Wind-hydro coordination using autoencoders to perform space dimension reduction and speed up evolutionary processes

Vladimiro Miranda, *Fellow IEEE* and Luís Costa

Abstract – This paper reports the application of neural networks denoted "autoencoders" in order to reduce the dimension of the search space in complex optimization problems. This allows a more efficient search by meta-heuristic algorithms, with a reduction in computing time and an improvement in the quality of results. The technique, coined as *miranda,* is illustrated with an application of an EPSO (Evolutionary Particle Swarm Optimization) algorithm to problems of medium term wind-hydro coordination, where the operation of cascading river dams with pumping-storage capability must be combined with decisions on the available wind power generation, depending on tariffs and market prices. One shows that an EPSO running of a reduced space generated by an autoencoder with solutions evaluated in a reconstructed space runs many times faster to obtain the same results as an EPSO running in the original problem space.

***Index Terms* — Wind power, wind-hydro coordination, neural networks, autoencoders**

I. Introduction

METAHEURISTICS such as evolutionary or particle swarm algorithms have proven to be flexible in representing realistic features of real world problems and became therefore a valuable tool in discovering optimising solutions in many problems. However, the suffer from some drawbacks namely being demanding in terms of computing effort, especially if one is dealing with large scale problems, which is the case in so many areas in Power Systems. The difficulty of dealing with large scale problems using meta-heuristics lies not only in computation time but also in some difficulty in converging to an optimal solution. Both these factors (need to an early stopping and difficulty in converging) contribute to originate a performance of metaheuristics below what would be desirable in practice.

A typical large scale problem may be wind-hydro coordination, which becomes relevant with the emerging importance of wind generation in the generation portfolio of many countries, namely in Europe. To assess this importance, notice that in Iberia (joining together Portugal and Spain) a total of over 25000 MW of installed capacity in wind generation is foreseen for 2011, while the joint peak power in Portugal and Spain, was about 54000 MW in December 2007. Wind energy value is greatly enhanced if combined with pumped storage so that energy may be delivered to the market during hours of high price but the decision to store must be weighed against the price of selling directly at the moment it is produced in the wind parks. This problem has obvious similarities with hydro-thermal coordination in the presence of pumping storage facilities and is represented by a complex time dependent formulation if cascading river dams are present.

In hydro-thermal coordination, several techniques were used such as Lagrangian relaxation [1], Stochastic Dynamic Programing [2] or Dual Dynamic Programming [3]. Models with Genetic Algorithms and Evolutionary Programming have also been proposed [4]. Models for wind-hydro coordination have also been proposed [5], [6][7]. In this paper, we will aply to the numerical examples an EPSO (Evolutionary Particle Swarm Optimization) algorithm [8] to test problems emulating the wind-hydro coordination context, built with enough complexity to test the optimization techniques under judgment.

A medium term operation planning or the water resources requires an evaluation of the operation for a period of the order of magnitude of 1 year and estimates of water and wind availability, with the division of the planning period in sub-periods corresponding to different months and different load levels with different estimated energy costs. The dimension of the problem may be very large.

Feature reduction and feature selection techniques have been used to reduce the number of variables of a problem to a set of meaningful ones. One popular technique is Principal Component Analysis (PCA) [8]. This technique projects the data into a linear subspace with minimum information loss, by multiplying the data by the eigenvectors of the sample covariance matrix. A point is then represented by its coordinates along the directions of greatest variance in the data set.

However, when one is optimizing and one needs to evaluate solutions during the process, a feature reduction or selection process is not applicable because some or all variable values of the original space would be unknown and therefore the actual value of the objective function could not be calculated.

This paper presents an original idea of using an autoencoder neural network to generate a pair consisting of a function f and its inverse f^{-1}, allowing to map a space of

V. Miranda is with INESC Porto and also with FEUP, Faculty of Engineering of the University of Porto, Portugal (email: vmiranda@inescporto.pt).

L. Costa is with INESC Porto and also with FEUP, Faculty of Engineering of the University of Porto, Portugal (email: lcosta@inescporto.pt).

dimension m into a space of dimension n (with n < m) and to reconstruct the original variables. The optimization procedure may evolve in the reduced space but objective function evaluation is performed in the original space thanks to the inverse decoding function f^{-1}. This reduction of the dimension of the search space has a notorious beneficial effect in the performance of the optimization algorithm. The function f acts as an intelligent encoder of the chromosomes organized in the original space: the chromosomes used in the optimization technique are no longer designed by the user but "engineered" by an intelligent process to better suit the optimization process.

II. AUTOENCODERS

It is a known mathematical property of real valued spaces that it is possible to define functions that establish a one-to-one mapping between points of a space of dimension m and a space of dimension n (with n < m without loss of generality). The theorems supporting this assertion, however, do not indicate how to construct such functions.

An approximation may however be achievable through the use of a space reduction technique known as *autoencoders*, which are feedforward neural networks that are trained to reproduce the input space S in the output. If an inner layer has a reduced number of neurons n compared to the set of m inputs/outputs, this layer will effectively be encoding variables from S into a smaller dimension space S' (see Fig. 1).

