level the local consumption assumption correspond to those configurations having an unidirectional energy exchange between the $G-L$ portion (see dashed rectangles in Figure 8.2) and the $S-N$ portion of the graph.

The 26 feasible configurations implementing the local consumption assumption are highlighted in Figure 8.2 with a yellow background colour and will be referred to as LC configurations. Similarly, the remaining 8 configurations, which do not implement the local consumption assumption, are highlighted with a light blue background colour.

Moreover, it can be noted that the 26 LC configurations can be paired in 13 couples by considering the two following cases.

- *Case 1*: when $E_k^G E_k^L = 0$ the prosumer balance E_k^{GL} is said absolute and the edge between the vertices G and L is disabled.
- *Case 2*: when $E_k^G E_k^L \neq 0$ the prosumer balance E_k^{GL} is said relative and the edge between the vertices G and L is enabled.

In fact, among the 26 feasible LC configurations one half have an absolute prosumer balance, whereas the remaining 13 have a relative prosumer balance, and they can grouped in pairs differing only for the status of the edge between the vertices G and L.

In order to remark this concept a label is added in Figure 8.2 for each LC configuration. More precisely, each label includes the letter R or A, depending if the configuration implement a relative or an absolute balance condition, and a number

ranging from 0 to 12. The 0 label is used to refer the trivial configuration (*i.e.* no energy exchange among the prosumer, the ESS and the grid).

It is interesting to note that each pair can be represented by a single configuration by considering the two unidirectional vertices G and L as a single bidirectional virtual vertex GL (see dashed rectangles in Figure 8.2). This further simplification allows to represent all the MG configurations compliant with the local consumption assumption as graphs having three bidirectional vertices and three edges (see Figure 8.3).

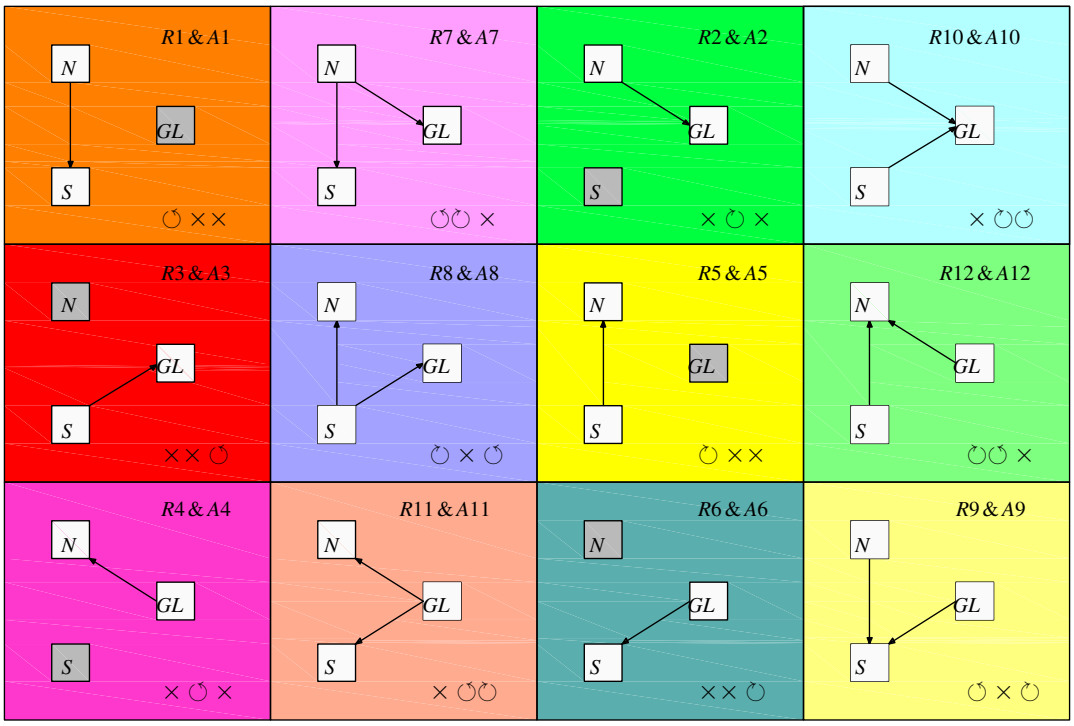


FIGURE 8.3: Summary of the 12 non trivial feasible LC configurations of the MG.

In this situation the energy balance constrain (8.3) can be simplified as follows

$$E_k^S + E_k^N + E_k^{GL} = 0, \qquad k = 1, 2, \ldots \tag{8.5}$$

With this representation each LC configuration is defined by the state of only three edges. Associating the symbols: $\times$, $\circlearrowleft$ or $\circlearrowright$, to disable state, clockwise enable state and counter-clockwise enable state, respectively, it is possible to completely

describe each configuration using a string of three symbols, representing the energy exchange between the network and the ESS, the network and the prosumer and the prosumer and the ESS, respectively, as shown in Figure 8.4 part (c). In fact, as well known from circuit theory, it is always possible to associate a graph with a circuit. In this perspective, the terms clockwise and counter-clockwise refer to the orientation of the virtual current flowing in the mesh composed by the three branches of the circuit associated to the graph edges. More precisely, an edge will be said clockwise oriented or counter-clockwise oriented if its power flow is concordant with the flow of the virtual current as shown in Figure 8.4 part (a) and (b), respectively.

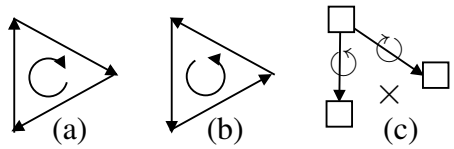


FIGURE 8.4: Loop current orientation: (a) Clockwise, (b) Counter-clockwise. (c) $(\circlearrowleft, \circlearrowright, \times)$ configuration. First, second and third symbols describe the status of the edges connecting the vertices $E^N - E^S$, $E^N - E^{GL}$ and $E^S - E^{GL}$, respectively.

8.2 An Effective Representation for MG Configurations

In this section it is shown that, if the LC assumption is adopted, it is always possible to reconstruct the operational configuration of the MG in the k-th slot of time based on the knowledge of E_k^{GL} and E_k^S. In fact the knowledge of E_k^{GL} and E_k^S allows to determine E_k^N using (8.5). Equation (8.5) defines a plane in the 3-dimensional space spanned by the three variables E^N, E^S and E^{GL}. Consequently, it is possible to define a new reference system having 2 coordinates on the plane defined by (8.5) and the

remaining one orthogonal to it.

$$\begin{cases} W = \frac{1}{6}\left(E^N + E^S + E^{GL}\right) \\ V = \frac{1}{6}\left(E^N + E^S - 2E^{GL}\right) \\ U = \frac{1}{6}\left(E^N - E^S\right) \end{cases} \tag{8.6}$$

The new variables U, V and W allow a simpler analysis of MG energy flows, however they don't have any particular physical meaning. In fact, solving (8.6) for E^N, E^S and E^{GL} and considering that $W = 0$ by construction (see (8.5)), the following equations are obtained

$$\begin{cases} E^S = -3U + V \\ E^N = +3U + V \\ E^{GL} = -2V \end{cases} \tag{8.7}$$

According to (8.7) it is clear that the three variables E^S, E^N and E^{GL} can be expressed in terms of the only two variables V and U.

Three split lines can be defined in the (V,U) plane by setting respectively $E^S = 0$, $E^N = 0$ and $E^{GL} = 0$

$$\begin{cases} E^S = 0 \\ E^N = 0 \\ E^{GL} = 0 \end{cases} \quad \Rightarrow \quad \begin{cases} V = +3U \\ V = -3U \\ V = 0 \end{cases} \tag{8.8}$$

Each split line cuts the (V,U) plane in two half-plane associated with positive and negative values of the variable which vanish on the split line itself. As shown in Figure 8.5, by means of different colours and considering the three split lines all together, a partition of the (V,U) plane into 6 sub regions and 6 half-line is obtained. For each sub region and half-line the sign of each variable E^S, E^N and E^{GN} is well defined. Note that on each half-line one and only one of the three variable is zero whereas in the origin all of the 3 variables vanish simultaneously.

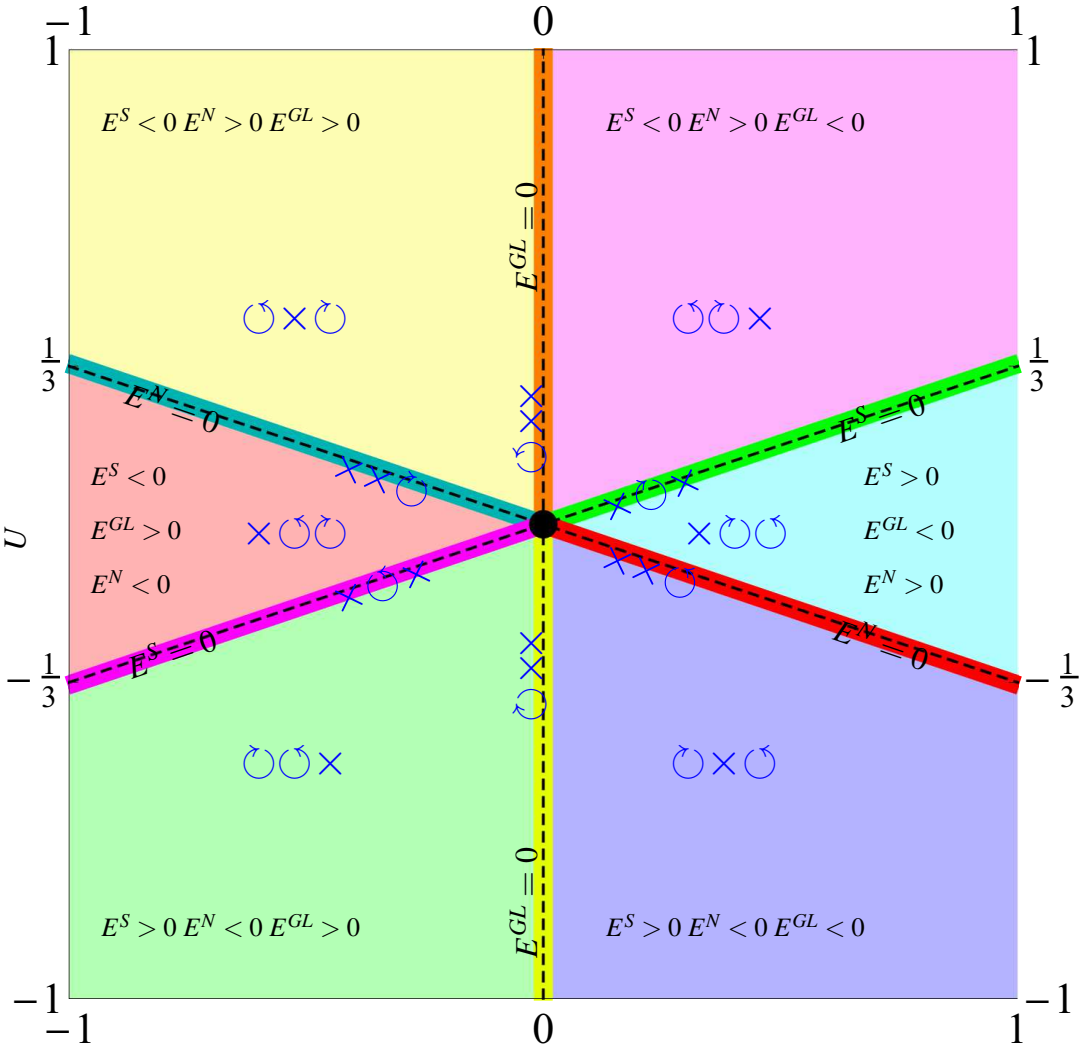


FIGURE 8.5: Representation of the (V,U) plane. The 12 (V,U) regions corresponding to each of the 12 configurations shown in Figure 8.3 are represented with different colours (corresponding to the ones adopted in Figure 8.3).

According with the signs that the variables E^S, E^N and E^{GN} assume in the 6 sub regions and on the 3 split lines of the (V,U) plane (see Figure 8.5), it is possible to map each of the 13 configurations in one sub region, over a split line or on the (V,U) plane origin. For example, let us consider the sub region on the right of Figure 8.5, which is characterized by the signs $E_k^S > 0$, $E_k^N > 0$ and $E_k^{GL} < 0$. The inequality $E_k^{GL} < 0$ implies that prosumer vertex *GL* is in overdemand condition.

Being simultaneously $E_k^N > 0$ and $E_k^S > 0$ both the *N* and *S* vertices must provide energy. Therefore, in order to avoid an unfeasible power flows redistribution, the edge interconnecting *N* and *S* must be disabled. The previous arguments completely describe the configurations $\times$ ↻↺.

Furthermore, it should be noted that for the points of the (V,U) plane lying on the splitting lines, the MG graph is not connected. For example for the points lying on the splitting line defined by $E^S = 0$, the *S* vertex is completely isolated, (*i.e.* the ESS is inactive). Finally, both the *S* and the *N* vertices are disconnected from the remainder of the grid in the origin of the (V,U) plane.

The previous arguments imply that at each non trivial point on the (V,U) plane corresponds unequivocally to one of the 12 configurations shown in Figure 8.3, whereas the plane origin corresponds to the trivial configuration $\times \times \times$. In order to remark this result each portion of the (V,U) plane in Figure 8.5 is provided with the string describing the corresponding configuration. The use of the (V,U) plane allows to determine the actual MG operational configuration and to visualize it on a suitable Human Machine Interface (HMI). Once the operative configuration is know, by using (8.1) and (8.2) the fraction of energy exchanged by each edge can be determined.

8.2.1 FIS based EMS Synthesis

In this chapter it is proposed a FIS-EMS based on a Mamdani FIS capable to estimate E_k^S for each time slot based on $\mathbf{u}_k$ [66].

The optimization procedure is aimed to the maximization of the overall profit gained by the energy trading with the main grid which includes the ESS degradation cost. The FIS-EMS does not consider any particular law restriction or contract constraint for the energy trade between the MG and the grid (*i.e.* on the E_k^N exchanged in each time slot) in order to consider each MG feasible configuration.

The synthesis of the FIS is performed off-line by a GA through a data driven approach [93]. Once trained, the EMS is able to take decisions in real time with low computational cost requirements. The FIS and GA taken into consideration are described in Chapter 3. The GA is used to determine an efficient rule base system, namely it is restricted on the only Θ_{RC} subset (see. Chapter 3-Sec. 3.1.1). The FIS Input array is defined by the current energy prices, the current ESS SOC and the energy balance. Moreover, in this study each linguistic variable involved in the problem is represented by a specific Term Set composed by 3 Fuzzy Sets.

The considered OF expressed in Monetary Unit (MU), is defined as

$$F = F_p - F_{ESS} = \sum_k P_k - p^{wear} \cdot \sum_k \left| E_k^S \right| \quad [\text{MU}] \tag{8.9}$$

where the summations are extended to the overall training simulation time. The term F_p represents the trading profit function (*i.e.* the summation of the overall incomes and expenses in the whole training interval as defined in (7.3)), whereas the second term is used to model the degradation cost of the ESS as a function of the energy exchanged in both charge and discharge conditions during the overall training simulation. The parameter p^{wear}, measured in [MU/kWh], represents the degradation cost associated with each kWh exchanged by the ESS. It allows to modulate the ESS degradation cost with respect to the energy trading profit F_p. The evaluation of F aims to point out if it could be potentially convenient for the prosumer (the owner of the MG) to invest on the installation of an ESS in order to trade energy with the main

grid. The FIS has been optimized in order to maximize the profit considering different values of p^{wear}, aimed at finding a threshold after which it becomes inconvenient to equip the MG with an ESS.

8.3 Simulation Settings and Results

The specific MG considered in this project is composed by a PV system, a system of aggregated loads and an ESS. The PV system and the aggregated loads are characterized by 19.95 kW and 8 kW peak power, respectively.

The ESS taken into consideration is a Li-ion SCiB battery pack produced by Toshiba (see Table 8.1). Its SoC ranges in $[0.15, 1]$.

TABLE 8.1: Toshiba SCiB module main characteristics.

Capacity	Rated Voltage	Current	Energy	Efficiency
80 (Ah)	300 (V)	8 (C-Rate)	24 (kWh)	1 (adim)

The dataset of aggregated power generation and consumption is shown in Figure 7.2 in Chapter 7. The time series cover one year with a sampling rate of 15 minutes. The dataset has been split in two disjoint sets taking the samples relative to the odd and the even days in order to build the TR and TS, respectively. The learning procedure is performed on the TR and the results are drawn on the TS.

The daily energy costs in sale and purchase are those shown in Figure 7.3 (see Chapter 7).

About the GA settings, these are shown in Table 8.2. The initial population is generated randomly. The optimization is arrested if the best individual doesn't improve by its $\gamma_{conv} = 5\%$ after $N_{conv} = 20$ generations or alternatively when it has been reached the maximum number of generation (*i.e.* $N_{stop} = 200$) with the purpose to extend as much as possible the GA exploration and exploitation performance. The OF is defined in (8.9). During each simulation the parameter p^{wear} is kept constant.

TABLE 8.2: GA Settings.

Individuals		**Operators Settings**		**Variables**		
					cardinality	type
N_{ind}	80	Γ_{select}	roulette wheel	Θ_{RC}	N_{rules}	integer
N_{elite}	10%	Γ_{cross}	scattered			
N_{cross}	$0.6 \cdot N_{ind}$	Γ_{small}	$\gamma_{freq} = 0.8$			
			$\gamma_{dist} = 10\%$			
			$\gamma_{dump} = 0.5$			
		Γ_{big}	$N_{scale} = 4$	$N_{rules} = N_{mf}^{N_{in}}$		
		Γ_{del}	$\gamma_{th} = F_{w/o\ ESS}$ (see Sec. 8.3.1)			
		Γ_{stop}	$N_{stop} = 200$			
			$\gamma_{conv} = 5\%$			
			$N_{conv} = 20$			

$N_{in} = 3$ and $N_{mf} = 5$.

8.3.1 Results

In this study the FIS EMS has been optimized on the TR previously described.

Changes in the value of the p^{wear} parameter, and thus on the OF, yield different FISs resulting from the learning procedure, as it is shown in Figure 8.6, where the TS performances are plotted as a function of p^{wear}.

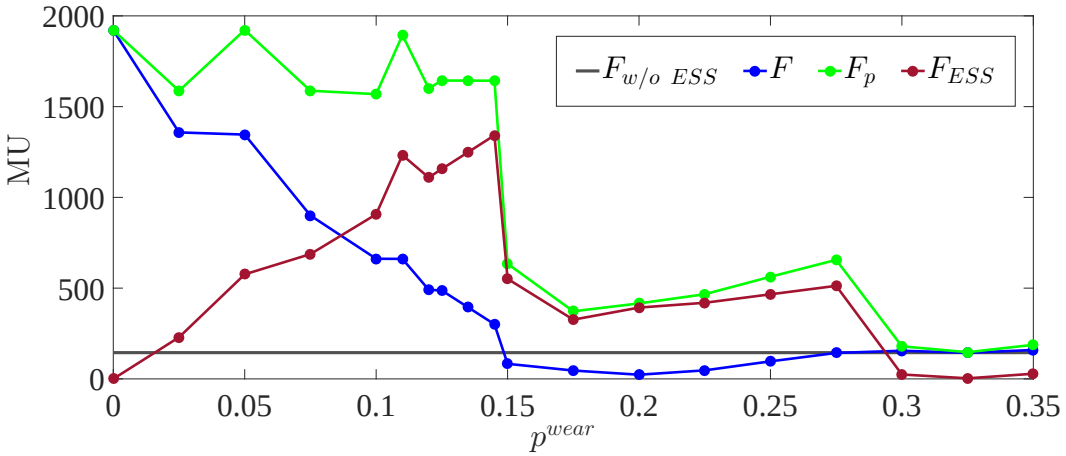


FIGURE 8.6: Comparison between the TS results obtained by the EMS optimized for different values of the p^{wear} parameter. The grey line represents the performance $F_{w/o\ ESS}$ of the EMS when considering a MG without ESS.

In this figure the values of F, F_p and F_{ESS} representing respectively the GA OF, the incomes generated by the energy trading and the battery degradation cost are

shown (see (8.9)). Moreover, the grey line represents the performance $F_{w/o\ ESS}$ of the EMS in the case the capacity of the ESS is set to zero (i.e. for a MG without storage capability). It can be noted that for values of p^{wear} ranging between 0 and 0.15 [MU/kWh] the ESS degradation cost F_{ESS} increases almost linearly. Being F_{ESS} roughly proportional to p^{wear} (see (8.9)) the ESS utilization (*i.e.* $\sum_k \left|E_k^S\right|$) is approximately constant in this range. Conversely, when p^{wear} is about 0.15 the ESS utilization has a significative decrease resulting in a consequent decrease of the energy trading profit F_p and consequently of the overall OF F, which decreases under $F_{w/o\ ESS}$, showing a not well trained FIS decision making strategy. The same inefficient behaviour can be found for values of p^{wear} ranging from 0.15 to 0.275. For higher values than 0.275 the OF reaches $F_{w/o\ ESS}$ and the overall sum of E_k^S is zero. This value shows when the EMS definitely stops to exploits the ESS because of its high degradation costs.

In Figure 8.7 (a) and (b) are shown the daily energy trading for $p^{wear} = 0$ and $p^{wear} = 0.15$ MU/kWh, respectively, with the purpose to exhibit the change of the EMS strategy. In these figures four periods have been highlighted referred to as 1^{st}, 2^{nd},3^{rd} and 4^{th} respectively. This distinction allows to focus on the different ESS utilization strategies during the considered day. It depends mostly on the trends of the energy prices since the OF in (8.9) is strictly related to them.

In Figure 8.8 (a) and (b) are reported the configurations that the MG assumed for $p^{wear} = 0$ and $p^{wear} = 0.15$ respectively, distinguishing the four periods. The configurations adopted are shown on the (V,U) plane introduced by (8.6) in Sec. 8.2. When $p^{wear} = 0$, in the 1^{st} and 4^{th} period (*i.e.* night and early morning), the EMS selects prevalently $(\circlearrowright, \times, \circlearrowleft)$ and $(\circlearrowleft, \circlearrowright, \times)$ configurations (see Figure 8.3). In fact, in this situation, when the energy price is higher in sale than in purchase the EMS buy as much as energy it can. When the ESS is almost full the energy stored in the ESS is sell to the main grid monetizing the price difference between buying and selling. This procedure is repeated until the pricing conditions continue to be favourable. In case of $p^{wear} = 0.15$, for the same periods the energy trading is strictly limited by the high

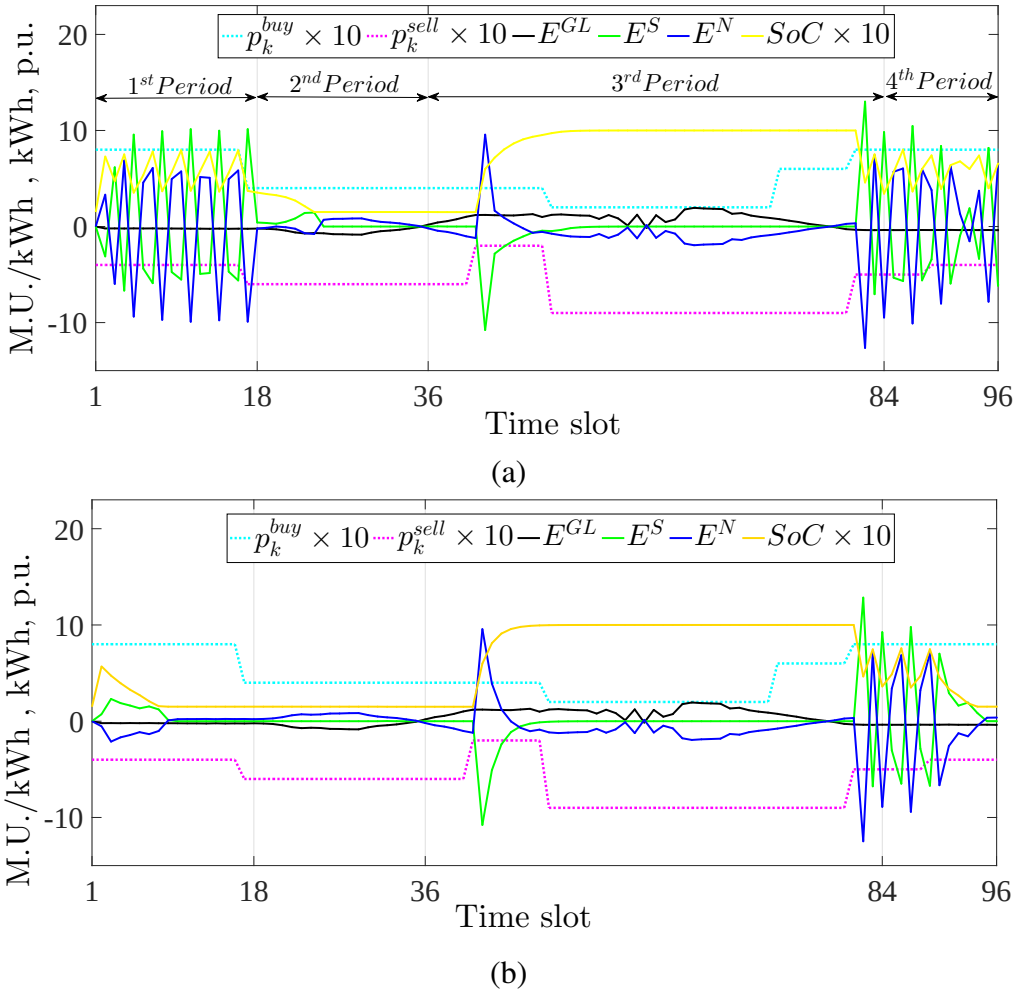


FIGURE 8.7: MG energy flows on the TS: (a) $p^{wear} = 0$, (b) $p^{wear} = 0.15$.

value of p^{wear} that penalizes the use of the ESS. In fact, especially in the 1^{st} period, the EMS prefers the selection of the configuration $(\times, \circlearrowleft, \times)$ and $(\circlearrowleft, \times, \circlearrowleft)$ instead of configuration $(\circlearrowleft, \circlearrowleft, \times)$ (see Figure 8.3). At the beginning of the 2^{nd} period, for $p^{wear} = 0$, the EMS takes advantage of the ESS in order to satisfy the prosumer over-demand and sale energy to the grid looking at the high energy prices both in sale and purchase (configuration $(\circlearrowleft, \times, \circlearrowleft)$). Afterwards, during the 2^{nd} period, the ESS is always disconnected because of the higher purchasing energy prices (configuration $(\times, \circlearrowleft, \times)$). Then, during this period the prosumer over-demand ($E^{GL} \leq 0$) is mainly satisfied by the main grid. Conversely, in case of $p^{wear} = 0.15$ it is adopted only the configuration $(\times, \circlearrowleft, \times)$ because the ESS was already discharged in the previous period. The purchase price is high during the 3^{rd} period. During the same period

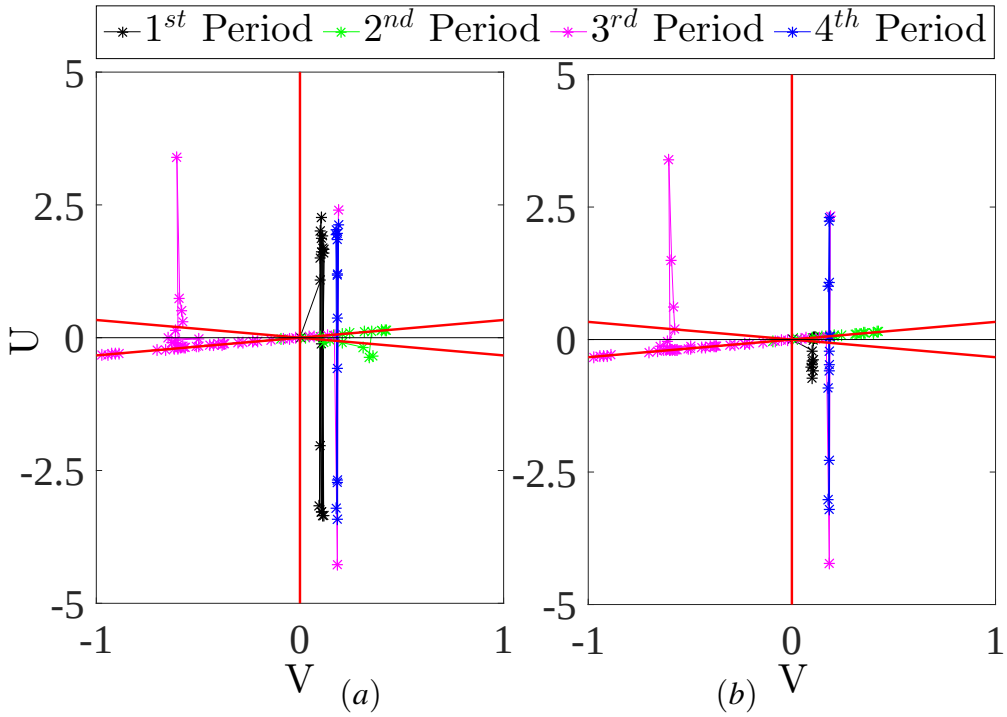


FIGURE 8.8: Representation in the (V,U) plane of the MG operational configuration adopted on the TS (see Figure 8.5): (a) $p^{wear} = 0$, (b) $p^{wear} = 0.15$.

the prosumer MG is in over generation ($E^{GL} \geq 0$) and the ESS has been completely charged ($SoCx10 = 10$) by purchasing energy from the grid at the very beginning of the 3^{rd} period when the purchase price is very low. For this reason, even if the sale price is quite low (in the 3^{rd} period), the EMS is forced to sell the overproduction to the main grid ($E^{N} \leq 0$). At the beginning of the 4^{th} period, when the sale price became high, the EMS starts to discharge the ESS, selling the energy to the main grid. This behaviour is consistent with the OF under examination (*i.e.* the maximization of the prosumer income).

