ISSN 1870-4069

Optimization of Energy Production in Oklahoma: A Variant of an Evolutionary Algorithm of Multi-Objective Optimization

Gustavo Adolfo Vargas Hakim, Efrén Mezura Montes

University of Veracruz,
Artificial Intelligence Research Center,
Mexico

vargashakimg@gmail.com, emezura@uv.mx

Abstract. A preliminary performance assessment of the well-known Non-Dominated Sorting Genetic Algorithm II and one of its variants to optimize the electrical power production in the state of Oklahoma is presented in this paper. Such variant has a chaotic model to generate the initial population. Solar, wind and natural gas power systems, the first two renewable energies, are considered in the problem of interest. Three conflicting objectives are optimized: (1) power production, (2) production costs and (3) CO_2 emissions. The spacing metric is computed to compare the performance of both variants. The obtained results suggest that the chaotic model for the initial population does not improve the performance of the original algorithm in this particular multi-objective optimization problem.

Keywords: evolutionary multi-objective optimization, energy production, renewable energies.

1 Introduction

Climate change is a very relevant problem that human kind is facing. The United Nations has warned that approximately a dozen years are left to limit climate change at 1.5 °C in order to prevent a world crisis [1]. Power production is one of the key elements to consider. Fossil fuels have been and continue to be widely used to produce electricity. These energy sources, when burned to generate heat to produce power, spreads nitrogen oxides and other contaminants that contribute to the smog and acid rain [2].

Several countries, including the United States, have started to adopt renewable energy sources, such as solar and wind energy, to produce power. In 2017, around the 11% of the total consumed energy in the US came from renewable energies [3]. With respect to the state of Oklahoma, according to the US department of energy, 50% of the annual energy production in 2016 comes from natural gas, 38% from coal, 10% comes from solar and wind power and 1% from hydro-energy [4]. Despite the US government position about climate change, as mentioned in [5], the renewable energy industry is expected to grow and play an important role in the energy production in some of the most populated states of the country.

Beyond the merely direct utilization of these power sources, an optimal distribution of the generation power between them is highly desirable. To optimize the power production problem different multi-objective evolutionary algorithms (MOEAs) have been adopted. The algorithm of interest in this paper is the Non-Dominated Sorting Genetic Algorithm II (NSGA-II) [6], which has been a popular choice to solve multi-objective power optimization problems. Wahlroos *et al.* [7] optimized a generation system in terms of CO_2 emissions, production costs and production adequacy, using NSGA-II. Wang and Zhou [8] utilized the same algorithm to optimize the emissions and energy-savings of a wind power system. Liu and Dongdong [9] optimized a multiple source power system considering its production cost and the amount of emissions it produces. Zhou and Sun [10] utilized this MOEA to optimize a hybrid energy system consisting of solar power and wind power. In this study, the results were compared with a modified version of this algorithm, called L-NSGA-II [11].

Motivated by the above mentioned, and particularly by [10], a multi-objective optimization problem is stated for the power production of Oklahoma and solved using NSGA-II and one of its variants (L-NSGA-II) with the aim to assess their performance in this new real-world instance.

The paper is organized as follows: Section 2 states the problem of interest, Section 3 details both, NSGA-II and the variant adopted in this work. After that, Section 4 includes the experiments and results and Section 5 presents the conclusions and future work.

2 Problem Statement

A multi-objective optimization problem can be defined, without loss of generality, as to: find a solution vector $\vec{x} = [x_1, x_2, \ldots, x_n]^T$, which minimizes $\vec{f}(\vec{x}) = [f_1(\vec{x}), f_2(\vec{x}), \ldots, f_m(\vec{x})]^T$, where each $x_j \in [lo_j, up_j]$.