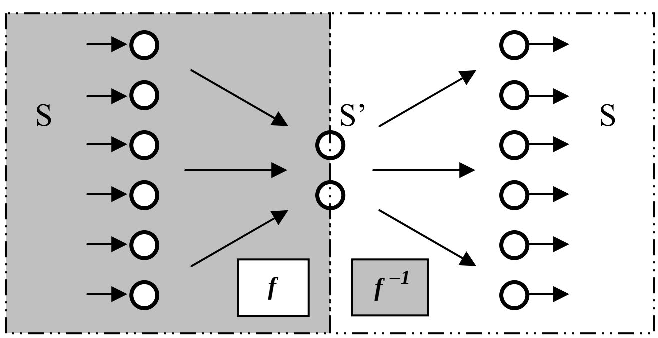

Fig. 1. An autoencoder neural network, with a *bottleneck* inner layer input and output layers of the same dimension and trained to reproduce the input variables in the output. In the inner layer one has a compressed set of values that encode, in a reduced dimension space S', the values in S.

This technique has been proposed in the past [10][11] with the purpose of using the reduced encoded values in S' to represent images in a compressed way, so that this representation would be subject then to distinct processing techniques such as identification and pattern recognition. For instance, face images could be identified and clustered according to sex, distinguished from non faces, etc [12].

The first half of the neural network approximates the function f that maps the input space to the space of compressed encoding S' while the second half approximates the inverse function f^{-1}. The quality of this encoding and decoding process depends on the quality of the training.

It has been found that it was much more difficult to optimize the weights in autoencoders with non-linear activation functions and multiple layers than with a single hidden layer and recent efforts were concentrated in devising schemes to achieve a more efficient training [13].

III. EPSO

EPSO – Evolutionary Particle Swarm Optimization, is a hybrid in concepts of Evolutionary Algorithms and Particle Swarm Optimization, first proposed in [8] and with an improved version in [14][15]. It is an Evolutionary Algorithm with an adaptive recombination operator inspired on the "movement rule" of PSO (Particle Swarm Optimization).

The movement rule of PSO generates a new individual as a weighted combination of parents, which are: a given individual in the population, the best ancestor of this individual and the best ancestor of the present generation. This may be seen as a form of intermediary recombination. In this type of recombination in evolutionary algorithms, a new individual is formed from a weighted mix of ancestors, and this weighted mix may vary in each space dimension. The mutation operator is only applied to the weights.

The recombination rule for EPSO is the following: given a particle $\boldsymbol{X}_i$, a new particle $\boldsymbol{X}_i^{new}$ results from

$$\boldsymbol{X}_i^{(k+1)} = \boldsymbol{X}_i^{(k)} + \boldsymbol{V}_i^{(k+1)}$$

$$\boldsymbol{V}_i^{(k+1)} = w_{i0}^* \boldsymbol{V}_i^{(k)} + w_{i1}^*(\boldsymbol{b}_i - \boldsymbol{X}_i) + w_{i2}^* C(\boldsymbol{b}_g^* - \boldsymbol{X}_i)$$

where the symbol * indicates that these parameters will undergo evolution under a mutation process, and

- $\boldsymbol{b}_i$ – best point found by the line of ancestors of individual i up to the current generation
- $\boldsymbol{b}_g$ – best overall point found by the swarm of particle in their past life up to the current generation
- $\boldsymbol{b}_g^* = \boldsymbol{b}_g + w_{i4}^* N(0,1)$ - particle in the neighborhood of $\boldsymbol{b}_g$.
- $\boldsymbol{X}_i^{(k)}$ – location of particle i at generation k
- $\boldsymbol{V}_i^{(k)} = \boldsymbol{X}_i^{(k)} - \boldsymbol{X}_i^{(k-1)}$ – "velocity" of $\boldsymbol{X}_i$ in generation k
- w_{i1} – weight of the *inertia* term (a new particle is created in the same direction as its previous couple of ancestors)
- w_{i2} – weight of the *memory* term (the new particle is attracted to the best position occupied by its ancestors)
- w_{i3} – weight of the *cooperation* or *information exchange* term (the new particle is attracted to the overall best-so-far found by the swarm).
- w_{i4} – weight affecting dispersion around the best-so-far
- $\boldsymbol{C}$ – a diagonal matrix with each element in the main diagonal being a binary variable equal to 1 with a given communication probability p and 0 with probability (1-p); in basic models, p = 1 but in advanced models p must be chosen from experiments and values of 0.7 < p < 0.8 have been shown to be optimal in many problems [14], although highly complex problems seem to require a very low non-zero value such as p < 0.2.

EPSO has been successful in a number of Power System applications [16][17].

IV. Optimizing in a Reduced Space

The original idea reported in this paper can be summarized in the following lines:

- First, launch an EPSO with individuals (particles) represented in S and will store the solutions that this search will discover, for a given number of iterations, in order to build up a training and a test set and with sampling hopefully more dense in preferable regions of space.
- Use an autoencoder neural network to generate a reduced dimension search space S' and a half-network decoder approximating f^{-1}, regenerating from each individual in space S' the corresponding individual in the original variable space S.
- Then launch an EPSO with individuals (particles) represented in S' and the movement rule will be applied in space S'.
- The evaluation of each new particle will be performed by decoding it (passing it through the f^{-1} half-network) and calculating an exact objective function value from the real variables produced in S.
- Based on this evaluation, selection will act on particles in space S'.
- When the process has reached an optimized value with particles in S', switch back to space S and continue with an EPSO there to further optimize the problem.

This process is illustrated in Fig. 2.

This process of optimization through ***M**ultiple **I**nput **R**eduction by **A**dopting **N**etworks **D**esigned as **A**utoencoders* will be called, with forgivable lack of modesty, the *miranda* method.

