Generation Expansion Planning under Wide-Scale RES Energy Penetration

K. Tigas, J. Mantzaris, G. Giannakidis, C. Nakos, N. Sakellaridis

Energy Systems Analysis Lab.
Energy Policy Division

Centre for Renewable Energy Sources and Savings
19th Km Marathonos Ave.
Pikermi, Attiki
GREECE

Contents

I. Overview

Renewable energy penetration in electricity is expected to have a spectacular growth in the forthcoming years. Power systems and the codes of their operation will be modified to take into account the specific characteristics of variable renewable energy operation (wind energy, photovoltaics). A number of issues arise in the context of developing new generation expansion planning methodologies, under a large scale penetration of renewable energy. A probabilistic approach is necessary due to the strong stochastic nature of variable (non dispatchable) renewables incorporating the statistics of the customer load and the non dispatchable RES generation. The level of penetration of variable renewable energy has to be calculated according to restrictions implied by the energy curtailment which can occur when the customer load is low and RES generation is high. Hydro pumped storage plants decrease this curtailment and consequently increase the level of RES penetration. In addition, fast reserve capacity is required to deal with big variations of variable RES energy generation. In summary the necessary storage and reserve capacity have to be calculated and incorporated in the costs of different scenarios related to expansion planning or operation simulation. In addition, grid costs mainly related to the Transmission System Expansion have to be calculated since they are related to the penetration of areas with high RES potential. In the present project an attempt is made to incorporate such costs in the TIMES model in order to have a more accurate approach of generation expansion scenarios, incorporating storage, reserve plants and Transmission System expansion in the cost of RES technologies. Using the results as an input a more detailed probabilistic approach is used to calculate the relevant production costs under a wide-scale penetration of RES. Methodologies presented here are based on the so called residual load duration curves approach.

II. Cost functions regarding renewable technologies in the TIMES model

Expansion planning of power systems under environmental constraints is leading to a rather complex techno-economic analysis involving large scale penetration of Renewable Energy Sources (RES) investments. RES investments economic analysis should normally include costs related to balancing units required due to variable renewables namely pumped storage plants and gas turbines and costs related to grid expansions necessary for a wide-scale penetration of renewables. In addition, to obtain an accurate least cost solution for RES penetration under a number of energy and environmental policy constraints one should take into consideration the utilization factor of renewables in a specific area mainly related to geographical position and meteorological conditions. Geographical information on economic RES potential can then be derived combining all the above parameters. To obtain an accurate least cost solution for RES penetration under a number of energy and environmental policy constraints it is necessary to develop geographical databases with RES potential which will be connected to the TIMES model describing areas including classes of the same RES technology with different costs. A TIMES model solution will then prioritize different classes of RES investments according to their real cost which would incorporate balancing units costs, grid expansion and connection costs together with utilization factors of RES in specific areas. To this end, a system of RES generation areas has to be established and a probabilistic approach can be used for calculating the expected energy production from each RES generation area incorporating all relevant statistics. Integration of dispersed generation plants into grids must be examined from the network point of view as the interconnection scheme affects both the investment cost and the operation of the power grid. TIMES modeling should include a multi region approach were areas of electricity generation from RES with similar characteristics can be modeled having as main parameter the cost of the interconnection of the area with the existing grid. A stepwise cost function can be established for each interconnection case to relate the level of transmission grid expansion with RES capacity interconnected. A necessary external step however is to provide as input the basic components of the grid for the time horizon of the solution and for the load levels forecasted in TIMES so that optimum generation-transmission combinations are established.

III. The Residual Load Duration Curve (RLDC)

In probabilistic production costing methodologies for dispatchable power plants [1], [2], there is a wide use of Load Duration Curves (LDCs) or Equivalent Load Duration Curves (ELDCs) which are the basis of the relevant expansion planning methodologies. For large penetration of non-dispatchable electricity plants (wind, pv, run of river hydro and CHP) a similar methodology can be used and the essential problem to be resolved is to calculate the remaining load (residual load) for dispatchable units (thermal, dam type hydro).

Non-dispatchable electricity generation depends on weather and other random parameters and therefore has a strong stochastic nature. Input data required for probabilistic generation expansion planning require a thorough statistical analysis of non dispatchable generation data combined with the relevant customer load data.

As a first step a stochastic analysis of non-dispatchable energy generation data is performed [7] and a number of scenarios for the exploitation of the economic potential of RES and CHP energy is developed according to the existing alternatives for a given future time period. Time series are determined based on a system of measuring historical data (wind masts, counters of wind turbines) simultaneously with the relevant customer load. These data represent a forecast of the production of variable RES units for a long run time period, and usually using one hour intervals. Further statistical analysis is then performed to obtain the cumulative distribution function (CDF) and the probability density function(PDF) of the different forms of non-dispatchable generation and the customer load.