Pareto dominance is used as a criterion to solve multi-objective optimization problems, and it is defined as follows: a solution vector $\vec{x} = [x_1, x_2, \ldots, x_n]^T$ is said to dominate $\vec{y} = [y_1, y_2, \ldots, y_n]^T$, denoted as $\vec{x} \prec \vec{y}$ if and only if $f_i(\vec{x}) \leq f_i(\vec{y})$ for all $i \in [1, \ldots, m]$ and $f_i(\vec{x}) < f_i(\vec{y})$ for at least one $i \in [1, \ldots, m]$.

A solution vector $\overrightarrow{x^*}$ is part of the Pareto Optimal Set P^* if there does not exist other solution $\vec{x}$ such that $\vec{x} \prec \overrightarrow{x^*}$. The Pareto Optimal front is then $PF^* = \{\vec{f}(\vec{x}) | \vec{x} \in P^*\}$.

The analyzed power production system, as previously mentioned, is based on solar, wind and natural gas energy. The three objectives to be considered are: (1) the overall power production, (2) the overall production cost, and (3) the overall CO_2 emission. The decision variables are the operation hours of the solar, wind and natural gas production systems, *hs*, *hw* and *hg*, respectively, and the natural gas power produced *Pg*, which is assumed without considering the ways to produce it.

2.1 Power Production Function

This objective function represents the entire, monthly amount of power produced by the three sources. This objective function is constructed as shown in Eq. 1:

$$P = f_s \boldsymbol{h_s} \overline{P_s} + f_w \boldsymbol{h_w} \overline{P_w} + f_g \boldsymbol{h_g} \boldsymbol{P_g}. \tag{1}$$

Decision variables are bolded. $\overline{P_s}$ and $\overline{P_w}$ are the estimation of the generated solar and wind power for any arbitrary month (explained later). P_g is the generated natural gas power, considered as decision variable (assuming it could be controlled). Constants f_s, f_w and f_g are the capacity factors of solar, wind and natural gas systems, respectively (see Table 1).

Table 1. Capacity factors (percentage of input power that is effectively transformed into electrical power) of the three studied energy sources.

Power Source	Capacity Factor (%)
Solar Power (f_s)	33
Wind Power (f_w)	43
Natural Gas Power (f_g)	87

To get all three objective functions to be minimized, this objective function, originally to be maximized, is transformed by using the concept of power relation as in Eq. 2:

$$PR_T = \frac{N * P_{SWG}}{f_s \boldsymbol{h_s} \overline{P_s} + f_w \boldsymbol{h_w} \overline{P_w} + f_g \boldsymbol{h_g} \boldsymbol{P_g}}, \tag{2}$$

where P_{SWG} is the current amount of energy produced by the three sources together. This value is approximately 7.07 MWh and it was measured for the month of August, 2018 [12]. N is an increasing factor. In an ideal context, where renewable power is reinforced, the increasing factor is greater than one, so that the total energy from solar, wind and natural gas power is increased. For this study, this total energy production is encouraged to be doubled, then $N = 2$.

2.2 Production Cost Function

The second objective function is the total, monthly production cost. Eq. 3 shows this objective to be minimized:

$$C = c_s \boldsymbol{h_s} \overline{P_s} + c_w \boldsymbol{h_w} \overline{P_w} + c_g \boldsymbol{h_g} \boldsymbol{P_g}, \tag{3}$$

where constants c_s, c_w and c_g are the production costs of solar, wind and natural gas systems, respectively (see Table 2).

Table 2. Production cost of solar, wind, and natural gas power systems.

Power Source	**Production Cost ($/MWh)**
Solar Power (c_s)	48.2
Wind Power (c_w)	33
Natural Gas Power (c_g)	15.5

Data from Tables 1 and 2 are approximations obtained from [13].

2.3 CO_2 Emissions Function

The third objective function is the monthly CO_2 emissions caused by the three energy sources. This objective function to be minimized is introduced in Eq. 4:

$$E = e_s \boldsymbol{h_s} \overline{P_s} + e_w \boldsymbol{h_w} \overline{P_w} + e_g \boldsymbol{h_g} \boldsymbol{P_g}, \tag{4}$$

where the CO_2 emissions rates for the solar, wind and natural gas systems are e_s, e_w and e_g, respectively (see Table 3).