8.4 Conclusion

In this chapter it has been proposed a novel approach on MG energy flows management study, based on a systematic classification of the operative configurations. To this aim the main elements of a MG and the possible electrical connections have

been represented synthetically by a graph made up of four vertices. Here each vertex defines a MG aggregated sub system (generation, load, ESS and the main grid), whereby the edges describe each possible connection among them. Using the above mentioned model together with the local consumption assumptions, it has been studied the MG energy flows redistribution in terms of energy exchange among three bidirectional vertices. Moreover, it has been demonstrated that it is possible to manage in real time the MG energy flows through a MISO EMS by controlling the ESS charge and discharge energy only. In this study it is assumed that the network can exchange energy without any constraints on E_K^N. The proposed EMS consists of a FIS that has been optimized through a GA in order to design an efficient decision making strategy. The optimization takes into consideration both the MG profit and the ESS wear and tear, evaluated by a suitable penalty parameter p^{wear}. Tests have been carried out with different increasing values of this parameter, until the EMS finally stops to exploit the ESS by converging to the solution achieved without energy storage capabilities. In this way it has been possible to focus on the absolute limits about the benefit of such MG-EMS paradigm in terms of profit generated and battery usage. Furthermore the solution performances have been studied by focusing on the overall energy exchanged by the ESS and the configurations adopted during the simulation.

Chapter 9

EMS Synthesis based on FIS-NSGA-II Paradigm

The MG EMS decision strategy proposed in this chapter aims to limit the stress of the main grid by enhancing the auto-consumption, possibly flattening the energy exchanged with the connected grid through a proper use of the ESS. The decision making system considered is the Mamdani type FIS introduced in Sec. 2.3 of Chapter 2. The FIS input vector is restricted to the current energy balance E_k^{GL}, the current SoC_k, and the last value of the energy exchanged with the grid E_{k-1}^N (see 7.2 in Chapter 7). The optimization procedure adopted consists in the implementation of a MO-GA, in particular a NSGA-II (see Sec. 3.3, Chapter 3).

9.1 Multi Objective OF Formulation

The efficiency of a given EMS strategy is evaluated defining suitable performance indices about the stress given to the grid and the MG auto-consumption.

The first term can be defined by the oscillation degree of the total energy exchanged with the grid along the whole period of interest. It is measured by means of the performance index F_N defined by

$$F_N = \frac{\sum_k (E_k^N - E_{k-1}^N)^2}{\sum_k (E_k^{GL} - E_{k-1}^{GL})^2}. \tag{9.1}$$

It should be noted that $F_N = 1$ when $E_k^N = -E_k^{GL}$ for each k-th time slot, *i.e.*, when the MG is devoid of any ESS and $E_k^S = 0$ for each time slot (see (7.1)). Consequently, when $0 < F_N < 1$ the less is the F_N value the more the EMS has effectively smoothed the E^N profile with respect to the same MG devoid of any ESS. Values of $F_N > 1$ correspond to solutions worse than that obtained without ESS. Therefore, latter solutions will be consequently discarded.

The MG auto-consumption is evaluated by the performance index F_A defined by

$$F_A = \sum_k \frac{|E_k^N|}{|E_k^{GL}|}. \tag{9.2}$$

Also in this case $F_A = 1$ when $E_k^S = 0$ in each time slot and similar considerations can be made.

Since neither (9.1) or (9.2) takes into consideration any information related to the ESS internal status it is useful to take into consideration a third performance index F_S designed in order to penalize those EMS strategies producing values of E_k^S outside of the safety SoC range (*i.e.* $[0.15, 0.85]$). The index F_S, constrained between $[0,1]$, has been formulated as the average of the *SoC* penalty function P_S extended to all the training samples.

$$F_S = \frac{1}{N}\sum_k^N (P_S)_k \tag{9.3}$$

The penalty function P_S is obtained as the tailoring of a penalty function described in Figure 14.2

$$P_S = \begin{cases} 1 + \frac{0.002-1}{0.15} SoC & \text{if } SoC \leq 0.15 \\ 0.002 + \frac{1-0.002}{1-0.85}(SoC - 0.85) & \text{if } SoC \geq 0.85 \\ \frac{1}{20}(1 - \sqrt[20]{cos((SoC-0.5)\pi)})) & \text{otherwise.} \end{cases} \tag{9.4}$$

The performance index F_S has a different role with respect to F_N and F_A. In fact, it measures the quality of the solution in terms of ESS safety, whereas F_N and

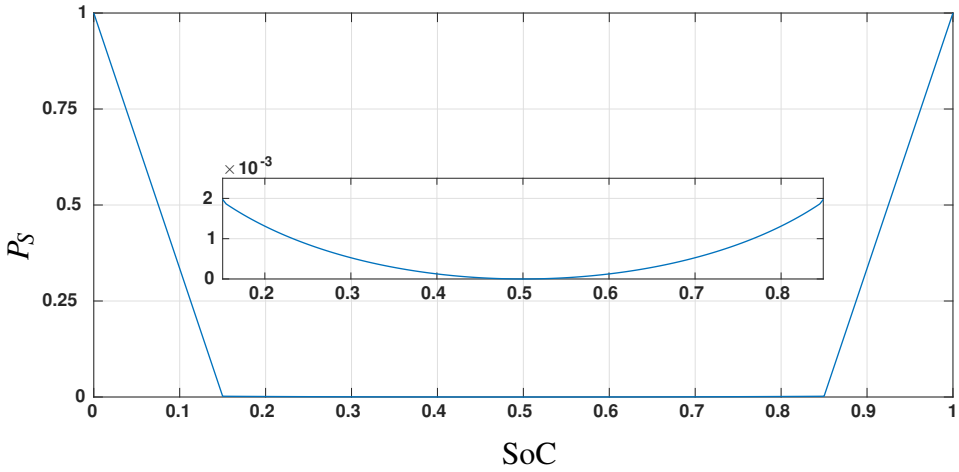


FIGURE 9.1: *SoC* penalty P_S.

F_A measures the quality of the solution in terms of grid stress which is the main requirement.

9.2 FIS Optimization Settings

The decision making systems has been implemented by means of a Mamdani FIS of 3 inputs (*i.e.* current energy balance, ESS *SoC* and the last value of the energy exchanged with the grid) and 1 output (*i.e.* the energy to exchange with the grid). Each linguistic variable has been normalized between [0 1] by considering their respective maximum energy that those variables can exchange (the *SoC* is already normalized by definition). The implemented FIS is a Mamdani type inference system where each term set is composed by 5 triangular MFs with an overlap of 25%. Further information about the FIS architecture are in Sec. 2.3.1 of Chapter 2.

The FIS rules will be automatically constructed by means of a training procedure on real data. More precisely, a custom implementation of a NSGA II will be used to assign the most proper consequent to each fuzzy rule in order to define an efficient rule base FIS. Said in other words, the FIS optimization is restricted to the RC parameters subset Θ_{RC} (see Chapter 3).

Regarding the OF(s) formulation, it has been chosen to take into consideration in

the optimization process only two OFs which comprehends F_N, F_A, F_S in the following manner

$$OF_I = (1-\alpha)F_N + \alpha F_S \tag{9.5}$$

$$OF_{II} = (1-\alpha)F_A + \alpha F_S \tag{9.6}$$

where α is a meta parameter considered equal to 0.5. This formulation allows to consider F_S without formulating a third OF.

This choice is attributable to the fact that it was observed that the use of F_S as independent OF forces the ESS to work around its minimum point (*i.e.* $SoC = 0.5$) reducing its operational range. Moreover, this formulation, by employing only two OFs, allows a more clear study and interpretation of the results.

The NSGA-II parameters are reported in Table 12.1. As can be seen the γ_{th} is set to 1 assuming that both F_S and F_A identify an efficient generated individual when they are lower than 1, whilst F_S is always constrained between 0 and 1.

TABLE 9.1: GA Settings.

Individuals		**Operators Settings**		**Variables**		
					cardinality	type
N_{ind}	80	Γ_{select}	tournament (3 ind.)	Θ_{RC}	N_{rules}	integer
N_{elite}	$0.1 \cdot N_{ind}$	Γ_{cross}	scattered			
N_{cross}	$0.6 \cdot N_{ind}$	Γ_{small}	$\gamma_{freq} = 0.8$			
			$\gamma_{dist} = 10\%$			
			$\gamma_{dump} = 0.5$			
		Γ_{big}	$N_{scale} = 4$	$N_{rules} = N_{mf}^{N_{in}}$		
		Γ_{del}	$\gamma_{th} = 1$ (for both the OFs)			
		Γ_{stop}	$N_{stop} = 500$			
			$\gamma_{conv} = 0.2\%$			
			$N_{conv} = 50$			

$N_{in} = 3$ and $N_{mf} = 5$.

9.3 Simulation Scenario

The specific MG here considered is connected to a PV system, a system of aggregated loads and an ESS. The PV generation and the aggregated loads are characterized by 19.95 kW and 8 kW peak power, respectively. The ESS taken into consideration is Li-ion SCiB battery pack produced by Toshiba (see Table 10.1).

TABLE 9.2: Toshiba SCiB module main characteristics.

Capacity	Rated Voltage	Current	Energy	Efficiency
80 (Ah)	300 (V)	8 (C-Rate)	24 (kWh)	1 (adim)

The overall dataset referred to the PV system and the loads is shown in Figure 7.2 in Chapter 7. The time series covers one year with a sampling rate of 15 minutes.

As in Sec. 8.3 dataset has been split in two disjoint sets taking the samples relative to the odd and the even days in order to build the TR and TS, respectively. The learning procedure is performed on theTR and the results are drawn on the TS.

9.4 Results

Once the optimization process described in Sec. 9.2 has been completed, the synthesized EMSs (*i.e.* last generation individuals of the procedure) which lay on the Pareto Front are simulated on the TS and plotted in Figure 9.2 on the OF_I-OF_{II} plane. In order to compare the behaviours of the best optimized EMSs, the three EMSs circled in green in Figure 9.2 have been taken into consideration. The three chosen solutions, named A, N and AN, present the best performance on OF_I, OF_{II} and a compromise between the two, respectively.

In Figure 9.3 (a), (b) and (c), are shown the ESS SoC profile managed by the MG EMS and the energy flows profiles of each aggregated element (*i.e.* $-E^S$, $-E^N$ and E^{GL} Test series) for solutions A, AN and N, respectively.

In each plot it is spanned the same period of 500 time slots over 17500 (*i.e.* around 5 days over 182).

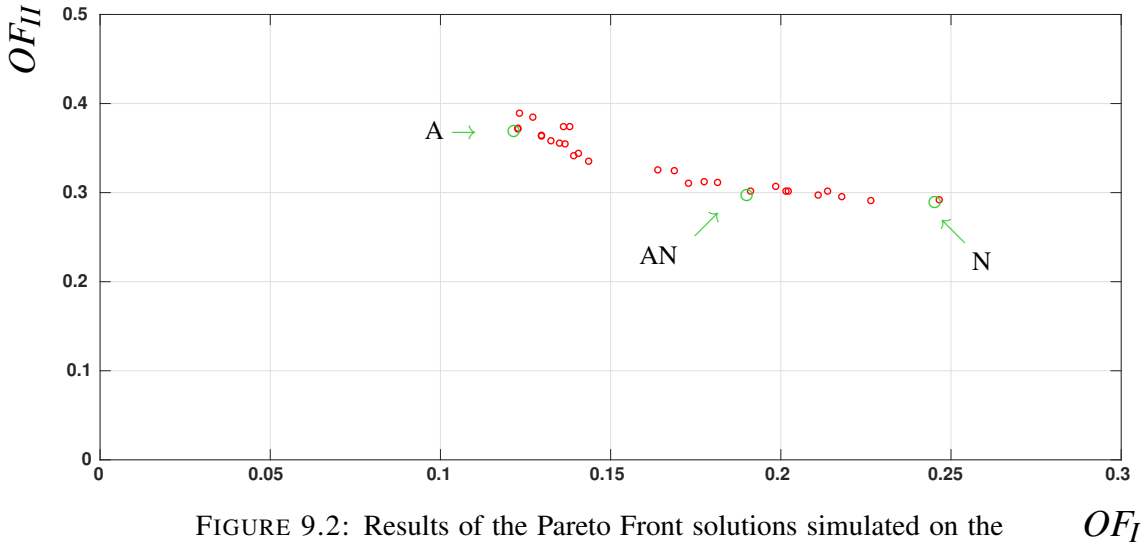


FIGURE 9.2: Results of the Pareto Front solutions simulated on the TS.

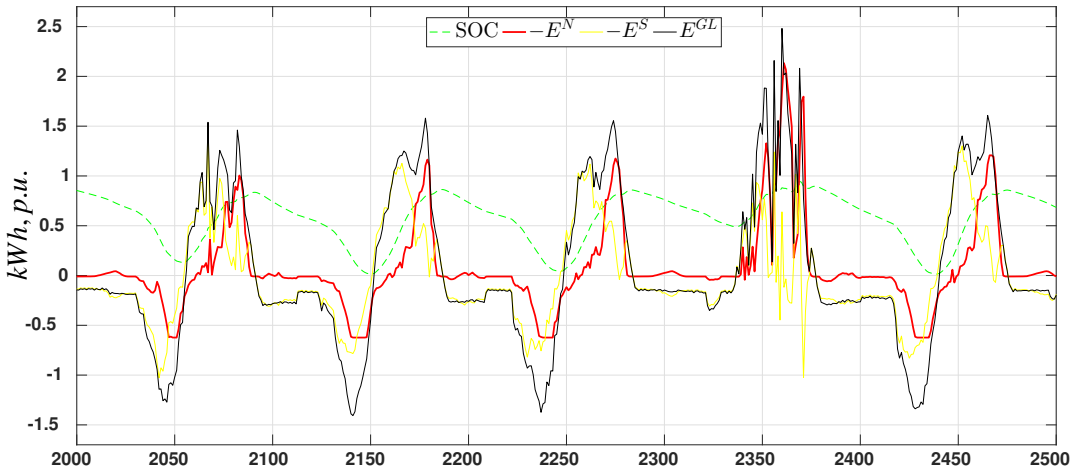
In order to better compare the performances of the three proposed solutions the $-E^N$ and the SoC profiles relative to the solutions N, A and AN are plotted together with the E^{GL} profile in Figure 9.4 (a) and (b), respectively.

The performance indices (*i.e.* F_N, F_A and F_S) achieved by the three proposed EMSs are collected in Table 9.3. On the same table are shown the maximum and

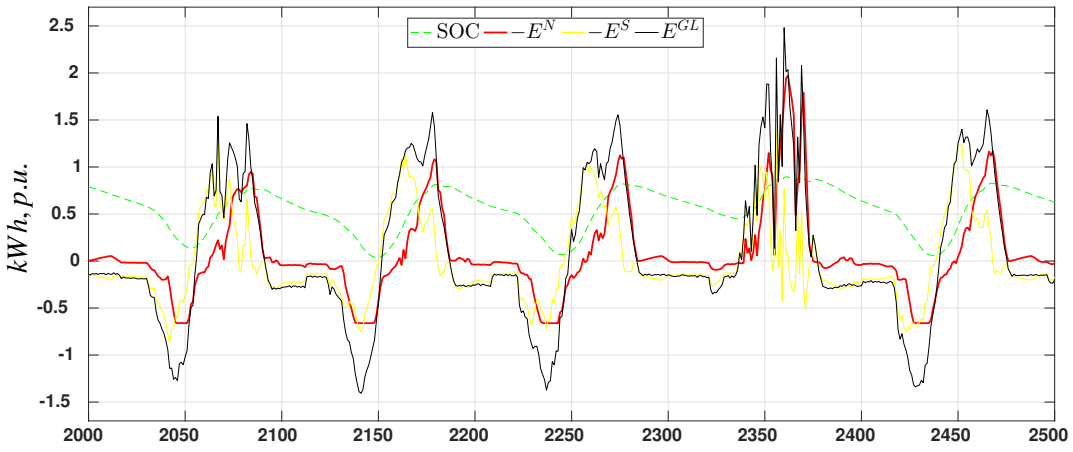
TABLE 9.3: Achieved performance indices on TS results for the A, AN and N solutions.

EMS	F_N	F_A	F_S	$\frac{max\{E^N\}}{min\{E^{GL}\}}$	$\frac{min\{E^N\}}{max\{E^{GL}\}}$
A	0.44	0.52	0.05	0.68	1.00
AN	0.31	0.58	0.04	0.53	0.99
N	0.21	0.71	0.03	0.65	1.00

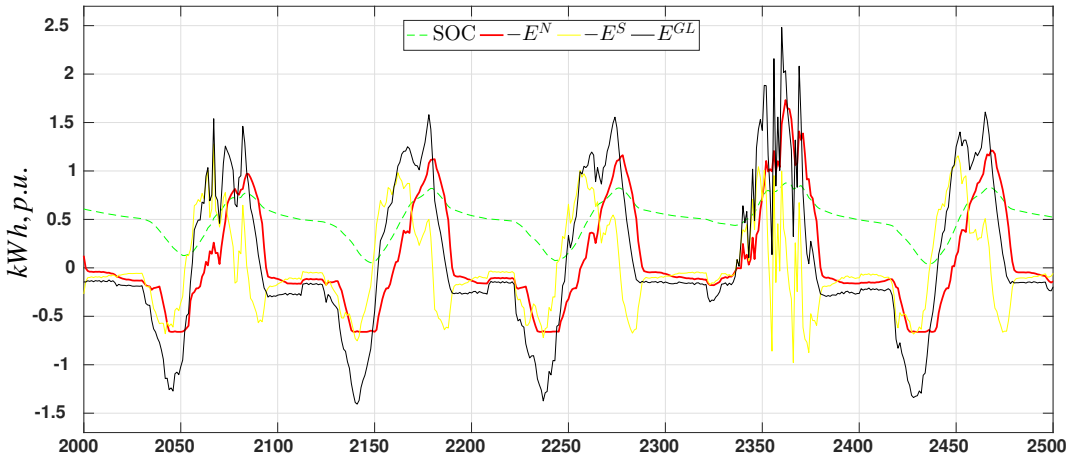
minimum values of the energy sold and bought from the network during the overall simulation normalized with respect to the higher values of E^{GL} in overproduction (*i.e.* $max\{E^{GL}\}$) and in overdemand (*i.e.* $min\{E^{GL}\}$), respectively. It should be noted that the maximum amount of energy sold corresponds to the maximum energy in overproduction, whereas the maximum amount of energy bought ranges from 53% to 68% of the maximum energy in overdemand. Hence, the proposed EMSs are able to perform peak shaving only when the prosumer is in overdemand. This is



(a)



(b)



(c)

FIGURE 9.3: MG energy flow profiles on the TS for solution A (a), AN (b), N (c), respectively.

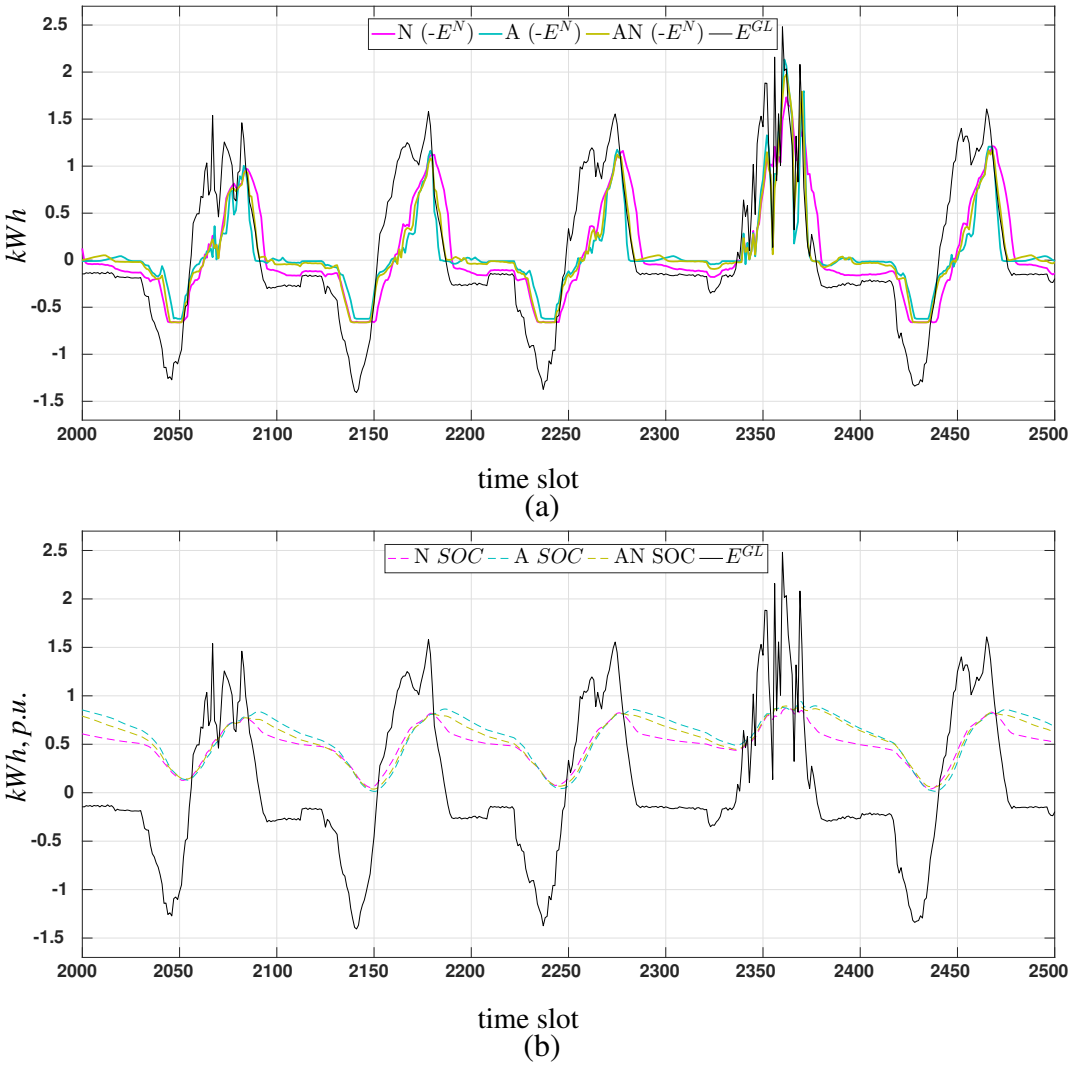


FIGURE 9.4: Comparison among N, A, AN solutions on the TS: (a) E^N, (b) SoC.

reasonable since the daily PV generation is usually greater than the aggregated daily energy demand.

From Figure 9.4 it can be seen that the main discrepancy among the three solutions occurs when E^{GL} is negative and approximately flat (*i.e.* overdemand during night periods). In fact, for solution A and AN, during these period, $E^N \simeq 0$ and the prosumer demand E^{GL} is almost completely satisfied by the ESS, whereas for solution N, the grid energy E^N is positive and the prosumer demand is split between the ESS and the grid. Indeed in solution N, the EMS relies on the ESS energy in order to smooth and delay the E^N peaks, working as a moving average filter of the energy balance (see Figure 9.3 (c)).

In conclusion, as expected solution A presents a better behaviour in terms of energy auto-consumption but it drives the ESS SoC to its limits, whereas solution N performs a lighter use of the ESS but it exchanges more energy with the grid. Solution AN results in a compromise between the solutions A and N.

9.5 Conclusion

In this chapter it has been proposed a novel EMS for a grid connected MG equipped with an ESS. It must satisfy the MG energy balance in real time considering timestamps of 15 mins. The EMS is based on a three inputs FIS and it has been designed in order to reduce the fluctuations of energy exchanged with the grid (*i.e.* the grid stress) and maximize the energy auto-consumptions by employing an efficient utilization of the ESS. The EMS modelling has been realized through a MO-GA based on NSGA-II method. The efficiency of the solutions found are studied selecting three of them which are considered the most representative.

The results show that the auto-consumption performance has been maximized between the 48% (best case) and the 28% (worst case). The grid stress has been reduced between the 44% (best) and 22% (worst). Moreover, in all cases there is a

considerable percentage of peak shaving relative to the maximum power requested from the grid greater than 30% [73].

Chapter 10

EMS FIS-GA Optimization Strategies Comparison

In this chapter a grid connected MG quipped with an ESS which is connected to a system of aggregated loads and a PV generation is considered.

The MG EMS is in charge to control and manage the power flows of the aggregated systems, while satisfying given constraints set by the main grid and assuring the services requested by the prosumer. In particular, it has been assumed that the energy balance cannot be curtailed or shifted and the ESS cannot exchange energy with the grid (see (7.2) (7.4) in Chapter 7).

The EMS decision making system is composed by two FISs, named FIS_{sell} and FIS_{buy} as illustrated in Figure 10.1. Both the FISs have in input the actual ESS SoC (SoC_k) and the next energy balance. Moreover, FIS_{sell} has the current sale energy price C_k^{sell} whereas FIS_{buy} is fed by the current purchasing energy price C_k^{buy} as illustrated in Figure 10.1. The FISs architecture is the same described in Sec. 2.3.1 in Chapter 2 and each term set is composed by 5 MFs.

The two FISs work alternatively, the first in case of over-production (*i.e.* $E_k^{GL} > 0$) the second in case of over-demand (*i.e.* $E_k^{GL} < 0$). More precisely, in case of over-demand (over-production) the output of FIS_{buy} (FIS_{sell}) represents the fraction of E_k^{GL} to be discharged from (charged in) the ESS, whereas in case over-production (over-demand) the FIS_{buy} (FIS_{sell}) output is assumed to be zero. Consequently, in order to estimate the amount of energy E_{k+1}^S to exchange with the ESS during the next

time slot the two FIS outputs are summed and then multiplied by E^{GL}_{k+1}, as shown in Figure 10.1.

In this chapter are compared different GA optimization strategies for the tuning of Θ_{MF}, Θ_{RW}, Θ_{RC} and Θ_{MH} subsets (see Sec. 3.1 and Sec. 3.2 of Chapter 3) of both FIS_{buy} and FIS_{sell} according with the profit OF P by considering a ToU energy price policy (see (7.3)).