The success of this idea depends of the gain in efficiency that one obtains from having the particles evolving in a space of reduced dimension. A few additional comments must be made.

First, it is true that the half-networks emulating f and f^{-1} only generate approximations to these functions. However, this is not important in the light that each point in S' is associated with a real solution in S and it is valued exactly (see Fig. 2).

Second, because the autoencoder is just an approximation and not a representation of the exact mappings S↔S', some information will eventually be lost. It is possible then that the exact optimum of the original problem may not be found in S' – but if the approximation is good enough, a near optimal solution or, at least, the location of the optimum will be found and an efficient post-optimization search may be launched, if necessary back in space S.

Third, the advantage of *miranda* will only be evident if the task of training the autoencoder becomes much smaller than the additional iterations needed by an algorithm searching in the original space S.

Notice that the training and test sets used to generate the autoencoder neural network are not obtained through random sampling.

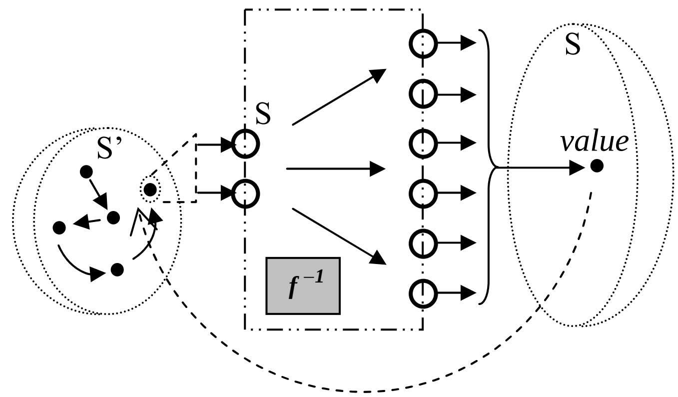

Fig. 2. Particles evolve in the reduced space S' but are decoded and evaluated in the original space S, influencing their selection in S'.

In fact, because the sampling is conducted using an evolutionary optimizing method, it becomes very likely that one will have a denser representation of the solution space in regions close to the optimum, which is a very desirable trait.

A partial answer to these points is given in the following section, where a complex problem of wind-hydro coordination that has been designed to serve as a test to the autoencoder hypothesis, illustrates clearly the advantages of the scheme.

V. Wind-Hydro Coordination in Planning

A. General description of the problem

An experimental confirmation of the potential of the method described above is given in this section, where a complex problem of wind-hydro coordination for medium term operation planning, designed to serve as a test to the autoencoder hypothesis, illustrates clearly the advantages of the scheme. This wind-hydro coordination planning problem is composed of an independent energy producer that owns a number of cascading hydro power plants, and also wind power plants that are treated as a single source (energy supplied through the transmission grid).

Load forecasts are admitted available, dividing each month of the planning period in peak and off-peak time steps. Wind and water inflow forecasts will be admitted to be available for the whole period of planning also. These forecasts are interpreted in terms of energy available for each time step.

To take in account the variability in wind, the data preparation process builds a set of wind power availability series derived from a set of historical wind series, representative of the wind behaviour in the region of interest. Water inflow variability is also treated by considering a set of historical water inflow series. This allows one to sample wind and water series and organize a Monte Carlo process that will estimate not only the expected value of the value of wind-hydro coordination but also its variance, which is an important indicator for risk..

Because the purpose of this paper is to demonstrate the potential and usefulness of the new technique, one will not devote much time to analyse the effects of uncertainties and concentrate on the optimization procedure instead.

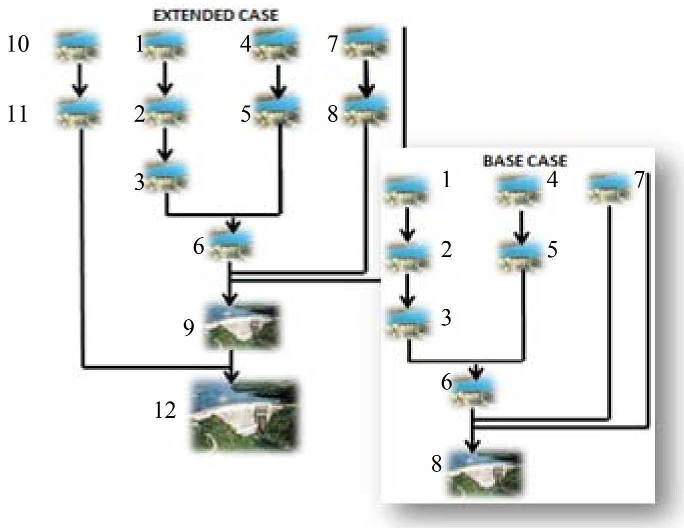

Fig. 3. Hydro power system scheme for the base case and extended case with cascading reservoirs marked with numbers.

Numerical results supporting the new method proposed will be presented for two cases of different dimensions, to illustrate the impact of *miranda* with problem size growth . The first case, with 8 reservoirs, has a river system presented in Fig. 3, where the cascading reservoirs are displayed, and has been built from [4]. This case will be called *base case* and will be used to test the methodologies proposed. The second case, with 12 reservoirs, will called *extended case* and will be used to test the robustness of the techniques in a larger problem than the base case. All reservoirs are admitted to be equipped with pumps allowing a certain amount of water to be moved upstream if convenient.

