CLUSTERING-BASED WIND POWER SCENARIO REDUCTION TECHNIQUE

J. Sumaili[1], H. Keko[1], V. Miranda[1], A. Botterud[2] and J. Wang[2]

[1]Instituto de Engenharia de Sistemas e Computadores do Porto (INESC Porto)
Faculdade de Engenharia, Universidade do Porto
Porto, Portugal
jean.sumaili@inescporto.pt, hrvoje.keko@inescporto.pt, vmiranda@inescporto.pt

[2]Argonne National Laboratory, Decision and Information Sciences Division
Argonne, IL, USA
abotterud@anl.gov, jianhui.wang@anl.gov

Abstract – This paper describes a new technique aimed at representing wind power forecasting uncertainty by a set of discrete scenarios able to characterize the probability density function of the wind power forecast. From an initial large set of sampled scenarios, a reduced discrete set of representative or focal scenarios associated with a probability of occurrence is created using clustering techniques. The advantage is that this allows reducing the computational burden in stochastic models that require scenario representation. The validity of the reduction methodology has been tested in a simplified Unit Commitment (UC) problem.

Keywords: wind power, uncertainty, scenario reduc`tion, probability.

1 INTRODUCTION

The continuous growth in wind power penetration increases variability and volatility in the power system, thus posing new challenges to power system management and planning. Wind power forecasts are getting more and more mature, offering better predictions generally characterized by a single-value forecast (or point forecast) for each look-ahead time horizon. To better employ the forecasts in practice, the users need additional information concerning the uncertainty of the forecasts. One way of presenting this uncertainty is using a discrete set of scenarios, sampled according to a probability density function associated to the forecasts, and therefore able to characterize the uncertainty according to the historical error distribution. However, in order to accurately capture the entire characteristic of the wind power forecasted during the prediction period, a decision maker may need to evaluate a high number of scenarios. This is a time-consuming and computationally demanding task – and the computational burden is usually not compatible with stochastic programming algorithms run in useful time.

This paper describes a new technique aimed at representing the uncertainty associated with wind power in a forecasting horizon by a reduced discrete set of scenarios, where each scenario becomes associated with the probability of a cluster that it represents. The approach is similar to the one presented in [1]. This new set of scenarios may then be used as input in computationally demanding stochastic problems (e.g. unit commitment, market bidding).

In the proposed methodology, one departs from a wind power scenario generator of Monte Carlo type, such as [2]. This method is used to generate a very large sample: a set of scenarios, which hopefully gives a good discrete approximation to the probability density function of wind power. Then, an evolutionary optimization algorithm [3] is used to cluster the scenarios, identifying the areas of high density, and replacing each cluster by its more representative (*focal*) scenario. The probability associated to each *focal scenario* will be given by the probability associated to its cluster within the sample.

The motivation for this work is to adequately represent the shape of the full probability density function with a reduced set of scenarios with assigned probabilities. This makes the reduction method adequate in risk assessment. In the literature, several scenario reduction methods exist, commonly based on a scenario tree construction method [4], [5]. In this work, a clustering approach is used instead that works directly on the scenarios generated by Monte Carlo type scenario generation identifying the area of maximum density according to a defined similarity criterion.

The reduced equivalent set is still a discrete representation of the probability density function of wind power scenarios, in a space whose dimension is equal to the number of time steps considered, The great advantage of having available the reduced set of focal scenarios is that it allows solving stochastic programming problems in a practical and industrial environment — and this opens the door for the management of the power system based on risk assessment models, something very much needed whenever the uncertainty associated to wind power leads to possibly large hedging costs.

This paper presents the details of the model developed and its application in a stochastic unit commitment model, validating the interest in the approach. Unit commitment is a major step in the decision process of system operators – and stochastic unit commitment models are one logical approach to modeling a system with large uncertainty present in data, such as the case of wind power prediction. One of the reasons why such methods have not been used in practice is the computational burden with scenarios, which must be associated to probabilities. This paper offers an innovative answer to this difficulty.