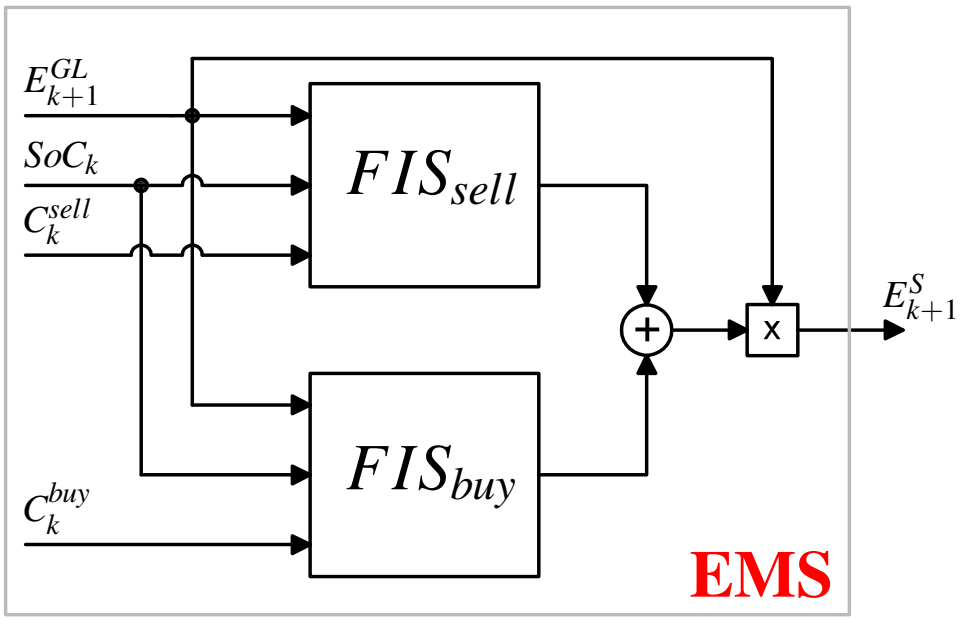


FIGURE 10.1: EMS scheme.

10.1 Optimization Strategies

Four different optimization strategies of the proposed Mamdani FIS decision making system have been considered for the synthesises of the MG EMS described in Chapter 7.

The four optimization strategies taken into consideration are described below:

- $S1$: An initial sub-optimization on Θ_{RC}, Θ_{MF} and Θ_{RW} is applied. Successively it is followed by a final optimization on Θ_{MH}.

- $S2$: An initial sub-optimization on Θ_{RC} followed by a simultaneous sub-optimization on Θ_{MF} and Θ_{RW} is applied. Finally the hierarchical optimization on Θ_{MH} is performed.

- $S3$: The same optimization process described in $S1$ is applied, except the fuzzy input E^{GL}_{k+1} (*i.e.* the energy balance) that is not considered in neither the FIS_{sell} nor the FIS_{buy}. Therefore the two FISs have two fuzzy inputs only.

- $S4$: The same optimization process described in $S2$ is applied, except the fuzzy input E^{GL}_{k+1}, as in the S3 case.

It is clear that $S2$ and $S4$ are organized as a sequence of sub-optimization processes unlike $S1$ and $S3$. For $S2$ and $S4$ the initial optimization on Θ_{RC} is motivated by the primary role of the rule base inference system (*i.e* decision making strategy) with respect to the tuning of the MFs (Θ_{MF}) and rule weights (Θ_{RW}). In fact, Θ_{RC} has more influence on the decisional surface shape, a mutation on a gene referred to Θ_{RC} is more relevant than a mutation on one gene associated to Θ_{MF} or Θ_{RW}. In all cases the hierarchical optimization is executed downstream of the optimization process with the purpose of simplify the FISs structure and delete those fuzzy sets which are not considered effective, necessary or not well tuned. The execution of $S3$ and $S4$ are considered to study the sensitivity of the proposed FIS-GA paradigm to the presence of the E^{GL}_{k+1} fuzzy input in terms of the generated profit and computational burden. The input simplification in $S3$ and $S4$ implies a reduction of variables in each subsets and consequently the complexity of the problem.

10.1.1 Simulation Settings

The specific MG considered in this study is composed by a PV generator, an aggregated load and an ESS. The PV generation and the aggregated load are characterized by 19.95 kW and 8 kW peak of power, respectively. The prosumer energy balance dataset and the daily energy prices are the same introduced in Sec. 7.3 of Chapter 7 (see Figure 7.2 and 7.3, respectively).

The prosumer dataset has been split in two disjoint sets taking the samples relative to the odd and the even days in order to build the TR and the TS, respectively. The learning procedure is performed on the TR and the results are drawn on the TS.

The ESS taken into consideration is Li-ion SCiB battery pack produced by Toshiba (see Table 10.1).

TABLE 10.1: Toshiba SCiB module main characteristics.

Capacity	Rated Voltage	Current	Energy	Efficiency
80 (Ah)	300 (V)	8 (C-Rate)	24 (kWh)	1 (adim)

In Table 12.1 are illustrated the GA parameter setting introduced in Sec. 3.2. The

TABLE 10.2: GA Settings.

Individuals		**Operators Settings**		**Variables**		
					cardinality	type
N_{ind}	80	Γ_{select}	tournament (3 ind.)	Θ_{RC}	$N_{rules} \cdot 2$	integer
N_{elite}	10%	Γ_{cross}	scattered	Θ_{MF}	$N_{MF} \cdot 2$	real
N_{cross}	60%	Γ_{small}	$\gamma_{freq} = 0.8$	Θ_{RW}	$N_{rules} \cdot 2$	real
			$\gamma_{dist} = 10\%$	Θ_{MH}	$N_{MH} \cdot 2$	binary
			$\gamma_{dump} = 0.5$			
		Γ_{big}	$N_{scale} = 4$	$N_{rules} = N_{mf}^{N_{in}}$		
		Γ_{del}	$\gamma_{th} = P_{low}$	$N_{MF} = N_{mf}(N_{in}+1)$		
		Γ_{stop}	$N_{stop} = 1000$	$N_{MH} = N_{mf}N_{in}$		
			$\gamma_{conv} = 0.2\%$			
			$N_{conv} = 50$			

For S1 and S2 $N_{in} = 3$, whereas for S3 and S4 $N_{in} = 2$. $N_{mf} = 5$.

number of individuals N_{cross} and N_{elite} subjected to the cross-over and the elitism operators are expressed in percentage relative to the size of the whole population having N_{ind} individuals. The GA parameters are set in order to reach a balance of both the exploration and the exploitation capabilities. The Θ_{RC} initial population genes are randomly generated between their respective upper and lower bounds. The Θ_{MF} subset is set according to Figure 3.1 (*i.e.* the MFs are equally spaced and symmetric with an overlap of the 50%). Each element of the subsets Θ_{RW} and Θ_{MH} is equal to one in order to give to all the rules the same degree of importance and to active

all the input fuzzy sets when the simulation starts. Summarizing, the individuals of G_0 differs from each other only for the genes associated with Θ_{RC}. The Θ_{MF} subset is composed by $3 \cdot 5 \cdot 4 \cdot 2 = 120$ real variables (*i.e.* genes) where 3, 5, 4 and 2 are the cardinality of the MFs, the term sets, the number of linguistic variables and the number of FISs (*i.e. sell* and *buy* FISs), respectively. The Θ_{RW} subset is composed by $5^3 \cdot 2 = 250$ real valued genes $\in (0,1]$ interval. The Θ_{RC} subset has the same Θ_{RW} number of genes but in this case they are discrete. Θ_{RC} genes assume values that range from 1 to 5 (*i.e.* the term set cardinality). Θ_{MH} is a binary set of size equal to the number of input fuzzy sets (*i.e.* $15 \cdot 2 = 30$ variables). Therefore, the overall Ω set owns 650 elements that represent the EMS parameters and the GA genes.

10.2 Results

In order to rate the results obtained by the four proposed GA-FIS solutions the upper and lower benchmark solution introduced Chapter 7.4 are calculated by considering the profit OF (7.3) and the same constraints regarding the respect of the energy balance and the exchange of energy between ESS and the grid (see (7.2) and (7.4)).

As in Chapter 7.4 the upper and lower benchmark solutions are referred with P_{upp} and P_{low}, respectively.

The simulations results are reported in Table 10.3, where the solutions $S1$, $S2$, $S3$ and $S4$ are compared with the benchmark solutions upper and lower. The simulations has been repeated 5 times with different randomly generated initial populations that differs only on Θ_{RC} as explained in 10.1.1. The values reported in Table 10.3 represent the average values and the standard deviation (in brackets). Each sub-optimization step of each simulation is analysed by simulating their respective best individual on the TS. In Table 10.3 for each sub-optimization step the rest of sub-sets to be optimized are indicated with the superscript '0', whereas those that were already optimized are denoted with the superscript '*'.

The simulation performances are rated in terms of profit (see (7.3)) and number of executed generations. These values are used to estimate the effectiveness and efficiency of the proposed EMS optimization paradigm, respectively.

Moreover, in Table 10.3 are shown the overall energy sold and bought expressed in terms of both $M.U.$ and kWh, the total ESS exploitation that is evaluated as the absolute sum of E_K^S extended to the overall simulation (*i.e.* $\sum_k \left|E_k^S\right|$), and finally the number of fuzzy sets deleted in each hierarchical optimization. At the bottom of Table 10.3 are shown the P_{low} and P_{upp} values. As it is shown in Table 10.3, all the solutions present a profit within the range $76 \div 78\%$ respect to the P_{upp} profit and around the 89% respect to the $P_{upp} - P_{low}$ interval. Simulations which privileges an EMS optimization of the Ω parameters developed in more stages (*i.e.* $S2$ and $S4$) present always greater profits respect to their corresponding optimizations developed in less stages, namely $S1$ and $S3$, respectively. However, these fragmented optimization processes show an increase on the profit constrained within the 2% of P_{upp}. Moreover, it is accompanied by an increase of the overall computational burden that is affected by a higher standard deviation as well.

The simultaneous approach followed in $S1$ and $S3$ shows a greater number of deleted fuzzy sets produced by the hierarchical optimization that, on average is almost double with respect to the other two solutions (*i.e.* 9.5 vs 5.5). This value suggests that, when the subsets $\Theta_{RC}, \Theta_{MF}, \Theta_{RW}$ are simultaneously optimized, the GA shows a poor ability in tuning the FIS nearby optimum or sub-optimum values. It can be noted that the optimization on the Θ_{RC} subset on $S4-a$ shows a profit less than the 3.5% respect to $S2-c$ but relatively much higher than the profit achieved by $S2-a$ (*i.e.* 102.2 Vs 96.2 M.U.). However, looking at the executed number of generations and the profit standard deviation , solution $S4-a$ produces the best performances in terms of efficiency and robustness.

On the contrary, simulation $S2-a$ shows that the optimization of the Θ_{RC} subset provided of the prosumer energy balance as input gives the worst results in terms of both efficiency and effectiveness. In fact, in this case the GA presents difficulties

in finding a good solution in a larger space. However, the subsequent optimizations performed on the remaining subsets bring the profit to higher values, more precisely to the best one. The fuzzy input energy balance increases the average profit less than 1.5% despite a decrease of their robustness and efficiency (*i.e.* higher # Generations, standard deviations on the profit and standard deviations on # Generations). Looking at the overall energy sold and the overall energy bought, all the solutions show values close to the upper benchmark solution that produce a reduction of the overall energy exchanged to the network greater than the %50 of of the overall energy exchanged by the lower benchmark solution.

10.3 Conclusion

In this work are compared four different optimization strategies of an MG EMS structured by two FIS in order to redistribute in an intelligent way the prosumer energy balance. The MG is supposed to be equipped with an ESS and to be connected to the main grid. The optimization, made on a TR through a GA, wants to maximize the profit generated by the energy trading with the main grid buffering the ESS during a given time period on a TS.

The EMS optimization parameters were grouped in four subsets representing the RCs association, the MFs shape, the rule weights and the deactivation of input fuzzy sets. It was also imposed that RCs association must be optimized at the first stage in order to immediately build an efficient inference system, so that the hierarchical optimization must be executed at the end to delete such input fuzzy sets that were inefficiently arranged. The adopted strategies differ in the order in which the optimization of such parameter subsets is applied. In particular, one strategy privileges a simultaneous optimization of the parameters relative to the RCs association, the MFs shape and the rule weights. The second strategy first define the FISs rule base inference system (*i.e.* the RCs association) and successively tunes simultaneously the MFs shape and the rule weights. Moreover, the simulations are repeated without

TABLE 10.3: Optimized EMSs TS results.

Simulation	Initial Population [subset]	Optimized Parameters [subset]	TS Profit [MU]	Energy Sold [kWh]	Energy Sold [MU]	Energy Bought [kWh]	Energy Bought [MU]	ESS exploit. [kWh]	# Deleted Fuzzy Set [adim]	# Generations [adim]
$S1-a$	$\Theta^0_{RC},\Theta^0_{MF},\Theta^0_{RW},\Theta^0_{MH}$	$\Theta_{RC},\Theta_{MF},\Theta_{RW}$	102.3(1.6)	961(14)	283(3)	349(12)	174(4)	1286(25)		650(211)+
$S1-b$	$\Theta^*_{RC},\Theta^*_{MF},\Theta^*_{RW},\Theta^0_{MH}$	Θ_{MH}	**104.2 (1)**	957(12)	283(2)	347(12)	172(3)	1290(24)	10(2)	98(40) = 748 (251)
$S2-a$	$\Theta^0_{RC},\Theta^0_{MF},\Theta^0_{RW},\Theta^0_{MH}$	Θ_{RC}	96.2(2.6)	945(10)	278(2)	334(10)	175(2)	1315(20)		748(368)+
$S2-b$	$\Theta^*_{RC},\Theta^0_{MF},\Theta^0_{RW},\Theta^0_{MH}$	Θ_{RW},Θ_{MF}	105.4(0.8)	944(10)	280(2)	332(10)	167(3)	1092(20)		346(110)+
$S2-c$	$\Theta^*_{RC},\Theta^*_{MF},\Theta^*_{RW},\Theta^0_{MH}$	Θ_{MH}	**105.8 (0.5)**	943(10)	280(2)	332(10)	166(2)	1320(20)	5(4)	75(40) = 1170 (518)
$S3-a$	$\Theta^0_{RC},\Theta^0_{MF},\Theta^0_{RW},\Theta^0_{MH}$	$\Theta_{RC},\Theta_{MF},\Theta_{RW}$	103.7(0.8)	967(13)	285(3)	356(13)	174(3)	1272(26)		512(188)+
$S3-b$	$\Theta^0_{MH},\Theta^*_{RC},\Theta^*_{MF},\Theta^*_{RW}$	Θ_{MH}	**104.0 (0.6)**	975(6)	286(1)	364(6)	175(2)	1256(13)	9(1)	68(11) = 581 (198)
$S4-a$	$\Theta^0_{RC},\Theta^0_{MF},\Theta^0_{RW},\Theta^0_{MH}$	Θ_{RC}	102.2(0.5)	969(1)	285(1)	359(1)	176(1)	1267(3)		277(85)+
$S4-b$	$\Theta^*_{RC},\Theta^0_{MF},\Theta^0_{RW},\Theta^0_{MH}$	Θ_{RW},Θ_{MF}	104.2(0.7)	971(3)	286(1)	363(7)	174(1)	1045(5)		378(223)+
$S4-c$	$\Theta^*_{RC},\Theta^*_{RW},\Theta^*_{RC},\Theta^0_{MH}$	Θ_{MH}	**104.3 (0.6)**	971(3)	286(1)	360(3)	174(6)	1263(4)	6(4)	52(4) = 707 (312)
P_{upp}	-	-	**135.1**	901	270	289	135	1405	-	-
P_{low}	-	-	**-144.6**	1599	411	997	556	0	-	-

considering the prosumer energy balance as fuzzy input in order to make a study on the sensitivity to this input. TS results show that the performances relative to the profit given by the optimized EMSs are close to the 80% compared to the upper bound solution (*i.e.* the optimal one). Moreover, the energy exchanged with the main grid is also reduced by more than the 40% respect with the lower bound solution (*i.e.* considering the same MG devoid of ESS). The hierarchical optimization, as it could be expected, works better when there is a large number of input term sets, or in other words when the GA is affected by a considerable number of variables which makes more difficult a fine tuning of the EMS.

Chapter 11

ANFIS EMS Synthesis by Clustering

In this chapter the problem under analysis about the synthesis of a decision making system has been switched to a supervised problem by considering an upper (and lower) benchmark solution evaluation. Its evaluation allowed the synthesis of an ANFIS based EMS by means of clustering techniques which aims to maximize the profit formula (7.3) by considering a ToU energy price policy. Three different clustering algorithms are compared for the EMS synthesis problem. By starting from the widely-known k-means algorithm, the others are mainly re-adaptations and/or extensions of it (still, well-known in literature) in order to explore and implement different (dis)similarity measures. In particular, besides k-means, the k-medians and k-medoids variants discussed in Chapter 4 Sec. 4.1.1 and Sec. 4.1.2 have been tested. Moreover, by taking advantage of the output space (*i.e.* the E^N_{opt} profile found through the upper benchmark solution), it is also investigated the clustering joint input-hyperplane space.

In this chapter, the EMS has to respect all the constraints introduced in Sec. 7.2 of Chapter 7. The OF taken into consideration is the profit formula (7.3) that in phase of validation and test is normalized as in (7.6) by considering the upper and lower benchmark solutions. These are calculated as in the example introduced in Sec. 7.4 (*i.e.* by means of a LP formulation).

Looking at the EMS architecture, the ANFIS is fed by an input vector constituted by 4 variables, namely the current energy balance E^{GL}_k, the current energy prices in sale and purchase (*i.e.* C^{sell}_k and C^{buy}_k) and the current ESS SoC Soc_k.

11.1 EMS Modelling Procedure

In this section it is explained in details how the ANFIS EMS introduced in Sec. 2.3.2 of Chapter 2 is efficiently modelled by means of a clustering algorithm. The ANFIS optimization process is explained from a generic point of view, valid for each clustering algorithm proposed in Chapter 4 Sec. 4.1, including their variants.

The clustering algorithm relies on a given dataset composed by E^{GL}, C^{buy} and C^{sell} time series. The whole dataset has been partitioned in TRs, VLs and TS in order to train the ANFISs, select the best for a given number of MFs and measure the performances of the optimum one, respectively. The dataset partition among TS, TRs and VLs is made on a daily base, namely all time slots associated with the same day will belong to a single set. More precisely, the whole dataset is firstly divided in two subsets having the same cardinality. The first subset constitutes the TS whereas the second one is partitioned in 5 different ways in order to constitute 5 different TR-VL pairs.

The clustering procedure exploits the generic TR_j dataset, namely the TR of the j-th TR-VL pair, together with the corresponding optimum profiles of $\left(SoC^{opt}\right)_j$ and $\left(E^{N,opt}\right)_j$ found through the LP optimization. For the sake of ease, the dataset read by the clustering algorithm is defined by $\Gamma_j^{cl} = \left\{E_j^{GL}, C_j^{buy}, C_j^{sell}, SoC_j^{opt}, E_j^{N,opt}\right\}$. It is clear that if the only EMS input space is considered, $\left(E^{N,opt}\right)_j$ will not take part as a feature for clustering and will be only used for the estimation of the hyperplanes coefficients defined by θ_j, once they are defined (see (4.6)). Conversely, if the clustering procedure is applied on the joint input-output space, the coefficients θ_j are updated after each Voronoi iteration together with the clusters.

The optimization process is ran for a number of clusters k ranging from $k = 2$ to $k = 25$ on the j-th dataset Γ_j^{cl}.

For each k, the clustering procedure is repeated for each Γ_j^{cl} (*i.e.* 5 times), returning the respective k representative vectors, covariance matrices and hyperplanes coefficients. These are used to model the ANFIS multivariate Gaussian MFs Φ_{kj} (*i.e.* rule

antecedent set) and the Sugeno hyperplanes (*i.e.* rule consequent set) represented by the θ_{kj} coefficient sets, respectively, according to (2.5) and (2.6), as introduced in Sec. 2.3.2. The previously described procedure results in 24×5 different ANFIS synthesis. More precisely, for each k value there exist 5 different ANFIS corresponding to the 5 different TRs. Each generated ANFIS is then simulated on its respective VL and its performance is evaluated according to (7.6). For each value of k, the best ANFIS among the 5 TR-VL pairs, named $ANFIS_k^{best}$, is selected whereas the remaining 4 are discarded. Finally, for each $ANFIS_k^{best}$ the MG EMS is simulated on the TS in order to evaluate the generalization capability.

The overall procedure of the EMS ANFIS design and optimization is illustrated in Algorithm 5.

Algorithm 3 EMS Training Procedure

procedure EMS DESIGN
  TS and $\langle TRs, VLs \rangle$ pairs partitioning
  for $j = 1$ to 5 **do** ▷ for each TR_j
    $\{E^{N,opt}, SoC^{opt}\}_j := \mathrm{LP}(TR_j)$ ▷ evaluation of the optimal solution on the j-th TR
    $\Gamma_j^{cl} := \{TR_j, SoC^{opt}, E^{N,opt}\}$ ▷ Clustering labelled pattern set definition
  end for
  for $k = 2$ to 25 **do** ▷ for each value of k
    for $j = 1$ to 5 **do** ▷ for each TR_j
      $\{\mu_{kj}, \mathbf{C}_{kj}, \theta_{kj}\} := \mathrm{clustering}(k, \Gamma_j^{cl})$ ▷ Application of the clustering algorithm
      $\Phi_{kj} := \{\mu_{kj}, \mathbf{C}_{kj}\}$ ▷ MF evaluation
      $ANFIS_{kj} := \{\Phi_{kj}, \theta_{kj}\}$ ▷ ANFIS synthesis
      simulation of $ANFIS_{kj}$ on $VL_j \rightarrow \bar{P}_{kj}$ ▷ MG simulation on VL_j and OF evaluation
    end for
    selection of $ANFIS_k^{best}$ according with $\bar{P}_{kj}$
    simulation of $ANFIS_k^{best}$ on the TS $\rightarrow P_k$
  end for
end procedure

11.2 Simulation Settings

In this study it has been considered a MG composed by the following energy systems: a PV generator of 19 kW, an aggregated load with a peak power around 8 kW, and ESS with an energy capacity of 24 kWh. For the ESS modelling it has been taken into consideration the Toshiba ESS SCiB module having a rated voltage of 300 V, a current rate of 8 C-Rate and a capacity of about 80 Ah.

The overall dataset is shown in Figure 11.1, together with the energy prices both in sale (positive) and purchase (negative). It is a part of the one shown in Figure 7.2 in Chapter 7, it is restricted to an overall period of 20 days.

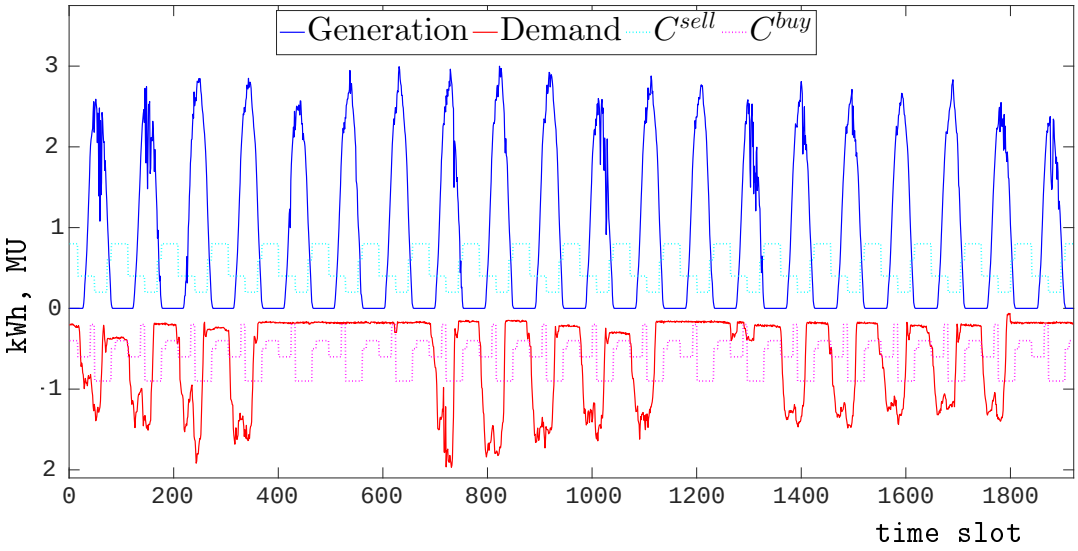


FIGURE 11.1: MG power production and demand of the overall dataset and ToU energy prices.

The even days are assigned to the TS, whereas the odd days are partitioned with a random selection between the TR and VL in order to form 5 different TR-VL partitions (see Sec. 11.1). It has been chosen to assign the 70% of the odd days to the TR and the remaining 30% to the VL.

The ANFIS optimization process is executed for each clustering algorithm introduced in Chapter 4 Sec. 4.1 and by varying the ε (see Sec. 4.2) value, namely the local approximation error influence, as follows:

- k-means for ε equal to 1, 0.75, 0.5 and 0.25.

- k-medians featured by Manhattan distance for ε equal to 1, 0.75, 0.5 and 0.25.
- k-medoids featured by Mahalanobis distance for ε equal to 1, 0.75, 0.5 and 0.25.

The clustering algorithms have been set with a maximum number of iterations equal to 50 and every execution of the algorithm (*i.e.* number of replicates) is repeated 20 times with a new, random, initial representatives selection.
The solution chosen by the clustering algorithm after each run is the one which minimizes its respective OF, namely the WCSS defined in (4.1) for k-means, WCSD for k-medians (4.2) and WCSD (4.3) for k-medoids.

11.3 Results

The simulation results have been studied by computing the OF $\bar{P}$ defined in (7.6) on the TS, whose optimal solution is shown in Figure 11.2.

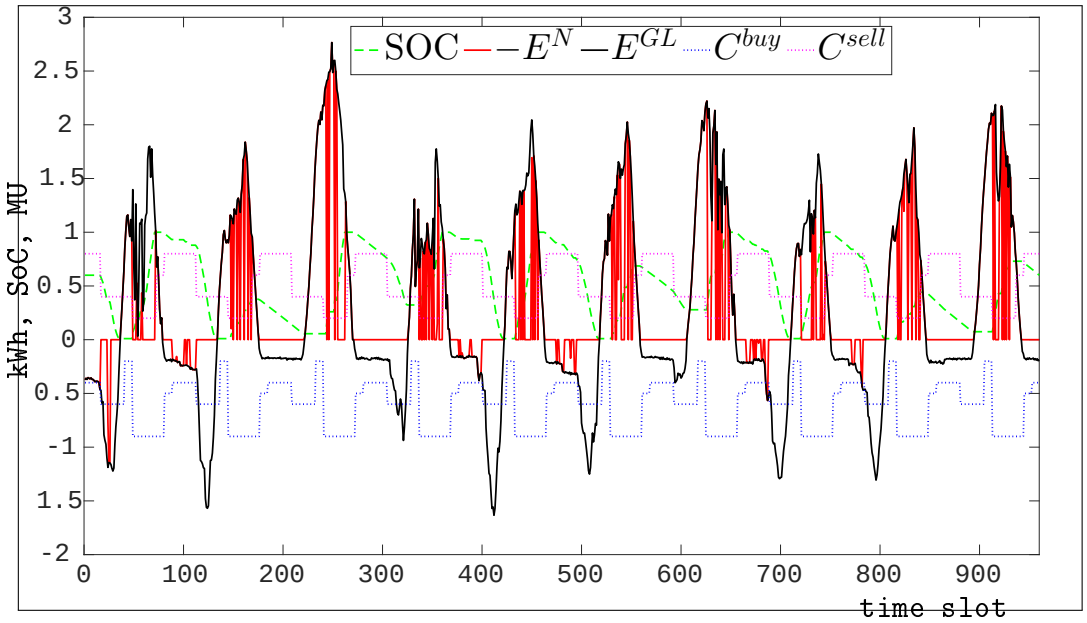


FIGURE 11.2: MG optimum solution energy flows, ESS *SoC* and energy prices computed on the TS by LP formulation.