The objective in both cases is to derive an operation plan that maximizes the profit obtained with the operation of the system throughout T time periods with different buying and selling energy prices, covering 1 year of operation. The operation plan will determine:

- Quantity of water to be turbined or pumped for each hydro power plant in each period of time and energy sold or used;
- Quantity of wind energy to be used for water pumping and the quantity of wind energy to be sold to the electric power system in each period.
- Detailed information about the amount of water storage in each reservoir and water storage capacity available for each period of time

The T time periods are divided in T/2 peak periods and T/2 off-peak periods. The *base case* has a horizon of 6 months (T=12) and the extended case of 12 months (T = 24). Six energy prices are defined for each period, also admitting average price forecasting based on market history:

- Hydroelectric energy selling price at peak and off-peak periods;
- Hydroelectric pumping price at peak and off-peak periods;
- Wind energy selling price at peak and off-peak periods.

The variables of this problem are defined in terms of water movement for each reservoir in each period. A positive value will indicate turbining while a negative value will indicate pumping in this period. Ecological spills or evaporation are not considered in this example but present no difficulty to the model. The models for the hydro power stations assume linearity with water flows in this numerical example, meaning that head influence in efficiency is neglected and reservoirs are assumed as being of high head, compared with the variation associated with reservoir depletion. This influence can also be easily incorporated in the model.

B. *The mathematical model*

The electric energy of hydro origin $H(t)$ generated is described by

$$H(t) = \sum_{n=1}^{N} K_n \left[h1_n(x_n) - h2_n(q_n + z_n) \right] \Delta T_t \qquad (4)$$

where:

K_n– Constant that includes the generator/pump efficiency, gravity acceleration, water density and unity conversion factors;

x_n – Volume of water available in reservoir n, period t;

q_n – Volume of water turbine or pumped by the hydro power plant n in period t;

z_n – Volume of water spilled in reservoir n in period t;

$h1_n(t)$ – Upstream water difference of level related to the sea level, in the reservoir n in period t;

$h2_n(t)$ – Downstream water difference of level related to the sea level, in the reservoir n in period t;

ΔT_t – duration of time interval t.

Constant K_n takes different values, one for pumping mode and a different one for the generation mode, representing the different efficiencies involved.

The available water volume for each reservoir is calculated for each period considering the variables associated to the reservoir, such as the natural affluences, the volume of water pumped or used in generation, the volume of water spilled and finally the already existing water volume, all of them in the previous period of time, and also considering the variables associated to the operation of the upstream reservoirs such as the quantity of water that was used for generation and now haves to be accommodated in the downstream reservoirs and also the water volume spilled from the upstream reservoirs. So in the wind-hydro coordination model, the procedure above is mathematically represented as dictated in equation (5):

$$x_n(t+1) = x_n(t) + y_n(t) + \sum_{k \in \Omega_n} \left[q_k(t) + z_k(t) \right] - q_n(t) - z_n(t) \qquad (5)$$

where:

$y_n(t)$ – Natural river inflow for reservoir n in period t;

$\Omega_n(t)$ – Set of all hydro reservoirs upstream of reservoir n.

Constraints reflect physical limitations such as the maximum and minimum quantity of water that can be pumped or turbined to or from a reservoir depending on the existing water volume in the reservoir and depending also on the operation of all other reservoirs, and reflecting some technical limitations such as the minimum ecological volume and the available wind energy. These are classical constraints and are represented by

$$x_{n,min} \le x_n(t) \le x_{n,max} \tag{6}$$

$$q_{n,min} \le q_n(t) \le q_{n,max} \tag{7}$$

where:

$x_{n,min}$ – minimum available water volume, corresponding to the ecological volume;

$x_{n,max}$ – Maximum available water volume in the hydro reservoir *n*, reservoir maximum capacity;

$q_{n,min}$ – Maximum available water volume that can be pumped into the reservoir *n*;

$q_{n,max}$ – Maximum available water volume in the hydro reservoir *n* that can be used for energy generation;

The energy generated at wind farms per period is estimated as *E(t)* – this is taken as data in the following. Its value per period is derived from the wind series and each wind farm production characteristic, which can be modeled separately from the optimization procedure. In fact, as there are no "reservoirs for wind", the generation forecast is a direct function of the wind forecast. One has

$$E(t) = E_V(t) + E_B(t) \tag{8}$$

where:

$E_V(t)$ – Wind energy sold to the grid in period *t*;

$E_B(t)$ – Wind energy used to pump water into the water reservoirs in period *t*.

The fitness function, or in other words the operation profit function for one scenario, is given by

$\phi = \sum_t L(t)$, with

$$L(t) = Prod(t) \times Sell\_pr_H(t) - Pump(t) \times Buy\_pr(t) + \\ + E_B(t) \times Buy\_pr(t) + E_V(T) \times Sell\_pr_E(t)$$

where:

Prod(t) – Positive q_n terms sum in period t;

Pump(t) – Negative q_n terms sum in period t;

Buy_pr(t) – Energy buying price for period t;

$Sell\_pr_H(t)$ – Hydroelectric energy selling price for period t;

$Sell\_pr_E(t)$ – Wind energy selling price for period t;

The expected economic value Φ over a set of M scenarios is given by

$$\Phi = \frac{1}{M} \sum_{k=1}^{M} \phi_k \tag{9}$$

C. Chromosomes and EPSO parameters

A chromosome representing an individual or a solution in the original space S has the structure illustrated in Fig. 4. In this figure only the object parameters (or natural variables of the problem) are identified and the strategic parameters (weights) should be added to obtain the complete representation.

