

PG0409

A Market Data Clustering Aimed to the Economic Analysis of an ESS-based Power Plant providing Ancillary Services

Simona Parrella¹, Vincenzo Galdi¹, Francesco Maria De Venezia¹, Vito Calderaro¹, Antonio Piccolo¹, Lucio Ippolito¹, Francesco Marzullo², Antonio Pascucci²

¹ University of Salerno - Department of Industrial Engineering
via Giovanni Paolo II, 132 - 84084 - Fisciano (SA) - Italy
Tel.: +39-089-964284
vgaldi@unisa.it

² TERNA RETE ITALIA S.p.A.
via Palmiano 101 - 00138 Roma - Italy

Abstract

The increasing use of Non-Programmable Renewable Energy Sources (NP-RES) in power systems determines very strong effects on the grid. The intrinsic uncertainty of the NP-RES causes strong power unbalance problems with consequent repercussions on the frequency and voltage stability; this requires greater use of balancing resources and Ancillary Services. One of the most promising solution to face the issues related to the spread of NP-RES is the use of Energy Storage Systems (ESSs), but their high cost makes ESSs still not very widespread, limiting its use to pilot plants. However, there may be particular market conditions that make it right to use the ESS to support the network.

In this paper, a hierarchical clustering method for the offers accepted on the electricity market is used to evaluate the affordability of ESSs able to provide ancillary services in different market areas. The proposed approach is applied to a historic series of Ancillary Services ex-ante market data, by using the Pearson coefficient to highlight the similarities between the historical sub-series of months. Then, by applying the Ward method, similar months are grouped together in order to extract the market characteristics of interest such as average price and share of hours of supply of resource in the market.

The presented methodology, formalized in a general way, is applied to a case study based on real data of the dispatching services market of the Southern Italy area where the penetration of NP-RES is particularly high.

Introduction

The new paradigm of the power system is based on increasingly widespread use of Non-Programmable Renewable Energy Sources (NP-RES) to replace carbon-intensive power sources reducing significantly global warming emission. Europe leads this process: for instance, in Italy from 2008 to 2017, photovoltaic (PV) and wind power plants have shown an average annual growth of 50% and 43%, respectively, [1], reaching at the end of 2017 more than 774.000 PV units (19.683 MW rated power) and 5.600 wind power units (9.766 MW rated power).

However, the growing development of NP-RES has also introduced significant issues in power system management, especially in the areas with high penetration of wind and solar power plant, changing the classical concept of the power system management, [2]. The traditional *one-directional* system (production → transmission → distribution → loads) is now being replaced by a more complex and integrated system characterized by *multi-directional power flows*, high volatility and low predictability [3].

To ensure the safe, reliable and resilient operation of the grid, it is necessary to guarantee, instant by instant, the balancing between the power generated by the production units and the power absorbed by the loads. The intrinsic uncertainty of NP-RES, highly dependent on weather conditions, and the impossibility of modulating their power produced leads to power imbalances that threaten the safe operation of the power grid. In fact, variability and uncertainty of renewable energy generation increase the cost of maintaining the short-term energy balance in power systems and, therefore, the increase of power regulation (*up reserve* and *down reserve*) used by the Transmission System Operator (TSO) [4], determining an economic impact that is evident from the prices of the Dispatching Services Market (DSM).

This situation determines very strong effects on the grid requiring a more efficient management and a wider use of Ancillary Services, especially in the areas with high penetration of wind and solar power plant [5, 6].

One of the most promising solution to face the issues related to the spread of NP-RES is the use of Energy Storage System (ESS). In fact, ESSs can provide multiple services and features to improve safety and reliability of power systems.

The installation of ESSs in areas of the country where the penetration of NP-RES is high, could be mitigating the volatility of renewable energy production and maintain the security and overall efficiency of the national electricity system, [7]. However, in order to select the proper storage technology able to support network services, it is advisable to analyze the characteristics of the ancillary services market. In fact, the large amount of available data and the dependence on non-recurring weather conditions that are very difficult to forecast, requires a careful pre-analysis.