Table 3. CO_2 emissions rates (related with global warming [14]) for solar, wind, and natural gas power systems.

Power Source	**CO_2 Emissions Rate (gr/KWh)**
Solar Power (e_s)	48.2
Wind Power (e_w)	33
Natural Gas Power (e_g)	15.5

The estimations of solar and wind power $\overline{P_s}$ and $\overline{P_w}$, depend on the solar radiation and wind speed, so climate predictions are required. Climate is hard to predict. Moura and de Almeida proposed a climate model prediction for Portugal, based on previous data [15]. A similar model is used for the state of Oklahoma but using applicable data. A dataset was built in Microsoft Excel with measurements of solar radiation and wind speeds for each month of each year from 2003 to 2017, provided by MESONET, an environmental monitoring station available in the state of Oklahoma [16].

The solar radiation measured by MESONET is in MJ/m^2. It is multiplied by the total solar panels area and divided by the total number of seconds in each month. Months with 30, 31 and 28 days were considered. The result unit is the Watt (W). The wind speed is measured in miles per hour and only needs to be converted to m/s.

3 NSGA-II and L-NSGA-II

NSGA-II is a genetic algorithm (GA) adapted to solve multi-objective optimization problems. Besides those GA elements (tournament selection and crossover and mutation operators), NSGA-II uses the so-called non-dominated sorting process to rank solutions based on Pareto dominance from the union of parents and offspring.

Those non-dominated solutions get rank 1 and they are separated from the aforementioned union.

Begin
  Generate an initial population *Pop* with **POPSIZE** solutions
  Evaluate each solution in the objective functions
  Apply non-dominated sorting to *Pop*
  While termination condition not reached
    Generate offspring population *Offs*
    Evaluate each offspring in the objective functions
    Apply non-dominated sorting to *Pop U Offs*
    Select, based on ranking and crowding distance, the next population from *Pop U Offs*
  End While
End

Fig. 1. NSGA-II general pseudocode.

From the remaining solutions those non-dominated are assigned rank 2 and so on. The next population is chosen based on ranking. Furthermore, a crowding-distance measure in the objective space is used to choose among solutions with the same ranking to get a population with the same size to start the next generation. A general pseudocode is presented in Figure 1.

L-NSGA-II keeps most of the original NSGA-II structure, except for adding a different population initialization method. Here, a hybrid chaotic model is defined for the initialization part. The usual initialization method is shown in Eq. 5:

$$x_j = lo_j + (up_j - lo_j)u, \tag{5}$$

where u is a random number with uniform distribution; lo_j and up_j are the variable boundaries. In L-NSGA-II, the random number u is substituted. First, for the current value of a counter k, there are two randomly found numbers (between 0 and 1) defined as part of the hybrid chaotic mapping model, u_k and r_k. The value of u for the next count $k + 1$ is defined in Eq. 6:

$$u_{k+1} = \mu u_k(1 - u_k), \tag{6}$$

where μ is a control variable, set as 0.5 for this study. The value of r for the step $k + 1$ is calculated depending on the value of r_k. This is displayed in Eq. 7:

$$r_{k+1} = \begin{cases} \dfrac{1}{1.001}(2r_k + 0.001u_k), & 0 \leq r_k \leq 0.5, \\ \dfrac{1}{1.001}[2(1 - r_k) + 0.001u_k], & 0.5 < r_k \leq 1. \end{cases} \tag{7}$$

Then, the initialization of a single variable of the next step $k + 1$ is as in Eq. 8:

$$x_j = lo_j + \left(up_j - lo_j\right)r_{k+1}. \tag{8}$$

According to the authors of this initialization proposal, this initialization model should contribute to the diversity of solutions in the Pareto front. Diversity is highly desired as it provides for more options to choose between advantages and disadvantages of each possible solution. It remains to be seen if this applies for the problem of interest in this paper.