In the second step the statistical analysis described in step one is used as input to calculate a Residual Load Duration Curve through the convolution of the customer load with non-dispatchable energy generation (hourly zones are used for small correlation between load and RES energy). If the probability density function of non dispatchable energy source j is simulated with the equation:

$$f_j(x) = \sum_{i=1}^{k} p_i \cdot \delta\left(x + C_i^j\right) \tag{1}$$

and C_i are the levels of non dispatchable generating capacity from energy source j with probabilities p_i and δ is the Dirac function, then the derived residual load curve[10] is :

$$\begin{aligned} I_{L-j}(x) &= Prob\left[L - C_j \geq x\right] = 1 - F_{L-j}(x) = 1 - \int_{y=-\infty}^{+\infty} \int_{z=-\infty}^{x+y} \mathrm{f_j}(\mathrm{y}) \cdot \mathrm{f_L}(\mathrm{z})\mathrm{dzdy} \\ &= \sum_{k=1}^{n} p_k^j \; I_L\left(x + C_k^j\right) \end{aligned} \tag{2}$$

L is the customer load, C_j is the level of generating capacity from variable renewable source j, *F* is the CDF (cumulative distribution function) of the convolution of the customer load with the non dispatchable generation source j. The above equation is valid if L and j are independent and the relevant correlation coefficient is small. If load and RES generation are correlated we consider 24 time zones per day and we apply the equation for the different zones.

The residual duration curve is then derived by considering conditional probability:

$$p(x) = \sum_{i=1}^{N} p(x/y_i)\, p(y_i) \tag{3}$$

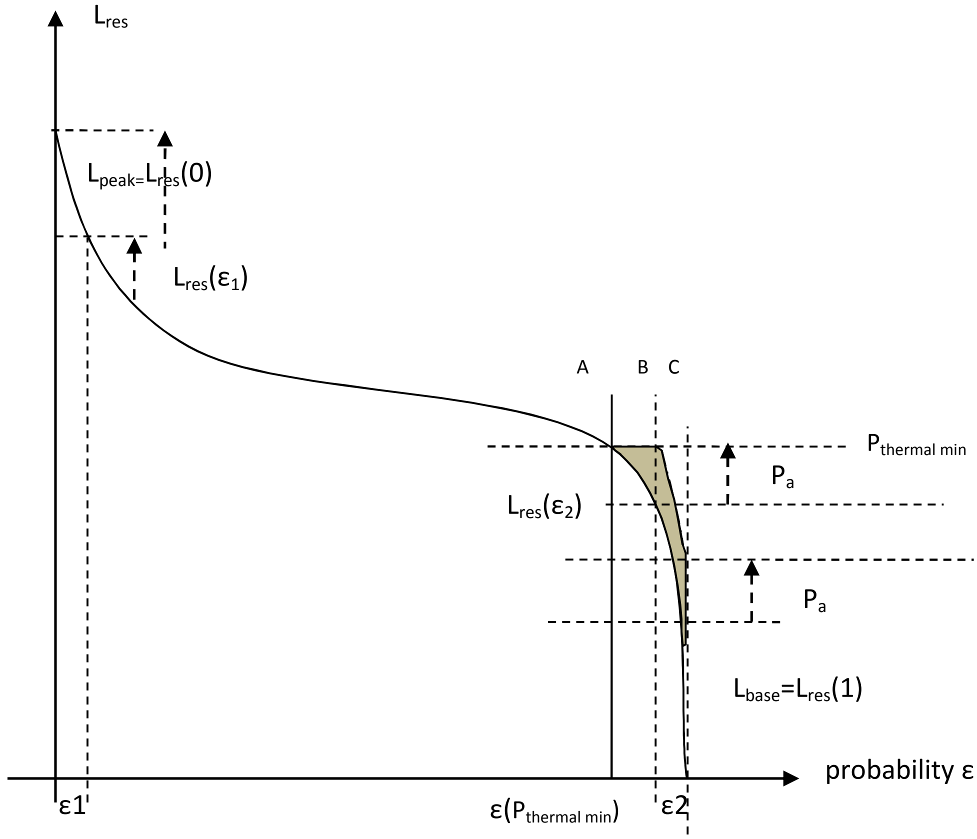

Figure 1 : Residual Load Duration Curve

Which for 24 hourly zones is :

$$p(x) = \sum_{i=1}^{N} p(x/h_i) \frac{1}{24} \tag{4}$$

Based on the “residual” load duration curve we can define a least cost combination of thermal plants and reservoir-type hydro plants. The level of penetration of non dispatchable energy [9] has to be a parameter of the solution in order to keep the level of non dispatchable electricity curtailment below a certain limit. Non dispatchable energy curtailment is related to the technical minimum of thermal power of the generation system and can be reduced either by selecting generation technologies with decreased technical minimum, or by using sufficient capacity of pumped storage plants P_a.

The pumped storage capacity normally is not balancing 100 % of the potential curtailement. For economic reasons the probability (or the number of hours) for curtailment is restricted by a parameter ε_2 with the equation :

$$Prob\left[L_{res} \geq \sum_i p_i C_i^{th-min}\right] = 1 - F_{L_{res}}(\sum_i p_i C_i^{th-min}) \geq \varepsilon_2 \tag{5}$$

L_{res} is the residual Load, $C_i^{th-min}$is the thermal minimum of thermal plant i and $p_i$is the availability of thermal plant i

Penetration of non-dispatchable energy plants can be increased if we increase the pumping capacity of pumped storage plants and satisfy the equation:

$$P_a = \sum_i p_i C_i^{pump} = \sum_i p_i C_i^{th-min} - L_{res}(\varepsilon_2) \tag{6}$$

where $C_i^{pump}$is the capacity of each pumped storage plant, p_i is the storage plant availability. L_{res} is the known residual load duration curve and P_a is the available pumping capacity which can be calculated from (6) since $p_i, C_i^{th-min},\ L_{res}(\varepsilon_2)$ can be calculated.