In this paper, a clustering algorithm for the ancillary services market ex-ante, based on the Ward methodology, is proposed.

The proposed approach is applied on a historic series of ancillary service ex-ante market data, by using the Pearson coefficient to highlight the similarities between the historical sub-series of the months of the year. Then, by applying the Ward method, similar months are grouped together to extract the market characteristics of interest such as, average price and share of hours of supply of resource in the market.

The proposed methodology has been applied to evaluate the economic feasibility of ESS power plant by using real data of the Italian Ancillary Services Market ex-ante of Central South and South Italian areas.

1. The Electricity Market in Italy

The electricity market, namely the place where transactions involving electricity are conducted, was set up in Italy as a result of Law no. 79 dated March 16, 1999 ("*Bersani Decree*") as part of the implementation of the EU directive on the creation of an internal energy market (Directive 96/92/EC repealed by Directive 2003/54/EC), [8].

The electricity market is divided into:

1. Day-Ahead Market (DAM)
2. Intra-Day Market (IDM)
3. Dispatching Services Market (DSM)

In the DAM and IDM - also referred to Energy Markets - producers, wholesalers and end customers, together with network operators, such as in Italy *Acquirente Unico* (AU) and *Gestore dei Servizi Energetici* (GSE), buy and sell wholesale quantities of electricity for the next day. These markets, which are managed by *Gestore dei Mercati Energetici* (GME), define system marginal prices at which the energy is traded. Whereas in DAM are defined the preliminary programs of production and withdrawal of each offer point for the next day, in the IDM market operators negotiate the offers to purchase and sell electricity for each hour of the following day, for the purpose of modifying the injection and withdrawal programs defined to satisfy new requirements not foreseen in the DAM.

In the DSM, the Italian TSO, *TERNA*, procures the resources it needs to manage and control the national electric system (solving intra-national congestions, creating energy reserves, real-time balancing, etc.), accepting bids/offers from market participants related to different reserve and balancing services [8].

The DSM is divided into:

- ex-ante DSM: 6 sub-sessions, where *TERNA* trades energy and balancing services in order to release congestions and to create reserve margins (secondary and tertiary reserve);
- Balancing Market (BM): 6 sub-sessions, where *TERNA* trades real-time balancing services to restore secondary/tertiary reserve and to maintain the grid balancing.

The DSM is divided into market zones defined considering:

- the transport capacity of the lines
- the location of the power production and absorption, [8].

In Italy the market areas have been redefined starting from Jan, 1st 2019 to consider new connections to Malta (already operational) and to Montenegro (under construction). However, the paper refers to market areas active until Dec 31st 2018 described in Table 1.

Table 1 - The Italian market zones as defined up to December 31st 2018.

Acronym	Market Zone Name	Geographical Area
NORD	Northern	Valle D'Aosta, Piemonte Liguria, Lombardia, Trentino Alto Adige, Veneto, Friuli Venezia Giulia, Emilia Romagna
CNOR	Centre - North	Toscana, Umbria, Marche
CSUD	Centre - South	Lazio, Abruzzo, Campania
SARD	Sardinia	
SUD	Southern	Molise, Puglia, Basilicata, Calabria
FOGN	Foggia	
BRNN	Brindisi	
ROSN	Rossano	
SICI	Sicily	
PRGP	Priolo G.	

To assess the cost of dispatching services in a market zone, it is useful to observe the differential price, namely the difference between rising prices (*TERNA's* energy purchase price) and falling prices (*TERNA's* energy sale price). In general, the more the difference is high, the more the situation is economically unfavorable since the energy that *TERNA* buys has a very high price compared to the price with it is sold; this condition shows a suffering of the network generally connected to a poor regulation power offered by the groups in production, as when the production from renewable is very high and variable. All this is confirmed by the data [9]: for example, with reference to the annual average price of the offers accepted on the ex-ante DSM in the period October 2015 - September 2018 reported in Figure 1, it is shown how in the CSUD, BRNN, FOGN and SOUTH areas, where most of the NP-RES energy production is installed, the differential price is very high, with a peak value in the CSUD area of 209 € per MWh, approximately the double of the average value of the other areas.