4 Experiments and Results

The optimization of the power production was conducted following the most accurate conditions that were possible, i.e., actual wind farms (412 wind turbines) and solar panels (20,000) that are currently in use in Oklahoma were considered. The decision variables were constrained due to the real limitations of the power production systems and the problem requirements. The hours of operation could not exceed the number of hours in a month. The maximum number of days considered was 28 (as February is the shortest month), equivalent to 672 hours. The minimum number of hours allowed was 240 hours. The maximum produced natural gas power was 7.07 MW and its minimum produced power was 4.88 MW. The boundaries of the decision variables are summarized in Table 4.

Table 4. Boundaries of the decision variables.

	h_s	h_w	h_g	P_g
Maximum	672 hr	672 hr	672 hr	7.07 MW
Minimum	240 hr	240 hr	240 hr	4.88 MW

The simulations of power production were executed using MATLAB using an Intel Core i7 processor. A previously constructed NSGA-II framework by Seshadri was utilized and modified to run the problem objectives and also to code L-NSGA-II [17].

Both algorithms were tested using 100 generations and an initial population of 20 individuals. Crossover and mutation were performed with 90% and 10% of probabilities, respectively, in both cases.

Yen and He defined several metrics to test the performance of MOEAs [18]. In this work the Spacing metric was chosen as it measures how diverse or well distributed are the solutions in a Pareto front. Eq. 9 describes the spacing metric:

$$S = \sqrt{\frac{1}{\bar{n}}\sum_{i=1}^{\bar{n}}(d_i - \bar{d})^2}, \tag{9}$$

where d_i is the Euclidean distance between a solution x_i and its nearest solution, $\bar{n}$ is the number of solutions in the Pareto front and $\bar{d}$ is the average Euclidean distance between solutions. A lower value indicates a better solution distribution.

For this paper, the experiments covered the month of May, 2022 and each algorithm was run 25 times. The spacing value was calculated for each one of the 25 fronts obtained per variant and the averages per variant are shown in Table 5.

Table 5. Average spacing for the Pareto Fronts of each variant for May, 2022.

Algorithm	Average Spacing
NSGA-II	6.87×10^8
L-NSGA-II	6.77×10^8

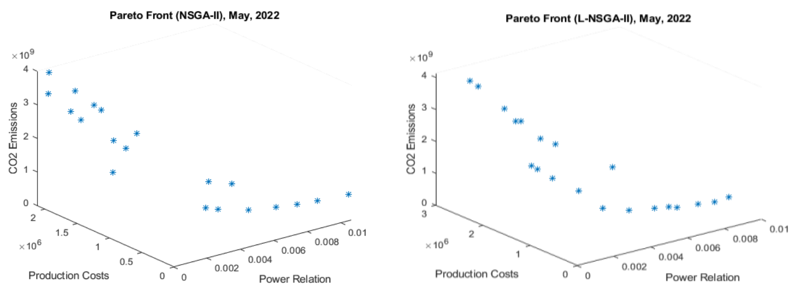

Fig. 2. Pareto front using NSGA-II (left) and using L-NSGA-II (right).

Considering the fact that none of the results samples fit the normal distribution (based on the Kolmogorov-Smirnov test), the Wilcoxon Signed-Rank test was computed, and its result ($p - value = 0.672$) indicated no significant differences between the two compared algorithms. Fig. 2 shows two Pareto fronts from both algorithms and a slightly better distribution provided by L-NSGA-II can be observed.

The decision-making of the best solution from the obtained Pareto front was based on the third objective function, i.e., CO_2 emissions. Such decision was based on the idea of getting the most environmentally friendly power production option. From the Pareto front in Fig. 2 (right), the individual that produced the lowest emissions was chosen (h_s = *361.06 h*, h_w = *419.79 h*, h_g = *251.6 h* and P_g = *6.837 MW*), which corresponds to the L-NSGA-II front. It is worth noticing that the operation hours number (h_g = *251.6 h*) of the natural gas system was the lowest of the three sources, reducing the CO_2 emissions.