2 METHODOLOGY

Given a set of M wind power scenarios sampled according to some probability density function describing the uncertainty in forecasts, the aim of the reduction process is to merge scenarios presenting high similarity in order to produce a new reduced set, containing representative scenarios, Each set of merged scenarios forms a cluster and is associated to some frequency within the sample – its representative or focal scenario is therefore associated with the same probability value.

Let us denote the forecasted wind power as a series $\mathrm{p}^{(\mathrm{m})} = \left\{\mathrm{P}_1^{(\mathrm{m})}, \dots, \mathrm{P}_\mathrm{T}^{(\mathrm{m})}\right\}$, for time horizon $\mathrm{t} = 1, \dots, \mathrm{T}$ in scenario $\mathrm{m} = 1, \dots, \mathrm{M}$. This is also a point in a space of dimension T.

The forecasted wind power scenarios $\mathrm{p}^{(\mathrm{m})}$ are normalized with respect to the installed power of the wind farm (or generator) in order to obtain the vectors $\mathrm{x}^{(\mathrm{m})}$, whose components lie in the range [0, 1]. Alternatively, this range can be changed to [0, 100], if the forecasted power is expressed as a percentage of the installed power.

Different metrics can be used to define the distance between two scenarios i and j. In this paper, the maximum deviation is used:

$$d(i,j) = \max_{1 \le t \le T} \left\| X_t^{(i)} - X_t^{(j)} \right\|, \qquad (1)$$

where $\mathrm{X}^{(\mathrm{i})}$ and $\mathrm{X}^{(\mathrm{j})}$ are the two scenarios being compared, and T is their dimension.

The process behind the reduction procedure described below begins with an extremely large set of wind power scenarios, generated from a Monte Carlo procedure; then, one finds, in the T-dimension space, the region with the highest probability density for the wind power. A cluster of scenarios in this region is defined and replaced by a best matching unit – its focal scenario, a representative element with a probability given as the probability of drawing any scenario within its cluster out of the whole sample set. This cluster is removed and the process is repeated until a stopping criterion is met. This way, a set of clusters is defined, each associated to a focal element together with a probability value of its cluster

For this method, it is necessary to define a tolerance for the scenario aggregation process. The value of this tolerance is a trade-off between accuracy and set reduction capability, which is related to the computational effort.

Figure 1 depicts the flowchart of the scenario reduction procedure. Once the admissible tolerance is defined, an Evolutionary Particle Swarm Optimization (EPSO) [3] is used to find the area of maximum density. EPSO, as a population-based method, relies on a set of "moving" solutions denoted as *particles* – in this case, a particle is a scenario or a vector in the T space. The EPSO algorithm has been proven successful in optimizing non-smooth and non-convex functions, however any optimization method may be applied here.