The results on the TS given by the procedure described in Algorithm 5 are shown in Figure 11.3, as a function of k, by reporting the respective value of $\bar{P}$ for each

$ANFIS_k^{best}$ generated. These are grouped by the different ε coefficient values, which range between 1 and 0.25 as defined in Sec. 11.2. It is possible to observe that k-medoids in most cases presents the best results, being its OF curves (in green) lower with respect to the other two competitors (k-means in red and k-medians in blue). Moreover, for lower values of ε, especially for $\varepsilon = 0.25$ (*i.e.* when the approximation error due to hyperplane prevails (see (4.7)), the k-medoids results appear to be more stable as k increases with respect to the other two algorithms (see Figure 11.3-d). k-means and k-medians show a smoother behaviour of $\bar{P}$, especially for $\varepsilon = 1$ (Figure 11.3-a), meaning that these are more stable when considering the input space only. In all cases their profit results are around 80% of the optimum solution.
In order to focus on the results reliability and sensitivity, the carried out tests have been repeated 10 times. In this study, for the sake of clarity, it has been decided to select only $ANFIS_k^{best}$ associated with the best OF result on the TS. The selected solutions are displayed through a box-plot representation in Figure 11.4 after being grouped by ε coefficient and clustering algorithm. Specifically, in Figure 11.4-a are reported the $ANFIS_k^{best}$ selected solution OF ($\bar{P}$) values; in Figure 11.4-b their respective number of clusters and, finally, in Figure 11.4-c are reported the Davies-Bouldin Index (DBI) values [94], which measures both intra-clusters compactness and inter-clusters separation.

As shown in Figure 11.4-a, best results are confirmed to be prevalently given by k-medoids. These are followed by the k-medians. On the other hand, k-medoids solutions present high level of variability both on the value of k and the DBI, as pointed by higher extension of the boxes (see Figure 11.4-b and c). Nevertheless, for $\varepsilon = 0.5$ the box-plots show a lower variance for $\bar{P}$, k and DBI. In k-means and k-medians solutions, considering the joint input-output space term seems to be scarcely significant. Only for k-means the joint input-output space term in OF yields a relevant improvement in terms of reduction of the DBI and the number of clusters. More into details, as far as k-medoids is concerned, by looking at the DBI it is possible to see

that the clustering problem is hard to solve[1], whereas the k-means and k-medians lead to better solutions. However, in terms of OF (see (7.6)) k-medoids overperforms the other two competitors. This is mainly due to the Mahalanobis distance which, by considering up to the second-order statistics (*i.e.* the covariance matrix), is aware of the clusters' shapes as well. Indeed, by using the Mahalanobis distance k-medoids effectively updates the covariance matrix at each iteration, conversely to the other two clustering-based ANFIS synthesis procedures, where the covariance matrix is evaluated once, at the end of the clustering procedure.

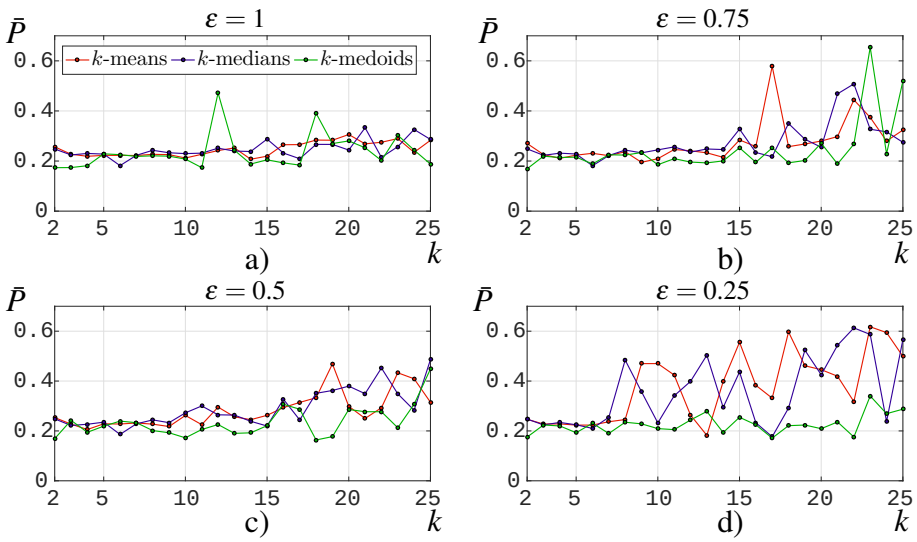


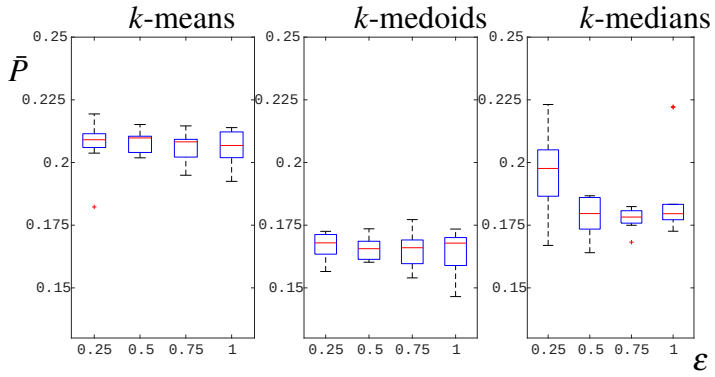
FIGURE 11.3: $ANFIS_k^{best}$ normalized profit values evaluated on TS as a function of k. Solutions are grouped according to ε.

11.4 Conclusion

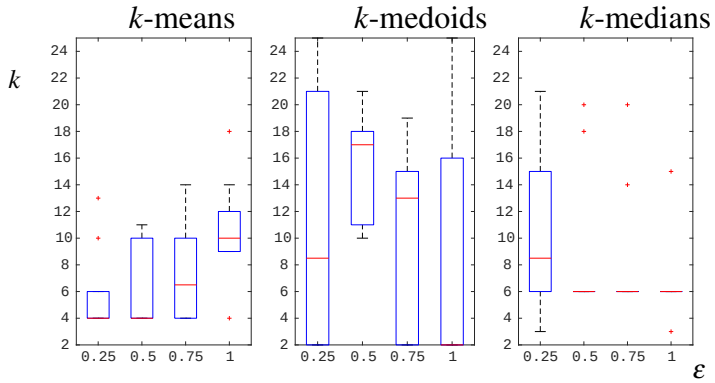
In this chapter it has been investigated a procedure for the synthesis of a MG EMS based on computational intelligence techniques. It performs a cluster analysis in order to find automatically the number and the location on the input space of the rules composing an ANFIS, as the core inference engine of the EMS.

The ANFIS synthesis is cast as a supervised machine learning problem, where patterns are input-output pairs, taking as ground-truth output values the ones coming

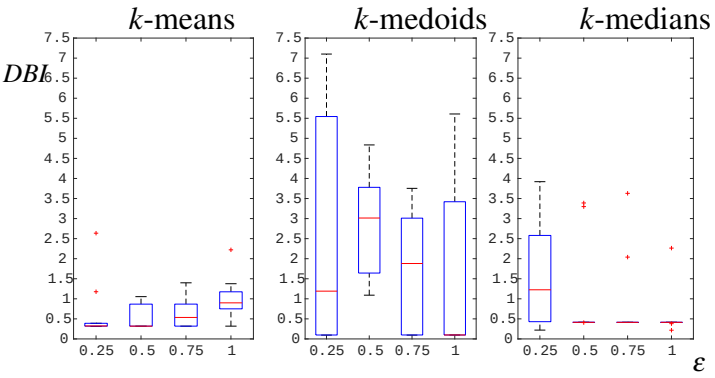
[1]Recall: the lower DBI, the better the partition. Indeed, a low DBI means low intra-variance (high compactness) and high inter-variance (high separation).



a)



b)



c)

FIGURE 11.4: Simulation results on TS considering the best solution of each run illustrated through a box-plot representation. Top and bottom box sides correspond to first and third quantiles, respectively, whereas the red dash corresponds to the median. Top and bottom whiskers extremities correspond to the maximum and minimum points not considered as outliers, respectively. These latter are marked with a red + symbol. In (a) are shown the OF values; in (b) the number of MFs; in (c) the DBI associated to the clustering solution which models each *ANFIS*.

from the benchmark solution obtained by a LP formulation.

In particular, three considered clustering algorithms have been compared, which differ for their respective (dis)similarity measures and the way clusters' representatives are evaluated. Besides, a clustering procedure in the joint input-output space has been investigated, evaluating its effectiveness when the weighting coefficient in the OF changes. Results show that adopting the Mahalanobis distance on the joint input-output space leads to profits way superior than k-means. This improvements, however, are paid with higher values of both the average number of clusters (*i.e.* model complexity) and its variance (*i.e.* algorithm robustness).

On the other hand, the adoption of the k-medians seems a good compromise in terms of both robustness and effectiveness of OF performance and partitions quality in terms of DBI.

Chapter 12

ANFIS EMS Supported by an ESN-Based Prediction Algorithm

In this study the EMS is in charge to define in real time how to rationally redistribute the current energy balance between the ESS and the grid, aiming to maximize the profit function defined in (7.3) assuming a ToU energy price policy. The core of the proposed EMS consists of an ANFIS decision making system synthesised through a clustering algorithm with an overall supervised learning procedure. In particular, the ANFIS takes advantage of the information given by the prediction on the prosumer energy balance made by a suitable ESN based prediction model. Since the EMS modelling relies on the upper and lower benchmark solutions, the OF taken into consideration is the normalized variant of P, *i.e.* $\bar{P}$ in (7.6).

In this chapter it is investigated the effectiveness of the energy balance prediction model in supporting the ANFIS EMS. In particular, the ANFIS synthesis has been repeated by considering different prediction time horizon extensions that ranges from 15 minutes up to 12 hours.

12.0.1 EMS Scheme

In this study, the overall EMS input array is defined by 7 variables. Considering the current k-th time slot, these are the current SoC measured by the BMS SoC_k, the

energy generation and demand at the last time slot E_k^G and E_k^L, the upcoming energy prices both in sale and purchase C_{k+1}^{sell} and C_{k+1}^{buy}. Moreover, the current day d_k, the current hour h_k and finally, a rough estimation of the future daily energy generation and energy demand in the current day $F_k^G = F^G(d_k) =$ and $F_{k+1}^L = F^L(d_k)$, respectively.

In Figure 12.1 it is shown in details the EMS scheme. It is composed by two main blocks, the ensemble of the ESN predictors and the ANFIS decision making system, the core of the EMS.

The ensemble of prediction models is composed by N_T ESN sub blocks equal to the prediction time horizon extension. In the prediction block, the i-th ESN sub block is in charge to compute the prediction about the energy balance that will occur after i time slots with respect to the current one ($\hat{E}_{k+i}^{GL}$). Moreover, each sub block is composed by 2 different and separately optimized ESNs, one dedicated to predict the generation and one the aggregated load. In Figure 12.1 have been distinguished and outlined their respective input arrays drawn in red and blue, respectively.

As shown in Figure 12.1 the prediction input array is: $\{\hat{E}_k^G, \hat{E}_k^L, d_k, h_k, F_k^G, F_k^L\}$. However, the input array of the ESNs dedicated to the generation energy system is restricted to: $\{\hat{E}_k^G, d_k, h_k, F_k^G\}$. Whereas the input array of the ESNs dedicated to the overall energy demand is: $\{\hat{E}_k^L, d_k, h_k, F_k^L\}$.

The prediction model block outputs the energy balance prediction sequence $\hat{\mathbf{E}}_k^{GL} = [\hat{E}_{k+1}^{GL}, ..., \hat{E}_{k+N_T}^{GL}]$ of length N_T. It is fitted by a LMSELR for the calculation of the two line coefficient, m_{k+1} and q_{k+1}. The LMSELR operator has the task to filter and summarize in two values the information given by the predictor (*i.e.* $\hat{\mathbf{E}}_k^{GL}$) in order to give to the ANFIS an estimation of the future prosumer energy trend.

Therefore, m_{k+1}, q_{k+1}, $\hat{E}_{k+1}^{GL}$ (*i.e.* the ESN_1 output), together with the current ESS SoC and the energy prices C_{k+1}^{sell} and C_{k+1}^{buy} forms the input data structure feeding the ANFIS, that is in charge of computing the energy to exchange with the grid.

It is important to underline that the adoption of this strategy allows to limit the ANFIS number of inputs to 6, regardless of the prediction time horizon (*i.e.* N_T).

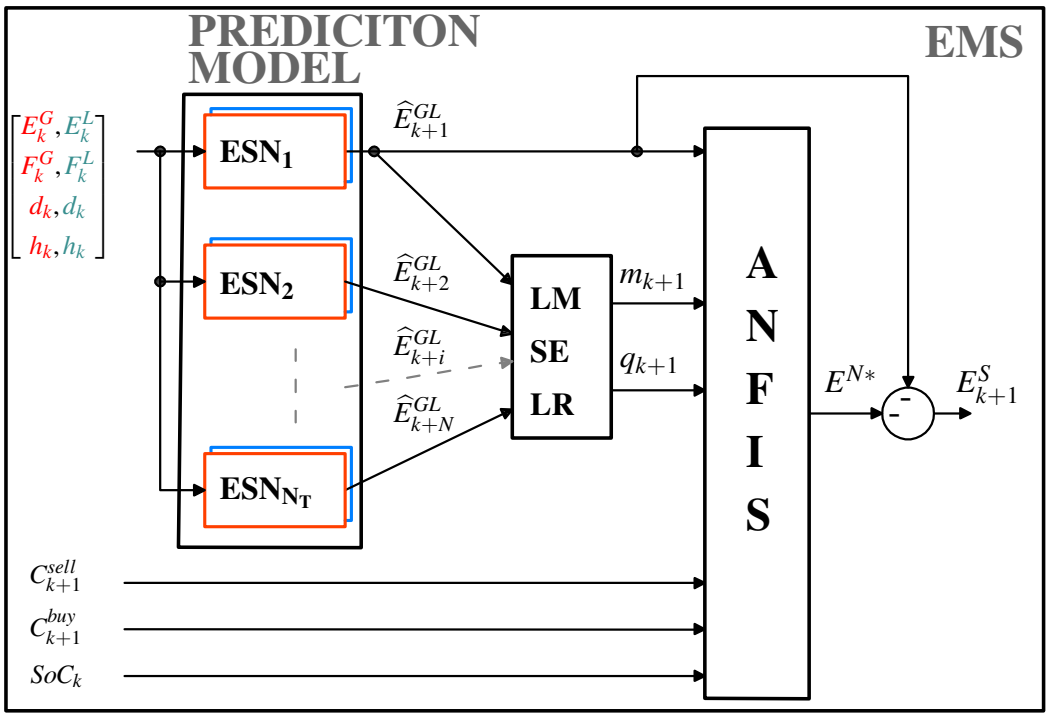


FIGURE 12.1: EMS model.

12.1 The EMS Synthesis Procedure

The overall ANFIS synthesis procedure, including the ESN based prediction model, is summarized in Alg. 5.

First of all, the dataset used is subject to multiple splits, following a *k-fold* cross validation is applied in order to create more TR-VL pairs, besides the generation of the TS.

Considering a TR-VL pair, it is evaluated the respective reference solution through a LP problem formulation on the TR. It is used for the estimation of SoC^{opt} and $E^{N,opt}$ (see Chapter 7.4 for further details). Following, given a time horizon N_T, it is defined the prediction model, named PM_{N_T} in Alg. 5. During the EMS synthesis procedure, N_T has been varied from 1 to $N_{T,max}$. In Alg. 5 it is illustrated that whenever N_T is updated, the new prediction model only requires to optimize those ESNs not yet synthesised on that given TR-VL pair by the GA. Therefore, it is possible to formulate the training pattern set $X^{TR} = \{\hat{\mathbf{E}}_k^{GL}(1); C^{sell}; C^{buy}; SoC^{opt}; m; q\}$ to feed the k-means clustering procedure and its corresponding label set $Y^{TR} = \{E^{N,opt}\}$. Successively,

the generated clusters are used to define the ANFIS MFs and the corresponding RC hyperplanes (*i.e.* ϕ and θ, respectively), since every pattern of X^{TR} is labelled by Y^{TR}. The clusters characterized by cardinalities equal or lower than the ANFIS input space dimension (*i.e.* 6) are discarded because of the impossibility to unambiguously define the associated RC hyperplane. Once the ANFIS is trained, it is performed together with its assigned prediction model on the MG system (see MG_{sim} function in Alg. 5) for the evaluation of $\bar{P}$ (see (7.6)). It is clear that the evaluation of $\bar{P}$ on the VL (and the TS) is carried out through the evaluation of the reference solutions (*i.e.* P^{opt} and $P^{w/o-ESS}$).

The ANFIS synthesis procedure by k-means clustering has been repeated by ranging K from 2 to 50. By employing each TR-VL pair to the same procedure, for every N_T it is selected the ANFIS (and its prediction model) which produces the lowest $\bar{P}$ on its corresponding VL. Those selected are then performed on the TS in order to examine the EMS efficiency by varying the prediction time horizon.

12.2 Simulation Settings

In this study it has been considered a MG composed by the following energy systems: a PV generator of 19 kW, an aggregated load with a peak of power of 8 kW, and an ESS with an energy capacity of 24 kWh. For the ESS modelling it has been taken into consideration the Toshiba ESS SCiB module. It has a rated voltage of 300 [V], a current rate of 8 [C-Rate], a Capacity of about 80 [Ah]. The BMS ensures that the ESS SoC is always between $0.15 - 1$ [p.u.].

The dataset used is given by the first 60 days of the dataset in Figure 7.2 in Chapter 7. The daily ToU energy price series are those shown in Figure 7.3.

The overall dataset is subjected to a *k-fold* cross validation which defines 3 TR-VL pairs and an unique TS for a total of 4 subsets of the same dimension (*i.e.* 15 days, which correspond to the 25% of the dataset). The TS is defined by a subset that comprehends all the days multiple of 4 in order to give it an uniform distribution. The

Algorithm 4 EMS Training Procedure

procedure EMS DESIGN
  TS & 3 TR-VL pairs definition
  for $p = 1$ to 3 **do** ▷ for each TR-VL pair
    $E^{N,opt(p)}, SoC^{opt(p)} := \mathrm{LP}(TR^p)$ ▷ bench.sol.
    $ESN_1^p = GA(VL^p, TR^p, 1)$ ▷ generate ESN_1
    $PM_1^p = ESN_1^p$ ▷ set the prediction model
    for $N_T = 2$ to $N_{T,max}$ **do** ▷ set the prediction T.H.
      $ESN_{N_T}^p = GA(VL^p, TR^p, N_T)$ ▷ generate ESN_{N_T}
      $PM_{N_T}^p = \left[PM_{(N_T-1)}^p, {ESN_{N_T}}^p\right]$ ▷ update PM_{N_T}
    end for
  end for
  for $N_T = 1$ to $N_{T,max}$ **do** ▷ set the prediction T.H.
    for $p = 1$ to 3 **do** ▷ for each TR-VL pair
      $\left\{\hat{\mathbf{E}}^{GL}, m, q\right\}^p = f(TR^p, PM_{N_T}^p)$
      $X^{TR,p} = \left\{\hat{\mathbf{E}}^{GL}(1); C^{sell}; C^{buy}; SoC^{opt}; m; q\right\}^p$
      $Y^{TR,p} = \left\{E^{N,opt}\right\}^p$ ▷ pattern & label set
      for $K = 2$ to 50 **do** ▷ set the num. of clusters
        $\{\mu, \mathbf{C}, \mathbf{\Theta}\} = k\text{-means}\ (K, X^{TR,p}, Y^{TR,p})$
        $\Phi := \{\mu, \mathbf{C}\}$ ▷ MF definition
        $ANFIS^{K,p} := \{\Phi, \mathbf{\Theta}\}$ ▷ ANFIS definition
        $\bar{P}^{K,p} = MG_{sim}(ANFIS^{Kp}, PM_{N_T}^p, VL^p)$
      end for
    end for
    $EMS^{N_T} = \{ANFIS^{K^*,p^*}, PM_{N_T}^{p^*} | K^*, p^* = \arg\min_{K,p}\{\bar{P}^{K,p}\}\}$
  end for
end procedure

rest of the days have been randomly assigned to the other 3 subsets. Every generated TR is composed by 2 subsets (*i.e.* 30 days, the 50% of the overall dataset) unlike the VL that has been composed by only one subset.

Every k-means clustering algorithm execution is repeated 20 times (*i.e.* number of replicates). The initialization involves a random selection of the centroids. Each replicate is characterized by a maximum number of iterations equal to 1000. However, in case of the WCSS (see. (4.1) in Sec. 4.1 of Chapter 4) does not decrease by 0.001%, the execution is arrested in advance.

The GA settings for the ESNs synthesis are illustrated in Table 12.1, where are defined the ranges of each ESN optimization parameter (*i.e.* gene) introduced in Sec 5.1. The GA initial population is randomly generated and the evolution process, beside considering a number of maximum generation equal to 60, is arrested in advance if the best individual fitness does not improve by $\gamma_{conv} = 0.1\%$ for at least $N_{conv} = 10$ generations. The rest of the GA parameters are set as in Chapter 10.

Different ANFIS EMSs have been modelled in this study by adopting the procedure described in Sec. 12.1 by increasing the prediction time horizon. The EMSs designed are studied and compared by evaluating $\bar{P}$ on the TS in order to observe and quantify the improvements given by the prediction.

TABLE 12.1: GA main parameters setting for the ESN design.

GA Parameter	**Value**	**GA Variable**	**Lower Bound-Upper Bounds**
Max. generations N_{stop}	60	λ	0.2-1.3
Population N_{ind}	40 individuals	N	20-1000
Elite N_{elite}	4 individuals	ρ	0.02-0.5
Cross-over Γ_{cross} type	scattered	α	0-0.8
Cross-over fraction N_{cross}	$0.6 \cdot N_{ind}$	$\bar{\psi}$	0.01-1
Selection type Γ_{select}	tournament (2 indiv.)		

12.3 Results

The results have been studied considering different time horizons for the prediction model by ranging N_T between 1 (*i.e.* 15 mins.) and 46 (*i.e.* 12 hours). It is clear that for $N_T = 1$ the prediction on just E^{GL}_{k+1} excludes the m and q evaluation. Since for $N_T = 2$ $\hat{E^{GL}}_{k+1} \equiv q$ it is necessary only the estimation of α. The prediction accuracy of each synthesised ESN on the TS is reported in Figure 12.2 and Figure 12.3 for the energy production and the energy demand, respectively. The plots show the Root Mean Square Error (RMSE) of each ESN according to their own assigned time slot to predict.

In Figure 12.4 are shown the EMSs results in function of the prediction model time horizon. The results show that the best EMSs (*i.e.* with the higher $\bar{P}$ on the TS) are generated for the highest N_T. In fact, for $N_T \leq 33$, except for just few solutions, $\bar{P}$ is between about $0.19 - 0.16$ whereas, for $N_T \geq 37$ (*i.e.* more than 9 hours), $\bar{P}$ ranges between $0.12 - 0.13$.

In Figure 12.5 are plotted the number of rules of each ANFIS of the EMS selected in function of its respective N_T. Looking at the figure, it can be observed that those solutions featured by a high number of rules (*i.e.* > 35) often have the best performances compared to those next to them (see both Figure 12.5 and 12.4 for $N_T = 20, 24, 25, 28$).

In Figure 12.6 it is illustrated the TS energy profiles together with the the ESS SoC and the EMS output of the best solution found for $N_T = 43$. The absence of any overlap of the EMS output and the E^N profiles (*i.e.* the possibility that the amount of energy set to be exchanged by the EMS and that one effectively exchanged) emphasizes the difficulties of the EMS in managing the ESS when it reaches its SoC limits, that in any case are monitored and respected thanks to the ESS BMS. This point can be explained by the fact that in the formulation of the optimum reference solution it has not been considered any penalty cost function when the ESS SoC is close to its limits. This is proven in Figure 12.7 where it is plot the optimum solution on the TS,

showing that the ESS is often completely charged or discharged.

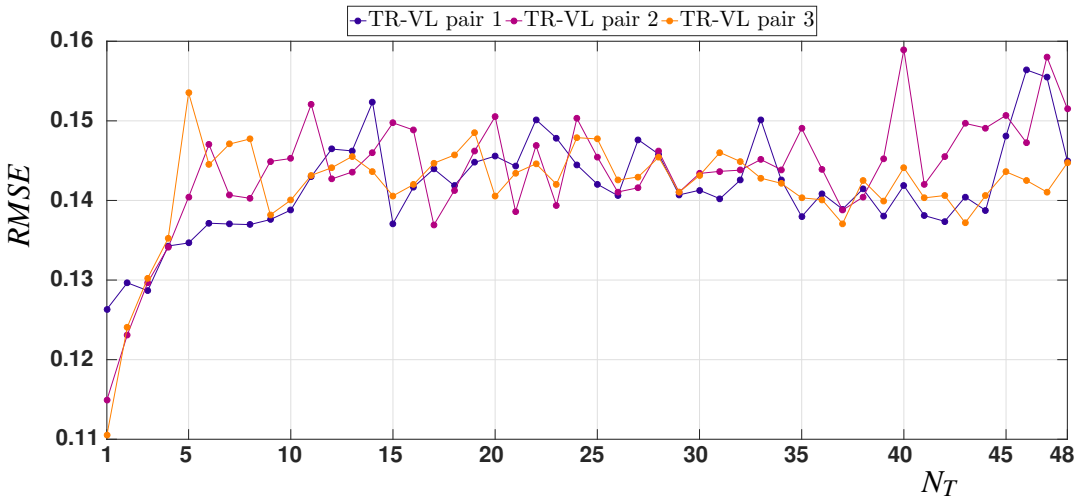


FIGURE 12.2: ESNs RMSE energy generation prediction on the TS for each TR-VL pair.

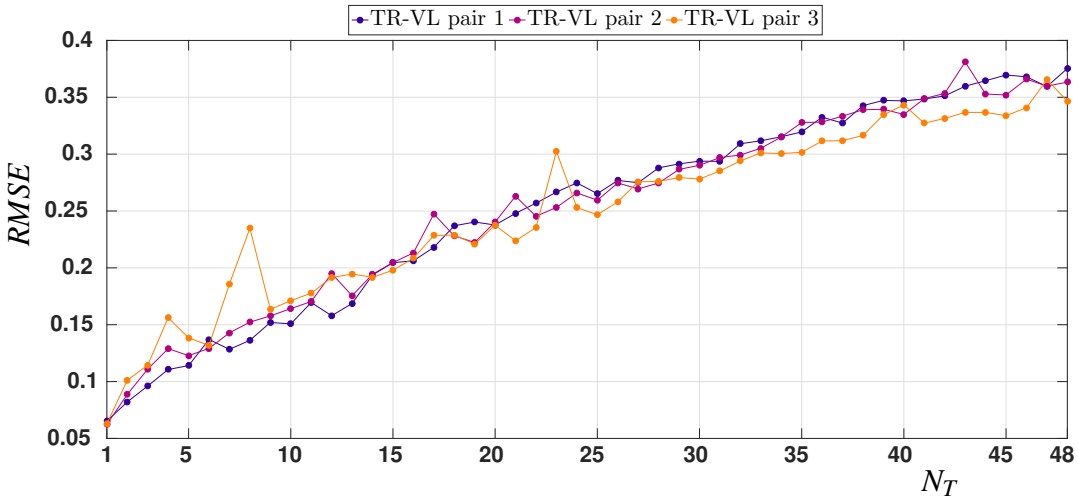


FIGURE 12.3: ESNs energy demand prediction RMSE on the TS for each TR-VL pair as function of N_T.