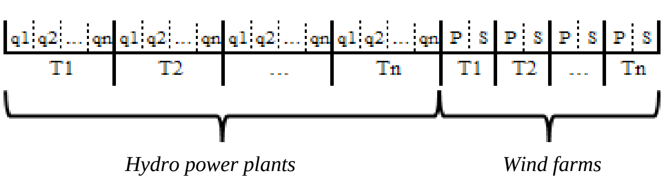

Fig. 4. Structure of a chromosome with its object parameters

In this figure, T1 to Tn represent the time periods, q1 to qn represent the turbined (or pumped) water per each hydro power station, P is the wind energy used to pump water upstream and S is the wind energy directly sold to the grid.

When using EPSO a population of 20 was adopted in all experiments, all weights were randomly initialized in the interval]0,1], the stopping criterion was a fixed number of fitness function evaluations (400000), the learning rate was set to $\tau = 0.6$ and the communication probability was set to 0.7. Constraints were dealt with a penalty technique.

D. Autoencoder structure

The autoencoders used were three-layer feed-forward neural networsk with the middle layer of half the size of the outer (input and output) layers. The first layer neurons had linear activation functions and the other neurons had symmetric sigmoid activation functions such as

$$y = \frac{2}{1+e^{-2sx}} - 1, \quad -1 \le y \le 1 \tag{10}$$

where x is the sum of neuron inputs, y is the neuron output and s a parameter regulating the shape of the sigmoid.

The input and output layers have the dimension of S, equal to the number of variables of the problem. In the base case S = 120 and S' = 60; in the extended case S = 336 and S' = 168.

E. Dealing with constraints in S'

One must realize that the meaning of the variables in space S' (the output of the neurons in the middle layer) is virtually unknown, but constraints associated with variable limits must be enforced in this space. The strategy adopted has been to observe in the autoencoder training set the values assumed by the variables in S'. From this observation, limits are defined for these variables taking in account the minimum and maximum values registered in the training set. All other constraints in the problem are enforced by applying penalties (negative) to the fitness function, calculated after the application of the autoencoder half representing f^{-1}.

VI. Optimizing One Scenario

This section presents the results of the optimization in one scenario taking one wind series and one water inflow series as well as one price series. The following experiments were conducted, all with a stopping criterion of a maximum of 400,000 fitness function evaluations:

A. Run an EPSO in full space dimension S for full optimization to the limit of computing effort set.

B. Take a set of the solutions discovered by the previous process to train the autoencoder in offline mode; then, run an EPSO in a reduced space S' until the stopping criterion is met

C. Take a set of the solutions discovered by the previous process to train the autoencoder in offline mode; then, run an EPSO in a reduced space S' for a given number of iterations; finally, switch to space S and continue the optimization until the stopping criterion is met.

D. Start an EPSO in Space S for a number of iterations; with a subset of the solutions discovered, train an autoencoder in online mode; then, run an EPSO in a reduced space S' for a given number of iterations; finally, switch to space S and continue the optimization until the stopping criterion is met.

Experiment A defines the benchmark. Experiment B is meant to demonstrate that the auto-encoder trained with a set of points in the region of interest (the neighborhood of the optimum, discovered by in Experiment A) does accelerate the EPSO algorithm by defining a search space of reduced dimension. The third experiment shows that post-optimization is possible after the action of the auto-encoder. Experiment D demonstrates that even a preliminary sampling allows the autoencoder, trained in online model, to accelerate the optimization procedure.

A. Results

All results presented in this section are average results obtained after running a number of times each experiment, in order to eliminate the effect of a random non-representative result. The number of fitness function evaluations is equal to the number of iterations times the double of number of particles.

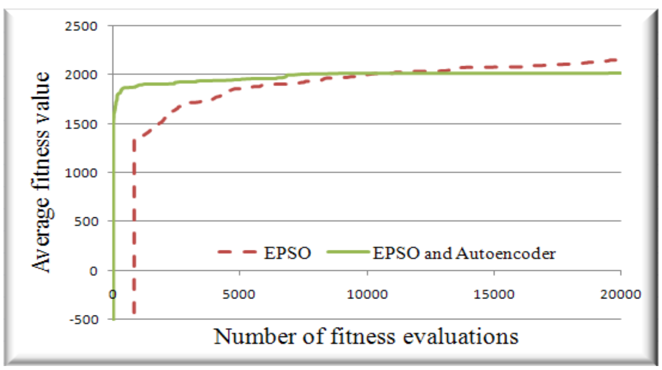

Fig. 5. Comparison of the evolution of the fitness function value in Experiments A and B (average of 5 runs) in the base case.

Fig. 5 displays a comparison between Experiment A and B in the base case. Two things are of immediate perception: the optimization in space S' discovers feasible solutions very quickly with already good values but if the optimization process is prolonged then eventually the EPSO running in space S eventually catches up and provides a better result. The interpretation is the following: the *miranda* technique accelerates the process thanks to a good representation of the solution space in the region of interest, as a consequence of the biases sampling obtained from the previous action of the EPSO run on S. However, the auto-encoding process leads to some loss of information (the training process is not perfect, there is some inevitable error) and the information on the exact location of the optimum may have been "erased", although the region of interest has been located.

The efficiency of the autoencoder in driving the swarm into the feasible region can be also appreciated in Table 1. The efficiency of *miranda* may be appreciated in Table 2, where a speed up of 16 times is obtained to reach the same fitness value with and without the autoencoder.