The remaining RES energy curtailment is then :

$$E_{curt} = 1/2(1 - \varepsilon_2) \cdot (\textstyle\sum_i p_i C_i^{th-min} - L_{res}(1)) \tag{7}$$

Based on similar considerations it is possible to calculate the additional thermal peak units capacity required in order to maintain a Loss of Load Probability equal to ε_1:

$$Prob\left[L_{res} \geq \sum_i p_i C_i^{peak}\right] = 1 - F_{L_{res}}(\sum_i p_i C_i^{peak}) \geq \varepsilon_1 \tag{8}$$

or

Additional thermal peak load capacity must in total satisfy the equation:

$$\sum_i p_i C_i^{peak} = L_{res}(\varepsilon_1) - \sum_j p_j\, C_j \tag{9}$$

C_j is the capacity of the thermal power plant j which has availability p_j

Since pumped storage units must operate for generation the peak units capacity is calculated as follows :

$$P_{peak} = \sum_i p_i C_i^{peak} = L_{res}(\varepsilon_1) - \sum_j p_j C_j^{base} - \sum_j p_j C_j^{intermediate} - \sum_j p_j C_j^{hydro} - \sum_k p_k C_k^{pump} \tag{10}$$

Regarding the production from storage plants it is assumed that they are exclusively used as balancing units to avoid RES energy curtailment thus pumped storage production must be equal to the shadowed area of the triangle in Figure 1 multiplied by their coefficient of performance :

$$E_{pp} = \sum_k n_k^a p_k C_k^{pump} \,(0.5(\varepsilon_2 - \varepsilon(P_{th\_min})) + (1 - \varepsilon_2)) = \sum_k n_k^a p_k C_k^{pump} \,(1 - 0.5(\varepsilon(P_{th\_min}) + \varepsilon_2)) \quad (11)$$

IV. Incorporating the RLDC in the TIMES model for the assessment of balancing units costs

The TIMES model can be used to calculate least cost generation expansion planning scenarios putting as constraint national environmental policies related to CO_2 emissions reduction, having as input scenarios of economic growth and evolution of energy technologies and incorporating competition of electricity with other energy sources (such as natural gas) in end use consumption.

The basic consideration made to incorporate the RLDC in TIMES is that customer load and non-dispatchable electricity generation are given in a number of seasonal time slices, including four values in each, namely peak hour, mean day, mean night, trough hour. Based on seasonal statistics, load factors for wind, pv, run of river hydro and CHP can be calculated for each time slice. Available power for mean day and mean night are derived from energy values by dividing the energy values with the corresponding duration (hours) of each time slice. The residual load duration curve presented on Figure 3 is then formulated in a simplified form based on equation (12).

$$L_{res} = L_{cust} - P_{wind} - P_{pv} - P_H - P_{CHP} \quad (12)$$

If E is the expected value (mean) of a random variable for a given time slice then:

$$E(L_{res}) = E(L_{cust}) - E(P_{wind}) - E(P_{pv}) - E(P_H) - E(P_{CHP}) \quad (13)$$

It is important to note that due to the fact that the residual load is approximately a Gaussian distribution, the mean value of the L_{res} lies at the centre of each interval. Then a piecewise linear curve is formulated using peak and trough hourly values at the edges and hourly mean values of L_{res} $L_{cust}, P_{wind}, P_{pv}, P_H, P_{CHP}$ at the center of each interval (Figure 3).

The available storage and peak load capacity required are :

$$P_a = \sum_i p_i C_i^{pump} = \sum_i p_i C_i^{th-min} - L_{res}(\varepsilon_2)$$
$$P_{peak} = \sum_i p_i C_i^{peak} = L_{res}(\varepsilon_1) - \sum_j p_j C_j^{base} - \sum_j p_j C_j^{intermediate} - \sum_j p_j C_j^{hydro} - \sum_k p_k C_k^{pump} \quad (14)$$

To use the linear solvers of TIMES the residual load duration curve is linearized and from equation (16) after defining the time intervals for ε_2, ε_1 we can obtain $L_{res}(1)$, $L_{res}(0)$ and $L_{res}(\varepsilon_2)$, $L_{res}(\varepsilon_1)$ using :

$$L = L_j + (\varepsilon - \varepsilon_j)\frac{L_{j+1} - L_j}{\varepsilon_{j+1} - \varepsilon_j} \quad (15)$$

Storage plants production must be equal to the area of polygon ABDE multiplied by the coefficient of efficiency for storage plants production. Energy curtailment is equal to the area of triangle BCD.

$$E_{curt} = 1/2(1-\varepsilon_2)\cdot(\sum_i p_i C_i^{th-min} - L_{res}(1))$$
$$E_{pp} = \sum_k n_k^a p_k C_k^{pump}\,(0.5(\varepsilon_2 - \varepsilon(P_{th\_min})) + (1-\varepsilon_2)) = \sum_k n_k^a p_k C_k^{pump}\,(1-0.5(\varepsilon\left(\left(P_{th\_min}\right)+\varepsilon_2\right)) \quad (16)$$