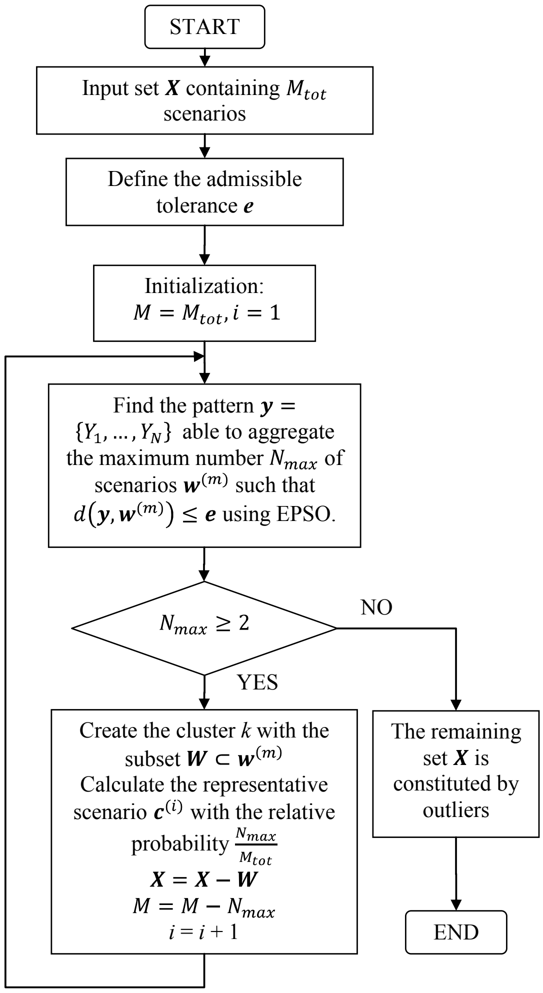

Figure 1. Flow-chart of the scenario reduction procedure

The objective function (to be maximized) defines a number of scenarios presenting an error (defined in terms of a distance or similarity measure) lower than the defined tolerance with respect to a current solution (a particle). The scenario with the highest aggregation capability is included in the initial set of particles. After some iterations, the scenarios aggregated by the best particle are then isolated from the initial set of scenarios and grouped into a cluster. The process is repeated while the maximum density area contains more than one scenario. The remaining scenarios are considered to be outliers.

After having defined a cluster and identified its members, there are two options for choosing the representative scenario $\boldsymbol{c}^{(i)}$ for the cluster. An option is to choose an existing scenario from the cluster, for instance the one with the minimum distance with respect to all the other components of the cluster. This scenario becomes the focal scenario and represents the cluster.

Another option is to calculate the mean value of each cluster to get the centroid of the cluster. This scenario is thus produced by averaging of the cluster members.

One should be careful in choosing the options for centroid calculation. Averaging the members of the cluster may artificially introduce smoothing effects in the new "artificial" scenario. This "average" scenario would not have been generated by the initial scenario generation process, so its behavior may not be in the line with the original probabilistic model that captured the predicted behavior of the wind power.

On the other hand, if a cluster member is chosen as the representative scenario, the total expected value may deviate from the expected value of the whole set of scenarios. The preferred method for the focal scenario calculation thus depends on the purpose of the scenario reduction.

Finally, after choosing the representative scenarios and the corresponding probabilities, the continuous probability density function of wind power initially represented by a large discrete sample, becomes represented by a much smaller discrete set of focal scenarios.

3 CASE STUDIES

3.1 Scenario Characteristics

The initial sets of scenarios used in this paper refer to a day-ahead wind power forecasts in 2006, for a wind farm located in the state of Illinois. Time series of day-ahead deterministic point forecasts were obtained from the National Renewable Energy Laboratory's Eastern Wind Integration and Transmission Study [6].

To conceive wind power scenarios, wind power data (forecasted and realized) were used to train the uncertainty estimation model, as well as to generate scenarios of the forecasted wind power, according to the methodology introduced by Pinson et al. [2], which is equivalent to producing scenarios under a Monte Carlo process. Figure 2 illustrates the appearance of 1000 scenarios generated in this way.