TABLE 1 – NO. OF ITERATIONS NECESSARY TO DISCOVER THE FIRST FEASIBLE SOLUTION IN EXPERIMENTS A AND B IN THE BASE CASE

	Iteration no.	Fitness value
Exp. A - EPSO	838	1352,03
Exp. B – EPSO/Autoencoder	13	1570,78

TABLE 2 – COMPUTING EFFORT TO REACH THE SAME FITNESS VALUE (1820) IN EXPERIMENTS A AND B IN THE BASE CASE

	Iterations	Fit. Evaluat.
Exp. A - EPSO	4512	180480
Exp. B – EPSO/Autoencoder	285	15400

Experiment B also teaches that a post-optimization is necessary because of the loss of information in the process of training the autoencoder. This is the principle behind Experiments C and D.

Fig. 6 displays the result of switching back to space S after 250 iterations in the base case. The superiority of the process tested in Experiment C is obvious: the post-optimization is effective because the optimization in space S' left the swarm in the region of the optimum.

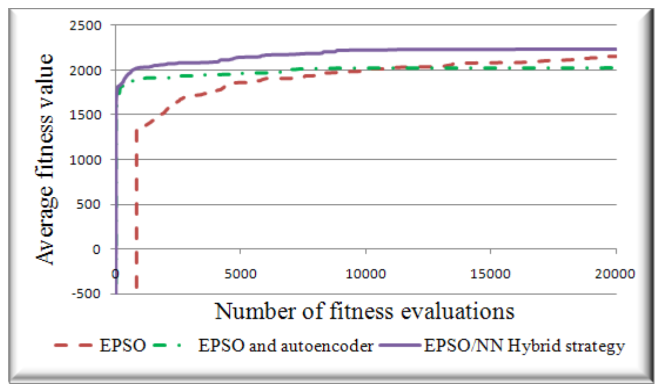

Fig. 6. Comparison of Experiments A, B and C for the base case (5 runs).

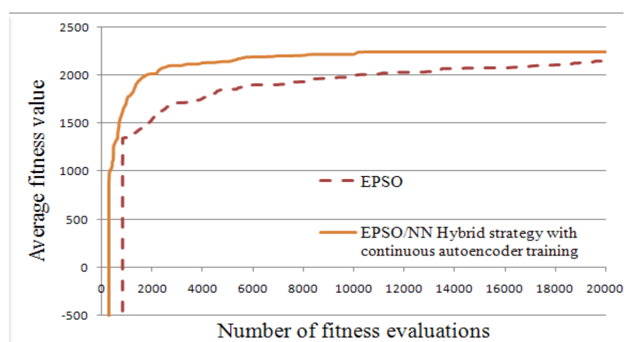

Fig. 7. Comparison of Experiments A and D in the base case (5 runs).

In Fig. 7 one may appreciate the performance of the strategy tested in Experiment D, by training the autoencoder with the solutions discovered by an EPSO in S during the first iterations. Although it takes a bit longer to reach a feasible solution (not found in the preliminary iterations is S) still the process performs much better than using only an EPSO in the original space S and reaches solutions of comparable quality with Experiment C, where the autoencoder was trained offline with a "better" training set.

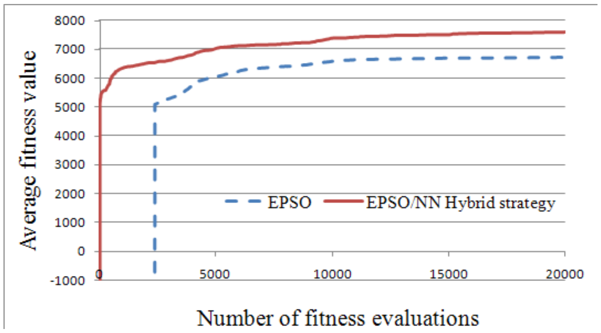

Fig. 8. Comparison of Experiments A and C for the extended case (5 runs)

Fig. 8 displays results for the extended case for Experiments A and C. The superiority of the EPSO/autoencoder approach is obvious.

Tables 3 and 4 again illustrate the superiority of the new technique over the simple use of a meta-heuristic, which takes advantage of an early discovery of the feasible domain and of the region of interest where the optimum is located.

TABLE 3 – NO. OF ITERATIONS NECESSARY TO DISCOVER THE FIRST FEASIBLE SOLUTION IN EXPERIMENTS A AND B IN THE EXTENDED CASE

	Iteration no.	Fitness value
A - EPSO	2389	4871,15
C – EPSO/Offline Autoencod.	26	5106,21

TABLE 4 – COMPUTING EFFORT TO REACH THE SAME FINAL FITNESS VALUE IN EXPERIMENTS A AND B IN THE BASE CASE

	Iteration no.	Fit. Evaluat.
A - EPSO	20000	800000
C – EPSO/Offline Autoencod.	3379	135180

In Fig. 9, finally, one may verify that the online training of the autoencoder (Experiment D) continues to reveal superiority relative to the simple adoption of a meta-heuristic in the original space S. Naturally, the discovery of feasible solutions is delayed when compared with Experiment C but there is still a remarkable gain.

It must be mentioned that although training the autoencoder consumed some computation time that must be added to the computational effort of the Experiment D, in all cases it resulted nevertheless in a net gain, whose relevance depends and grows with the complexity of the fitness function.