Table 1 : Pumped storage capacity required, calculated with equations (14)-(16) for three scenarios

Pumped Storage Capacity (MW)							
	2020	**2025**	**2030**	**2035**	**2040**	**2045**	**2050**
Base-new	1600	1618	1858	1858	1858	1858	2138
GHG-p60	1600	1600	2232,491	2232	2902,773	2902	3195,602
REN-p60	1600	1608,11	3583,336	4138,846	4412,048	4412	5143,479

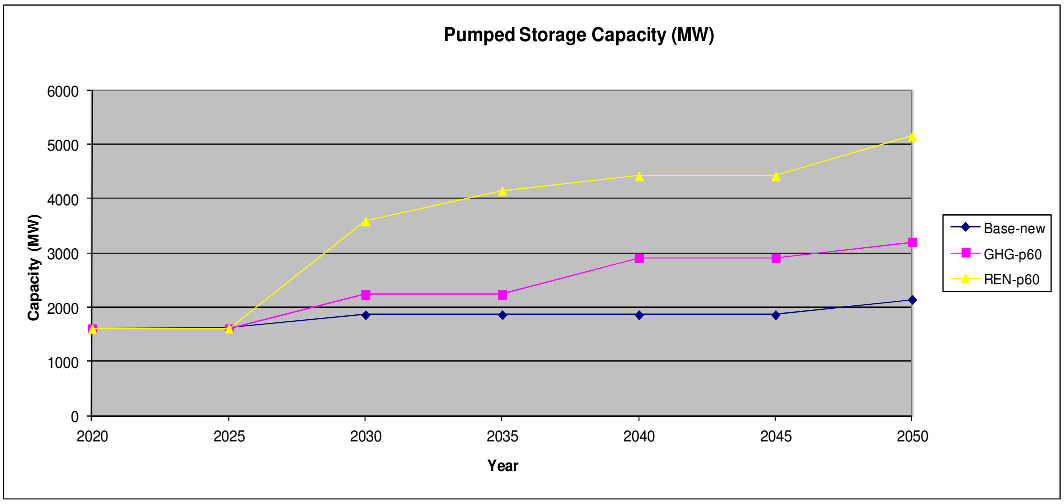

Figure 2 : Pumped Storage Capacity Required in the Greek generation system

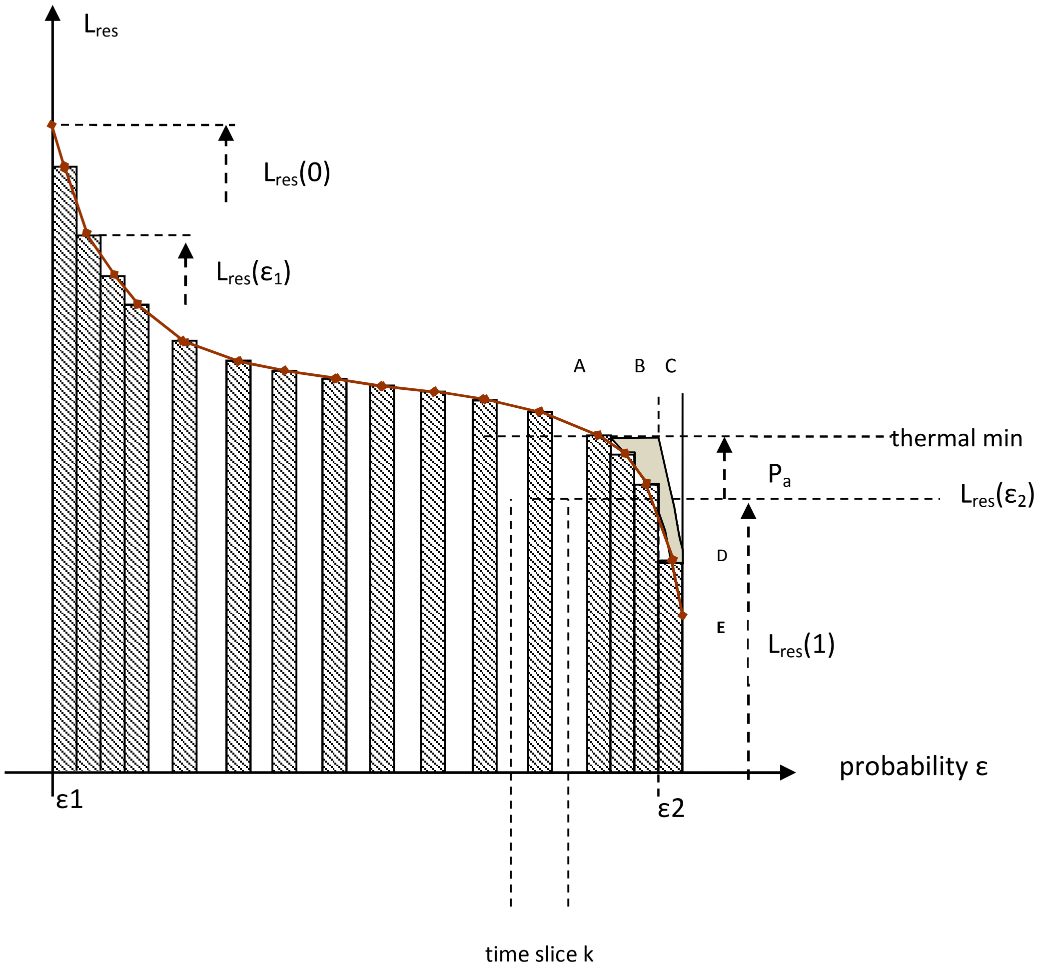

Figure 3 : Residual Load Duration Curve in TIMES model

Table 2 : Pumping required to minimize energy curtailment

	Pumping (GWh)						
	2020	**2025**	**2030**	**2035**	**2040**	**2045**	**2050**