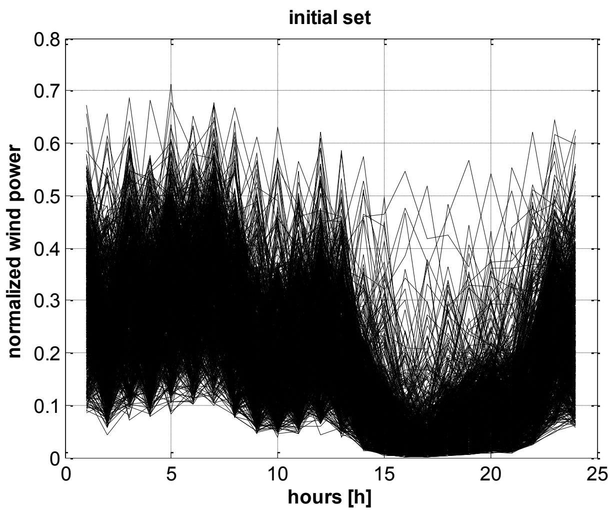

Figure 2. Initial set of scenarios

3.2 Results

The data set of wind power scenarios consists of M = 1000 scenarios of day-ahead forecasted wind power, generated as previously mentioned (Figure 2). Figure 3 and Figure 4 present the focal scenarios and the set of outliers obtained by the reduction procedure for the analyzed day considering e = 0.20.

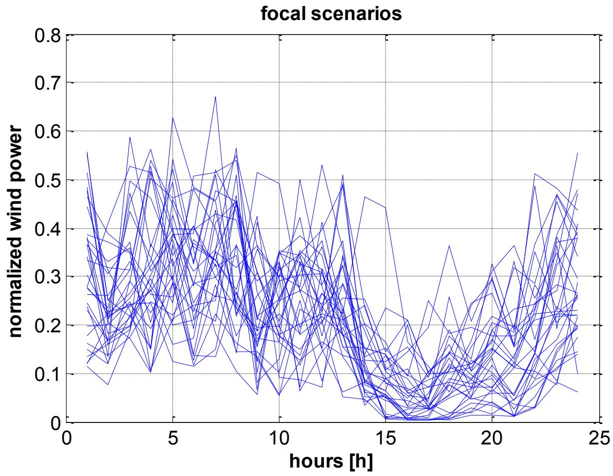

Figure 3. Focal scenarios, each representative of one cluster

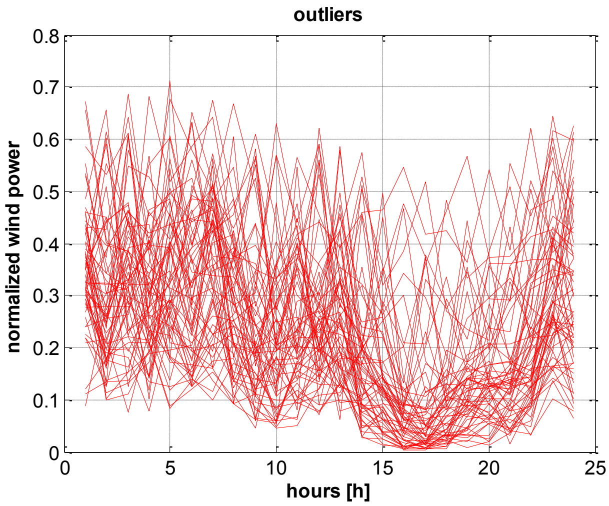

Figure 4. Outliers, i.e. scenarios that could not be clustered with any other, within the tolerance defined.

The probability associated to each focal scenario and outliers is shown in Figure 5, which presents also the cumulative probability associated to the clusters, when ordered by probability value from high to low value. From the initial 1000 scenarios, EPSO distilled 31 clusters grouping at least 2 scenarios and 64 outliers. The 31 focal scenarios contain about 93.6% of the initial set of scenarios while the outliers represent the remaining 6.4%. Using another metric for characterizing the similarity or dissimilarity among the original scenarios will lead to different cluster composition.

The results show how useful the method can be – it quantifies risk: one can keep, for further processing a selected number of focal scenarios and have information on the risk level accepted because of that choice: the risk of not including, for further consideration, a representative behavior or pattern of the wind power series.

For instance, in this example, by accepting all the 51 focal scenarios one would be having a confidence level of about 93.6% of having included a representation of the actual outcome, or running a 6.4% risk of having missed it.