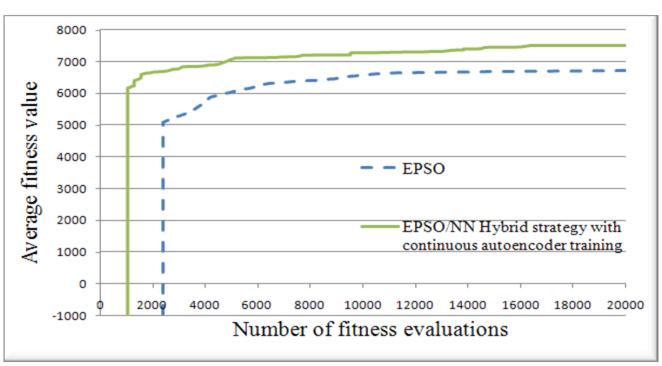

Fig. 9. Comparison of Experiments A and D for the extended case (5 runs).

VII. VALUE OF WIND-HYDRO COORDINATION

The application of *miranda* to the optimization of wind-hydro coordination, with the speed up in calculations achieved, allows the assembling of a Monte Carlo process to reach a probabilistic evaluation of the value of wind-hydro coordination, when compared with an independent operation of the hydro and the wind generation.

The Monte Carlo process is organized by sampling series of wind, water inflows and energy prices. In order to compare the cases of coordinated and uncoordinated operation of wind and hydro generation, calculations have also been made for the hydro system placing an artificial wind series at constant zero value (zero wind scenario) and adding the cost of the full sale of wind generation to the grid at the hours of generation.

Table 5 presents the average result for 20 wind series scenarios of the comparison for the extended case of a strategy of wind-hydro coordination with the sum of independent wind and hydro strategies with no coordination – meaning that wind energy is sold to the grid only and that the hydro power stations owner uses pumping when useful by buying from the grid. One admits in these trials that one has pumping in all hydro power stations in both cases.

TABLE 5 – AVERAGE OPERATION PLAN PROFIT VALUES FOR 20 DIFFERENT WIND SCENARIOS

	Operation plan Profit
Without W-H coordination	5653,8
With W-H coordination	7423,23

This result represents the expected 1769 units per year value of wind-hydro coordination. To assess the risk of a coordination strategy, one displays in Fig. 10 the histogram of the profit form a coordination strategy, as a result of the scenario simulation. Admitting that all the scenarios (wind series) have similar probability, one may therefore calculate a risk index associated to an expected gain from a coordination strategy over a non-coordinated strategy. For instance, there is a risk of 15% of not benefiting more than 1478 units in one year and the probability of gaining in one year more than 2093 units is estimated at 10%.

The marginal value of pumping at a given power station can be also evaluated. Table 6 compares the average results in 20 wind series scenarios from a wind-hydro coordination strategy, when having or not pumping ability in power station 6in the extended case.

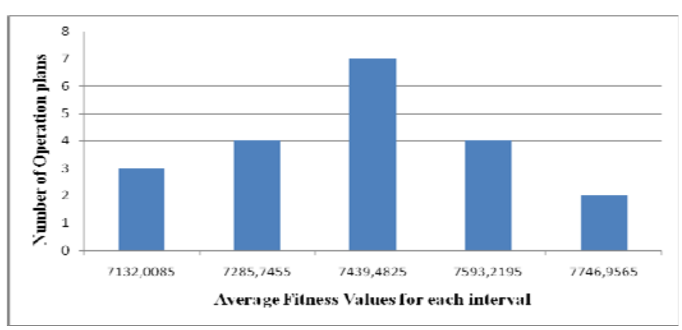

Fig. 10. Distribution of the yearly results for a wind-hydro coordination strategy for 20 different wind series scenarios.

TABLE 6 – COORDINATION PROFIT FOR BOTH WITH AND WITHOUT PUMPING ABILITY SCENARIOS

	Coordination plan Profit
With pumping ability	7520,96
Without pumping ability	6978,7

The yearly additional value of having pumping in a particular power station may then serve to justify a possible investment in power station 6, in a wind-hydro coordination strategy. These calculations are difficult and time consuming and this underlines the usefulness of techniques that speed up simulations. The adoption of autoencoders, by accelerating the algorithms considerably, makes these calculations feasible.

VIII. CONCLUSIONS

Meta-heuristics or population-based methods are known to lose efficiency in large scale problems: the convergence becomes slow when the number of variables is large and the computing effort to reach the optimum becomes heavy. One of such problems is the wind-hydro coordination in medium term operation planning, where several elements of complexity are present, namely the spatial and temporal dependence introduced by the cascading hydro power stations and the need to represent a large set of time steps. This paper presents a novel method to evaluate the added value of a wind-hydro coordination strategy when compared with an independent operation of hydro and wind generation systems. Wind-hydro coordination leads to the concept of a market agent (or a partnership of agents) operating it may extract added value from the renewable resources. This analysis must necessarily be probabilistic, given the uncertainties associated to the renewable energies availability and also to the energy prices, and requires a considerable number of simulations of a diversity of scenarios. Therefore, any technique that speeds up computations becomes extremely valuable.

This is the case with the Multiple Input Reduction by Adopting Networks Designed as Autoencoders approach, or *miranda*, proposed in this paper as a very novel process that achieves problem optimization by organizing meta-heuristic searches in an equivalent reduced dimension search space. The cleverness of the method lies in the fact that the evolutionary process acts upon individuals represented by chromosomes that are not designed ad-hoc by a human but instead result from an intelligent coding achieved by a first half of an autoencoder, while the fitness function is evaluated by decoding the intelligent chromosomes with the second half of the autoencoder. Because the clever chromosomes are represented in a space of reduced dimension, the optimization process becomes much more efficient.