12.4 Conclusions

In this chapter it is investigated the effectiveness in considering a prosumer energy balance prediction model to support an ANFIS residential MG EMS. The EMS has the aim to redistribute effectively the current energy balance in real time in order to maximize the profit generated by the energy exchanged with the grid. The prediction model proposed is composed by several MISO ESNs, every one dedicated to predict

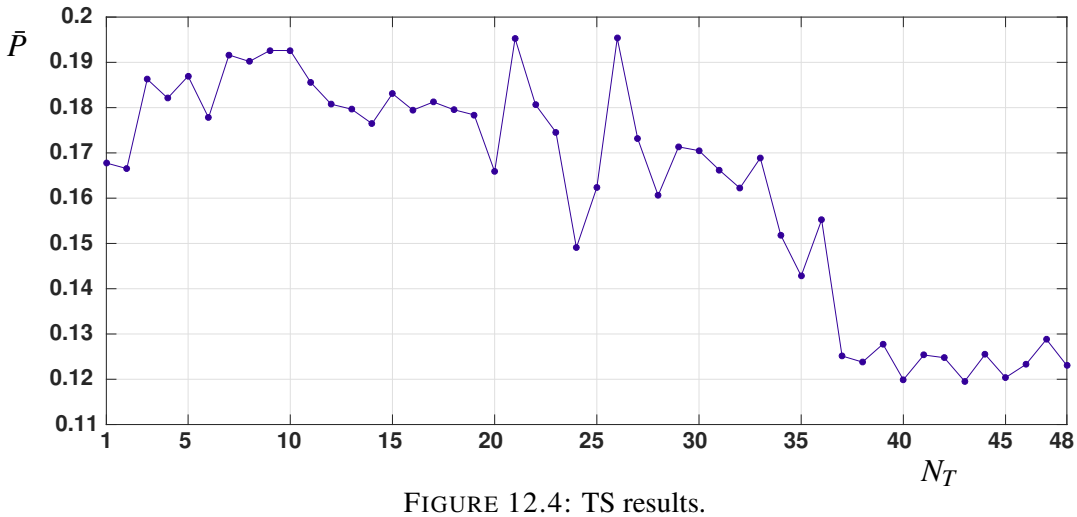


FIGURE 12.4: TS results.

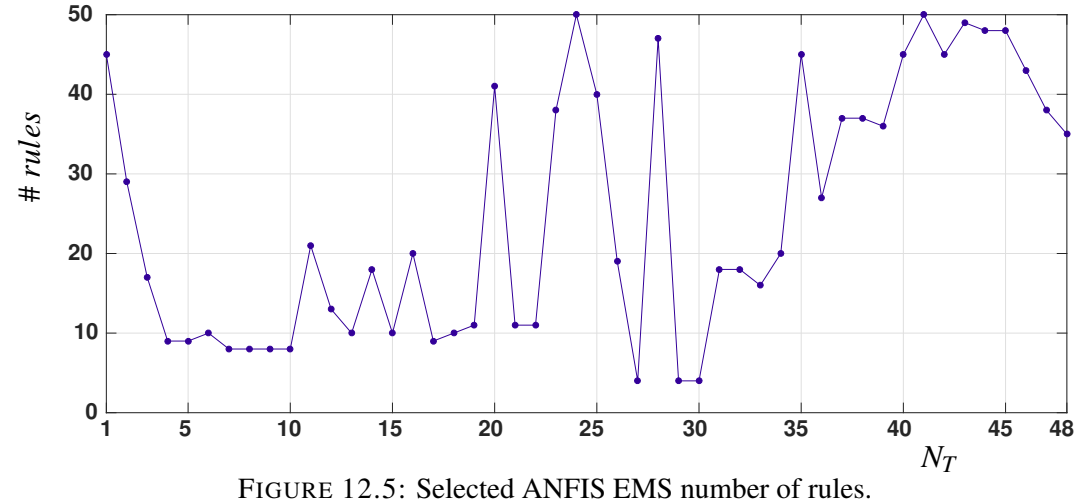


FIGURE 12.5: Selected ANFIS EMS number of rules.

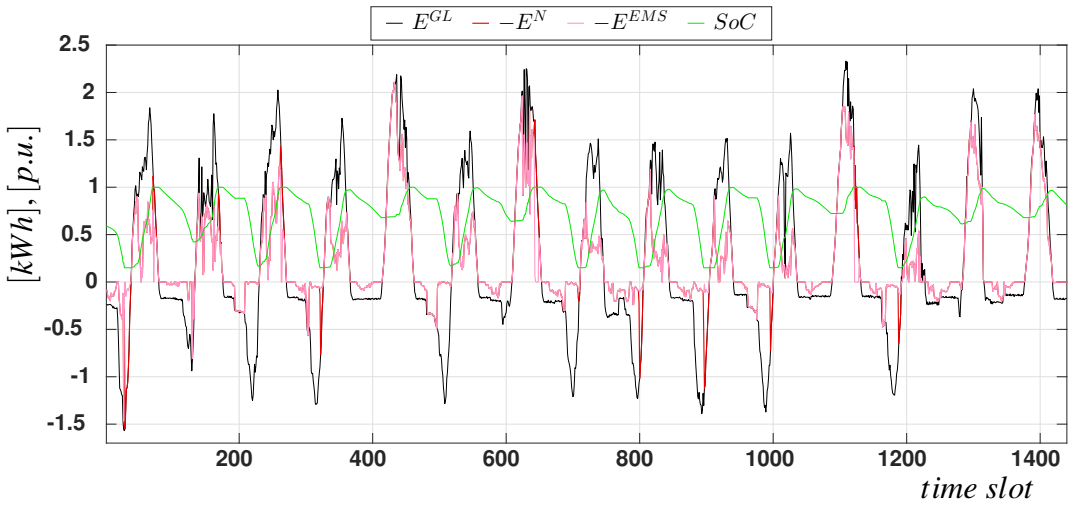


FIGURE 12.6: TS selected solution energy profiles, the EMS output and the ESS SoC profile. The considered solution is supported with a prediction model with a time horizon $N_t = 43$.

the energy produced (or demanded) by the prosumer a given number of time slots after the current one. The ESNs are separately optimized through a GA. Instead, the

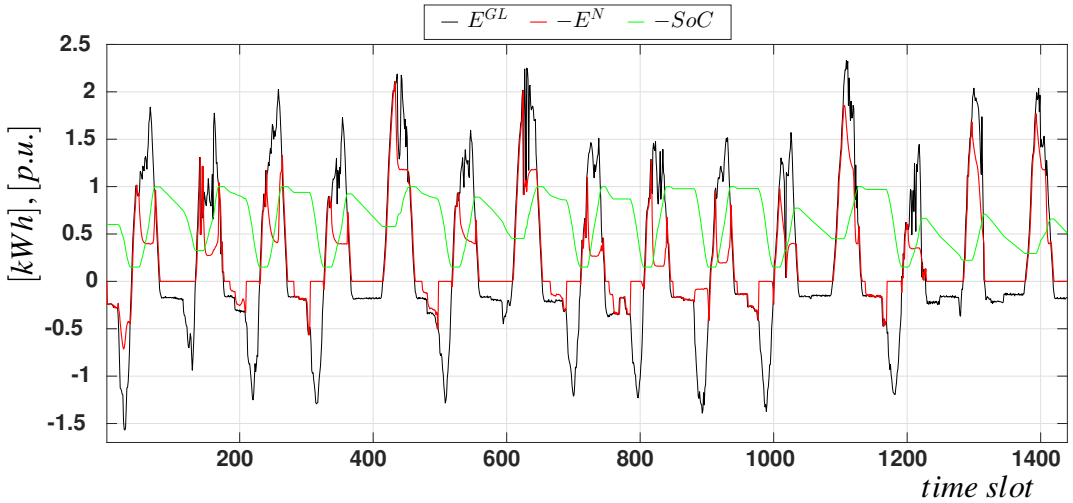


FIGURE 12.7: Optimum TS energy profiles and the ESS SoC profile.

ANFIS is synthesised by means of a k-means clustering algorithm taking advantage of a reference solution found with a LP problem formulation. The results obtained shows that the ANFIS EMS performances increases for prediction time horizons greater than ten hours. The EMS reaches interesting results, close to the optimal benchmarking solution (by $86-88\%$), with a performance increase around 30% with respect to those solutions characterized by shorter prediction time horizons (*i.e.* less than 4 hours). Obviously, the performance improvements are paid in terms of a greater structural complexity (number of rules) of the ANFIS.

Chapter 13

ANFIS EMS by Hyperplane Clustering

This chapter deals with the synthesis of the ANFIS model (see Sec. 2.3.2 of Chapter 2) synthesised by clustering techniques and Min-Max classifier. The ANFIS model is supposed to be implemented on a MG EMS that is in charge to efficiently manage the prosumer energy balance in order to maximize the profit generated by the energy trading with the grid by assuming a ToU energy price policy. The prosumer is supposed to be equipped with a PV plant and an ESS capable of both storing and delivering energy from/to the connected grid as formulated in Chapter 7.

The implementation of clustering techniques for the ANFIS sinthesis implies the evaluation of the benchmark solutions as explained in Chapter 7.4 and consider the normalized function $\bar{P}$ (7.6) of the profit function P (7.3).

In Figure 14.1 it is shown the ANFIS EMS architecture that outputs the energy to be exchanged with the ESS in the next time slot. As described in figure, the ANFIS input array is defined by the current energy balance given by the PV generation and the prosumer energy demand (see (7.1)), the current ESS SoC and the energy prices in sale and purchase.

In this chapter, is investigated an ANFIS EMS modelled in a fully data driven manner using the joint contribution of hyperplane clustering and Neurofuzzy Min-Max classifier (see Chapter 4 Sec. 4.3). For the first clustering stage, all the three k-means-based hyperplane clustering variants introduced in Sec. 4.4 are implemented

(*i.e.* partitional, divisive and agglomerative). After the clustering stage, a (Pruning) Adaptive Resolution Min-Max Classifier [84] is responsible of properly refining the ANFIS MFs in order to avoid ineffective overlapping. The three clustering variants are compared in terms of effectiveness and efficiency together with the benefits of applying the classification stage after the clustering stage.

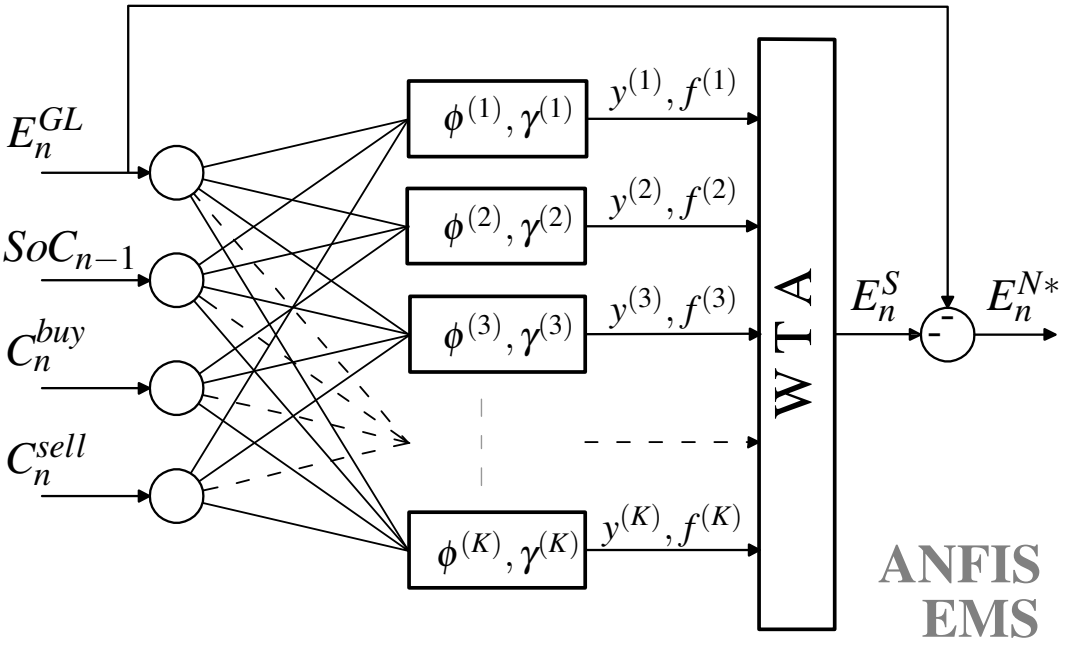


FIGURE 13.1: ANFIS EMS scheme.

13.1 The EMS Synthesis Procedure

According to the ANFIS synthesis procedure by clustering techniques described in Chapter 4, in Alg. 5 is summarized the ANFIS based EMS synthesis procedure for a given hyperplane clustering algorithm. It is set by considering one of the three variants introduced in Sec. 4.1 of Chapter 4, and the PARC classifier described in Sec. 4.3 set with a specific value of λ (see (4.9)). The synthesis procedure starts by splitting the full dataset into three non-overlapping subsets, namely TR, VL and TS. In particular, 3 TR-VL pairs are generated in a 3-*fold* cross-validation fashion. The synthesis procedure ends with the selection of the ANFIS EMS deemed as the best one, namely the one which presents the minimum value of $\bar{P}$ on its VL, followed by

TS performance evaluation.

Considering the p-th TR-VL pair, namely TR_p and VL_p, for each of the two sets, the respective benchmark solutions are evaluated through a MILP problem formulation. As regards TR_p, `bmrk1()` outputs only the optimum profiles $SoC_{TR_p}^{opt}$ and $E_{TR_p}^{N,opt}$. Conversely, for VL_p, `bmrk2()` outputs the corresponding profit values $P_{VL_p}^{opt}$, $P_{VL_p}^{w/o}$. The latter two are used for evaluating (7.6) on the validation phase, thus for the selection of the best solution. The evaluation of the $SoC_{TR_p}^{opt}$ and $E_{TR_p}^{N,opt}$ on the TR accomplishes for the definition of the training pattern set $\mathbf{X}_{TR_p} = \left[E_{TR_p}^{GL}\ C_{TR_p}^{sell}\ C_{TR_p}^{buy}\ SoC_{TR_p}^{opt}\right]$ labelled with $\mathbf{y}_{TR_p} = \left[E_{TR_p}^{N,opt}\right]$. Looking at Alg. 5, after `bmrk2()`, the labelled pattern set $\langle \mathbf{X}_{TR_p}; \mathbf{y}_{TR_p} \rangle$ is fed to the clustering algorithm that produces $k_{\max} - 1$ distinct partitions. Specifically, for partitional clustering, the algorithm must be executed independently $k_{max} - 1$ times by changing the number of clusters $k \in [2, k_{\max}]$. Conversely, for hierarchical clustering algorithms, $k-1$ solutions are automatically returned due to the layer-wise interpretation of the resulting dendrogram, where each layer is the set of leaf nodes after each iteration, namely after each merging (splitting) for the agglomerative (divisive) variant. Regardless of how clustering solutions are built, each solution is defined by a number of k complementary subsets of the form $\langle \mathbf{X}_i, \mathbf{y}_i \rangle$ such that $\bigcup_{i=1}^{k} \mathbf{X}_i = \mathbf{X}_{TR_p}$ and $\bigcup_{i=1}^{k} \mathbf{y}_i = \mathbf{y}_{TR_p}$.

Every solution generated is then subject to the classifier which possibly splits the clusters in N_{HB} hyperboxes, thus rising the total number of final MFs. It is worth remarking that hyperboxes characterized by a number of patterns less or equal to the input dimensionality ($m = 4$ in this case, see Figure 14.1) are discarded since it is impossible to unambiguously define the associated hyperplane. Successively, each ANFIS is performed on the MG exploiting its corresponding VL. The simulation function `sim()` in Alg. 5 returns the $\bar{P}$ value. The ANFIS characterized by the lowest $\bar{P}$ on the VL (EMS^* in Alg. 5) is then selected as the best one and is finally simulated on the TS to report the EMS performance synthesized by a given clustering algorithm and λ value.

Algorithm 5 EMS Training Procedure

procedure EMS DESIGN

  Clustering algorithm selection

  PARC classification λ parameter setting

  TS & 3 TR-VL pairs definition: $\{\mathbf{TR};\mathbf{VL}\}_{p=1}^{3}$

  for $p = 1$ to 3 **do** ▷ for each TR-VL pair

    $\left\{E_{TR_p}^{N,opt}, SoC_{TR_p}^{opt}\right\} = bmrk1(\mathbf{TR}_p)$ ▷ optimum profiles

    $\left\{P_{VL_p}^{opt}, P_{VL_p}^{w/o}\right\} = bmrk2(\mathbf{VL}_p)$ ▷ benchmark performance evaluation

    $\left\{\mathbf{X}_{TR_p};\mathbf{Y}_{TR_p}\right\} =$

      $= \left\{E_{TR_p}^{GL} C_{TR_p}^{sell}; C_{TR_p}^{buy}; SoC_{TR_p}^{opt}; E_{TR_p}^{N,opt}\right\}$

    $\left\{\{\mathbf{X}_i;\mathbf{y}_i\}_{i=1}^{k}\right\}_{k=2}^{k_{max}} = clustering(k_{\max}, \mathbf{X}_{TR_p}, \mathbf{Y}_{TR_p})$

    for $k = 2$ to k_{max} **do** ▷ select the clustering layer

      $\{\mathbf{X}_i;\mathbf{y}_i\}_{i=1}^{N_{HB}} = classifier(\{\mathbf{X}_i;\mathbf{y}_i\}_{i=1}^{k})$

      $EMS_{p,k} = EMS\left(\{\mathbf{X}_i;\mathbf{y}_i\}_{i=1}^{N_{HB}}\right)$ ▷ EMS def.

      $\bar{P}_{p,l} = sim(EMS_{p,k}, \mathbf{VL}_p, P_{VL_p}^{opt}, P_{VL_p}^{w/o})$ ▷ VL $\bar{P}$

    end for

  end for

  $EMS^* = \{EMS_{p^*,k^*} | p^*, k^* = \arg\min_{p,k}\{\bar{P}_{p,k}\}\}$ ▷ EMS selection

  $\left\{P_{TS}^{opt}, P_{TS}^{w/o}\right\} = bmrk2(\mathbf{TS})$ ▷ benchmark performance

  $\bar{P} = sim(EMS^*, \mathbf{TS}, P_{TS}^{opt}, P_{TS}^{w/o})$ ▷ EMS performance evaluation on the TS

end procedure

13.2 Simulation Settings

The ANFIS EMS is synthesized by considering a prosumer grid-connected MG equipped with a PV generator of 19 $[kW]$, an aggregated load with a peak of power of 8 $[kW]$, and an ESS with an energy capacity of 24 $[kWh]$. For the ESS modelling it has been taken into consideration the Toshiba ESS SCiB module. It has a rated voltage of 300 $[V]$, a current rate of 8 $[C-Rate]$, a capacity of about 80 $[Ah]$. The BMS is in charge to limit and control the ESS SoC between $0.15-1$ $[p.u.]$ by correcting the EMS output E^{N*} with the aim to always respect the prosumer energy balance.
As in Chapter 12, the dataset covers an overall period of 60 days sampled with a 15 minutes time period. It is defined by the first 60 days of the dataset shown in Figure 7.2 in Chapter 7. Moreover, the daily ToU energy prices series are those shown in Figure 7.3 in the same chapter.

The dataset partitioning is subject to a 3-*fold* cross validation which produces 3 TR-VL pairs and a unique TS for a total of 4 subsets. These are characterized by the same length (*i.e.*, 15 days, namely 25% of the overall dataset). In particular, the TS is defined by a subset that comprehends all the days multiple of 4 in order to give it an uniform distribution. The rest of the days have been randomly assigned to the other 3 subsets. For each TR-VL pair 2 of the 3 generated subsets are dedicated to the TR definition (*i.e.*, 30 days, namely 50% of the overall dataset) whereas the other subset is assigned to the VL.

Several simulations are set in order to study the effectiveness of the proposed EMS ANFIS synthesis by hyperplane clustering, in particular by considering the three different k-means-based clustering algorithms in Sec. 4.1.
All the clustering variants are set in order to return clustering solutions that range from 2 to $k_{\max} = 50$ clusters. Thus, the agglomerative variant exploits the partitional variant for building the bottom layer of $k = k_{\max} = 50$ and subsequent layers are built by merging; the divisive variant starts from the entire dataset and iteratively runs 2-means on the worst leaf node in each layer until a layer with $k = k_{\max} = 50$ clusters

is obtained. For the sake of comparison, $k_{\max} - 1$ independent runs of the partitional variant have been performed, by varying $k \in [2, 50]$. The k-means has been set up for all variants as follows: every k-means execution has been repeated 20 times since the algorithm initialization has been performed randomly; each execution is arrested by a maximum number of iterations equal to 1000 with an early-stop criterion in case points-to-clusters assignments stop changing.

The clustering algorithms were separately applied in the procedure introduced in Sec. 13.1 by considering in (4.7) $\varepsilon = 0$ (*i.e.* hyperplane clustering) for the ANFIS EMS synthesis. For a thorough investigation, every procedure has been repeated by ranging the classifier complexity parameter λ in $[0, 1)$.

In the next Section, these results will also be compared with the very same procedure but without considering the PARC classifier, in order to address the benefits of applying such classification system right after the clustering phase in real-world scenarios. Moreover, once determined a suitable value of λ for which the three clustering algorithms, along with PARC, lead to better ANFIS performances, the joint input-hyperplanes space will be investigated; that is, whilst λ will be kept constant to that value, ε will range in $(0, 1]$.

13.3 Computational Results

At the end of the previous Section, the two-fold tests suite has been introduced. Aim of the first tests suite is to discuss the results obtained by the described EMS synthesis procedure, along with a comparison of the procedure itself, but without considering the PARC classifier in the synthesis chain.

Figure 13.2 shows the $\bar{P}$ results on the TS of the EMS modelled as function of the classifier λ parameter (see (4.9)) and grouped by clustering algorithm (namely, partitional, divisive and agglomerative). Instead, Figure 13.3 and 13.4 regard the EMS complexity, namely the number of fuzzy rules (Figure 13.3) and the number of clusters (Figure 13.4) on the top of which said rules have been built. Figure 13.2, 13.3

and 13.4 are organized as follows: the results sketched in the rightmost plot have been obtained by considering the PARC classifier after the clustering phase and, for a thorough investigation, several values for λ have been considered. Conversely, the results sketched in the leftmost plot have been obtained by removing the PARC classifier after the clustering phase. Filled dots correspond to the average value among five runs (due to intrinsic randomness in the overall procedure), whiskers indicate the standard deviation, top and bottom arrows indicate the maximum and minimum value, respectively.

By looking at Figure 13.2 it is possible to see that employing a classifier after the clustering phase leads to improvements in terms of profit optimization. Specifically, without the PARC classifier the gap between the benchmark solution and the ANFIS solution is in range $20-25\%$, whereas with the PARC classifier in most of the cases (regardless of the specific algorithm adopted) said gap is in range $10-15\%$, with the divisive clustering algorithm reaching the best result (8% for $\lambda = 0.4$). As λ increases, a slight decay is observed. In turn, especially for $\lambda \geq 0.6$ the resulting ANFISs are featured by a lower complexity (*i.e.*, low number of fuzzy rules – or hyperplanes), as reported in Figure 13.3. These results are perfectly coherent with the PARC classifier behaviour (see Sec. 4.3).

Whilst in terms of profit a clear winner among the three clustering algorithms does not emerge, from Figure 13.3 it is possible to see that the agglomerative variant, for almost every λ, leads to an EMS with lower complexity.

The performance boost obtained by considering the PARC classifier, however, have to be paid with a generally higher number of fuzzy rules. This phenomenon is clear by comparing Figure 13.3 with Figure 13.4. If one does not consider the PARC classifier, the ANFIS is synthesized by considering the clustering results only, therefore there will be as many rules as there are clusters (rightmost plots). The PARC classifier, in its strive of designing non-overlapping hyperboxes, will return a number of hyperboxes generally higher than the number of clusters (leftmost plots). This behaviour can also be seen in the toy problem explained in Sec. 4.3 of Chapter 4:

whilst the number of returned clusters is 3, the number of resulting hyperboxes after the PARC execution is 4.

Another aspect which should be considered for comparison is the computational burden (*i.e.*, running time) needed to synthesize the EMS among the three clustering algorithms. The software has been implemented in MATLAB® R2018a on a workstation equipped with two Intel® Xeon® 6-cores CPUs at 2.40GHz. Table 13.1 shows the running times for training and performing a single decision, in the upper and lower parts, respectively. It confirms that adding the PARC classifier increases the computational complexity in terms of running time for synthesizing the EMS. Among the three clustering algorithms, the two hierarchical variants drastically overperform the partitional variant, with the divisive variant being the fastest. The lower part of Table 13.1 highlights the time that the EMS spends in order to take one decision. The decision running times slightly increase by adding the PARC classifier. This result is coherent with Figure 13.3 and 13.6: PARC might increase the number of rules, hence it might increase the ANFIS complexity. Indeed, the computational complexity for taking one decision goes as $\mathcal{O}(mr)$, where m is the input dimensionality and r is the number of rules.

TABLE 13.1: Running Times for EMS Synthesis and Real-Time Decision.

Running Time for the EMS Synthesis [in min.]			
	Partitional	Divisive	Agglomerative
with PARC	237	37	50
w/o PARC	202	5	10
Running Time for the EMS Decision [in sec.]			
	Partitional	Divisive	Agglomerative
with PARC	0.006	0.006	0.006
w/o PARC	0.004	0.005	0.004

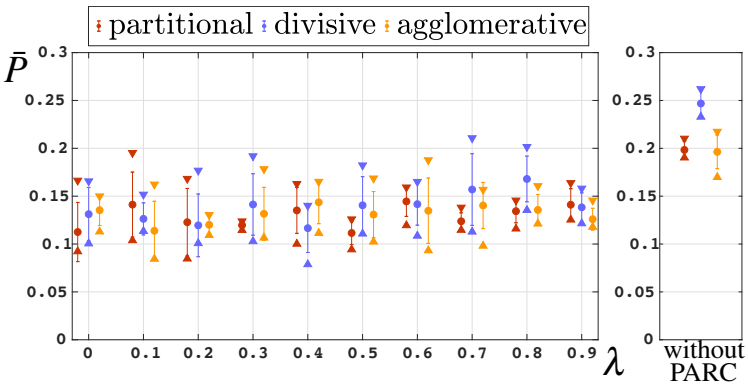


FIGURE 13.2: $\bar{P}$ results on the Test Set ($\varepsilon = 0$).

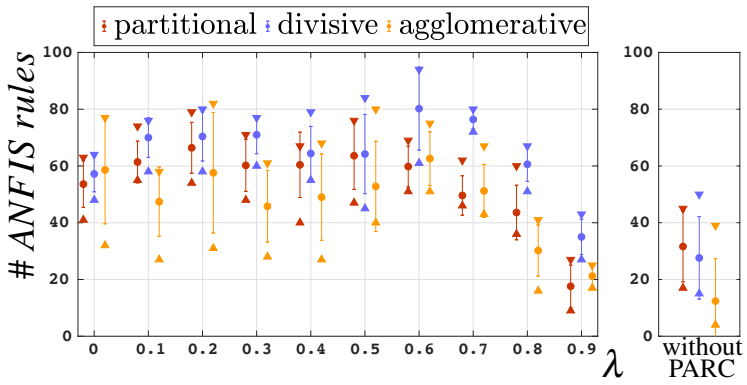


FIGURE 13.3: ANFIS number of rules for each solution found ($\varepsilon = 0$).