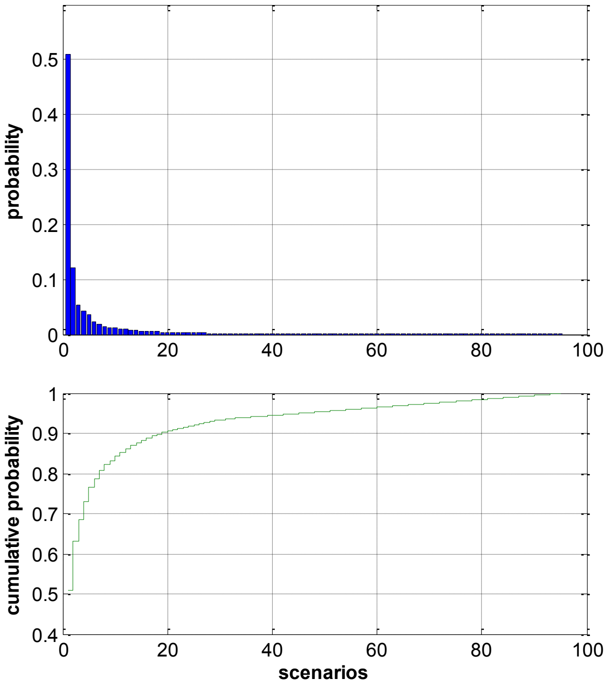

Figure 5. Probability associated with each focal scenario and with each outlier (when this set is ordered, left to right, from large to small probability, with value on the left axis) and cumulative probability.

4 VALIDATION PROBLEM

In order to validate the presented method, a simplified unit commitment problem was used. The unit commitment formulation applied here is based on [7] and represents a relaxed version including wind generators, presented in **Erro! A origem da referência não foi encontrada.**:[8]

$$\min \sum_{s=1}^{S} prob_s \cdot \left\{ \sum_{t=1}^{T} \sum_{j=1}^{J} c_j^{p,s}(t) + \sum_{t=1}^{T} C_{ENS} \cdot PNS_s(t) \right\} \quad (2)$$

subject to

$$c_j^{p,s}(t) = A_j v_j(t) + \sum_{l=1}^{L} C_{l,j} \delta_{l,j}^{s}(t) \quad \forall j, \forall t, \forall s \quad (3)$$

$$0 \le \delta_{l,j}^{s}(t) \le \Delta_{l,j} \quad \forall l, \forall j, \forall t, \forall s \quad (4)$$

$$p_j(t) = P_{m,j} \cdot v_j(t) + \sum_{l=1}^{L} \delta_{l,j}^{s}(t) \quad \forall j, \forall t, \forall s \quad (5)$$

$$P_{m,j} \cdot v_j(t) \le p_j(t) \le P_{M,j} \quad (6)$$

$$\|p_j(t) - p_j(t-1)\| \le P_{ramp,j} \quad \forall t \quad (7)$$

$$p_j(0) = P_{init,j} \quad (8)$$

$$\sum_{j=1}^{J} p_j(t) + P_{wind,s} + PNS_s(t) = P_D(t) \quad \forall t, \forall s \quad (9)$$

where

s, S scenario number and total number of scenarios, respectively
$prob_s$ probability associated to the scenario s
t, T time step and total number of time steps
j, J generator number and total number of generators
v_j binary variable defining whether the generator j is commited in time step t
$P_{m,j}, P_{M,j}$ minimum and maximum operating power for the generator j, respectively
$P_{init,j}$ initial operating power for the generator j (at the time step $t = 0$)
$p_j(t)$ total operating power of generator j in period t
$P_{ramp,j}$ maximum allowed change of operating power between subsequent periods for generator j
$c_j^{p,s}(t)$ operating costs for a generator j, for time step t and scenario $s \in S$
C_{ENS} costs of unserved energy
PNS_s(t) unserved power for the time step t and scenario s; multiplied by time step duration (1 hour) gives the unserved energy during that hour
$P_D(t)$ system load (demand) for period t
$P_{wind,s}(t)$ forecasted wind power in scenario s and period t

Due to the complex nature of the landscape of its objective function and because of the presence of integer variables, unit commitment problems represent a class of very difficult problems even for evolutionary meta-heuristics. In a full stochastic formulation, even if state of the art solvers are used, solving the full problem as presented in **Erro! A origem da referência não foi encontrada.**[8] with a complete large set of Monte Carlo generated scenarios rapidly becomes an infeasible task. Thus, in order to perform the comparison of the initial set of scenarios with the reduced one, a number of constraints have been eliminated from the complete formulation. Furthermore, wind curtailment was not considered in this formulation.