Base-new	875,835	1315,306	1583,914	1180,632	1231,315	1201,403	1631,208
GHG-p60	1120,548	1541,701	2589,067	2209,35	3244,958	2543,646	3265,435
REN-p60	1120,548	1613,095	6013,327	6513,032	6517,124	6690,293	7625,006

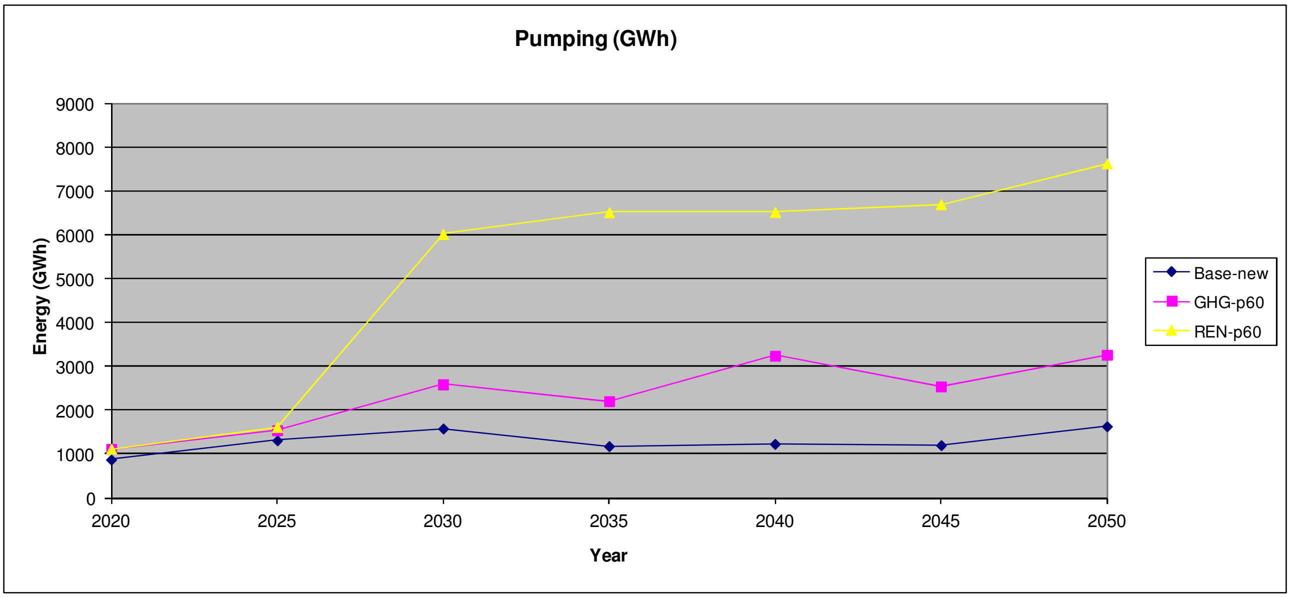

Figure 4 : Pumping required to minimize energy curtailment in the Greek generation system

Table 3 : Energy Curtailment in the Greek Generation System

Energy Curtailment (GWh)							
	2020	2025	2030	2035	2040	2045	2050
Base-new	80,35207	182,8493	174,823	120,7889	175,1433	253,7246	339,1494
GHG-p60	129,4365	241,8589	337,9753	364,8847	498,6121	407,4515	574,04
REN-p60	379,1192	252,3341	390,4794	502,2403	548,7968	601,3383	608,4288