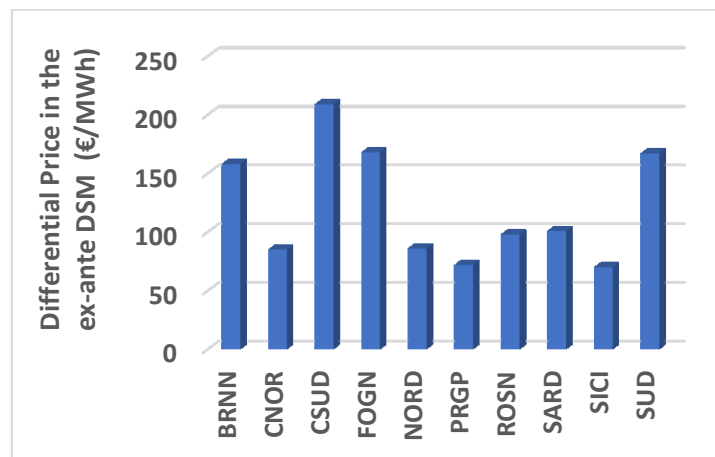


Figure 1 - Average Differential Price in the ex-ante DSM for the different market zones.

2. The Proposed Methodology

In this Section, the proposed methodology will be defined. In order to develop a procedure useful for assessing the sustainability of an investment in storage systems for network service erection based on the analysis of ex-ante DSM market data, considering the large amount of data to be analysed, the proposed approach achieves data clusterisation with the aim of aggregating similar market periods for each of which it makes sense to use an average price.

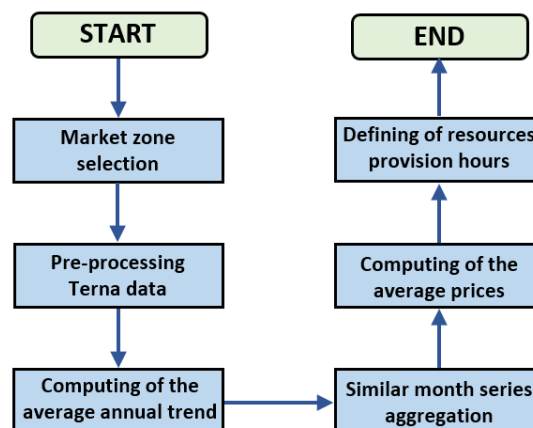


Figure 2 - The flowchart of the proposed methodology.

Figure 2 shows the flowchart of our procedure applied to the time series of the offers accepted on the ex-ante DSM. The procedure can be applied for one or more market areas in order to identify the different average prices of the offers and the hours of resource provision in the market. In such a way, it is possible to economically evaluate ESS investments and their profitability in each market zone.

After the market zones selection, the input time series are filtered removing noise by smoothing. In particular, in the second step, starting from *TERNA* data, the input time series of the offers accepted on the ex-ante DSM, are pre-processed through a moving average algorithm. In particular, a 3-period moving average is used, [10]:

$$x_j^* = \frac{x_{j-1} + x_j + x_{j+1}}{3} \quad \forall j = 2, K, d$$

where d is the length of the input time series and x_j^* is the j -th element of the filtered one.