Table 1. Generator data

generator	a c/MW³	b c/MW²	c c/MW	P_m MW	P_M MW	P_{ramp} MW	P_{init} MW
G1	0.50	20	500	80	150	150	150
G2	0.25	40	500	20	100	50	100
G3	0.25	50	500	40	200	40	40
G4	0.50	35	500	0	100	50	50
G5	0.50	80	200	0	120	100	0

Table 1 gives the test system data. The generator costs are modeled as quadratic function of generator operating power:

$$c_j(p) = a_j p^2 + b_j p + c_j \quad (10)$$

The cost of unserved energy is set to $C_{ENS} = 5 \cdot 10^5$. A large value for this cost is chosen so that the algorithm results in using most of the available generator capacities.

The comparisons have been performed on a representative day using day-ahead predictions. i.e. the unit commitment time horizon was 24 hours. The problem was formulated as a mixed integer linear problem and solved using the ILOG CPLEX solver, using its Python interface to formulate it.

The formulation of total costs in a stochastic model represents the expected value for the objective function. To compare the performance of the stochastic model with a plain model using scenarios with no classification (sometimes wrongly denoted as "equally probable") a set of "blindly chosen" 95 samples was compared to 31 focal scenarios obtained by the clustering process plus the 64 outliers, so that the same number of scenarios is used.

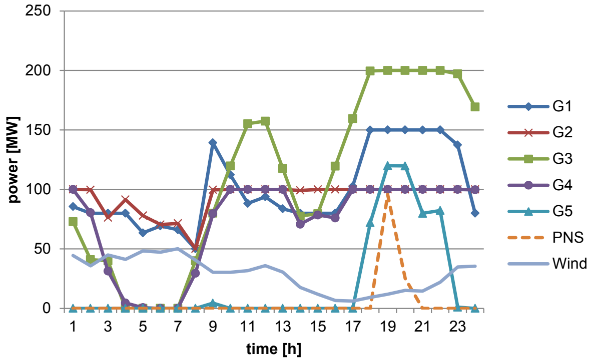

Figure 6. Unit commitment results for the representative day (expected values for the full scenario set – 1000 scenarios)

Two versions of the clustering algorithm have been tested: one where the focal scenario representing the cluster is chosen from the initial set of scenarios, and another one where the focal scenario is a cluster centroid obtained by averaging the cluster members.

Figure 6 illustrates the unit commitment results depicting the expected values of generator productions obtained by running the algorithm with the full set of 1000 scenarios.

For the three reduced sets, i.e. two clustered scenario sets and a set of blindly chosen 95 scenarios, a comparison is presented in Table 2. This table presents the maximum difference between expected hourly values of generator production:

$$\max \Delta p_j(t) = \left\| E(p_j(t)) - E(\, p_j^{reduced}(t)) \right\| \qquad (11)$$

Table 2. Maximum absolute difference between expected values of generator production of the complete set and the reduced sets

	Clustered set Existing scenario as centroid	Clustered set Average as centroid	Blind scenario selection
	MW	MW	MW
G1	8.44	4.25	16.61
G2	3.82	3.95	32.06
G3	3.91	4.00	40.00
G4	4.86	3.95	10.36
G5	11.22	3.40	6.40
PNS	8.71	0.12	0.70
Wind	11.22	0.00	3.91

The sum of the expected value differences $\sum_{t=1}^{T} \Delta p_j(t)$ is 215 MW for the clustering-based scenario reduction with the focal scenario chosen from the initial set, 63 MW for the clustering method with averaged centroids and 231 MW for the "blind" selection method. Both clustering techniques have superior performance compared to the "blind" scenario selection method.