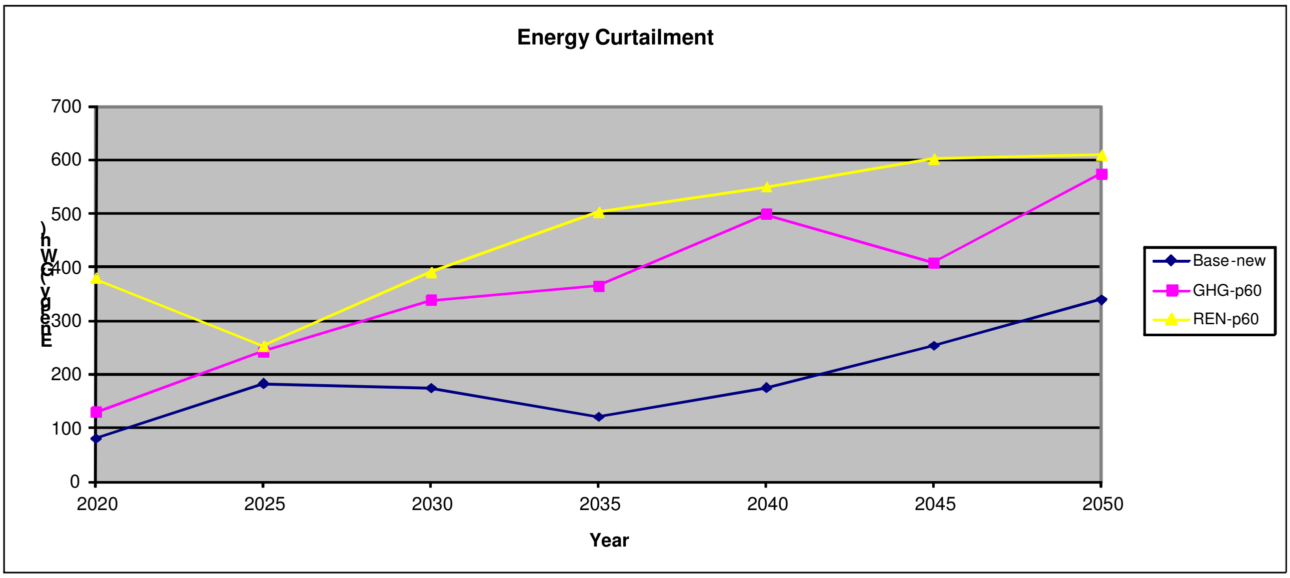

Figure 5 : Energy Curtailment in the Greek generation system after the use of pumped storage plants

V. Probabilistic Production Costing using the RLDC

A probabilistic simulation model was developed to calculate the operation cost of the generation system defined in the previous step. The methodology is using equations (1)-(9) for the calculation of the Residual Load Duration Curves through weekly or daily convolution of customer load with non-dispatchable RES generation using as input hourly based time series. Following, convolution of weekly or daily RLDCs with dispatchable thermal and hydro plants is used to define, the expected energy generation, the hours of operation of each block of the thermal plants and the relevant probabilistic production cost. The methodology used allows multi-block and multi-state simulation of thermal units.

The Long Run Average Cost of the electricity produced is calculated using the equation :

$$[LAVC_j] = f_{inv} C_j \frac{i(i+1)^N}{(i+1)^N - 1} + f_{O\&M} C_j + \sum_{nb=1}^{n} (v_{CO_2} + v_{O\&M} + v_{fuel})[E_{nb}] \qquad (17)$$

Equation (17) refers to a power plant and f is used for fixed costs, v for variable costs, C for plant capacity and $[E_{nb}]$ is the expected value of the energy produced by each block of a thermal unit. Equation (17) is used to assess the influence of large-scale penetration of RES energy to the cost of electricity generation and the economics of the conventional plants (Figure 3). The solution of the present methodology can be adjusted according to the level of the desired long-run average cost of electricity generation derived from (17).

Storage capacity required and RES energy curtailment for each time interval (week, day) are calculated using equations (5) - (11). Reserve capacity related to peak load is calculated through convolution of the Residual Load with the dispatchable plants outage PDFs.

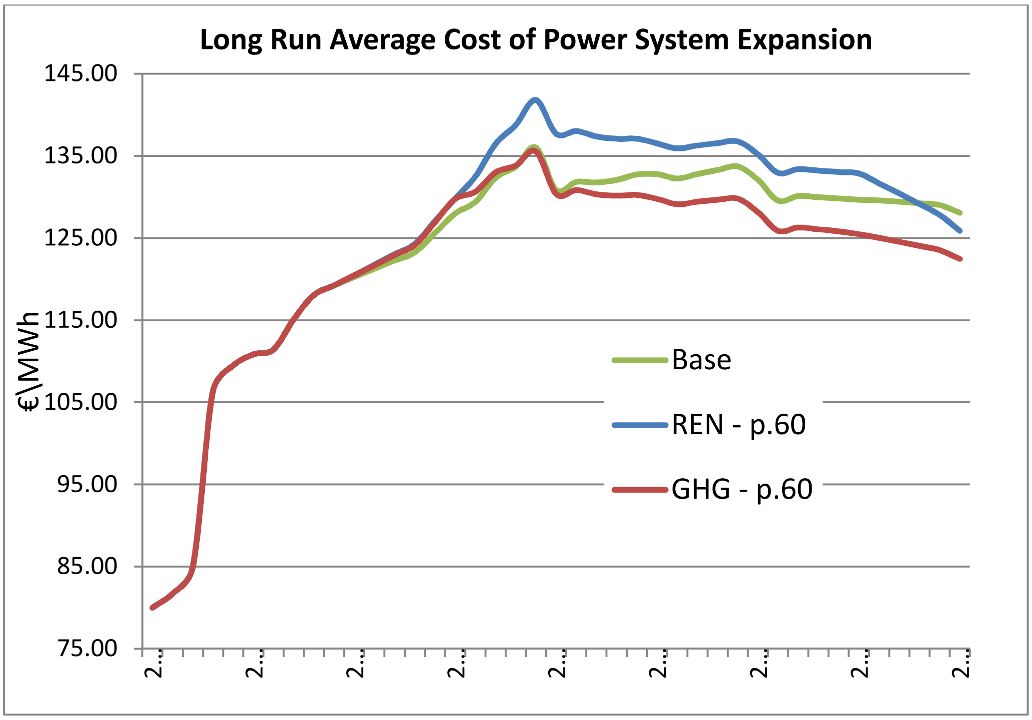

Figure 6 : Long Run Average Cost of the Greek Power System Expansion for three RES penetration scenarios