In the third step, the aim is to obtain the average year trend of the offers accepted on the ex-ante DSM through similarity assessment between the times series. Very often, the Pearson coefficients (or correlation coefficients) are used. Given two time series:

$${}^1X = \{{}^1X_1, {}^1X_2, \dots, {}^1X_T\} \text{ and } {}^2X = \{{}^2X_1, {}^2X_2, \dots, {}^2X_T\}$$

the Pearson coefficient is defined as [10]:

$$R({}^1X, {}^2X) = \frac{\sum_{t=1}^T ({}^1X_t - {}^1X_m)({}^2X_t - {}^2X_m)}{\sqrt{\sum_{t=1}^T ({}^1X_t - {}^1X_m)^2} \sqrt{\sum_{t=1}^T ({}^2X_t - {}^2X_m)^2}}$$

where 1X_m and 2X_m are the average value of the first and second data series, respectively:

$${}^1X_m = \frac{1}{T} \sum_{t=1}^T {}^1X_t \text{ and } {}^2X_m = \frac{1}{T} \sum_{t=1}^T {}^2X_t$$

It is worth to note that the Pearson coefficient can only assume values between -1 and +1: in particular, the positive sign indicates that the two variables increase or decrease together (positive linear relationship), while the negative sign indicates that with the increase of a variable the other decreases and vice versa.

In the proposed methodology, we indicate as S_i the annual time series of the offers accepted on the ex-ante DSM related to the year $i = 1, 2, \dots, n$; each time series S_i , therefore, consists of the month time series G_n, F_n, \dots, D_n (January, February, ..., December) of the n -th year.

$$S_1 = \{G_1, F_1, \dots, D_1\}, S_2 = \{G_2, F_2, \dots, D_2\}, K, S_n = \{G_n, F_n, \dots, D_n\}$$

To obtain the average annual trend of the offers accepted on the ex-ante DSM, (step 3 of Figure 3) it is necessary to assess the dissimilarity between identical months belonging to different years. For instance, considering the January time series for different years, the related Pearson coefficients matrix R_G is defined as follow:

$$R_G = \begin{pmatrix} 1 & R(G_1, G_2) & R(G_1, G_3) & \dots & R(G_1, G_n) \\ R(G_2, G_1) & 1 & R(G_2, G_3) & \dots & R(G_2, G_n) \\ \dots & \dots & \dots & \dots & \dots \\ R(G_n, G_1) & R(G_n, G_2) & R(G_n, G_3) & \dots & 1 \end{pmatrix}$$

Then, the January's average month G_M is defined as follow:

$$G_M = \frac{1}{n} (G_{1,M} + G_{2,M} + \dots + G_{n,M})$$

where the terms $G_{i,M}$ are the pairs that have maximum Pearson coefficients along the line. Finally, by repeating this procedure for each month, the average year trend is obtained:

$$Year_{avg} = \{G_M, F_M, \dots, D_M\}$$

The following step is the clustering procedure aimed to group similar month time series within the average annual trend. In this paper, we use the hierarchical clustering approach proposed in [119] consisting in a four-step procedure to identify different clusters:

1. identification and union of the most similar elements, i.e. those at the smallest distance in the distance matrix (according to the *Ward method*), so as to form the first group. At this point, there are $n-1$ groups, in which, one is formed by two objects and the others $n-2$ by a single object;
2. determination of a new distance matrix (the new matrix dimension is reduced by 1), obtained by calculating the distance of the obtained group with respect to the other groups;
3. identification of the couple with the smallest distance (according to the *Ward method*) and grouping in a single cluster;
4. repetition of steps 2) and 3) until all the elements are united in a single cluster.

It is important to define a measure of similarity between the objects. However, the Pearson coefficients are not a similarity measure; in order to obtain a distance measure, starting from the Pearson coefficients, it is possible to use the relationship defined by Golay et al., [12]:

$$D(^1X, ^2X) = \sqrt{2[1 - R(^1X, ^2X)]}$$

At each iteration of the clustering procedure, we have different clustering solutions: in order to define the proper cluster number and quantify the goodness of the obtained clustering solution, it is necessary to compute the Silhouette coefficient as defined in the following. Considering the time series X_i within the cluster C_k , a measure of the intra-cluster variance $a(X_i)$ is obtained by measuring the average distance of X_i from the other time series of the cluster. The separation between different clusters is defined by measuring the minimum mean distance $b(X_i)$ between X_i from the other clusters, [10]. More in detail:

$$a(X_i) = \frac{1}{n_k - 1} \sum_{j=1, j \neq i}^{n_k} D(X_i, X_j) \quad b(X_i) = \min_{k' \neq k} \left\{ \frac{1}{n_{k'}} \sum_{j \in C_{k'}} D(X_i, X_j) \right\}$$

where n_k is the element number in the cluster C_k . The Silhouette coefficient for each series is defined below, and it ranges from -1 to +1.