The formulation of this particular UC problem is sensitive to the differences in expected value of wind power. If the focal scenario is chosen among the initial scenario set, then differences in expected values of wind power occur. Thus, it is not surprising that the clustering method with representative scenarios obtained as cluster averages performs better than the clustering where existent scenarios are chosen as focal scenarios. However, this observation may not be valid for a different problem.

To summarize, the use of a reduced set of scenarios allows introducing more realistic restrictions in the unit commitment formulation, while keeping the realistic shape of the wind forecast uncertainty and reducing the computational effort.

In order to evaluate to computational effort reduction, CPU time variation with respect to the number of scenarios is reported in Table 3.

Table 3. Evaluation of the computational effort

Number of scenarios	1000	500	200	100	50
CPU time [s]	80.8	30.6	11.4	4.8	2.8

Notice that reducing the number of scenarios by 10 from 1000 to 100 scenarios, the CPU time has been divided by about 16. In a bigger problem, e.g. with a larger number of generators the reduction can be even more. These results have been achieved using a Core i7-640M processor.

5 CONCLUSION

The representation of the uncertainty in wind power forecasting may be accomplished by a sampling procedure of the Monte Carlo type. However, to have a sufficiently accurate representation of the probability density function for the forecast, with adequate coverage of the forecasting space, one needs to sample and generate an extremely large set of elementary scenarios.

The size of this set is incompatible with its use as input in a stochastic programming algorithm in an application on optimal unit commitment, because of the excessive computing effort that it would demand.

The solution offered in this paper can replace the large set of elementary scenarios by a smaller set of clusters, each cluster being replaced by a representative member (focal scenario) associated to the probability of the subset it represents.

Therefore, from an initial set of scenarios it is possible to reduce the number of scenarios to representative scenarios characterized by different probabilities. The reduction capability depends on the admitted error, the metric used to define it and the scenario dispersion. The choice of the metric and the admissible error has to be defined in accordance with the problem and the level of accuracy desired by the decision-maker.

The number of focal scenarios to be retained becomes associated to a risk index, which will be useful when assessing risks and the need for hedging.

This new method gives a strong contribution to promote the use, by system operators, of stochastic optimization models when planning the operation of their systems, by allowing this problem to become tailored to the convenient size that computation resources will allow, while keeping control of the risks of misrepresentation incurred. In the future work, a more detailed evaluation of scenario reduction performance in a number of applications is envisioned.

APPENDIX

Short reference to the EPSO method

The optimization method used in the discovery of the scenarios with maximum aggregation power was EPSO. for Evolutionary Particle Swarm Optimization.

EPSO is a hybrid in concepts of EA and PSO, first proposed in [3] and with an improved version in [9] and [10]. It is an Evolutionary Algorithm with an adaptive recombination operator inspired in the "movement rule" of PSO (Particle Swarm Optimization). This rule generates a new individual (also denoted as particle) in the population or swarm as a weighted combination of parents, which are: a given individual, its most recent ancestor, its best ancestor during a given period of the history of the particle and the best ancestor of the present generation of the population, also during a given period of the evolutionary history of the swarm. This may be seen as a form of intermediary recombination, by which 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, therefore forming a self-adaptive recombination operator.