Table 4 : Long Run Average Cost of the Greek Power System Expansion

	GHG - p.60	Base	REN - p.60		GHG - p.60	Base	REN - p.60
2010	80,01	80,01	80,01	**2031**	130,84	131,83	138,02
2011	81,73	81,73	81,73	**2032**	130,32	131,76	137,37
2012	84,93	84,93	84,93	**2033**	130,15	132,09	137,08
2013	106,39	106,39	106,39	**2034**	130,25	132,74	137,09
2014	109,49	109,49	109,49	**2035**	129,77	132,79	136,52
2015	110,83	110,83	110,83	**2036**	129,12	132,25	135,91
2016	111,43	111,43	111,43	**2037**	129,42	132,79	136,27
2017	115,11	115,11	115,11	**2038**	129,65	133,28	136,56
2018	118,06	118,06	118,06	**2039**	129,80	133,71	136,77
2019	119,22	119,22	119,22	**2040**	128,16	132,10	135,19
2020	120,46	120,24	120,49	**2041**	125,87	129,55	132,93
2021	121,69	121,25	121,76	**2042**	126,29	130,12	133,39
2022	122,93	122,26	123,03	**2043**	126,06	129,98	133,22
2023	124,16	123,28	124,30	**2044**	125,82	129,82	133,04
2024	126,94	125,62	127,05	**2045**	125,48	129,65	132,86
2025	129,76	127,98	129,82	**2046**	125,02	129,57	131,66
2026	130,64	129,40	132,49	**2047**	124,55	129,43	130,43
2027	132,99	132,33	136,46	**2048**	124,05	129,25	129,14
2028	133,83	133,70	138,76	**2049**	123,54	129,02	127,82
2029	135,55	136,01	141,82	**2050**	122,46	128,08	125,88
2030	130,36	130,86	137,70				

Finally it is necessary to calculate reserve required to maintain a constant index of reliability during the residual Load hourly variations. The principle behind these reserve calculations is based on a methodology presented in [3]. According to this methodology it is assumed that a load shedding incident can happen in one of the following ways. The variation of the amount L_{res} which is known for time intervals of one hour is greater than the unknown reserve required R_h also calculated on an hourly basis. The previous case can happen together with a generator trip. Finally the previous case can happen together with two generator trips and then three or four but since the probability of an outage of three or more generators simultaneously is very small the formulation of the problem can be restricted to the outage of two generators. We assume that the probability for the variation of the amount L_{res} to exceed the reserve required R_h in the time interval h can be modeled with a normal distribution :

$$Prob[L_{res} \geq R_h\ ] = 1 - \Phi\left(\frac{R_h}{\sigma_{L_{res},h}}\right) \tag{18}$$

The probability of RES generation shedding in the time interval h when the residual Load is greater than the spinning reserve R_h or there is just one generator trip and the residual Load is greater than the spinning reserve is then :

$$PLSNO_h = \left(\prod_{i=1}^{G}\left(1 - FOP_{i,h}\right)\right) \times \left(1 - \Phi\left(\frac{R_k}{\sigma_{L_{res},h}}\right)\right) \tag{19}$$

$$+\sum_{i=1}^{G} FOP_{i,h} \times \left(\prod_{\substack{j=1 \\ j\neq i}}^{G} \left(1 - FOP_{j,h}\right) \right) \times \left(1 - \Phi\left(\frac{R_h - Cn\alpha fo_{i,h}}{\sigma_{L_{res},h}}\right)\right)$$

where $FOP_{i,h}$ is the forced outage probability of generator i in the time interval hand $Cn\alpha fo_{i,h}$ is the not available capacity of generator i in the time interval h. After extending (18),(19) to all possible cases and maintaining a constant probability of load shedding for all k we can solve the final set of equations (19), (20) and (21) and calculate R_h. It is obvious that for the time interval h the sum of the rump rates of the available generating units must be equal to R_h.

$$PLSFO_{k,h} = \left(\prod_{\substack{i=1 \\ i\neq k}}^{G} \left(1 - FOP_{i,h}\right) \right) \times \left(1 - \Phi\left(\frac{R_h - Pn\alpha fo_{k,h}}{\sigma_{L_{res},h}}\right)\right)$$
$$+\sum_{\substack{i=1 \\ i\neq k}}^{G} FOP_{i,h} \times \left(\prod_{\substack{j=1 \\ j\neq i \\ j\neq k}}^{G} \left(1 - FOP_{j,h}\right) \right) \times \left(1 - \Phi\left(\frac{R_h - Pn\alpha fo_{k,h} - Pn\alpha fo_{i,h}}{\sigma_{L_{res},h}}\right)\right) \tag{20}$$

$$PLS_h = PLSNO_h + \frac{1}{2}(Hr) \times \left[FOP_{1,h}\ FOP_{2,h}\ \dots\dots\dots FOP_{G,h}\right] \times \begin{bmatrix} PLSFO_{1,h} - PLSNO_h \\ PLSFO_{2,h} - PLSNO_h \\ \vdots \\ \vdots \\ PLSFO_{G,h} - PLSNO_h \end{bmatrix} \tag{21}$$