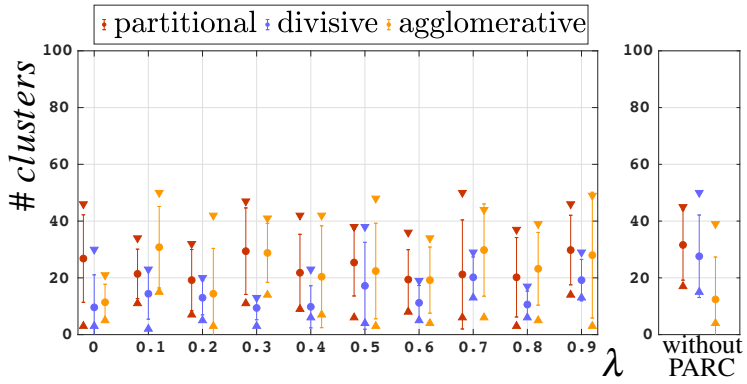


FIGURE 13.4: Number of clusters ($\varepsilon = 0$) before PARC step.

The second tests suite, conversely, aims at addressing the joint input-hyperplanes space. After selecting $\lambda = 0.4$, the weighting term ε has been varied in range $[0.2, 1]$ with step size 0.2. Figure 13.5, 13.6 and 13.7 are the joint input-hyperplanes space counterparts of Figure 13.2, 13.3 and 13.4, respectively.

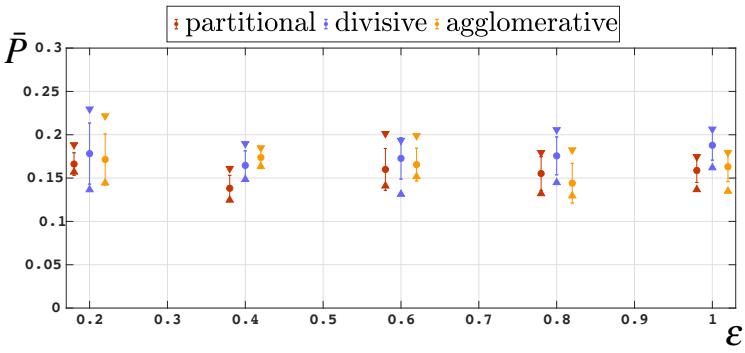


FIGURE 13.5: $\bar{P}$ results on the Test Set as function of the input-hyperplanes space ε weight ($\lambda = 0.4$).

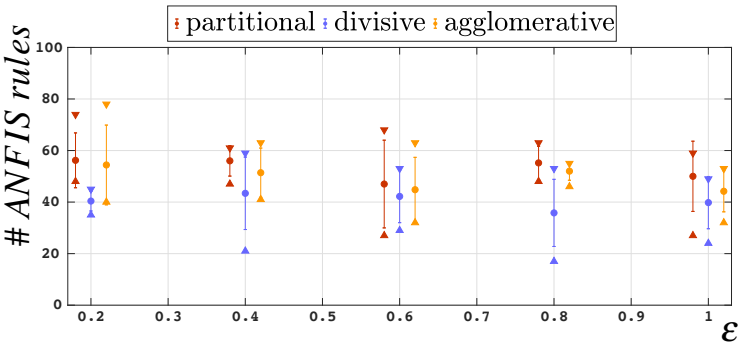


FIGURE 13.6: ANFIS number of rules for each solution found as function of the input-hyperplanes space ε weight ($\lambda = 0.4$).

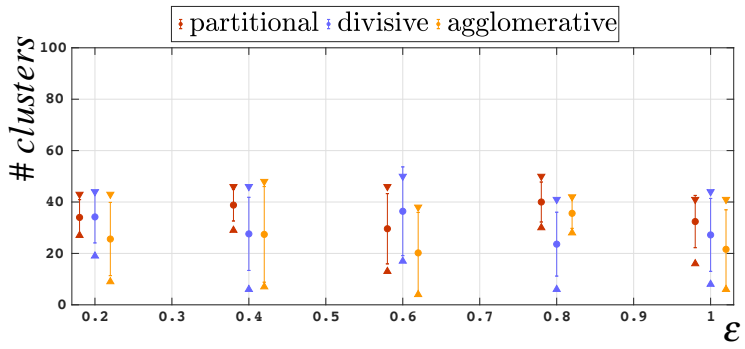


FIGURE 13.7: Number of clusters before PARC step as function of the input-hyperplanes space ε weight ($\lambda = 0.4$).

As far as the profit gap minimization is concerned (*i.e.*, Figure 13.2 vs. Figure 13.5), introducing the input space leads to lower performances with respect to hyperplane clustering, regardless of its importance within the dissimilarity measure given by ε. Considering the joint input-hyperplanes space, most of the solutions are

within $15-20\%$ from the MILP solution, whereas using a pure hyperplane clustering approach most of the solutions are within $10-15\%$ from the MILP solution. On the other hand, the joint input-hyperplanes space leads to slightly less complex ANFISs. Indeed, by considering the number of rules (*i.e.*, Figure 13.3 vs. Figure 13.6), in the joint input-hyperplanes space several models with less than 50 rules can be found, whereas the hyperplane space counterpart features at least 50 rules (on average). However, the number of clusters on the top of which said rules have been built seems higher in the joint input-hyperplanes space (on average between 30 and 40 clusters in many cases) with respect to the hyperplane clustering approach (maximum 30 clusters on average).

For the sake of comparison, the three tables in Table 13.2 summarizes the best results obtained from the two test suites: the upper table contains the results obtained with the hyperplane clustering and PARC using $\lambda = 0.4$ (the most promising value according to Figure 13.2); the table in the middle contains the results obtained with the hyperplane clustering without PARC; the last table contains the results obtained with PARC ($\lambda = 0.4$) and the joint input-hyperplane space clustering using $\varepsilon = 0.4$ (the most promising value according to Figure 13.5).

13.4 Conclusion

In this chapter a novel data-driven procedure for the synthesis of a real time MG EMS based on computational intelligence techniques has been proposed. As in previous studies, the EMS consists of an ANFIS model trained by a clustering algorithm and relying on a properly formulated benchmark solution. This study focused on the advantage of defining the ANFIS by hyperplane clustering supported by a PARC classification algorithm. Moreover, three k-means variants are investigated, namely partitional, hierarchical agglomerative and hierarchical divisive. PARC, by designing proper hyperboxes around eventually overlapping clusters, allows a more accurate partitioning which in turn leads to a more effective ANFIS MFs synthesis.

TABLE 13.2: Comparison Table between Hyperplane Clustering with and without PARC and Joint Input-Hyperplane Space Clustering.

	Hyperplane Clustering ($\lambda = 0.4$)			
	mean	stdev	max	min
partitional	0.135	0.024	0.163	0.100
divisive	0.09	0.140	0.025	0.11
agglomerative	0.144	0.022	0.165	0.111
	Hyperplane Clustering (no PARC classifier)			
	mean	stdev	max	min
partitional	0.198	0.009	0.210	0.190
divisive	0.233	0.262	0.013	0.24
agglomerative	0.196	0.018	0.218	0.10
	Input-Hyperplane Clustering($\lambda = 0.4$, $\varepsilon = 0.4$)			
	mean	stdev	max	min
partitional	0.138	0.015	0.161	0.125
divisive	0.165	0.01	0.190	0.148
agglomerative	0.14	0.009	0.185	0.163

In order to evaluate the expected improvements, it has been compared with two alternative synthesis procedures for the same problem. The former does not consider the classification stage after clustering, therefore without taking into account possible overlapping between clusters in the input space. The latter explores the joint input-hyperplanes space but considering the classification stage as well.

Obtained results can be summarized in three key points: first, the classifier support leads to a performance improvements (profit optimization) in the hyperplane space that is around 50% with respect to not considering the classifier. Second, by considering the joint input-hyperplanes space, the performance improvements are smaller, hence privileging the cluster compactness in the input space is not necessary due to the classifier implementation. Finally, the improvements due to the classifier are also underlined by comparing with past studies where the classifier was not considered in the input-hyperplanes clustering. In fact, in [83] the profit gap with respect to the optimum value was around 20%, against 15% obtained here.

However, such improvements due to the PARC classifier must be paid with higher ANFIS complexity, namely the number of rules, since the number of hyperboxes

is generally higher than the number of clusters. Nevertheless, the complexity of the obtained EMS (number of rules composing the ANFIS) is still suitable to be implemented and run in real time on inexpensive microcontroller-based devices.

Chapter 14

EMS Synthesis of a FC Station MG

In this chapter it is considered a grid-connected MG equipped with an ESS and a PV plant for supporting a public FC station. The FC station is characterized by only one power pole and it is featured by a FC power supply of 50 kW. In order to reduce the cost of the kWh bought and limit the peak of power exchanged with the grid, the AC-DC converter is designed to only buy energy from the grid (*i.e.* the power flow towards the grid is unidirectional) and the maximum power exchanged is limited to 20 kW. The MG is equipped with a EMS that must be able to intelligently redistribute in real time the energy systems power flows in order to satisfy the FC demand and support the local auto-consumption taking advantage of the local PV generation. The aim of this study is focused on assessing the effectiveness the FC station MG paradigm by considering the synthesis of a suitable EMS decision-making system based on machine learning techniques. In particular, the EMS consists of an ANFIS supported by a PV prediction algorithm (see Sec 5.2 in Chapter 5). The EMS synthesis procedure is mostly the same as that proposed in Chapter 12. In particular, the ANFIS is modelled by means of the k-means divisive variant applied on the hyperplane space and it is followed by the additional clusters partitioning given by the PARC MinMax classifier. The prediction system instead is restricted to the only PV generation. It is optimized by considering the same procedure adopted in Chapter 12. The labelled pattern sets used for the EMS synthesis has been realized by means of the formulation of an optimal (upper) benchmark solution (see Chapter 13) aimed to

fulfill the FC energy demand and maximize the PV exploitation. Moreover, by taking into consideration the problem so formulated, the procedure has been repeated by implementing different PV systems by changing the maximum power generation of the same time series in Sec. 7.3 of Chapter 7 in order to detect the the most suitable PV production range (*i.e.* size).

14.1 Re-utilization of 2^{nd} Life Batteries

In [95] it is discussed that after the 2018 the first commercial PEVs generation and Plug In Hybrid Electric Vehicle (PHEV) will dismiss their own battery packs. Owing to autonomy, safety and security reasons, the dismissed PEV and PHEV batteries are characterized by a drop of efficiencies around the 20% and 10%, respectively. In [95] it is sated in particular that the strong requirements on PEV batteries suggest their use for other stationary applications.

Moreover, the study in [96] on EV lifecycle distinguishes the EV emissions in four stages: battery production, vehicle production, use and end of life. The authors affirm that most of the emissions are released during the use phase since a huge portion of the (electric) energy generation for batteries recharging mainly comes from the fossil fuel combustion.

Since FC stations are suggested to be in public areas, namely close to parking zones, the installation of a PV system, or plants dedicated to other RESs in general, can be a key solution for supporting the FC service. It would reduce the overall energy demanded to the grid (*i.e.* the installation of new fossil fuel or gas power plants) and increase the local penetration of low emission energy resources.

However, since (also) the energy production from RESs is stochastic in nature, it is necessary to design a suitable electric system in charge to effectively manage the local power flows in order to minimize the stress on the grid and improve the local auto-consumption.

For this reason the utilization of 2^{nd} life batteries as auxiliary power unit can support the developing of FC stations equipped with a small sized generation unit. It would reduce the FC station operational costs and PEVs $CO2$ emissions as well.

14.2 MG Problem Abstraction

Also in this study the formulation of the MG problem is mostly based on the hypotheses introduced in Chapter 7 considering some modifications that are explained below in this section. In Figure 14.1 it is shown the MG architecture under analysis. As shown in figure, at the given k-th time slot the energy balance on the MG back-

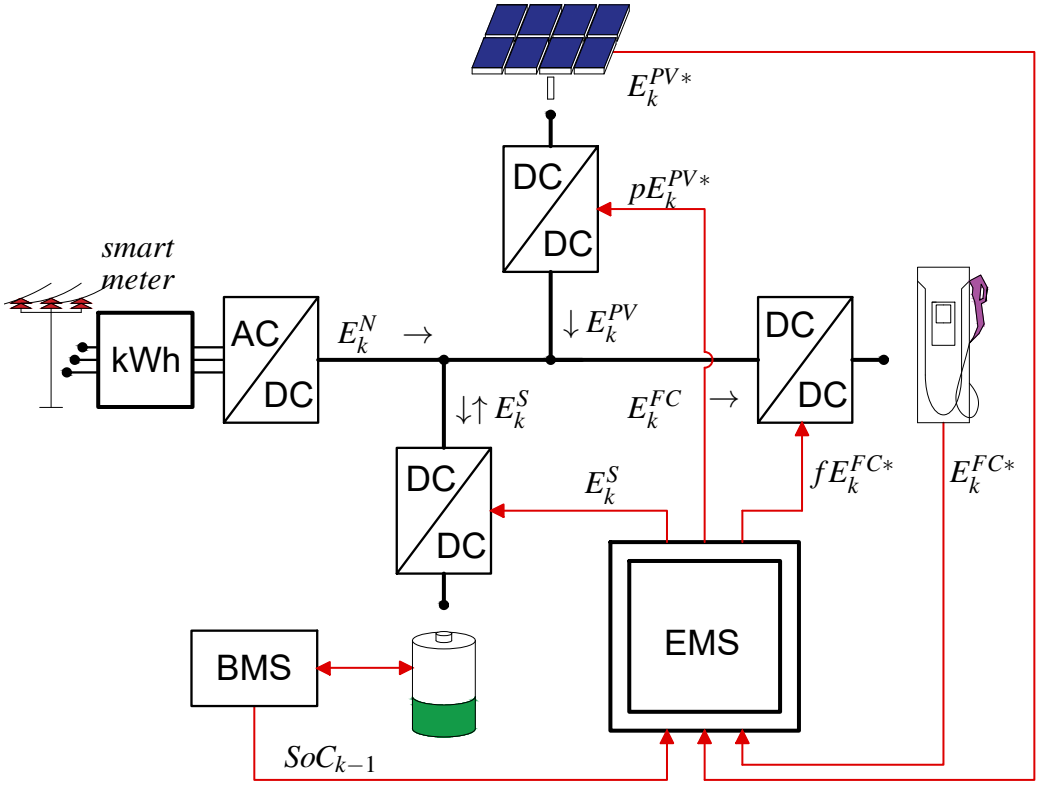


FIGURE 14.1: MG architecture. Signal wires in red, power lines in black.

bone is defined by the sum of the energies exchanged with the three energy systems (FC station power pole, PV generator and the ESS) and the grid:

$$E_k^{FC} + E_k^{PV} + E_k^S + E_k^N = 0, \quad \text{for } k = 1, 2, ... \tag{14.1}$$

where E_k^{FC} (≤ 0) is the energy provided by the DC bus to the FC station, E_k^S is the energy exchanged with the ESS (≥ 0 in case of discharge), E_k^N (≥ 0) is the energy bought from the grid and E_k^{PV} (≥ 0) the energy actually exploited by the PV system. The relations between the PV production and its actual use (*i.e.* E_k^{PV}) and the PEV energy demand with the E_k^{FC} value are the followings

$$E_k^{PV} = p_k \cdot \eta_{PV} E_k^{PV*}, \;\; p_k \in [0,1] \text{ for } k = 1,2,... \tag{14.2}$$

$$E_k^{FC} = f_k \cdot \frac{1}{\eta_{FC}} E_k^{FC*}, \;\; f_k \in \{0,1\} \text{ for } k = 1,2,... \tag{14.3}$$

Where E_k^{PV*} is the maximum PV production at the given time slot, η_{PV} the corresponding DC-DC power converter energy efficiency and p_k a real number $\in [0,1]$ that defines the portion of E_k^{PV*} actually exploited. Similarly, E_k^{FC*} is referred to the PEV energy demand, η_{FC} takes into consideration the FC station power losses given by the energy conversion and transmission. f is a binary value that states if the charging demand is going to be fulfilled ($f_k = 1$) or not ($f_k = 0$).

The MG ESS SoC is monitored and controlled by the BMS. It respects the following relations

$$SoC_{k+1} = SoC_k + \delta_k, \;\; \text{for } k = 1,2,... \tag{14.4}$$

$$SoC_k = SoC_{in}, \;\; \text{for } k = 1 \tag{14.5}$$

$$SoC^{min} \leq SoC_k \leq SoC^{max}, \;\; \text{for } k = 1,2,... \tag{14.6}$$

$$\delta_k = \begin{cases} if\ E_k^S \geq 0, & \delta soc_k = \frac{E_k^S}{\eta_{dch} C^{max}} \quad discharge \\ else\ E_k^S < 0, & \delta_k = \frac{E_k^S \eta_{ch}}{C^{max}} \qquad charge \end{cases} \tag{14.7}$$

where δ_k is the SoC variation due to the E_k^S energy exchange and the ESS charge (discharge) efficiency η_{ch} (η_{dch}).

The EMS role is to fulfill the PEV energy demand by exploiting the local production of the connected PV system (*i.e.* f_k and p_k have to be kept equal to 1). However, it is important to note that these two objectives are in conflict to each other. Indeed, for the maximization of the PV utilization the ESS must always be able to accept energy (*i.e.* its SoC must be kept at its minimum) since it is not assumed that the MG can sell energy to the grid. On the other hand, the ESS must be maintained at a high SoC to maximize the fulfillment of the recharge demand. For this reason, a data-driven approach has been applied in order to train an efficient EMS decision-making system by means of machine learning techniques in order to efficiently manage (in real time) the ESS SoC and the energy exchanged with the other MG elements.

As shown in Figure 14.1 the EMS input array is composed by the upcoming energy demand E_{k+1}^{FC*}, the energy generation E_{k+1}^{PV*} and the current ESS SoC. [1] Instead, the EMS output are f_k, p_k, E_k^S and E_k^N. These are communicated in real time to their corresponding power converters in order to respect the energy balance in (14.1).

At a first sight, the EMS decision-making system could be formulated as a Multi Input Multi Output (MIMO) system. However, on the basis of some assumptions, a MISO decision-making system restricted to the evaluation of E_{k+1}^S has been implemented. Indeed, after the decision-making system outputs E_{k+1}^S, the EMS here formulated evaluates if the FC demand can be fulfilled by considering the energy generation availability E_{k+1}^{PV*} together with the maximum energy to be exchanged with the grid. In case the FC demand can be fulfilled f_k is set to 1. Therefore, given E_{k+1}^S, E_{k+1}^{FC} and the energy balance in (14.1), the f_{k+1} value is maximized respect with E_{k+1}^N which in any case must be kept positive.

Moreover, the EMS decision-making system, in addition to considering the fulfillment of the FC demand and the use of PV, should avoid to bring the SoC ESS

[1] The (very) low computational effort of the EMS decision-making system (see Chapter 13) permits to assume to know the PV generation and FC demand at the upcoming time slot.

close to its respective upper and lower limits in order to properly use it and for a proper training of the algorithm proposed. For this reason, the battery SoC penalty function P^{SoC} plot in Figure 14.2 has been taken into account.

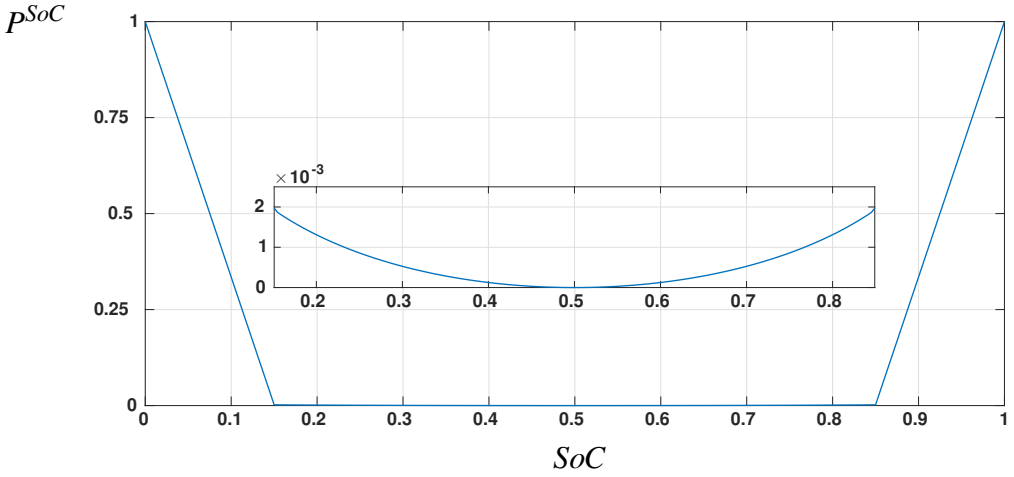


FIGURE 14.2: ESS SoC penalty function.

14.3 Synthesis and Performance Evaluation Benchmark Solutions

In this study a DP tool [39] has been used on a dataset defined by the E^{PV*} and E^{FC*} time series extended on more days in order to evaluate three different benchmark solutions which differ from the OF formulation. In Sec. 14.2 three EMS targets have been introduced: the fulfillment of the FC demand, the utilization of the PV generation and the ESS stress. These targets can be respectively represented by the following cost functions

$$F^{FC} = \frac{\sum_k^N (1 - f_k)}{N} \tag{14.8}$$

$$F^{PV} = \frac{\sum_k^N (1 - p_k) \cdot \frac{E_k^{PV*}}{12.5}}{N} \tag{14.9}$$

$$F^{SoC} = \frac{\sum_k^N P^{SoC}(SoC_k)}{N} \tag{14.10}$$

Where 12.5 kWh corresponds to the characteristic FC demand during one time slot and N is the time series length. The OF of each benchmark solution correspond to the sum of the cost functions (14.8), (14.9) and (14.10) which are respectively multiplied by a properly assigned meta-parameter.

The first benchmark solution is called *optimal* (denoted with the abbreviation "opt" in superscript). It is used for the training of the EMS decision-making system, *i.e.*, for the definition of a labelled pattern set. The optimal OF is formulated as follows

$$F^{opt} = \frac{1}{3} \cdot F^{FC} + \frac{1}{3} \cdot F^{PV} + \frac{1}{3} \cdot F^{SoC} \tag{14.11}$$

The meta-parameters are set equal to $1/3$ in order to attribute the same importance to each cost function.

The evaluation of the optimal solution allows to evaluate the time series of the ESS SoC SoC^{opt} and the energy profile $E^{S,opt}$. Indeed, the calculation of SoC^{opt}, together with the time series E^{PV*} and E^{FC*}, complete the pattern set definition referred to the EMS input array (see Figure 14.1in Sec. 14.2), whereas $E^{S,opt}$ is the corresponding label set.

The other two benchmark solutions are called *bmkFC* and *bmkPV* (denoted with the abbreviation "bmkFC" and "bmkPV" in superscript, respectively). They are used to evaluate the EMS performance during the validation phase, i.e., for a proper EMS OF formulation. The *bmkFC* meta-parameters are set in order to maximize the FC demand fulfillment discouraging the utilization of the PV energy generation as follows

$$F^{bmkFC} = \frac{1}{2} F^{FC,bmkFC} - 10^{-3} F^{PV,bmkFC} + \frac{1}{2} F^{SoC,bmkFC} \tag{14.12}$$

Hence, it is an upper benchmark solution compared to the FC fulfillment as well as a lower benchmark solution considering the PV exploitation. The *bmkPV* solution instead is formulated by permuting F^{FC} and F^{PV} in (14.12), namely

$$F^{bmkPV} = -10^{-3} F^{FC,bmkPV} + \frac{1}{2} F^{PV,bmkPV} + \frac{1}{2} F^{SoC,bmkPV} \tag{14.13}$$

Conversely to (14.12), it maximizes the PV utilization (upper benchmark) minimizing the FC demand fulfillment (lower benchmark).

The calculation of *bmkFC* and *bmkPV* allow to properly normalize in $[0;1]$ the EMS OF. Indeed, considering a given EMS featured by the function costs F^{FC}, F^{PV} and F^{SoC} on a given simulation scenario (*i.e.*, dataset), the OF is formulated as follows

$$F = \frac{1}{3}\frac{F^{FC} - F^{FC,bmkPV}}{F^{FC,bmkFC} - F^{FC,bmkPV}} + \frac{1}{3}\frac{F^{PV} - F^{PV,bmkFC}}{F^{PV,bmkPV} - F^{PV,bmkFC}} + \frac{1}{3}F^{SoC} \qquad (14.14)$$

where the *bmkFC* and *bmkPV* values, as F^{FC} and F^{FC}, are referred to the same dataset.

14.4 FC Demand Profiling

The prior knowledge of FC stations charging profiles (*i.e.* E^{FC*}) is important for the investigation of the FC station location and facility sizing. A proper formulation of FC load profiles helps to understand the impact these will have on the distribution grids and to have a clearer view on the investment statement relating to FC infrastructures. Since the development and distribution of EVs is not yet mature, informations about FC energy demand profiles and their impact on a given local area requires to be simulated with the support of suitable stochastic, deterministic tools which are able to effectively represent the scenario under analysis. In fact, the use of datasets available in literature can be misleading since they could not reflect the scenario under analysis.

Several methods for reproducing PEV FC demand profiles are proposed in literature. In [97] for instance, a stochastic model is proposed to investigate the impacts of FC stations on the distribution transformers loading and on the voltage profiles of the distribution grid power lines. The work proposed in [98] analyse and compare

the behaviours and habits of future PEV drivers respect with conventional ICE vehicle drivers. The authors developed an algorithm based on a mixed deterministic and probabilistic approach in order to investigate about which factors may influence the request of FC from PEV commuters. By simulating a commuter PEVs fleet, the algorithm is able to generate a daily commuter fleet demand profile with a per minutes granularity.

Starting from [98], a tool able to generate a public FC station energy demand profile (*i.e.* a FC station daily energy demand dataset) has been developed. It is set in order to fairly represent the scenario in which the MG is installed, namely the "Smart Columbus" project [99] of the city of Columbus-Ohio in the U.S. The proposed tool consists in the profiling of an EV commuters fleet that covers a given area in which a certain number of public FC stations are distributed. It can be summarized in two main phases. The first is dedicated to the setting of the commuter fleet habits and PEV features parameters, whereas the second deals with the PEVs fleet simulation.

14.4.1 Commuters Feature Parameters Settings

Here following are introduced the feature settings of the PEV commuters fleet according to the "Smart Columbus" project scenario [99]. In the area are supposed to be installed 30 public 50 kW FC stations. For safe and security reasons, the FC brings the plugged-in EV ESS SoC up to 0.8 $p.u.$. The commuters and PEVs features are evaluated by setting the feature mean value and standard deviation considering a Gaussian distribution probability or rather by setting their interval of existence in case it is considered a uniform distribution probability.

As shown in Tab. 14.1 the PEVs are distinguished in three types, *small*, *medium* and *large*. Each of them is featured by a corresponding Consumption Rate (CR) and ESS size. These are expressed in kWh/km and kWh, respectively. In table are indicated the penetration of each EV type respect with the fleet and the range of the ESS size and CR features which are characterized by an uniform distribution.

TABLE 14.1: Simulation scenario, EV features

EV type	penetration	Capacity $[kWh]$	CR $[kWh/km]$
Small	30%	$10-20$	$0.10-0.13$
Medium	60%	$20-30$	$0.15-0.18$
Large	10%	$30-40$	$0.20-0.25$

TABLE 14.2: Simulation scenario, commuter behaviour parameters

Commuter	Penetration	Time dep. $[h]$	Park. time $[h]$
Part-time	16%	11 ± 2.5	4 ± 0.5
Full-time	64%	8 ± 1.5	8 ± 1
Non reg.	20%	11 ± 6	4 ± 1

Commuters have also been divided into three categories: *part-time*, *full-time* and *non regular* workers (see Tab.14.2). For each category is associated a commuter time departure and a parking time interval. While the PEV is parked, the commuter has the chance to charge its PEV by means of a 6 kW slow charge station. Moreover, each commuter is featured by the (slow) home charge opportunity, the distance to cover during the trip and the EV average speed. In particular, the 90% of the commuters are equipped with home charge but the 10% forget to use it during the night. In case of the commuter forgets, the initial SoC is set to 0.6 $p.u$ considering an uncertainty of 0.1 $p.u$. The commuter average speed is set to 50 ± 10 km/h whilst the commuter average covered distance is 40 ± 7 km.