The recombination rule for EPSO is the following: given a particle X_i, a new particle X_i^{new} results from

$$X_i^{(k+1)} = X_i^{(k)} + V_i^{(k+1)} \quad \text{(A1)}$$

$$V_i^{(k+1)} = w_{i1}^* V_i^{(k)} + w_{i2}^*(b_i - X_i) + w_{i3}^* P\left(b_g^* - X_i\right) \quad \text{(A2)}$$

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

b_i best point found by the line of ancestors of individual i up to the current generation;

b_g best overall point found by the swarm of particle in its past life up to the current generation;

$b_g^* = b_g + w_{i4}^* N(0,1) \Rightarrow$ particle in the neighborhood of b_g

$X_i^{(k)}$ location of particle i at generation k,

$V_i^{(k)} = X_i^{(k)} - X_i^{(k-1)} \Rightarrow$ "velocity" of 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;

P is 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 [10], although highly complex problems seem to require a very low non-zero value such as p < 0.2.

Weights w_{ik}^* are mutated in each iteration according to $w_{ik}^* = w_{ik}[logN(0,1)]^\tau$, k = 1,…,3
and $w_{i4}^* = w_{i4} + \sigma N(0,1)$
where

$N(0,1)$ is a random variable that follows a Gaussian distribution with zero mean and unit variance

$logN(0,1)$ is a random variable that follows a Lognormal distribution associated with a Gaussian $N(0,1)$

τ and σ are externally fixed learning parameters that control the amplitude of mutations.

The EPSO algorithm ensures that each particle generates a number of offspring with mutated weights using (A1) and (A2). Among the descendants of each particle, one is selected to integrate a new generation of the population. Any selection operator may be applied but most practical cases have adopted an offspring of 2 and elitism.

ACKNOWLEDGEMENTS

This paper was supported by UChicago Argonne, LLC, under Contract No. DE-AC02-06CH11357 with the U.S. Department of Energy.

J. Sumaili and H. Keko are funded by Fundação para a Ciência e a Tecnologia (FCT) within "Ciencia 2008" program and PhD Scholarship SFRH/BD/43087/2008, respectively.

REFERENCES

[1] E. Canestrelli, S. Giove, "Scenarios Identification for Financial Time Series", in Current Topics in quantitative Finance, E. Canestrelli editor, Springer, 1999

[2] P. Pinson. H. Madsen. H.A. Nielsen. Papaefthymiou and B. Klöckl, "From Probabilistic Forecasts to Statistical Scenarios of Short-term Wind Power Production", Wind Energy 2009, 12, pp. 51 – 62, 2009

[3] V. Miranda and N. Fonseca, “EPSO – Best-of-two-worlds metaheuristic applied to power system problems” in Proceedings of the World Congress on Computational Intelligence (WCCI 2002), CEC – Conference on Evolutionary Computing, Honolulu, (Hawaii), USA, May 2002.

[4] J. Dupačová, N. Gröwe-Kuska, and W. Römisch, “Scenario Reduction in Stochastic Programming: An Approach Using Probability Metrics,” Math. Program, Ser. A, vol. 95, pp. 493-511, 2003.

[5] N. Growe-Kuska, H. Heitsch, and W. Romisch, “Scenario reduction and scenario tree construction for power management problems,” in Power Tech Conference Proceedings, 2003 IEEE Bologna, vol. 3, Vol.3, 2003.

[6] Eastern Wind Integration and Transmission Study (EWITS), National Renewable Energy Laboratory (NREL). http://www.nrel.gov/wind/systemsintegration/ewits.html.

[7] M. Carrion and J. Arroyo, “A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem,” Power Systems, IEEE Transactions on, vol. 21, no. 3, pp. 1371-1378, 2006.

[8] J. Wang, A. Botterud, R. Bessa, H. Keko, L. Carvalho, D. Issicaba, J. Sumaili, V. Miranda, “Representing Wind Power Forecasting Uncertainty in Unit Commitment,” Applied Energy, 2011, in press (available online),.

[9] 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.

[10] V. Miranda. H. Keko and A. Jaramillo, “Stochastic Star Communication Topology in Evolutionary Particle Swarms (EPSO)”, International Journal of Computational Intelligence Research, vol. 4, no.2, 2008.