Table 5 : Maximum reserve capacity required, calculated with equations (19)-(21) for three scenarios of expansion of the Greek generation system

Reserve for Hourly Load Variations (MW)				
	2020	**2030**	**2040**	**2050**
Base-new	1170	1230	1400	1720
GHG-p60	1170	1360	2030	2530
REN-p60	1170	1490	2210	2710

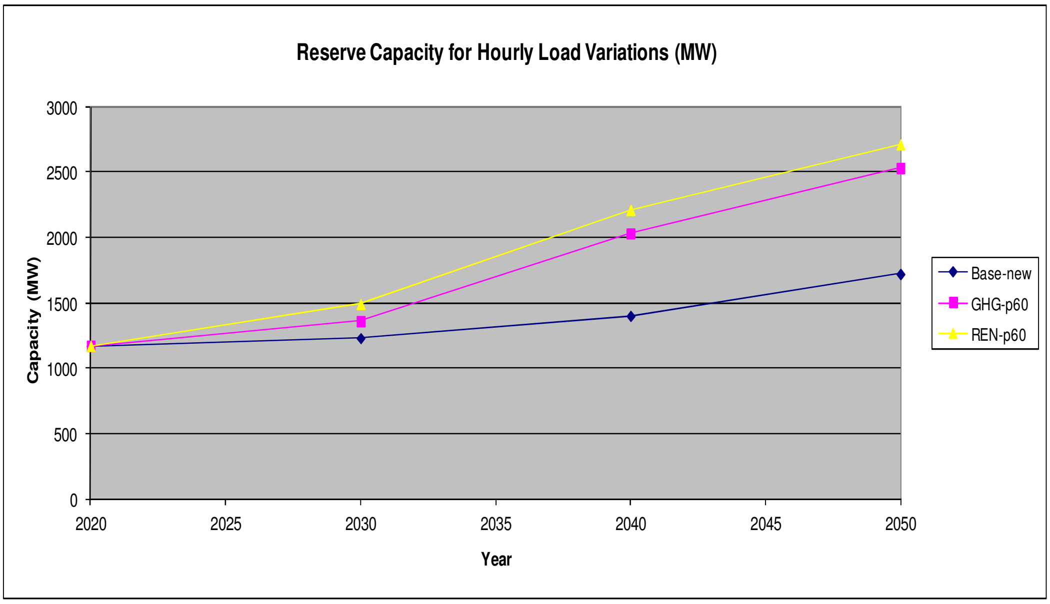

Figure 7 : Maximum reserve capacity required, calculated with equations (19)-(21) for three scenarios of expansion of the Greek generation system

VI. References

[1] R. R. Booth, “Power System Simulation Model based on Probability Analysis” IEEE Transactions, Vol. PAS-91, pp.62-69, 1972

[2] K.F. Schenk, R.B. Misra, S. Vassos, W.Wen, “A new Method for the Evaluation of Expected Energy Generation and Loss of Load Probability”, IEEE Transactions on Power Apparatus and Systems, Vol. PAS-103, No. 2, February 1984

[3] R. Doherty, M. O’Malley “A New Approach to Quantify Reserve Demand in Systems with Significant Installed Wind Capacity”, IEEE Transactions on Power Systems, Vol. 20, No. 2, May 2005

[4] B.C. Ummels, E. Pelgrum, W.L. Kling, “Integration of Large Scale Wind Power and use of Energy Storage in the Netherlands’ Electricity Supply”, IET Renewable Power Generation, Vol.2 No 1, pp. 34-46, 2008

[5] A. Rencher “Methods of Multivariate Analysis”, 2nd edition, Wiley, 2002

[6] A. Stuart, J. Ord "Kendal’s Advanced Theory of Statistics," vol. 1 Distribution Theory, Wiley, 2004

[7] P. Chaviaropoulos, K.Rossis, S. Tentzerakis, D. Foussekis, “Analysis of Wind Enrgy Generation Data and Development of Scenarios for the Penetration of Wind Energy into the Greek Interconnected Electricity Generation System”, Centre for Renewable Energy Sources and Energy Saving on behalf of the Greek Regulatory Authority for Energy, CRES.WIND.RAE2A/06, December 2006

[8] S. Vassos, K. Tigas, J. Mantzaris "Probability Production Costing under wide-scale penetration of Renewable Energy Sources," Center for Renewable Energy Sources and Energy Saving, Nov. 2009.

[9] H. Bayem, L. Capely, F. Dufourd, M. Petit “Probabilistic study of the maximum penetration rate of renewable energy in an island network”, IEEE Bucharest Power Tech Conference, Bucharest, Romania, June 28th - July 2nd, 2009

[10] J. Kabouris, "Penetration of Renewable Energy Sources in Electricity Generation Systems," Ph.D. dissertation, Dept. Electrical. Eng., National Technical Univ. of Athens, 1992.