The commuter range of anxiety is represented by SoC outward and return thresholds set to 0.3 $p.u$. and 0.4 $p.u$., respectively. When the commuter alerts a low SoC, namely a SoC lower than the corresponding threshold, he deviates to the first available FC station. The path deviation increases the travelling distance, in outward or return, by

$$path\ deviation\ = \frac{\alpha \cdot SoC^{thr} \cdot Capacity}{CR}\ km \tag{14.15}$$

where the parameter α is set to 0.85 and SoC^{thr} is the range of anxiety SoC threshold.

14.4.2 PEV Commuters Fleet Simulation

The simulation phase is organized as follows. Each commuter drives from home to a destination point (*i.e.* outward path), parks the PEV for a given time interval and after come back to home. During the driving, if the PEV SoC goes under the given SoC threshold, the commuter must book and deviate to a FC station. The commuters are simulated in parallel in order to properly consider the FC station occupancy. The time frame resolution is of 1 min. For the sake of simplicity, the proposed tool does not take into consideration the FC stations location, traffic congestion and roads topology. Therefore, when the commuter has to book the FC station he does not consider any FC station proximity and, if more FC station are available, it is randomly chosen. In Alg.6 are listed the commuter actions that are summarized in *drive* from home to the destination point, *park* and *drive* again from destination point to home. In Alg.7, is registered the distance covered and the distance still to cover and when the the commuter alert a low SoC of his EV and thus must deviate to a FC station. In Alg.8 instead are described the deviation to the FC station and the (fast) charging action. During the simulation, all the FC station energy demand profiles are traced, however, just one FC profile will be used for the dataset generation.

As a demonstration, in Figure 14.3 is plot a daily trace of the FC stations energy demand considering a scenario of 30 FC stations and a fleet of 2400 PEVs.

Algorithm 6 EV commuters fleet Procedure

1: **procedure** CONSIDERING ONE SINGLE COMMUTER
2: $C = Commuter(P)$ ▷ commuter selection from the fleet
3: $t = C.timedeparture$ ▷ time departure from home
4: $SoC^{thr} = C.SoC^{outward}$
5: $[C,t] = drive(C,FCS,t,SoC^{thr})$ ▷ Drive to work
6: $[C,t] = park(C,t)$ ▷ Parking
7: $SoC^{thr} = C.SoC^{return}$
8: $[C,t] = drive(C,FCS,t,SoC^{thr})$ ▷ Drive to home
9: $Commuter(P) = C$ ▷ refresh commuters fleet
10: **end procedure**

Algorithm 7 Drive Function

procedure $[C,t] = drive(C, FCS, t, SoC^{thr})$
  $d^{dc} = 0$ ▷ distance covered
  $d^{d2c} = C.dist2Cover$ ▷ distance to cover
  $v = C.speed$ ▷ Commuter speed
  $CR = C.CR$ ▷ Commuter Consumption Rate
  $SoC_{in} = C.EV.SoC$ ▷ Commuter EV SoC
  $Cap = C.EV.Capacity$ ▷ Commuter EV Battery Cap.
  $t = t + \delta t$ ▷ current time
  $SoC = max(SoC^{thr}, SoC_{in} - \frac{CR \cdot d^{d2c}}{Cap})$
  $\delta t = \frac{(SoC_{in} - SoC)Cap}{CR \cdot v}$
  $t = t + \delta t$
  $d = \delta t \cdot v$
  if $SoC = SoC^{thr}$ **then**
    $[C,t] = deviation2FCst(C,t,FCs,SoC)$
    $SoC = SoC - \frac{(d^{d2c} - d)CR}{Cap}$ ▷ complete the path
    $t = t + SoC - \frac{(d^{d2c} - d)}{v})$ ▷ time arrival
  end if
end procedure

Algorithm 8 Drive to the first available FC station

procedure $[C,t] = deviation2FCst(C,t,FCs,SoC)$
  $CR = C.CR$ ▷ Commuter Consumption Rate
  $v = C.speed$ ▷ Commuter speed
  $Cap = C.EV.Capacity$ ▷ Commuter EV Battery Cap.
  $v = C.speed$ ▷ Commuter speed
  $CR = C.CR$ ▷ Commuter Consumption Rate
  δt^{wait}: wait for the first available FCst.
  $t = t + \frac{(1-\alpha)SoC \cdot Cap}{CR \cdot v} + \delta t^{wait}$
  $SoC = \alpha SoC$ ▷ SoC once reached the FCst.
  $t = t + \frac{(0.8 - SoC)Cap}{P^{FC}\eta^{FC}} + \delta t^{wait}$
  $SoC = 0.8$ ▷ SoC Fast Charging limit
end procedure

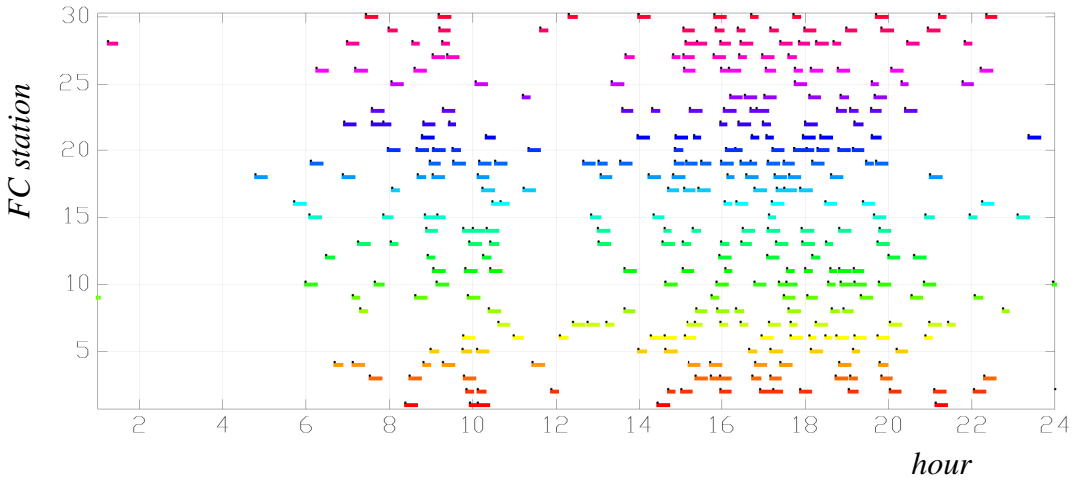


FIGURE 14.3: The fleet FC demand traces considering a scenario of 2400 PEV commuters and the installation of 30 FC stations. Each raw correspond to a 50 kW FC station trace.

14.5 Energy Management System Design

The MG EMS decision-making system modelled is a MISO model in charge to define E^S in real time (see Sec.14.2). The EMS has the same properties of that proposed in Chapter 12. Namely, it can be distinguished in two main parts, the EMS ANFIS decision making system block and the prosumer prediction system which has the role to support the ANFIS decisions. However, the application of a ESNs-based prediction algorithm based on the FC energy demand is assumed to be not effective because of the stochastic and intermittent nature of the load as stated in Sec.14.4. For this reason, the future trend on the FC demand is defined by the probability of finding a plugged EV h_k at the current hour by means of the generation of a daily histogram referred to the presence of PEV on the FC station. The histogram is generated by considering the FC demand on the TR. In Figure 14.4 a detailed scheme of the overall decision making system is shown. As described in the figure, the ANFIS input array is composed by the next energy generation E_k^{PV*}, the next energy demand E_k^{FC*}, the current ESS SoC SoC_k, h_k^{FC} and the trend line coefficients m_k^{PV}, q_k^{PV} referred to the future PV generation (for further details see Sec. 14.2). Moreover, as in Chapter 12, the input array comprehends also the current hour $hour_k$ and a raw daily estimation of the weather forecast W_k (a signal discretized in 4 states). These two signals are

used for the PV prediction.

The EMS procedure of synthesis is the same of Chapter 12 and Chapter 13. The ANFIS is synthesised by means of the the model proposed in Chapter 13. In particular, the divisive variant applied on the hyperplane space and supported by the MinMax classifier. In addition, in the phase of training different prediction time horizons on the PV generation has been tested in order to better explore the EMS performance.

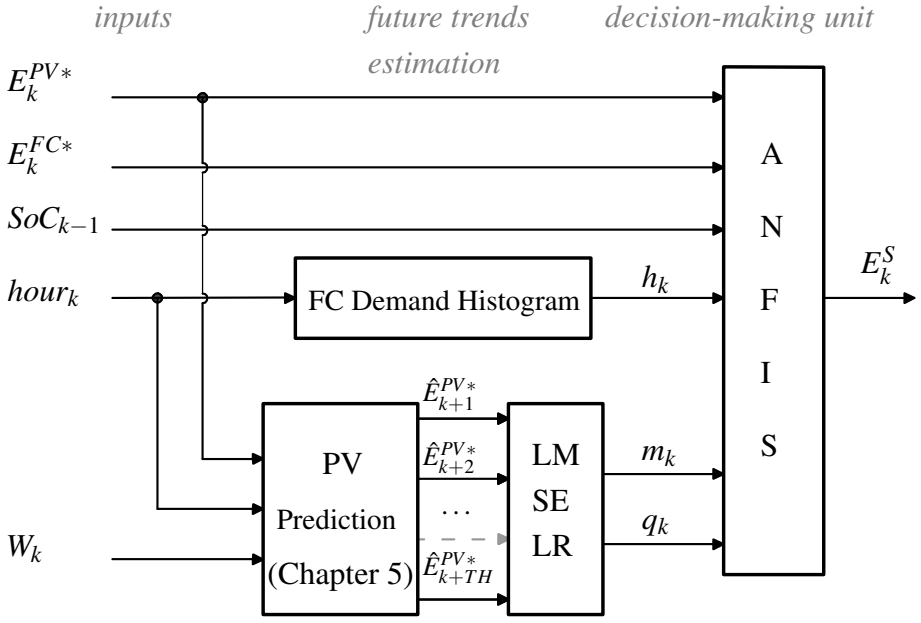


FIGURE 14.4: EMS decision-making system architecture.

14.6 Simulation Settings

The MG ESS and the EV battery pack (*i.e.* the FC load) modelling are referred to [100] and [101]. In these manuscripts the authors stated that the efficiency of a EV ESS ion-lithium cell, which depends about several factors (*e.g.* the particular type of cell), can be assumed approximately around 95%, however the overall battery pack has an efficiency (η_{FC}) of 90% considering the power losses given by transmission, the BMS and the power converter. Instead, 2^{nd} use batteries undergo degradations that reduce the overall capacity by the 80% and charge-discharge efficiencies ($\eta_{dch} -$

η_{ch}) to 80%. The MG ESS is composed by two dismissed battery pack for an overall capacity of 39 kWh. Its SoC must be keep within the 15% and 95%.

Several simulation scenarios have been keep into consideration in order to better study the problem under analysis. As regarding the FC station, 3 different dataset are defined. These are generated by considering a commuter fleet of 1600, 2400 and 3200 PEVs, keeping fixed the rest of the parameters introduced in 14.4. About the PV generation, a real PV generation dataset (the same used in Chapter 12 and Chapter 13) has been used for generating 5 different PV profiles which differ by their respective peak of power equal to 10 kW, 20 kW, 30 kW, 40 kW and 50 kW. In order to accomplish it, the series has been firstly normalized in $[0,1]$ and then multiplied by the given peak of power which would establish its size.

The definition of more datasets allowed the formulation of $3 \times 5 = 15$ different scenarios whose each of them is represented by properly defined FC station and PV system. For each scenario, it has been applied the decision-making system synthesis procedure summarized in Sec. 12.1 of Chapter 12 which relies on the definitions of 3 different TR-VL pairs and a single TS by means of a k-fold cross validation. Each dataset covers an overall period of 60 days sampled with a 15 minutes frequency that is partitioned in 4 subsets of the same dimension (*i.e.* 15 days, which correspond to the 25% of the dataset) with the purpose to attribute 30 days to the TR and 15 to both VL and TS. The prediction model time horizon of the PV generation has been ranged from 4 time slots to 96 (one day) considering the following values: 4, 6, 8, 12, 16, 24, 38, 50, 64, 80 and 96. The EMS performances associated to a given scenario are evaluated on the corresponding TS and then compared with the *optimal* solution introduced in Sec.14.3. The comparison is focused on the FC load fulfillment, the PV effective utilization, the ANFIS complexity and the characteristic prediction time horizon selected.

14.7 Results

In Tab. 14.3 and 14.4 are illustrated the overall cumulative energy demand and energy generation of the FC station and PV system with respect to the TS of each given simulation scenario. In Figure 14.5 are plot the histograms of each FC demand. In Figure 14.6 a comparison between the 10 kW PV generation TS profile and the ESNs prediction is shown. For the sake of simplicity, the prediction curves plotted are referred to those ESNs generated on the first TR-VL pair with prediction time horizons equal to those set to the prediction models generation. All the ESNs perform a RMSEs on the TS (the error is evaluated by considering the PV generation profile normalized in $[0;1]$) always lower than the 8% and greater than the 4%.

TABLE 14.3: FC demand on the TS for each scenario.

Commuter fleet $[\#PEVs]$	1600	2400	3200
Overall FC demand $[kWh]$	800	1637	1950
Overall FC demand $[\#]$	64	131	156

TABLE 14.4: PV overall energy generation on the TS in function of the PV peak of power (size) considering each scenario.

PV size $[kW]$	10	20	30	40	50
Energy gen. $[kWh]$	850	1701	2551	3401	4251

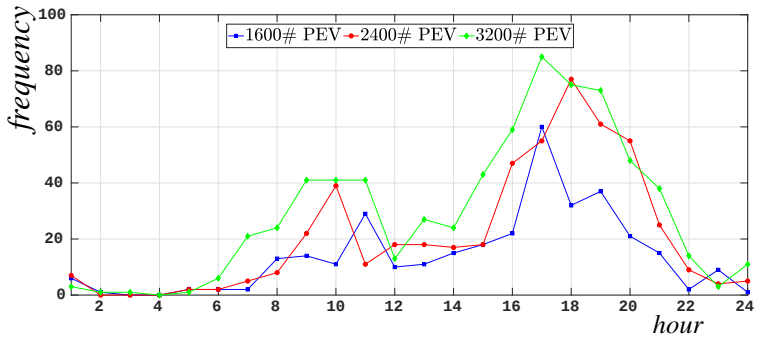


FIGURE 14.5: Histograms referred to the FC demand considering commuters fleets of 1600, 2400 and 3200 PEVs, respectively.

In order to better observe the robustness of the proposed methodology especially about what concerns the ANFIS generation (*i.e.* the *clustering* and *MinMax* application), the ANFIS synthesis procedure has been repeated 5 times.

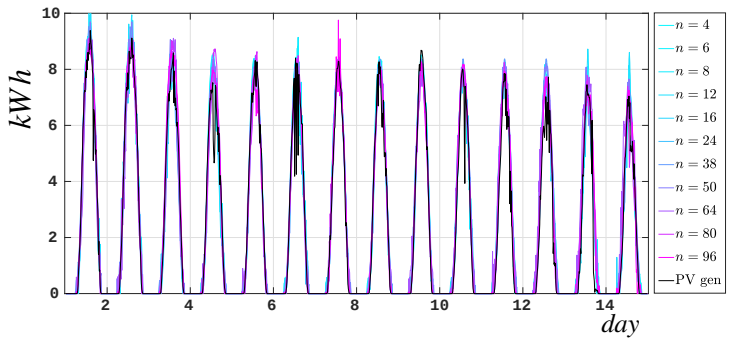


FIGURE 14.6: TS generation profile with a power peak of 10 kW. In black the signal to predict, in cool colours the profiles output by the ESN in function of the prediction time horizon n $[\#\, time\ slots]$.

The EMS performances have been compared on the TS with the *optimal* solution and the overall FC demand and PV generation data which are reported in Tab. 14.3 and 14.4. In Tab. 14.5 are shown the number of not-fulfilled FC demand of the *optimal* solution respect with the overall FC demand. Following, in Tab. 14.8 are shown the respective ANFIS-based EMS decision-making system results. These are represented in terms of mean value and standard deviation. Tab. 14.7 and Tab. 14.8 show the amount of not used PV energy generation compared to the total production (see Tab. 14.4) considering the *optimal* solutions and the synthesised EMS decision-making systems, respectively.

		PV size $[kW]$				
		10	20	30	40	50
# PEVs	1600	0.02	0.00	0.00	0.00	0.00
	2400	0.03	0.02	0.02	0.02	0.02
	3200	0.03	0.01	0.01	0.00	0.00

TABLE 14.5: FC performance on of the *optimal* solution with respect to the overall demand. In dark green the values close to 0. As the performance worsens, the colour fades to yellow.

Results show that in most of cases, neither the *optimal* solution satisfies the overall FC demand. The difference between the EMS performances and the *optimal* solution on the FC demand fulfillment are equal or less than the 10%. In addition, for PV size over 20 kW the EMS performances are equal or less than the 5%.

		PV size $[kW]$				
		10	20	30	40	50
# PEVs	1600	0.03 0.01	0.02 0.01	0.01 0.01	0.00 0.00	0.01 0.01
	2400	0.13 0.00	0.07 0.01	0.06 0.00	0.05 0.01	0.06 0.01
	3200	0.12 0.01	0.08 0.00	0.05 0.01	0.02 0.00	0.01 0.00

TABLE 14.6: FC performance mean value and standard deviation of the EMS solutions with respect to the overall FC demand.

		PV size $[kW]$				
		10	20	30	40	50
# PEVs	1600	0.13	0.40	0.56	0.66	0.72
	2400	0.01	0.16	0.33	0.44	0.53
	3200	0.02	0.17	0.30	0.40	0.48

TABLE 14.7: PV performance on of the *optimal* solution with respect to the overall generation. In dark blue the values close to 0. As the performance worsens, the colour fades to pink.

		PV size $[kW]$				
		10	20	30	40	50
# PEVs	1600	0.24 0.02	0.55 0.03	0.67 0.00	0.74 0.00	0.79 0.00
	2400	0.07 0.00	0.27 0.02	0.42 0.00	0.54 0.01	0.61 0.00
	3200	0.17 0.04	0.29 0.01	0.40 0.01	0.51 0.00	0.58 0.01

TABLE 14.8: PV performance on of the EMS solution with respect to the overall generation.

		PV size $[kW]$				
		10	20	30	40	50
# PEVs	1600	4.8(1.1)	37.6(24.1)	14.0(4.5)	96.0(0.0)	13.6(20.4)
	2400	4.0(0.0)	44.8(26.3)	67.2(7.2)	73.6(33.2)	83.2(7.1)
	3200	56.8(35.5)	32.0(17.9)	77.2(16.7)	38.0(0.0)	24.4(18.6)

TABLE 14.9: Prediction system time horizons.

		PV size $[kW]$				
		10	20	30	40	50
# PEVs	1600	21.2(3.5)	25.6(15.3)	14.0(11.2)	12.0(0.0)	27.6(10.4)
	2400	4.0(0.0)	15.2(4.5)	2.8(0.5)	12.4(9.6)	6.2(8.49)
	3200	35.6(4.98)	22.8(5.2)	11.4(2.3)	12.8(1.1)	17.0(11.0)

TABLE 14.10: ANFIS decision-making unit numbers of rules.

About the PV utilization, it is important to note that it is conditioned by the FC demand (*i.e.* the PEVs fleet magnitude) since the PV production can serve only for the auto-consumption. The performances of the *optimal* solutions show that with the increasing of the PV size the auto-consumption of energy decrease up to the 30%. Compared to them, most of the EMS performances are approximately equal or lower than the 10%.

About the selection of the prediction models and the ANFISs, in Tab. 14.9 and Tab. 14.10 are shown the mean values and standard deviation of the prediction time horizon (expressed in terms of # time slots) and the ANFIS number of rules considering each simulation scenario. The results show that in most of cases the PV future trends estimation does not show a preferred prediction time horizon extension. About the number of rules instead, these are less than 40, therefore the ANFIS complexity is very limited.

For the sake of clarity, in Figure 14.7 and 14.9 are illustrated the energy flows of the *optimal* solution and of an ANFIS-EMS solution considering a scenario of 2400 PEVs and a PV size of 30 kW which are followed by the corresponding ESS state of energy profiles in Figure 14.8 and Figure 14.10. In figures are shown the FC energy demand fulfilled (in green) and not (in red), the PV generation used (in blue) and not exploited (in orange).

14.8 Conclusions

In this chapter an ANFIS based decision-making system applied to a grid-connected FC station MG EMS is investigated. The MG is assumed to be equipped with an ESS composed by 2 dismissed PEV battery packs and is connected to a PV system in order to reduce the PEVs $CO2$ emissions. In addition, with the aim to reduce the stress given to the distribution grid, the MG can only buy energy considering a maximum power exchange of 20 kW. The decision-making system model used is inspired to

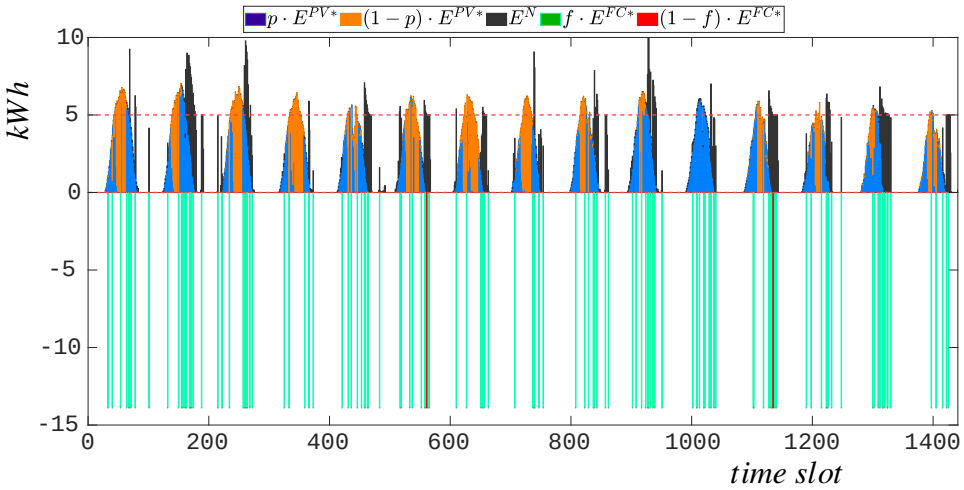


FIGURE 14.7: Energy flows of the *optimal* solution referred to the TS considering the following simulation scenario: PV size 30 kW, fleet of 2400 PEVs. The dotted line is referred to the maximum energy that can be exchanged with the main grid in a time slot.

FIGURE 14.8: Energy state of the ESS considering the *optimal* solution in Figure 14.7.

models already proposed in Chapter 12 and Chapter 13 for what concerns its architecture and the procedure of synthesis, respectively. For a fair study of the model, the EMS performance are evaluated by properly formulating a set of simulation scenarios referred to the "Smart Columbus" project. These have been reproduced by means of the creation of a suitable tool in charge to simulate a PEV commuter fleet in a given urban area. It enabled to generate the load profile of a FC station over a period of 2 months. The EMS performances have been compared respect with a properly formulated optimal benchmark solution found with a DP tool which has been applied

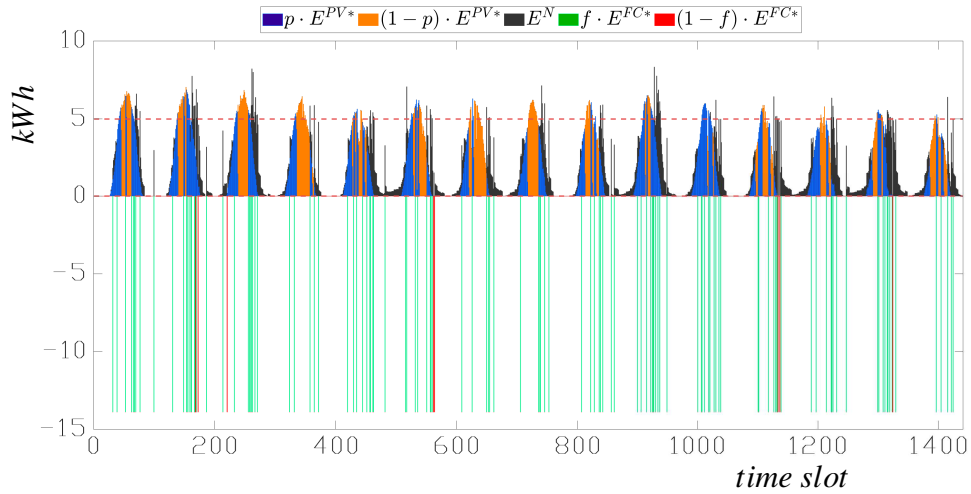


FIGURE 14.9: Energy flows of one ANFIS-EMS solution referred to the TS considering the following simulation scenario: PV size 30 kW, fleet of 2400 PEVs.

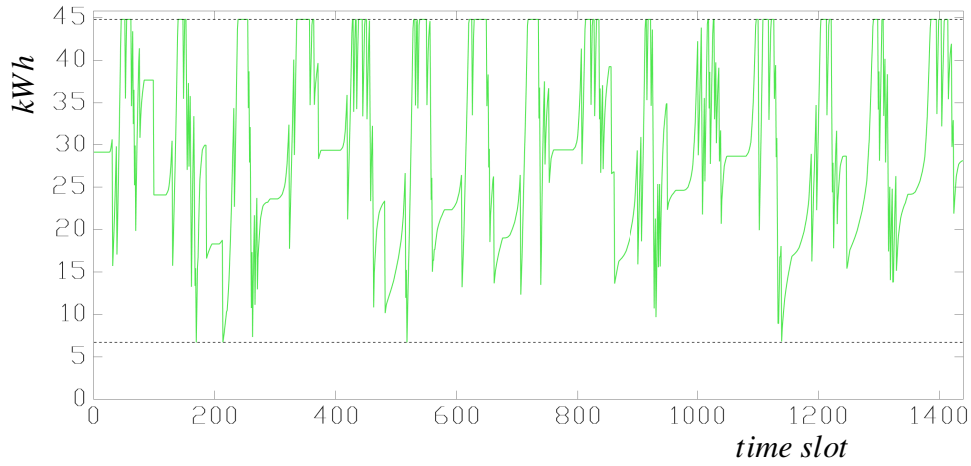


FIGURE 14.10: Energy state of the ESS considering the ANFIS-EMS solution in Figure 14.9.

by assuming to know a priori the overall energy systems time series. Very interesting results have been obtained. In particular, for PV generation with peaks of powers higher than 20 $[kW]$ the EMS performances on the FC demand fulfillment are very close to the *optimal* solution (lower than 5%). The performance on the PV utilization are approximately around the 10% respect with the optimal benchmark solution.

Chapter 15

Conclusions

The capillary expansion of MGs in the distribution grid is contributing to the growth of a new player in the energy market known as prosumer. It is a special customer equipped with a local electric grid able to connect, manage and control local DG systems, flexible loads and auxiliary power units such as battery ESSs and small sized controllable generators. Grid-connected MGs are proving to be very effective solutions since they have the capability to manage and control those loads and (distributed) power generators featured by stochastic and intermittent behaviours by means of the application of DR programs in agreement with the DSO. Furthermore, the MGs development would help to preserve the overall distribution grid stability, reduce the transmission energy losses and support a fair integration and development of the DG.

The MG EMS is responsible of DR programs execution, it is in charge to monitor, control and distribute in real time the MG energy flows taking advantage of the energy systems flexibility and the auxiliary power units. The modelling of a suitable real time EMS decision making system is a topic widely discussed in literature. Indeed, the (direct) application of deterministic methods based on LP, MILP or DP formulation results in most of cases not effective due to the volatility of the electrical loads and the DG. However, deterministic methods can be useful for the training of "intelligent" algorithms through the adoption of data driven approaches and heuristics. In particular, the application of machine learning techniques on real data has been revealed a successful solution.

The main topic of this thesis is focused on the modelling of a real time fuzzy-based decision making system able to intelligently redistribute the MG energy flows and the energy exchanged with the distribution grid. Indeed, FISs are characterized by a high degree of reliability given by their rule-based structure which allows a better human-machine cooperation.