The application of this original technique to two realistic wind-hydro coordination problems is made feasible by the speed up achieved with thenovel technique. The results presented fully demonstrate the interest of the technique, which is of general application.

IX. ACKNOWLEDGMENT

This work is partially integrated in the project EPSO POSC/EEA-ESE/60980/2004 financed by FCT, Portugal.

X. REFERENCES

[1] Yan, H. Luh, P.B. Guan, X. Rogan, P.M., “Scheduling of hydrothermal power systems”, IEEE Transactions on Power Systems, Volume: 8, Issue: 3, page(s): 1358-1365, Aug 1993

[2] Sherkat, V.R. Campo, R. Moslehi, K. Lo, E.O., “Stochastic Long-Term Hydrothermal Optimization for a Multireservoir System”, IEEE Transactions on Power Apparatus and Systems, Volume: PAS-104, Issue: 8, page(s): 2040-2050, Aug. 1985

[3] Gorenstin, B.G. Campodonico, N.M. da Costa, J.P. Pereira, M.V.F., “Stochastic optimization of a hydro-thermal system including networkconstraints”, IEEE Transactions on Power Systems, Volume: 7, no.2, page(s): 791-797, May 1992

[4] V. Miranda, D. Srinivasan, L. Proença, “Evolutionary computation in Power Systems”, Int. J. in Electrical Power and En. Systems, Jan 1998

[5] Zoumas, C.E.; Bakirtzis, A.G.; Theocharis, J.B.; Petridis, V., “A genetic algorithm solution approach to the hydrothermal coordination problem”, *IEEE Transactions on Power Systems,* page(s): 1356- 1364, Volume: 19, Issue: 3, Aug. 2004

[6] D. Castronuovo and J. A. Peças Lopes, “On the Optimization of the Daily Operation of a Wind-Hydro Power Plant”, IEEE Trans. On Power Systems, VOL. 19, NO. 3, AUGUST 2004

[7] D. Castronuovo and J. A. Peças Lopes, “Bounding Active Power Generation of a Wind-Hydro Power Plant”, Proceedings of the International Conference on Probabilistic Methods Applied to Power Systems, Iowa State University, Ames, Iowa, September 12-16, 2004

[8] V. Miranda and N. Fonseca, "EPSO – Best-of-Two-Worlds Meta-Heuristic Applied To Power System Problems ", *Proc. of WCCI 2002 - World Congress on Computational Intelligence - CEC -Conference on Evolutionary Computing*, Honolulu, Hawaii, U.S.A., May, 2002

[9] Jolliffe I.T., “Principal Component Analysis”, Springer Series in Statistics, Springer, NY, 2002

[10] G.W. Cottrell, P. Munro and D. Zipse**r, “**Learning internal representations from gray-scale images: An example of extensional programming”, Proceedings of the Ninth Annual Conference of the Cognitive Science Society, Seattle (WA), USA, 1987

[11] M.K. Fleming, G.W. Cottrell, “Categorization of faces using unsupervised feature extraction”, Proceedings of IJCNN - International Joint Conference on Neural Networks, vol. 2, pp. 65-70, San Diego (CA), USA, 17-21 Jun 1990

[12] B. Golomb and T. Sejnowski, “Sex Recognition from Faces Using Neural Networks”, in A. Murray (ed.), “Applications of Neural Networks”, pp. 71-92, Kluwer Academic Publishers, 1995

[13] G. E. Hinton and R. R. Salakhutdinov, “Reducing the Dimensionality of Data with Neural Networks”, Science, Vol. 313, no. 5786, pp. 504 – 507, July 2006

[14] V. Miranda, H. Keko and A. Jaramillo,“EPSO: Evolutionary Particle Swarms”, Ch. 6 in “*Advances in Evolutionary Computing for System Design*”, L. Jain et al. Eds., Springer, series: Studies In Computational Intelligence, ISBN 978-3-540-72376-9, Volume 66, pp. 139-168, 2007.

[15] Vladimiro Miranda, Hrvoje Keko and Álvaro Jaramillo Duque, “Stochastic Star Communication Topology in Evolutionary Particle Swarms (EPSO)”, Int. J. of Comp. Intelligence Research, no.2, 2008

[16] V. Miranda and N. Fonseca, “EPSO - Evolutionary Particle Swarm Optimization, a New Algorithm with Applications in Power Systems”, Proceedings of the IEEE Transmission and Distribution Asia-Pacific Conference 2002, Yokohama, Japan, Oct 2002

[17] H. Mori and Y. Komatsu, “A Hybrid Method of Optimal Data Mining And Artificial Neural Network for Voltage Stability Assessment”, *Proc. of IEEE St. Petersburg PowerTech Conference*, Russia, June 2005

VI. BIOGRAPHIES

Vladimiro Miranda (M’90-SM’04-F’05) received his graduation, Ph.D. and Agregado degrees from the Faculty of Engineering of the University of Porto, Portugal (FEUP) in 1977, 1982 and 1991, all in Electrical Engineering. In 1981 he joined FEUP and currently holds the position of Full Professor. He is also a researcher at INESC since 1985 and is currently Director of INESC Porto, an advanced research institute in Portugal. He has authored many papers and been responsible for many projects in areas related with the application of Computational Intelligence to Power Systems

Luís Costa is a M.Sc. student at FEUP in Portugal and a young researcher at INESC Porto in the area of Power Systems.