$$S(X_i) = \frac{b(X_i) - a(X_i)}{\min\{a(X_i), b(X_i)\}}$$

The higher is the Silhouette coefficient the more the time series belongs to a given cluster. Finally, through the calculation of the average Silhouette value, it is possible to have a global indication of the goodness of the obtained clustering solution.

$$S = \frac{1}{K} \sum_{K=1}^K \frac{1}{n_k} \sum_{i \in C_k} S(X_i)$$

The clustering procedure is implemented in an iterative algorithm that stops when the Silhouette coefficient computed at the current iteration is equal to the silhouette coefficient computed at previous iteration. The clustering procedure ends when this condition is achieved, then the cluster number and their elements are defined.

In order to apply this clustering procedure for the zone market prices application, consider the month time series in the average year trend:

$$G_M = \{G_1, G_2, \dots, G_{31}\}, F_M = \{F_1, F_2, \dots, F_{28}\}, K, D_M = \{D_1, D_2, \dots, D_{31}\}$$

Each time series consist of a number of points coinciding with the number of days in the same month time series. In order to apply the clustering procedure, it is necessary to calculate the Pearson coefficients matrix R_T :

$$R_T = \begin{pmatrix} 1 & R(G_M, F_M) & R(G_M, M_M) & \dots & R(G_M, D_M) \\ R(F_M, G_M) & 1 & R(F_M, M_M) & \dots & R(F_M, D_M) \\ \dots & \dots & \dots & \dots & \dots \\ R(D_M, G_M) & R(D_M, F_M) & R(D_M, M_M) & \dots & 1 \end{pmatrix}$$

Starting from the R_T matrix and by applying the Xavier-Golay relationship, we calculate the dissimilarity matrix for grouping the most similar month time series.

After the clustering procedure, in the fifth step we proceed with the identification of the average offer price in each cluster: we compute the average value of all offers related to the same day of different months within a single cluster. After that, the average value of the results obtained for a single day are computed to find a single representative price of the cluster.

Finally, we have computed the hours of resources provision in the DSM: they can be directly calculated from the data available on the *TERNA* website.

An hourly offer day can be modeled in a vector of 24 points, where each element represents a price. In particular, if there have been no resources provisioning at the i -th hour, the i -th cell of the vector is empty. Therefore, by calculating the number of no-empty cells, the resources provision hours within a day can be calculated.

3. Case Study

We use as input time series of the proposed cluster methodology the offers accepted on the ex-ante DSM from October 2015 to September 2018, [9]. Moreover, in order to verify the effectiveness of the proposed methodology, we choose the Internal Rate of Return (IRR) as metric to evaluate the economic profitability of four different ESS battery technologies located in market areas where the penetration of NP-RES is higher. Table 2 compare the most relevant features for the four considered ESS technologies, [13, 14]. ESS rated power and energy values, used for the economic analysis, are extracted from Terna public reports related to the first and second year of ESS experimentation in south of Italy, [13].

3.1. Clustering Results and Economic Analysis

Clustering analysis have highlighted that market areas where the penetration of NP-RES is higher, and thus the investment for ESS could be the more profitable, are CSUD, SUD and

FOGN. Hence, in this Section, we summarize obtained results only for these market zones.