5 Conclusions and Future Work

A comparison of a variant of the popular NSGA-II algorithm with the original to solve one instance of the power production in Oklahoma was presented in this paper. The variant was L-NSGA-II, which had a chaotic model to generate the initial population. The spacing metric was used for comparison purposes between the two algorithms when solving one instance of the problem related with one month (May 2022). The statistical results obtained suggested that the way L-NSGA-II generates the initial population does not produce any significantly distribution improvement in the Pareto front. The preference handling also showed that, when preferring the objective related to CO_2 emissions, an environmentally friendly solution can be obtained.

Future work consists on testing both variants in more problem instances and using other metrics like hyper-volume. Moreover, preference handling can be used within the search to focus only on well-distributed solutions favoring low CO_2 emissions.

Appendix

The codes used and the climate database constructed for this paper can be found and downloaded following the link below: http://drive.google.com/drive/folders/1eNfMthw7v-i-gK_UkQx_zW6Wk3zFgt6J?usp=sharing

References

1. The Guardian: https://www.theguardian.com/environment/2018/oct/08/global-warming-must-not-exceed-15c-warns-landmark-un-report. (2018)
2. US EPA: https://www.epa.gov/nutrientpollution/sources-and-solutions-fossil-fuels. (2019)
3. US EIA: https://www.eia.gov/energyexplained/?page=renewable_home. (2019)
4. US Department of Energy: https://www.energy.gov/sites/prod/files/2016/09/f33/OK_Energy%20Sector%20Risk%20Profile.pdf. (2019)
5. Rödl & Partner: https://www.roedl.com/insights/erneuerbare-energien/2017-05/renewable-energy-trump-administration. (2019)
6. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A Fast and Elitist Multi-objective Genetic Algorithm; NSGA-II. IEEE Transactions on Evolutionary Computation. 6(2), pp. 182–197. (2002)
7. Wahlroos, M., Jääskeläinen, J., Hirvonen, J.: Optimisation of an Energy System in Finland using NSGA-II Evolutionary Algorithm. 2018 15th International Conference on the European Energy Market (EEM). (2018)
8. Wang, J., Zhou, Y.: Multi-objective Dynamic Unit Commitment Optimization for Energy-saving and Emission Reduction with Wind Power. 2015 5th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT). (2015)
9. Liu, T., Dongdong, Z..: Multi-Objective Optimal Calculation for Integrated Local Area Energy System Based on NSGA-II Algorithm. 2019 IEEE International Conference on Energy Internet (ICEI). (2019)
10. Zhou, T., Sun, W.: Optimization of Wind-PV Hybrid Power System based on Interactive Multi-objective Optimization Algorithm. 2012 International Conference on Measurement, Information and Control. (2012)
11. Liu, T., Gao, X., Wang, L.: Multi-objective Optimization Method using and improved NSGA-II algorithm for oil-gas production process. Journal of the Taiwan Institute of Chemical Engineers, 57, pp. 42–53. (2015)
12. US EIA: https://www.eia.gov/state/?sid=OK#tabs-4. (2019)
13. IRENA: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2018/Jan/IRENA_2017_Power_Costs_2018.pdf. (2019)
14. NASA: https://climate.nasa.gov/vital-signs/carbon-dioxide/. (2019)

15. Moura, P., De Almeida, A.: Multi-Objective Optimization of a Mixed Renewable System with Demand-Side Management. Renewable and Sustainable Energy Reviews. 14(6), pp. 1461–1468. (2010)
16. MESONET: https://www.mesonet.org/index.php/weather/mesonet_averages_maps #y=average&m=12&p=wspd_mx&d=false. (2019)
17. Seshadri, A.: https://www.mathworks.com/matlabcentral/fileexchange/10429-nsga-ii-a-multi-objective-optimization-algorithm. (2019)
18. Yen, G., He, Z.: Performance Metrics Ensemble for Multi-Objective Evolutionary Algorithms. IEEE Transactions on Evolutionary Computation, 18(1), pp. 131–144. (2013)