In the studies carried out a prosumer grid-connected MG equipped with a PV system and a battery ESS has been taken into account. This scenario has been formulated considering two different case of study which differ for the type of the user. The first case study concerns a unit of aggregated loads connected to the distribution grid of a given area of Rome. In the second case it is considered a public FC station where important restrictions are assumed on the connection with the grid. In particular, the grid can only feed the MG and its maximum power is much lower than the FC rated power.

Different OF formulations and synthesis procedures have been investigated for the EMS modelling. Regarding the OF formulation, for the first case of study it has been considered the profit generated by the energy exchange with the grid by assuming a ToU energy price policy. In addition, it has also been considered the maximization the MG auto-consumption and the reduction of the energy oscillations with the distribution grid by means of the formulation of a multi-objective optimization problem. For the FC station MG instead, the OF is formulated in order to assure the FC service and improve the PV auto-consumption.

Regarding the FIS synthesis procedures, the first studies leverage on the modelling of a Mamandani-type FIS optimized by means of a GA and NSGA-II in case of a multi-objective optimization problem. However, the effectiveness and computational effort of the FISs synthesis by EAs is strictly related with the number of inputs and the number of MFs for each Term Set. For this reason, other approaches have been explored. In particular, by formulating the optimization problem as a supervised one, first order Takagi-Sugeno type FISs (*i.e.* ANFIS) have been synthesised

by means of *k*-means based clustering algorithms in order to better partition the input domain (*i.e.* the ANFIS MFs generation). A first study has been focused on the implementation and comparison of different dissimilarity functions in the clustering algorithm. In particular, the mahanalobis distance, the manhattan distance and, obviously, the squared euclidean distance have been used, respectively. Following, hierarchical clustering variants have been implemented in order to accelerate the overall EMS process of synthesis proving that, especially the divisive variant, drastically over-performs the simple partitional clustering. A study on the clustering space has been investigated as well. Starting from considering the pure input space, it has been extended to the joint input-hyperplane space where it has been proven to be very effective when the clustering algorithm is supported downstream by a classifier in charge to avoid incorrect overlaps of the fuzzy MFs. The possibility to increase the number of input variables allowed to consider the upstream support of prediction algorithms able to forecast the future trends of the loads and the DG as well. In this latter study, very interesting results have been obtained in case of profit maximization. In particular, the ANFIS synthesis by clustering on the hyperplane space with the support of a PARC classifier and the prediction on the prosumer energy systems brought to performances close to 10% the optimal solution. As concerns the FC station case of study, performances are very close to the optimal solution both in terms of the internal use of the PV and in the fulfillment of the FC demand, considering the most relevant simulation scenarios.

Besides the implementation of rolling horizon strategies, future studies will concern the implementation of other machine learning models for the EMS synthesis in order to better compare and analyse the results obtained so far. Multi Layer Perceptron NNs, ESNs and Support Vector Machines have been proven to be very interesting and successful models. Furthermore, the installation of a small sized controlled generator in support of the MG is another interesting topic, as well as the application of DR services for intelligently reshaping and schedule flexible loads such as the (slow) charge of PEVs, washing machines and heat pumps.

Chapter 16

List of Pubblications

- "Multi Objective Optimization of a Fuzzy Logic Controller for Energy Management in Microgrids", Stefano Leonori, Enrico De Santis, Antonello Rizzi and F. M. Frattale Mascioli, 2016 IEEE Congress on Evolutionary Computation (CEC), July 2016, pp. 319-326.

- "Optimization of a Microgrid Energy Management System based on a Fuzzy Logic Controller", Stefano Leonori, Enrico De Santis, Antonello Rizzi and F. M. Frattale Mascioli, IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society, Oct. 2016, pp 6615-6620.

- "An Optimized Microgrid Energy Management System based on FIS-MO-GA Paradigm", Stefano Leonori, Maurizio Paschero, Antonello Rizzi and F. M. Frattale Mascioli, July 2017, pp 1-6.

- "FIS Synthesis by Clustering for Microgrid Energy Management Systems", Stefano Leonori, Maurizio Paschero, Antonello Rizzi and F. M. Frattale Mascioli, Neural Advances in Processing Nonlinear Dynamic Signals, Springer International Publishing, Jan 2019, pp 61-71.

- "ANFIS Synthesis by Clustering for Microgrids EMS Design", Stefano Leonori, Alessio Martino, Antonello Rizzi, Fabio Massimo Frattale Mascioli, Proceedings of the 9th International Joint Conference on Computational Intelligence - Volume 1 IJCCI, Jul. 2017, pp 328-337.

- "Microgrid Energy Management by ANFIS Supported by an ESN Based Prediction Algorithm", Stefano Leonori, Antonello Rizzi, Maurizio Paschero, Fabio Massimo Frattale Mascioli, 2018 International Joint Conference on Neural Networks (IJCNN), Jul. 2018.

- "ANFIS Microgrid Energy Management System Synthesis by Hyperplane Clustering Supported by Neurofuzzy Min-Max Classifier", Stefano Leonori, Alessio Martino, Antonello Rizzi and F. M. Frattale Mascioli, IEEE Transaction on Emerging Topics, in press.

Bibliography

[1] F. M. Gatta, *Impianti Elettrici Vol. 1*. Società Editrice Esculapio, 2014.

[2] R. Shah, N. Mithulananthan, R. Bansal, and V. Ramachandaramurthy, "A review of key power system stability challenges for large-scale pv integration", *Renewable and Sustainable Energy Reviews*, vol. 41, pp. 1423 –1436, 2015, ISSN: 1364-0321.

[3] D. Streimikiene, "The impact of international ghg trading regimes on penetration of new energy technologies and feasibility to implement eu energy and climate package targets", *Renewable and Sustainable Energy Reviews*, vol. 16, no. 4, pp. 2172 –2177, 2012, ISSN: 1364-0321.

[4] P. Denholm, M. O'Connell, G. Brinkman, and J. Jorgenson, *Overgeneration from solar energy in California: a field guide to the duck chart*. National Renewable Energy Laboratory Golden, CO, 2015.

[5] X. Luo, J. Wang, M. Dooner, and J. Clarke, "Overview of current development in electrical energy storage technologies and the application potential in power system operation", *Applied Energy*, vol. 137, pp. 511 –536, 2015, ISSN: 0306-2619.

[6] B. Zakeri and S. Syri, "Electrical energy storage systems: A comparative life cycle cost analysis", *Renewable and Sustainable Energy Reviews*, vol. 42, pp. 569 –596, 2015, ISSN: 1364-0321.

[7] H. L. Ferreira, R. Garde, G. Fulli, W. Kling, and J. P. Lopes, "Characterisation of electrical energy storage technologies", *Energy*, vol. 53, pp. 288 –298, 2013, ISSN: 0360-5442.

[8] M. Luzi, M. Paschero, A. Rizzi, E. Maiorino, and F. M. Frattale Mascioli, "A novel neural networks ensemble approach for modeling electrochemical cells", *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–12, 2018, ISSN: 2162-237X.

[9] A. Sciarretta and L. Guzzella, "Control of hybrid electric vehicles", *IEEE Control Systems Magazine*, vol. 27, no. 2, pp. 60–70, 2007, ISSN: 1066-033X.

[10] C. Botsford and A. Szczepanek, "Fast charging vs. slow charging: Pros and cons for the new age of electric vehicles", Apr. 2009.

[11] H. Farzin, M. Fotuhi-Firuzabad, and M. Moeini-Aghtaie, "A practical scheme to involve degradation cost of lithium-ion batteries in vehicle-to-grid applications", *IEEE Transactions on Sustainable Energy*, vol. 7, no. 4, pp. 1730–1738, 2016, ISSN: 1949-3029.

[12] C. Botsford and A. Szczepanek, "Fast charging vs. slow charging: Pros and cons for the new age of electric vehicles", in *International Battery Hybrid Fuel Cell Electric Vehicle Symposium*, 2009.

[13] R. Scholer, "Dc charging and standards for plug-in electric vehicles", SAE Technical Paper, Tech. Rep., 2013.

[14] M. Ahmadi, N. Mithulananthan, and R. Sharma, "A review on topologies for fast charging stations for electric vehicles", in *2016 IEEE International Conference on Power System Technology (POWERCON)*, 2016, pp. 1–6.

[15] D. Steen and L. A. Tuan, "Fast charging of electric buses in distribution systems", in *2017 IEEE Manchester PowerTech*, 2017, pp. 1–6.

[16] M. Ahmadi, N. Mithulananthan, and R. Sharma, "A review on topologies for fast charging stations for electric vehicles", in *2016 IEEE International Conference on Power System Technology (POWERCON)*, 2016, pp. 1–6.

[17] F. Wu and R. Sioshansi, "A stochastic flow-capturing model to optimize the location of fast-charging stations with uncertain electric vehicle flows", *Transportation Research Part D: Transport and Environment*, vol. 53, pp. 354 –376, 2017, ISSN: 1361-9209.

[18] A. Zegers and H. Brunner, "Tso-dso interaction: An overview of current interaction between transmission and distribution system operators and an assessment of their cooperation in smart grids", *International Smart Grid Action Network (ISAGN) Discussion Paper Annex*, vol. 6, 2014.

[19] Q Qdr, "Benefits of demand response in electricity markets and recommendations for achieving them", *US Dept. Energy, Washington, DC, USA, Tech. Rep*, 2006.

[20] M. Albadi and E. El-Saadany, "A summary of demand response in electricity markets", *Electric Power Systems Research*, vol. 78, no. 11, pp. 1989 –1996, 2008, ISSN: 0378-7796.

[21] R. Deng, Z. Yang, M. Chow, and J. Chen, "A survey on demand response in smart grids: Mathematical models and approaches", *IEEE Transactions on Industrial Informatics*, vol. 11, no. 3, pp. 570–582, 2015, ISSN: 1551-3203.

[22] E. M. Lightner and S. E. Widergren, "An orderly transition to a transformed lt;newline/ gt;electricity system", *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 3–10, 2010, ISSN: 1949-3053.

[23] X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid — the new and improved power grid: A survey", *IEEE Communications Surveys Tutorials*, vol. 14, no. 4, pp. 944–980, 2012, ISSN: 1553-877X.

[24] T. Dragicevic, J. C. Vasquez, J. M. Guerrero, and D. Skrlec, "Advanced lvdc electrical power architectures and microgrids: A step toward a new generation of power distribution networks.", *IEEE Electrification Magazine*, vol. 2, no. 1, pp. 54–65, 2014, ISSN: 2325-5897.

[25] D. T. Ton and M. A. Smith, "The u.s. department of energy's microgrid initiative", *The Electricity Journal*, vol. 25, no. 8, pp. 84 –94, 2012, ISSN: 1040-6190.

[26] A. Ali, W. Li, R. Hussain, X. He, B. W. Williams, and A. H. Memon, "Overview of current microgrid policies, incentives and barriers in the european union, united states and china", *Sustainability*, vol. 9, no. 7, 2017, ISSN: 2071-1050.

[27] P. Mancarella and G. Chicco, "Real-time demand response from energy shifting in distributed multi-generation", *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 1928–1938, 2013, ISSN: 1949-3053.

[28] D. E. Olivares, A. Mehrizi-Sani, A. H. Etemadi, C. A. Cañizares, R. Iravani, M. Kazerani, A. H. Hajimiragha, O. Gomis-Bellmunt, M. Saeedifard, R. Palma-Behnke, G. A. Jiménez-Estévez, and N. D. Hatziargyriou, "Trends in microgrid control", *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1905–1919, 2014, ISSN: 1949-3053.

[29] B. T. Patterson, "Dc, come home: Dc microgrids and the birth of the "enernet"", *IEEE Power and Energy Magazine*, vol. 10, no. 6, pp. 60–69, 2012, ISSN: 1540-7977.

[30] L. Mariam, M. Basu, and M. F. Conlon, "Microgrid: Architecture, policy and future trends", *Renewable and Sustainable Energy Reviews*, vol. 64, pp. 477 –489, 2016, ISSN: 1364-0321.

[31] Z. Chen, J. M. Guerrero, and F. Blaabjerg, "A review of the state of the art of power electronics for wind turbines", *IEEE Transactions on Power Electronics*, vol. 24, no. 8, pp. 1859–1875, 2009, ISSN: 0885-8993.

[32] R. S. Sutton and A. G. Barto, *Introduction to Reinforcement Learning*, 1st. Cambridge, MA, USA: MIT Press, 1998, ISBN: 0262193981.

[33] K. Murty, *Linear programming*. Wiley, 1983.

[34] A. Richards and J. How, "Mixed-integer programming for control", in *American Control Conference, 2005. Proceedings of the 2005*, IEEE, 2005, pp. 2676–2683.

[35] D. Liberzon, *Calculus of variations and optimal control theory: a concise introduction*. Princeton University Press, 2011.

[36] M. Fagiani, M. Severini, M. Valenti, F. Ferracuti, L. Ciabattoni, and S. Squartini, "Rempy: A comprehensive software framework for residential energy management", *Energy and Buildings*, vol. 171, pp. 131 –143, 2018, ISSN: 0378-7788.

[37] ®IBM, *Ibm ilog cplex optimization studio*, https://www.ibm.com/products/ilog-cplex-optimization-studio,, 2018.

[38] ®Matlab, *Intlinprog, mixed-integer linear programming (milp)*, https://it.mathworks.com/help/optim/ug/intlinprog.html,, 2018.

[39] O. Sundstrom and L. Guzzella, "A generic dynamic programming matlab function", in *2009 IEEE Control Applications, (CCA) Intelligent Control, (ISIC)*, 2009, pp. 1625–1630.

[40] S. L. Arun and M. P. Selvan, "Intelligent residential energy management system for dynamic demand response in smart buildings", *IEEE Systems Journal*, vol. 12, no. 2, pp. 1329–1340, 2018, ISSN: 1932-8184.

[41] A. Scalfati, D. Iannuzzi, M. Fantauzzi, and M. Roscia, "Optimal sizing of distributed energy resources in smart microgrids: A mixed integer linear programming formulation", in *2017 IEEE 6th International Conference on Renewable Energy Research and Applications (ICRERA)*, 2017, pp. 568–573.

[42] A. C. Luna, N. L. Diaz, M. Graells, J. C. Vasquez, and J. M. Guerrero, "Mixed-integer-linear-programming-based energy management system for

hybrid pv-wind-battery microgrids: Modeling, design, and experimental verification", *IEEE Transactions on Power Electronics*, vol. 32, no. 4, pp. 2769–2783, 2017, ISSN: 0885-8993.

[43] X. Zeng and J. Wang, "A parallel hybrid electric vehicle energy management strategy using stochastic model predictive control with road grade preview", *IEEE Transactions on Control Systems Technology*, vol. 23, no. 6, pp. 2416–2423, 2015, ISSN: 1063-6536.

[44] J. Liu, H. Chen, W. Zhang, B. Yurkovich, and G. Rizzoni, "Energy management problems under uncertainties for grid-connected microgrids: A chance constrained programming approach", *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–1, 2016, ISSN: 1949-3053.

[45] W. L. Theo, J. S. Lim, W. S. Ho, H. Hashim, and C. T. Lee, "Review of distributed generation (dg) system planning and optimisation techniques: Comparison of numerical and mathematical modelling methods", *Renewable and Sustainable Energy Reviews*, vol. 67, pp. 531 –573, 2017, ISSN: 1364-0321.

[46] S. Parhizi, H. Lotfi, A. Khodaei, and S. Bahramirad, "State of the art in research on microgrids: A review", *IEEE Access*, vol. 3, pp. 890–925, 2015, ISSN: 2169-3536.

[47] Y. Riffonneau, S. Bacha, F. Barruel, and S. Ploix, "Optimal power flow management for grid connected pv systems with batteries", *IEEE Transactions on Sustainable Energy*, vol. 2, no. 3, pp. 309–320, 2011, ISSN: 1949-3029.

[48] R. Palma-Behnke, C. Benavides, F. Lanas, B. Severino, L. Reyes, J. Llanos, and D. Sáez, "A microgrid energy management system based on the rolling horizon strategy", *IEEE Transactions on Smart Grid*, vol. 4, no. 2, pp. 996–1006, 2013, ISSN: 1949-3053.

[49] M. H. K. Tushar, C. Assi, M. Maier, and M. F. Uddin, "Smart microgrids: Optimal joint scheduling for electric vehicles and home appliances", *IEEE*

Transactions on Smart Grid, vol. 5, no. 1, pp. 239–250, 2014, ISSN: 1949-3053.

[50] J. Soares, C. Lobo, M. Silva, H. Morais, and Z. Vale, "Relaxation of non-convex problem as an initial solution of meta-heuristics for energy resource management", in *2015 IEEE Power Energy Society General Meeting*, 2015, pp. 1–5.

[51] F. Lezama, E. Sucar, J. Soares, Z. Vale, and E. Munoz de Cote, "Differential evolution strategies for large-scale energy resource management in smart grids", in *Genetic and Evolutionary Computation Conference*, 2017.

[52] G. L. Storti, M. Paschero, A. Rizzi, and F. M. F. Mascioli, "Comparison between time-constrained and time-unconstrained optimization for power losses minimization in smart grids using genetic algorithms", *Neurocomputing*, vol. 170, pp. 353 –367, 2015, Advances on Biological Rhythmic Pattern Generation: Experiments, Algorithms and Applications Selected Papers from the 2013 International Conference on Intelligence Science and Big Data Engineering (IScIDE 2013) Computational Energy Management in Smart Grids, ISSN: 0925-2312.

[53] A. Chaouachi, R. M. Kamel, R. Andoulsi, and K. Nagasaka, "Multiobjective intelligent energy management for a microgrid", *IEEE Transactions on Industrial Electronics*, vol. 60, no. 4, pp. 1688–1699, 2013, ISSN: 0278-0046.

[54] F. M. Bianchi, E. Maiorino, M. C. Kampffmeyer, A. Rizzi, and R. Jenssen, "An overview and comparative analysis of recurrent neural networks for short term load forecasting", *arXiv preprint arXiv:1705.04378*, 2017.

[55] J. Ma and X. Ma, "A review of forecasting algorithms and energy management strategies for microgrids", *Systems Science & Control Engineering*, vol. 6, no. 1, pp. 237–248, 2018.

[56] D. Arcos-Aviles, F. Guinjoan, J. Pascual, L. Marroyo, P. Sanchis, R. Gordillo, P. Ayala, and M. P. Marietta, "A review of fuzzy-based residential grid-connected microgrid energy management strategies for grid power profile smoothing", in *Energy Sustainability in Built and Urban Environments*, Springer, 2019, pp. 165–199.

[57] L. Zadeh, "Fuzzy sets", *Information and Control*, vol. 8, no. 3, pp. 338 –353, 1965, ISSN: 0019-9958.

[58] C. A. Peña-Reyes, "Evolutionary fuzzy modeling human diagnostic decisions", *Annals of the New York Academy of Sciences*, vol. 1020, 190—211, 2004, ISSN: 0077-8923.

[59] S. Guillaume, "Designing fuzzy inference systems from data: An interpretability-oriented review", *IEEE Transactions on Fuzzy Systems*, vol. 9, no. 3, pp. 426–443, 2001, ISSN: 1063-6706.

[60] E. De Santis, A. Rizzi, and A. Sadeghian, "Hierarchical genetic optimization of a fuzzy logic system for energy flows management in microgrids", *Applied Soft Computing Journal*, vol. 60, pp. 135–149, 2017.

[61] L. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning—i", *Information Sciences*, vol. 8, no. 3, pp. 199 –249, 1975, ISSN: 0020-0255.

[62] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-15, no. 1, pp. 116–132, 1985, ISSN: 0018-9472.

[63] M. Gupta and J. Qi, "Theory of t-norms and fuzzy inference methods", *Fuzzy Sets and Systems*, vol. 40, no. 3, pp. 431 –450, 1991, Fuzzy Logic and Uncertainty Modelling, ISSN: 0165-0114.

[64] E. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller", *International Journal of Man-Machine Studies*, vol. 7, no. 1, pp. 1 –13, 1975, ISSN: 0020-7373.

[65] J.-S. Jang, "Anfis: Adaptive-network-based fuzzy inference system", *IEEE transactions on systems, man, and cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.

[66] O. Cord *et al.*, *Genetic fuzzy systems: evolutionary tuning and learning of fuzzy knowledge bases*. World Scientific, 2001, vol. 19.

[67] J. R. Jang and C.-T. Sun, "Neuro-fuzzy modeling and control", *Proceedings of the IEEE*, vol. 83, no. 3, pp. 378–406, 1995, ISSN: 0018-9219.

[68] A. Rizzi, F. M. Frattale Mascioli, and G. Martinelli, "Automatic training of anfis networks", in *Fuzzy Systems Conference Proceedings, 1999. FUZZ-IEEE '99. 1999 IEEE International*, vol. 3, 1999, 1655–1660 vol.3.

[69] M. Panella, A. Rizzi, F. M. Frattale Mascioli, and G. Martinelli, "Anfis synthesis by hyperplane clustering", in *Proceedings Joint 9th IFSA World Congress and 20th NAFIPS International Conference (Cat. No. 01TH8569)*, vol. 1, 2001, 340–345 vol.1.

[70] O. Cordón, F. Herrera, F. Hoffmann, and L. Magdalena, *Genetic Fuzzy Systems: Evolutionary Tuning and Learning of Fuzzy Knowledge Bases*, 1^{st}, ser. Advances in Fuzzy Systems–Applications and Theory. World Scientific Publishing Co. Pte. Ltd., 2001, vol. 19.

[71] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii", *Evolutionary Computation, IEEE Transactions on*, vol. 6, no. 2, pp. 182–197, 2002, ISSN: 1089-778X.

[72] S. Guillaume, "Designing fuzzy inference systems from data: An interpretability-oriented review", *IEEE transactions on fuzzy systems*, vol. 9, no. 3, pp. 426–443, 2001.

[73] S. Leonori, M. Paschero, A. Rizzi, and F. M. Frattale Mascioli, "An optimized microgrid energy management system based on fis-mo-ga paradigm", in *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2017, pp. 1–6.

[74] S. Leonori, E. De Santis, A. Rizzi, and F. M. Frattale Mascioli, "Optimization of a microgrid energy management system based on a fuzzy logic controller", in *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*, 2016, pp. 6615–6620.

[75] ——, "Multi objective optimization of a fuzzy logic controller for energy management in microgrids", in *2016 IEEE Congress on Evolutionary Computation (CEC)*, 2016, pp. 319–326.

[76] J. MacQueen, "Some methods for classification and analysis of multivariate observations", in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Oakland, CA, USA., vol. 1, 1967, pp. 281–297.

[77] S. Lloyd, "Least squares quantization in pcm", *IEEE transactions on information theory*, vol. 28, no. 2, pp. 129–137, 1982.

[78] D. Aloise, A. Deshpande, P. Hansen, and P. Popat, "Np-hardness of euclidean sum-of-squares clustering", *Machine learning*, vol. 75, no. 2, pp. 245–248, 2009.

[79] P. S. Bradley, O. L. Mangasarian, and W. N. Street, "Clustering via concave minimization", in *Advances in neural information processing systems*, 1997, pp. 368–374.

[80] L. Kaufman and P. Rousseeuw, "Clustering by means of medoids", *Statistical data analysis based on the L1-norm and related methods*, 1987.

[81] H.-S. Park and C.-H. Jun, "A simple and fast algorithm for k-medoids clustering", *Expert systems with applications*, vol. 36, no. 2, pp. 3336–3341, 2009.

[82] P. C. Mahalanobis, "On the generalised distance in statistics", *Proceedings of the National Institute of Sciences of India, 1936*, pp. 49–55, 1936.

[83] S. Leonori, A. Martino, A. Rizzi, and F. M. F. Mascioli, "Anfis synthesis by clustering for microgrids ems design", in *Proceedings of the 9th International Joint Conference on Computational Intelligence - Volume 1: IJCCI,*, INSTICC, SciTePress, 2017, pp. 328–337, ISBN: 978-989-758-274-5.

[84] A. Rizzi, M. Panella, and F. M. Frattale Mascioli, "Adaptive resolution min-max classifiers", *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 402–414, 2002.

[85] S. Haykin, *Neural Networks: A Comprehensive Foundation*, ser. International edition. Prentice Hall, 1999, ISBN: 9780139083853.

[86] A. Rosato, R. Altilio, R. Araneo, and M. Panella, "Prediction in photovoltaic power by neural networks", *Energies*, vol. 10, no. 7, 2017, ISSN: 1996-1073.

[87] F. M. Bianchi, E. D. Santis, A. Rizzi, and A. Sadeghian, "Short-term electric load forecasting using echo state networks and pca decomposition", *IEEE Access*, vol. 3, pp. 1931–1943, 2015, ISSN: 2169-3536.

[88] M. Lukoševičius, "A practical guide to applying echo state networks", in *Neural Networks: Tricks of the Trade: Second Edition*, G. Montavon, G. B. Orr, and K.-R. Müller, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 659–686, ISBN: 978-3-642-35289-8.

[89] H. Jaeger, "The "echo state" approach to analysing and training recurrent neural networks-with an erratum note", *Bonn, Germany: German National Research Center for Information Technology GMD Technical Report*, vol. 148, no. 34, p. 13, 2001.

[90] H. Jaeger, "Echo state network", *Scholarpedia*, vol. 2, no. 9, p. 2330, 2007, revision #183563.

[91] C. Chen, S. Duan, T. Cai, B. Liu, and G. Hu, "Smart energy management system for optimal microgrid economic operation", *IET Renewable Power Generation*, vol. 5, no. 3, pp. 258–267, 2011, ISSN: 1752-1416.

[92] E. D. Santis, A. Rizzi, A. Sadeghiany, and F. M. F. Mascioli, "Genetic optimization of a fuzzy control system for energy flow management in microgrids", in *2013 Joint IFSA World Congress and NAFIPS Annual Meeting (IFSA/NAFIPS)*, 2013, pp. 418–423. DOI: 10.1109/IFSA-NAFIPS.2013.6608437.

[93] D. Whitley, "A genetic algorithm tutorial", *Statistics and Computing*, vol. 4, no. 2, pp. 65–85, 1994, ISSN: 1573-1375.

[94] D. L. Davies and D. W. Bouldin, "A cluster separation measure", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 2, pp. 224–227, 1979, ISSN: 0162-8828.

[95] J. Neubauer and A. Pesaran, "The ability of battery second use strategies to impact plug-in electric vehicle prices and serve utility energy storage applications", *Journal of Power Sources*, vol. 196, no. 23, pp. 10351 –10 358, 2011, ISSN: 0378-7753.

[96] L. A.-W. Ellingsen, B. Singh, and A. H. Strømman, "The size and range effect: Lifecycle greenhouse gas emissions of electric vehicles", *Environmental Research Letters*, vol. 11, no. 5, p. 054 010, 2016.

[97] K. Yunus, H. Z. D. L. Parra, and M. Reza, "Distribution grid impact of plug-in electric vehicles charging at fast charging stations using stochastic charging model", in *Proceedings of the 2011 14th European Conference on Power Electronics and Applications*, 2011, pp. 1–11.

[98] G. Celli, G. G. Soma, F. Pilo, F. Lacu, S. Mocci, and N. Natale, "Aggregated electric vehicles load profiles with fast charging stations", in *2014 Power Systems Computation Conference*, 2014, pp. 1–7.

[99] M. The City of Columbus Andre J. Ginther, "Beyond traffic: The smart cyty challenge (phase 2)", in *Volume 1: Technical Application*, 2016. [Online]. Available: https://www.columbus.gov/WorkArea/DownloadAsset.aspx?id=2147487896.

[100] C. Heymans, S. B. Walker, S. B. Young, and M. Fowler, "Economic analysis of second use electric vehicle batteries for residential energy storage and load-levelling", *Energy Policy*, vol. 71, pp. 22 –30, 2014, ISSN: 0301-4215.

[101] J. Neubauer and A. Pesaran, "The ability of battery second use strategies to impact plug-in electric vehicle prices and serve utility energy storage applications", *Journal of Power Sources*, vol. 196, no. 23, pp. 10351 –10 358, 2011, ISSN: 0378-7753.