CSUD market zone: the proposed procedure allows obtaining four clusters characterized by Silhouette coefficient equal to 0,94, as showed in Table 3. A clustering solution with five clusters shows lower performance with a Silhouette coefficient equal to 0,86

Table 2 - Comparison between ESS battery technologies.

Technology	Lead-Acid	NaS	Lithium-ions	Flow Battery
Rated energy [MWh]	80			
Rated power [MW]	12			
Batteries lifetime [cycles]	750	4500	2750	12000
Charging time [hours]	10			
Discharging time [hours]	7,5			
ESS lifetime [years]	15			
Annual charging/discharging cycles [cycles]	50	300	183,3	800
Annual ESS charging hours [hours]	500	3000	1833,3	8000
Annual ESS discharging hours [hours]	375	2250	1375	6000
Average charging power [MW]	8			
Average discharging power [MW]	10,67			
ESS cost [€/kWh]	264	688	1320	506
Total investment cost [€]	21.120.000	55.000.000	105.600.000	40.480.000

Table 3, - Clustering results for the CSUD area.

Clusters	Month	Average price [€/MWh]		Resources provision hour [%]
		Go up	Go down	
1	June	268,73	20,42	64
2	October	151,65	16,67	47
3	January, April, March, August	212,72	15,27	61
4	February, May, July, September, November, December	200,85	16,67	55

Table 4 - Economic analysis for the CSUD area.

	Lead-Acid	NaS	Lithium-ions	Flow Battery
Annual revenue [€]	993.240	4.870.208	3.151.195	6.787.075
O&M cost [€]	6006			
IRR	-4%	4%	-9%	15%

Table 5 - Clustering results for the FOGN area.

Clusters	Months	Average price [€/MWh]		Resources provision hour [%]
		Go up	Go down	
1	July	123,59	13,93	22

2	August, September	122,49	23,65	13
3	June	290,25	22,78	38
4	January, March, April, May, October, November and December	132,25	21,03	34
5	February	156,15	17,22	35

FOGN market zone: the proposed procedure allows obtaining four clusters characterized by Silhouette coefficient equal to 0,84, as showed in Table 5. A clustering solution with five clusters shows lower performance with a Silhouette coefficient equal to 0,78.

Table 6 - Economic analysis for the FOGN area.

	Lead-Acid	NaS	Lithium-ions	Flow Battery
Annual revenue [€]	931.399,57	3.195.409,4	2.157.353,07	3.602.976,53
O&M cost [€]	6006			
IRR	-5%	-2%	-12%	4%

SUD market zone: the proposed procedure allows obtaining four clusters characterized by Silhouette coefficient equal to 0,92, as showed in Table 7. A clustering solution with five clusters shows lower performance with a Silhouette coefficient equal to 0,81

Table 7 - Clustering results for the SUD area.

Clusters	Months	Average price [€/MWh]		Resources provision hours [%]
		Go up	Go down	
1	October	217,51	0,88	20%
2	December	201,24	0,03	28%
3	January, February	150,89	0,03	31%
4	March, June, July e August	129,25	0,06	30%
5	April, May, September and November	158,10	1,32	38%

Table 8 - Economic analysis for the SUD area.

	Lead-Acid	NaS	Lithium-ions	Flow Battery
Annual revenue [€]	817.973,44	3.822.701,6	2.037.094,7	4.562.701,97
O&M cost [€]	6006			
IRR	-6%	1%	-13%	7%

Obtained results show that the CSUD area is characterized by the higher resources provision hours supplied and the purchased energy than other areas, while the SUD area is characterized by having very low sales prices. Therefore, the economic analysis showed that the lead-acid and lithium-ion storage battery based ESSs present a negative IRR in all the considered market areas. The NaS battery based ESS, instead, is characterized by positive IRR in the CSUD and SUD areas while it is negative in the FOGN area. Finally, the flow battery based ESS, due to the highest operating cycle number, allows having positive IRR in all the market zones.

4. Conclusions

A methodology for the technical and economic evaluation of ESSs able to provide ancillary services within the Italian electricity market is presented. In order to identify the most economically convenient ESS technology in different market areas, the features of the DSM market have been analyzed through a hierarchical clustering procedure. This approach, using the Pearson coefficient matrix, has been applied to 2015-2018 time series of the ex-ante DSM data.

Results of the clustering procedure are used to evaluate the economic feasibility of lead-acid, NaS, lithium-ions, and flow battery based ESSs in the CSUD, FOGN and SUD market areas, where the penetration of NP-RES is higher. Assuming, a fifteen-year investment lifetime, results has been show that lead-acid and lithium-ion based ESSs present a negative IRR for all the market areas, whereas NaS and flow battery based ESSs are characterized by positive IRR due to the high number of operating cycles.

References

- [1] Terna, "Renewable Energy Sources". Available on-line at: <https://www.terna.it/en-gb/sistemaelettrico/fontirinnovabili.aspx>.
- [2] S. Favuzza, M. G. Ippolito, A. Madonia, F. Massaro, E. M. Carlini, C. Quaciari, "Impact of non-programmable renewable sources in Day-Ahead Market in Sicily. Reasons to make more flexible and smarter the national transmission grid," in IEEE PES Innovative Smart Grid Technologies Europe, 2014, pp.1-5.
- [3] A. Richter, E. van der Laan, W. Ketter, K. Valogianni, "Transitioning from the traditional to the smart grid: Lessons learned from closed-loop supply chains," in International Conference on Smart Grid Technology, Economics and Policies, 2012, pp. 1-7.
- [4] A. Calabria, R. Capata, M. Di Veroli, S. Sangiorgio, "Energy market impact of renewable energy source," in IEEE International Conference on Environment and Electrical Engineering, 2015, pp.1-6.
- [5] Massa G., Gross G., Galdi V., Piccolo A., Dispersed Voltage Control in Microgrids, in IEEE Transactions on Power Systems, 2016, Vol. 31, n. 5, Pages: 3950 - 3960, DOI: 10.1109/TPWRS.2015.2503393
- [6] Lamberti, F., Calderaro, V., Galdi, V., Piccolo, A., Ancillary services provided by residential ESSs in LV networks: Assessing the opportunity costs, in Proc. of IEEE Manchester PowerTech, 2017
- [7] Nan Li, Kory W. Hedman, "Economic Assessment of Energy Storage in Systems with High Levels of Renewable Resources," in IEEE Transactions on Sustainable Energy, vol. 6, no. 3, pp. 1103-1111, July 2015.
- [8] Terna. Electricity Market. Available on-line at: <http://www.terna.it/en-gb/sistemaelettrico/mercatoelettrico.aspx>.
- [9] Terna. Dispatching. Available on-line at: [https://www.terna.it/it-it/sistemaelettrico/dispacciamento/ valoridelmercatoeiservizididispacciamento.aspx](https://www.terna.it/it-it/sistemaelettrico/dispacciamento/valoridelmercatoeiservizididispacciamento.aspx)
- [10] H. Lütkepohl, "New Introduction to Multiple Time Series Analysis," Springer, 2005.
- [11] SungSik Park, Young B. Park, "Photovoltaic power data analysis using hierarchical clustering," in Int. Conference on Information Networking, 2018, pp.727-731.
- [12] X. Golay, S. Kollias, G. Stoll, D. Meier, A. Valavanis, "A New Correlation-Based Fuzzy Logic Clustering Algorithm for FMRI," in Magnetic Resonance in Medicine, vol. 40, no. 2, pp. 249-260, August 1998.

- [13] Terna, Energy Storage Systems. Available on-line at: <https://www.terna.it/en-gb/sistemaelettrico/progettipilotadiaccumulo.aspx>
- [14] IRENA – International Renewable Energy Agency, “Electricity Storage and Renewables: Costs and Markets to 2030,” October 2017. Available on-line at: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2017/Oct/IRENA_Electricity_Storage_Costs_2017.pdf